Algorithms for Intelligent Systems
Series Editors: Jagdish Chand Bansal · Kusum Deep · Atulya K. Nagar

Geetam Singh Tomar
Narendra S. Chaudhari
Jorge Luis V. Barbosa
Mahesh Kumar Aghwariya *Editors*

International Conference on Intelligent Computing and Smart Communication 2019

Proceedings of ICSC 2019

Springer

Algorithms for Intelligent Systems

Series Editors

Jagdish Chand Bansal, Department of Mathematics, South Asian University, New Delhi, Delhi, India
Kusum Deep, Department of Mathematics, Indian Institute of Technology Roorkee, Roorkee, Uttarakhand, India
Atulya K. Nagar, Department of Mathematics and Computer Science, Liverpool Hope University, Liverpool, UK

This book series publishes research on the analysis and development of algorithms for intelligent systems with their applications to various real world problems. It covers research related to autonomous agents, multi-agent systems, behavioral modeling, reinforcement learning, game theory, mechanism design, machine learning, meta-heuristic search, optimization, planning and scheduling, artificial neural networks, evolutionary computation, swarm intelligence and other algorithms for intelligent systems.

The book series includes recent advancements, modification and applications of the artificial neural networks, evolutionary computation, swarm intelligence, artificial immune systems, fuzzy system, autonomous and multi agent systems, machine learning and other intelligent systems related areas. The material will be beneficial for the graduate students, post-graduate students as well as the researchers who want a broader view of advances in algorithms for intelligent systems. The contents will also be useful to the researchers from other fields who have no knowledge of the power of intelligent systems, e.g. the researchers in the field of bioinformatics, biochemists, mechanical and chemical engineers, economists, musicians and medical practitioners.

The series publishes monographs, edited volumes, advanced textbooks and selected proceedings.

More information about this series at http://www.springer.com/series/16171

Geetam Singh Tomar · Narendra S. Chaudhari ·
Jorge Luis V. Barbosa ·
Mahesh Kumar Aghwariya
Editors

International Conference on Intelligent Computing and Smart Communication 2019

Proceedings of ICSC 2019

Springer

Editors
Geetam Singh Tomar
Machine Intelligence Research Labs
Gwalior, India

THDC-Institute of Hydropower Engineering and Technology
Uttarakhand, India

Jorge Luis V. Barbosa
Applied Computing Graduate Program
University of Vale do Rio dos Sinos
Sao Leopoldo, Rio Grande do Sul, Brazil

Narendra S. Chaudhari
Indian Institute of Technology Indore
Indore, Madhya Pradesh, India

Uttarakhand Technical University
Uttarakhand, India

Mahesh Kumar Aghwariya
Department of Electronics and Communication Engineering
THDC-Institute of Hydropower Engineering and Technology
Tehri, Uttarakhand, India

ISSN 2524-7565 ISSN 2524-7573 (electronic)
Algorithms for Intelligent Systems
ISBN 978-981-15-0632-1 ISBN 978-981-15-0633-8 (eBook)
https://doi.org/10.1007/978-981-15-0633-8

© Springer Nature Singapore Pte Ltd. 2020
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Preface

ICSC stands for Intelligent Computing and Smart Communication. It is the first international conference of the series ICSC, jointly organized by THDC Institute of Hydropower Engineering and Technology, New Tehri, Uttarakhand, India, and School of Engineering, Cochin University of Science and Technology, Cochin, Kerala, India, during April 19–21, 2019, at THDC-IHET, Tehri.

The main focus of ICSC 2019 is to provide an opportunity for researchers to meet and discuss the latest solutions, scientific results, and methods in solving intriguing problems in various fields of computing and communication along with real-world applications. ICSC 2019 attracted a wide spectrum of thought-provoking research papers on various aspects of Intelligent Computing and Smart Communication with umpteen applications, theories, and techniques. A total of 164 quality research papers are selected for publication after the peer review process in the form of proceedings of ICSC 2019.

We are sure that the research findings in the novel papers contained in this proceeding will be much fruitful and may inspire more and more researchers to work in various fields of Intelligent Computing and Smart Communication. The topics that are presented in this proceedings are fuzzy logic and fuzzy controller, artificial neural network, machine learning, different optimization techniques, Big Data analysis, networks and cyber security, information security, grid computing, data mining and clustering, computational seismology, robotics and automation, embedded system design, digital and analog communications, RF and wireless communication, VLSI technologies, signal and image processing, remote sensing, wireless sensor networks, luminescence, designing and analysis using IoT, communications and information security, MIC, MMIC, MEMS/NEMS devices, and optoelectronics. Therefore, this proceedings must provide an excellent platform to explore the assorted various computational and communication techniques.

The editors would like to express their sincere gratitude to general chairs, plenary speakers, invited speakers, reviewers, Technical Programme Committee members, International Advisory Committee members, and Local Organizing Committee members of ICSC 2019; without whose support, the quality and

standards of the conference could not be maintained. Special thanks to the Springer and its team for this valuable publication.

Over and above, we would like to express our deepest sense of gratitude to THDC-IHET, Tehri, for hosting this conference. Also, sincere thanks to all the sponsors of ICSC 2019.

Gwalior/Uttarakhand, India — Prof. Geetam Singh Tomar
Indore/Uttarakhand, India — Prof. Narendra S. Chaudhari
Sao Leopoldo, Brazil — Prof. Jorge Luis V. Barbosa
Tehri, India — Mahesh Kumar Aghwariya

About This Book

The proceedings of ICSC 2019 will serve as an academic bonanza for scientists and researchers working in the field of computing and communication. This book contains theoretical as well as practical aspects using fuzzy logic, neural networks, VLSI design, microwave engineering, embedded circuit design swarm intelligence algorithms, etc., with many applications under the umbrella of computational and communication engineering. This book is beneficial for the young as well as experienced researchers dealing across complex and intricate real-world problems for which finding a solution by traditional methods is a difficult task.

The different application areas covered in the proceedings are fuzzy logic and fuzzy controller, artificial neural network, machine learning, different optimization techniques, Big Data analysis, networks and cyber security, information security, grid computing, data mining and clustering, computational seismology, robotics and automation, embedded system design, digital and analog communications, RF and wireless communication, VLSI technologies, signal and image processing, remote sensing, wireless sensor networks, luminescence, designing using IoT, communications and information security, MIC, MMIC, MEMS/NEMS devices, optoelectronics, etc. This will surely helpful for the researchers/scientists working in similar fields of optimization.

Contents

A Survey of Fuzzy Logic Inference System and Other Computing Techniques for Agricultural Diseases 1
Bhavna Chilwal and P. K. Mishra

Credit Card Fraud Detection Using Correlation-Based Feature Extraction and Ensemble of Learners 7
I. Sumaiya Thaseen and K. Lavanya

A Model for Predicting Occurrence of Leaf Blast Disease in Rice Crop by Using Fuzzy Logic Techniques 19
Bhavna Chilwal and P. K. Mishra

Improving the Effectiveness of Moving Target Defenses by Amplifying Randomization 27
Vaishali Kansal and Mayank Dave

Remaining Life Assessment of Solid Insulation in Power Transformer Using Fuzzy Inference System (FIS) 37
Deepak Kanumuri, Veena Sharma and O. P. Rahi

Comparison of Fuzzy Logic Based MPPT of Grid-Connected Solar PV System with Different MPPT 47
Ashutosh Tiwari and Ashwani Kumar

Comparative Study of Different Classification Models on Benchmark Dataset of Handwritten Meitei Mayek Characters 61
Deena Hijam and Sarat Saharia

A Literature Review on Energy-Efficient Routing Protocols for Heterogeneous WSNs 73
Isha Pant and S. K. Verma

Location-Based Proactive Handoff Mechanism in Mobile Ad Hoc Network 85
D. Kalyani, Somula Ramasubbareddy, K. Govinda and V. Kumar

PSO-Based Improved DV-Hop Localization Algorithm in Wireless Sensor Networks 95
Gaurav Sharma and Arunacharam Rajesh

Countermeasures Against Variants of Wormhole in Wireless Sensor Networks: A Review 105
Manish Patel, Akshai Aggarwal and Nirbhay Chaubey

Adaptive Backup Power Management in Ad Hoc Wireless Network 115
Ganesh Gupta, Vivek Jaglan and Ashok K. Raghav

Spatial Correlation Based Outlier Detection in Clustered Wireless Sensor Network 127
Robin Kamboj and Vrinda Gupta

Guidelines for an Effective Network Forensic System 137
Rajni Ranjan Singh and Deepak Singh Tomar

Handling Incomplete and Delayed Information Using Optimal Scheduling of Big Data Stream 147
Ravi Kishan Surapaneni, Sailaja Nimmagadda and Roja Rani Govada

Twitter Sentimental Analytics Using Hive and Flume 159
Rupesh Kumar Mishra, Suman Lata and Soni Kumari

KKG-512: A New Approach for Kryptos Key Generation of Size 512 Bits Using Plaintext 167
Kamal Kumar Gola, Gulista Khan, Ashish Joshi and Rahul Rathore

HD-MAABE: Hierarchical Distributed Multi-Authority Attribute Based Encryption for Enabling Open Access to Shared Organizational Data 183
Reetu Gupta, Priyesh Kanungo and Nirmal Dagdee

A Unified Platform for Crisis Mapping Using Web Enabled Crowdsourcing Powered by Knowledge Management 195
A. Vijaya Krishna, Somula Ramasubbareddy and K. Govinda

Web Image Authentication Using Embedding Invisible Watermarking 207
T. Aditya Sai Srinivas, Somula Ramasubbareddy, K. Govinda and S. S. Manivannan

Frequent Item Set, Sequential Pattern Mining and Sequence Prediction: Structures and Algorithms 219
Soumonos Mukherjee and R. Rajkumar

Study of AdaBoost and Gradient Boosting Algorithms for Predictive Analytics 235
Pritika Bahad and Preeti Saxena

Enhancing Privacy and Security in Medical Information with AES and DES 245
Nikhil Khandare, Omkar Dalvi, Valmik Nikam and Anala Pandit

A Comprehensive Review on Unsupervised Feature Selection Algorithms 255
Anala A. Pandit, Bhakti Pimpale and Shiksha Dubey

On the Prerequisite of Coprimes in Double Hashing 267
Vivek Kumar

Multilingual Machine Translation Generic Framework with Sanskrit Language as Interlingua 271
Promila Bahadur

Machine Learning Approach for Crop Yield Prediction Emphasis on K-Medoid Clustering and Preprocessing 287
Huma Khan and S. M. Ghosh

Combined Approach to Classify Human Emotions Based on the Hand Gesture 301
Samta Jain Goyal, Arvind Kumar Upadhyay and Rakesh Singh Jadon

Detecting Gene Modules Using a Subspace Extraction Technique 311
Pooja Sharma, D. K. Bhattacharyya and Jugal K Kalita

A Web Portal to Calculate Codon Adaptation Index (CAI) with Organism Specific Reference Set of High Expression Genes for Diverse Bacteria Species 319
Piyali Sen, Abdul Waris, Suvendra Kumar Ray and Siddhartha Sankar Satapathy

Blockchain-Based Transparent and Secure Decentralized Algorithm 327
Shreya Sudhakaran, Sunil Kumar, Priya Ranjan and Malay Ranjan Tripathy

Prediction of Cancer Diagnosis Patients from Fine-Needle Aspirates Using Machine Learning 337
Deepak Mehta and Chaman Verma

Recognition of Facial Expression Based on the Position of Hands Surrounding the Face Through Median Filter 349
Samta Jain Goyal, Arvind Kumar Upadhyay and Rakesh Singh Jadon

Secure Sharing of Location Data Using Elliptic Curve Cryptography 359
Nikhil B. Khandare and Narendra S. Chaudhari

Cyberbullying Checker: Online Bully Content Detection Using Hybrid Supervised Learning 371
Akshi Kumar and Nitin Sachdeva

Location-Wise News Headlines Classification and Sentiment Analysis: A Deep Learning Approach 383
Ashwin Kabra and Seema Shrawne

Review of Plagiarism Detection Technique in Source Code 393
Anala A. Pandit and Gaurav Toksha

Study on the Future of Enterprise Communication by Cloud Session Border Controllers (SBC) 407
Siddarth Kaul and Anuj Jain

Task Scheduling Based on Hybrid Algorithm for Cloud Computing 415
A. Vijaya Krishna, Somula Ramasubbareddy and K. Govinda

An Integrated Approach for Botnet Detection and Prediction Using Honeynet and Socialnet Data 423
Mahesh Banerjee, Bhavna Agarwal and S. D. Samantaray

Teaching–Learning-Based Functional Link Artificial Neural Network for Short-Term Electrical Load Forecasting 433
Rudra Narayan Pandey, Sarat Mishra and Sudhansu Kumar Mishra

An Enhanced K-Means MSOINN Based Clustering Over Neo4j with an Application to Weather Analysis 451
K. Lavanya, Rani Kashyap, S. Anjana and Sumaiya Thasneen

Proactive Preventive and Evidence-Based Artificial Intelligene Models: Future Healthcare 463
Kamal Kr. Sharma, Shivaji D. Pawar and Bandana Bali

Utilization of Artificial Neural Network for the Protection of Power Transformer 473
Mudita Banerjee and Anita Khosla

Analysing Tweets for Text and Image Features to Detect Fake News Using Ensemble Learning 479
Priyanka Meel, Harsh Agrawal, Mansi Agrawal and Archit Goyal

A Patchy Ground Antenna for Wide Band Transmission in S-Band Application 489
Anurag Saxena, Vinod Kumar Singh and Ashutosh Kumar Singh

Electromagnetic Scattering From Two Crossed Ferromagnetic Microwires 495
Tarun Kumar, Rajeev Kamal and Abhinav Sharma

Performance Analysis of Wearable Textile Antenna Under Different Conditions for WLAN and C-Band Applications 503
Ashok Yadav, Vinod Kumar Singh and Himanshu Mohan

Chip Resistor Loaded Fractal Frequency Selective Surface Based Miniaturized Broadband Microwave Absorber from 2 to 18 GHz 513
Tanveer Kaur Suchu, Arpit Sahu, Ravi Panwar and Rajesh Khanna

Design and Analysis of Triple Split Ring Resonator-Based Polarization-Insensitive, Multiband Metamaterial Absorber 523
Arpit Sahu, Ravi Yadav, Trivesh Kumar and Ravi Panwar

Design and Analysis of Slot Loaded C-Shaped Trapezoidal Microstrip Antenna 533
Rishabh Kumar Baudh, Mahendra Kumar, Ravi Kant Prasad and Sonal Sahu

Calculation of SAR on Human Brain Model and Analysis of the Effect of Various Dielectric Shielding Materials to Reduce SAR 545
Pravesh Chaudhary and Ravika Vijay

Equivalent Circuit Model Analysis of Single Square Loop FSS for Transmission Mechanism 551
Rahul Shukla and Garima Tiwari

Design of a Dual Annular Ring CPW Fed UWB Antenna for Wireless Applications 559
Hitanshu Katiyar, Alaknanda Ashok, Abhinav Sharma and Tarun Kumar

Path Loss Calculation at 900 MHz and 2.4 GHz in WBAN 565
Purnima K. Sharma, T. Vijay Sai and Dinesh Sharma

Ladder Mat Shape Microstrip Patch Antenna for X Band 575
Nivedita Dash and Sunil Kumar Singh

Frequency Agile Slotted Diagonally Sliced Elliptically Polarized Square Patch Antenna 583
Shipra Bhatia and M. V. Deepak Nair

Drone Shape Slot Array Microstrip Patch Antenna for X-Band 589
Palak Jain and Sunil Kumar Singh

An Optimal Design of Split-Ring Resonator and Electronic Waste Composite-Based Cost-Effective Microwave Absorber for Low Observable Applications 597
Arpit Sahu and Ravi Panwar

Design of Rectangular Microstrip Patch Antenna and Array for Broadband Applications 605
Vivek Kumar, D. K. Parsediya and Anamika Gupta

Solar Rectenna to Power Wireless Sensors and Implanted Electronic Applications 615
B. Naresh, Vinod Kumar Singh and V. K. Sharma

Design and Performance Enhancement of Wearable Textile Antenna 625
Monika Parmar, Mohit Gaharwar and Jaydeep Dewangan

Wideband E-Shaped Planar Antenna for Cellular, GPS, and Wireless Applications 633
Shobit Agarwal, Umair Rafique and Vasu Jain

Radial Velocity Distortion Reduction for NLFM-Based Radar System Using a Notch Filter 643
Ankur Thakur, Salman Raju Talluri and Sandeep Kumar

Design of Compact Two-Element MIMO F-Antenna 657
Anamika Gupta, Laxmi Shrivastav and Santosh Sharma

Chebyshev Polynomials-Based Pulse Compression Waveform for Modern Microwave Radars 665
Ankur Thakur, Salman Raju Talluri and P. K. Verma

Wideband MIMO Antenna with Diverse Polarization and Radiation Pattern for 5G Applications 677
Dinesh Sharma and E. Kusuma Kumari

Study the Effect of Dielectric Constant on Notch-Loaded U-Shaped Microstrip Patch Antenna 689
Saurabh Singh, Sudhanshu Verma and Rishabh Kumar Baudh

Triple-Band Reconstruct Circular Aperture Loaded CDRA Deploying $HE_{11\delta}$ and $HE_{21\delta}$ Modes for Wireless Applications 697
Ajay Kumar Dwivedi, Anand Sharma and Ashutosh Kumar Singh

Tumor Detection in Multilayer Brain Phantom Model by Symmetrical-Shaped DGS Rectangular Microstrip Patch Antenna 705
Hemant Kumar Gupta, Raghavendra Sharma and Vandana Vikas Thakre

Design and Optimization of S-Shape Multi-resonant Microstrip Antenna 713
Mohit Gaharwar, Shivam Tyagi and Aditya Boggavarapu

Sigma-Structured Microstrip Antenna for Harvesting Energy for Low-Power Devices 721
Bharat Bhushan Khare, Vinod Kumar Singh and Anurag Saxena

Nodal Pricing Analysis of Distribution System with Wind Power and D-STATCOM for Realistic ZIP and RIC Loads 729
Banothu Sridhar and Ashwani Kumar

Enhancement of Bandwidth and Gain of a Slotted E-Shaped Patch Antenna 749
Yatharth Shankar Misra and Ramesh Kumar Verma

Circularly Polarized Compact Monopole Antenna with an Offset Microstrip Feedline for C-Band Applications 759
Akansha Yadav, Sudhanshu Verma and Saurabh Singh

An H-Shaped Reconfigurable Slot Antenna Using SIW Technology 769
Rishi Raj Singh and Akhilesh Mohan

Frequency Scanning SIW Leaky Wave Horn Antenna for Wireless Application 779
Mahesh Kumar Aghwariya, Tanvi Agarwal and Ragini Sharma

Randomizing Vigenere Cipher in Polyalphabetic Substitution Scheme 789
Prabha Elizabeth Varghese and Latha R. Nair

Novel Tool to Determine Kinetic Parameters of Thermoluminescence (TL) Glow Curve—CGCD:$CaZrO_3$: Eu^{3+}, Tb^{3+} 795
Shubha Tripathi, Vikarm Awate, K. K. Kushwah, Ratnesh Tiwari and Nigama Prasan Sahoo

Determination of Spectroscopic Parameters via Judd–Ofelt Analysis of Eu^{3+} Doped $La_2Zr_2O_7$ Phosphor 805
Neha Dubey, Jagjeet Kaur, Vikas Dubey and Manish Kumar Mishra

CGCD Technique for Thermoluminescence Studies of $Y_2Zr_2O_7$:Eu^{3+} Phosphor 811
Seema Chopra, Ratnesh Tiwari, Anshu Gupta and A. K. Beliya

A Study of Thermally Induced Vibrations of Circular Plate of Nonuniform Thickness 815
Narender Kumar Sarswat, Vakul Bansal, Praveen Kumar and Mahesh Kumar Aghwariya

Industrial Motor Bearing Fault Detection Using Vibration Analysis 827
Rajul Misra, Kshitij Shinghal, Amit Saxena and Alok Agarwal

A New Modified Form of Murnaghan Thermodynamic Equation of State 841
Tanveer Ahmad Wani, B. K. Das, Babita Tripathi and Ikhlaq Ahmed Khan

A Study of Shear and Temperature on Vibrations of Infinite Plate . . . 849
Narender Kumar Sarswat

Voltage Stability Analysis of Wind Integrated Grid 857
Taruna Sharma and Omveer Singh

CDM-Based PID Controller with Filter Design for Performance Improvement of Two-Mass Drive System 867
Subrata Jana and Benjamin A. Shimray

Autonomous Vehicle Power Scavenging Analysis for Vehicular Ad Hoc Network 879
Garima Sharma, Praveen Kumar Singh and Laxmi Shrivastava

Cuckoo Search Algorithm and Ant Lion Optimizer for Optimal Allocation of TCSC and Voltage Stability Constrained Optimal Power Flow 889
Sheila Mahapatra, Nitin Malik and A. N. Jha

A New Approach to Smart Knock Detection Based Security System for Door Lock 907
Amit Chaurasia, Umesh Kumar Dwivedi, Amita Chaurasia and Shubham Kumar Jain

Effective Control Strategy in Microgrid 915
Ankur Maheshwari, Yog Raj Sood, Aashish Goyal, Mukesh Singh and Sumit Sharma

An Efficient Trust-Based Approach to Load Balanced Routing Enhanced by Virtual Machines in Vehicular Environment 925
Rakhi and G. L. Pahuja

Parameter Optimization of a Modified PID Controller Using Symbiotic Organisms Search for Magnetic Levitation Plant 937
D. S. Acharya and S. K. Mishra

Review of Literature—Analysis and Detection of Stress Using Facial Images 949
Ayusha Harbola and Ram Avtar Jaswal

A Survey on Text Detection from Document Images 961
M. Ravikumar and G. Shivakumar

Object Recognition Using SBMHF Features 973
M. Ravikumar, S. Sampathkumar, M. C. Prashanth and B. J. Shivaprasad

DWT Based Compression Algorithm on Acne Face Images 985
Garima Nain, Ashish Gupta and Rekha Gupta

Segmentation of Blood Vessels from Retinal Fundus Images Using Bird Swarm Algorithm and River Formation Dynamics Algorithm 995
Jyotika Pruthi, Shaveta Arora and Kavita Khanna

Image Processing for UAV Using Deep Convolutional Encoder–Decoder Networks with Symmetric Skip Connections on a System on Chip (SoC) 1009
Abhiraj Hinge, Pranav Garg and Neena Goveas

A Comparison of Different Filtering Strategies Used in Attribute Profiles for Hyperspectral Image Classification 1017
Arundhati Das, Kaushal Bhardwaj and Swarnajyoti Patra

Standard Statistical Feature Analysis of Image Features for Facial Images Using Principal Component Analysis and Its Comparative Study with Independent Component Analysis 1027
Bulbul Agrawal, Shradha Dubey and Manish Dixit

Indian Currency Recognition Using Radial Basis Function and Support Vector Machine 1049
Anand Upadhyay, Shashank Shukla and Verlyn Pinto

LDA- and QDA-Based Skin Lesion Melanoma Detection 1057
Anand Upadhyay, Arvind Chauhan and Darshan Kudtarkar

Design and Analysis of a Multilevel DC–DC Boost Converter 1065
Divesh Kumar, Dheeraj Kalra and Devendra Kumar

Improved Image Quality of Hybrid Underwater Image Restoration (IR) Using the Dark Channel Prior (DCP), Color Attribution Prior, and Contrast Stretching 1077
Anuradha Vashishtha and Jamvant Singh Kumare

Design of Low Voltage High-Speed Universal Logic Gates Using Different Models of CMOS Schmitt Trigger 1093
Bhavika Khanna, Raghav Gupta, Cherry Bhargav and Harpreet Singh Bedi

Design and Analysis of 3-Bit Shift Register Using MTCMOS Technique 1107
Bhupendra Sharma, Ashwani Kumar Yadav, Vaishali and Amit Chaurasia

Dynamic Power Reduction Techniques for CMOS Logics Using 45 nm Technology 1117
Rajesh Yadav, Reetu and Rekha Yadav

Design of an Optimal Integer Frequency Synthesizer for 5 GHz Frequency at 45 nm CMOS Technology 1127
Rekha Yadav and Sakashi Kaushik

Performance Evaluation of Hybrid RF/FSO Communication System in High-Speed Train 1139
Ankita Chauhan and Pankaj Verma

Real-Time Analysis of Low-Cost Software-Defined Radio Transceiver Using ZigBee Protocol 1151
Nikhil Marriwala, O. P. Sahu and Anil Vohra

Comparative Overview of Profit-Based Unit Commitment in Competitive Electricity Market 1171
Ayani Nandi and Vikram Kumar Kamboj

THD Analysis of New Multilevel Inverter Topology with Different Modulation Techniques 1197
Nikhil Agrawal, Praveen Bansal and Niraj Umale

Sensitivity-Based Adaptive Activity Mapping for Optimal Camera Calibration 1211
Shashank and S. Indu

Spectral–Spatial Active Learning with Attribute Profile for Hyperspectral Image Classification 1219
Kaushal Bhardwaj, Arundhati Das and Swarnajyoti Patra

Secrecy Performance Analysis of Hybrid-Amplify-and-Decode-Forward (HADF) Relaying Scheme Under Multi-hop Scenario 1231
Shweta Pal and Poonam Jindal

A Review on Photonic Crystal Fibers 1241
Arati Kumari Shah and Rajesh Kumar

Multiplier-Less Architecture for 4-Tap Daubechies Wavelet Filters Using Algebraic Integers 1251
Mohd. Rafi Lone and Najeeb-ud-Din Hakim

Bit Representation for Candidate Itemset Generation 1259
Carynthia Kharkongor and B. Nath

PROD: A Potential Rumour Origin Detection Model Using Supervised Machine Learning 1269
Akshi Kumar and Harshita Sharma

Double-Stage Sensing Detectors for Cognitive Radio Networks 1277
Ashish Bagwari, Jyotshana Kanti and Geetam Singh Tomar

Modified Soft Combination Scheme for Cooperative Sequential Detection Considering Fast-Fading in Cognitive Radios 1285
Mayank Sahu and Amit Baghel

Role of Chaos in Spread Spectrum Communication 1293
Devendra Kumar, Divesh Kumar and Dheeraj Kalra

Design Analysis of CG-CS LNA for Wideband Applications Using Noise Cancelation Technique 1311
Dheeraj Kalra, Devendra Kumar and Divesh Kumar

Performance Evaluation of Hybrid Renewable Energy System for Supplying Electricity to an Institution and a Hospital Using HOMER 1317
Ingudam Chitrasen Meitei, Amit Kumar Irungbam and Benjamin A. Shimray

Energy Scheduling of a Household with Integration of Renewable Energy Considering Different Dynamic Pricing Schemes 1327
Bhukya Balakrishna and Sandeep Kakran

Optimum Design of Photovoltaic System For a Medical Institute Using HOMER 1337
Ingudam Chitrasen Meitei, Thounaojam Bebekananda Singh, Konjengbam Denish, Huidrom Hilengamba Meetei and Naorem Ajesh Singh

Raising Concerns on High PV Penetration and Ancillary Services: A Review 1347
Rajesh Kumar, Aman Ganesh and Vipin Kumar

Analysis of 150 kW Grid-Connected Solar PV System Using Fuzzy Logic MPPT 1365
Ashutosh Tiwari and Ashwani Kumar

Clock System Architecture for Digital Circuits 1375
Amit Saxena, Kshitij Shinghal, Rajul Misra and Alok Agarwal

TCAD Modeling and Analysis of sub-30nm Strained Channel MOSFET 1383
Lalthanpuii Khiangte, Kuleen Kumar and Rudra Sankar Dhar

InGaAs MOSFET for High Power Applications 1389
Manoj Singh Adhikari, Vikalp Joshi and Raju Patel

Low Power Efficient $Si_{0.7}Ge_{0.3}$ Pocket Junction-Less DGTFET with Sensing Ability for Bio-species 1395
Suman Lata Tripathi and Shekhar Verma

Three-Layered Channel with Strained Si/SiGe/Si HOI MOSFET 1405
Lalthanpuii Khiangte and Rudra Sankar Dhar

Electric Field Modeling and Critiques of Dual-Halo Dual-Dielectric Triple-Material Surrounding-Gate MOSFET 1413
Prashant Kumar, Neeraj Gupta, Rashmi Gupta and Ganesh Gupta

Modeling and Logic Synthesis of Multifunctional and Universal 3 × 3 Reversible Gate for Nanoscale Applications 1423
Naira Nafees, Insha Manzoor, Majid Irfan Baba, Soha Maqbool Bhat, Vishal Puri and Suhaib Ahmed

Logic Design and Modeling of an Ultraefficient 3 × 3 Reversible Gate for Nanoscale Applications 1433
Insha Manzoor, Naira Nafees, Majid Irfan Baba, Soha Maqbool Bhat, Vishal Puri and Suhaib Ahmed

A Novel Reversible DSG Gate and Its Quantum Implementation 1443
Shaveta Thakral and Dipali Bansal

Performance Analysis of Various Embedded Linux Firmwares for ARM Architecture Based IoT Devices 1451
Mahendra Swain, Rajesh Singh, Md. Farukh Hashmi and Anita Gehlot

Landslides Detection in Prone Hilly Areas Using Raspberry Pi 1461
Prabin Kumar Das, Rajesh Singh, Anita Gehlot, Km. Vaishnavi Gupta and Arun Singh

Image and Video Capturing for Proper Hand Sanitation Surveillance in Hospitals Using Euphony—A Raspberry Pi and Arduino-Based Device 1475
Navjot Rathour, Rajesh Singh and Anita Gehlot

Industrial Hazard Prevention Using Raspberry Pi 1487
Prabin Kumar Das, Praveen Kumar Malik, Rajesh Singh, Anita Gehlot, Km Vaishnavi Gupta and Arun Singh

IoT-Based Traffic Management System Including Emergency Vehicle Priority 1501
Anmol Joshi, Naman Jain and Alok Pandey

Design and Development of IoT-Enabled Portable Healthcare Device for Rural Health Workers 1509
Saurabh Nautiyal and Rekha Devi

Head-Gesture-Based Human–Computer Interface for Disabled People 1525
Sumit Malik, Nitin Kumar and Anuj Jain

Intelligent and Hybrid Control Techniques for Robotic Manipulator 1535
Mandeep Singh and M. K. Shukla

Health Monitoring Gadgets 1547
Jasjit Singh, Ankur Kohli, Bhupendra Singh, Aishwarya Prasad Bhatoye and Sauhardh Prasad Bhatoye

Evaluating Input Stream for IoT Commodity Eyes 1553
Shubham Asthana, Rajiv Pandey and Archana Sahai

Application of Sensors Using IoT for Waste Management System 1565
Akhil Chitreddy, Kailash Gogineni, Vattikuti Anirudh, P. V. Akhilesh, Konda Krishna Vamsi and P. Swarna Latha

Nodal Price Determination for Radial Distribution System Using Load Flow Approach 1577
Ishita Jain and Ashwani Kumar

Smart Quadripod Walking Stick for the Aid and Security of Visually Challenged and Elderly People 1587
Amitejash Rout, Kshaunish Roy, Swapnil Chhatre, Rhythm Sahu and H. Parveen Sultana

Real-Time Tracking and Lane Line Detection Technique for an Autonomous Ground Vehicle System 1609
Tamal Datta, S. K. Mishra and S. K. Swain

Wireless Controlled Lake Cleaning System 1627
Sudhanshu Kumar, Saket Kumar, Rajkumar Viral and H. P. Singh

Multipurpose Voice Control-Based Command Instruction System 1641
Deepak Ranjan, Rajkumar Viral, Saket Kumar, Gaurav Yadav and Prateek Kumar

IoT-Based Cross-Functional Agribot 1657
Suraj Sudhakar and Sharmila Chidaravalli

A Novel Interfacing Scheme for Analog Sensors with AMLCD Using Raspberry Pi 1673
Peeyush Garg, Ajay Shankar, Mahipal Bhukya and Vinay Gupta

Author Index 1685

About the Editors

Prof. Geetam Singh Tomar received his undergraduate, graduate and Ph.D. degrees in Electronics Engineering from the Institution of Engineers Calcutta, MNREC Allahabad, and RGPV Bhopal, respectively. He has also received a second Ph.D. in Computer Engineering from the University of Bristol, UK, and a postdoc from the University of Kent, UK. He has worked with MITS and IIITM Gwalior, the University of Kent, UK, and the University of West Indies, Trinidad, and Tobago. He has served as the director of many reputed engineering institutes along with the additional charge of the Director of Machine Intelligence Research Labs, Gwalior, India. He also served in the Indian Air Force for 17 Years. Prof. Tomar has many sponsored research projects to his credit, received the International Plato Award for academic excellence from the IBC, Cambridge, UK, in 2009, has published more than 180 research papers, and holds two patents. Currently, he is working as the Director at THDC Institute of Hydropower Engineering and Technology, New Tehri, Uttarakhand, India.

Prof. Narendra S. Chaudhari received his B.Tech. (EE), M.Tech. (CSE), and Ph.D. (CSE) degrees from the IIT Bombay, India. He has worked with many national and international universities, including Nanyang Technological University, Singapore; Devi Ahilya University, Indore; and ITM University, Gwalior. He has served as the director of reputed institutes like the VNIT Nagpur and MNIT Bhopal. He has contributed significant research work on game AIs, novel neural networks models, context-free grammar parsing, and the graph isomorphism problem. He has more than 300 publications to his credit, including four authored books, five edited books, and 11 book chapters. Currently, he is serving as Vice-Chancellor of Uttarakhand Technical University, Dehradun, India.

Prof. Jorge Luis V. Barbosa received his M.Sc. and Ph.D. in Computer Science from the Federal University of Rio Grande do Sul, Porto Alegre, Brazil. He subsequently conducted postdoctoral studies at Sungkyunkwan University (SKKU), Suwon, South Korea. Jorge is currently a Full Professor at the Applied Computing Graduate Program (PPGCA) of the University of Vale do Rio dos Sinos

(UNISINOS), head of the university's Mobile Computing Lab (MOBILAB), and a researcher at the Brazilian Council for Scientific and Technological Development (CNPq). His main research interests are mobile and ubiquitous computing, context prediction using context histories, and ubiquitous computing applications mainly in health, u-accessibility and learning (u-learning).

Mahesh Kumar Aghwariya is an Assistant Professor at the Department of Electronics and Communication Engineering, THDC Institute of Hydropower Engineering and Technology, New Tehri, Uttarakhand, India. Over the past seven years, he has made notable contributions to research in microwave engineering. He has many papers in prominent national and international journals to his credit. His main area of interest is microwave and RF engineering.

A Survey of Fuzzy Logic Inference System and Other Computing Techniques for Agricultural Diseases



Bhavna Chilwal and P. K. Mishra

1 Introduction

Expert System is one of the achieved goals of AI. Fuzzy logic is the method of the Expert system which resembles the human brain in reasoning ability. Fuzzy logic has decision-making ability and used for prediction analysis in different domains. Agriculture is the most important sector as it provides food to the world and food is the basic necessity for humans. Agriculture provides food grains, vegetables, fruits. All of these grain crops are important for feeding the whole world. But there are some issues that affect crop productivity. 1. Diseases (fungal diseases) in crops is one of the issues because of which approx. 60–70% of crop loss happens worldwide. 2. Pest insects may become a difficult problem when they damage crops and food production, destroying livestock, also hazardous to human health. 3. Weeds are responsible for reducing farm productivity, they continuously try to absorb water, soil nutrients, and sunlight which is provided for the crop, this gives rise to a loss in crop yield. 4. Weather conditions play a major role in crop productivity but adverse weather conditions have a very bad impact on crops. So with the development of soft computing technologies, many scientists have applied their approaches in different fields as agriculture is one of the fields, where technologies like a neural network, machine learning, fuzzy logic, etc. are applied. These are used for crop management, pest management, disease management, and precision agriculture [1].

In this paper, different computing methods to solve problems in the agricultural field are discussed and section shows functions of the fuzzy logic in precision agriculture and its importance.

B. Chilwal (✉) · P. K. Mishra
Department of Computer Engineering, GBPUAT, Pantnagar, India
e-mail: bhavnachilwal@gmail.com

P. K. Mishra
e-mail: pkmishra.cs@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_1

1.1 Agricultural Diseases and Different Computing Techniques

The disease and pest problems in crop plants are very hazardous in crop production and they disturb the agricultural development. From many years' statistical reports from different food and agricultural organizations which showed that all over the world crop yield is lost around 10% by pests and 14% because of diseases. So by predicting and detecting these diseases, we can help farmers to take the right prevention and correct treatments in order to control the disease and obtain maximum benefits from crops. Different researchers are working on the impact of the disease in the agricultural sector.

Donatelli et al. [2]: A simulation model has been made up to provide a framework to incorporate the impacts of pests and diseases and this model will be used by the particular pathosystem to give the required solutions for diseases and improve the productivity. Various researchers have provided methods for classification and detection methods of different diseases in the agriculture domain, some are reviewed here. Sharifa and Khana [3] have proposed an approach called the Hybrid approach, for the detection and classification of citrus disease, to find its six different types in the agricultural field. McCallum et al. [4] took cassava as a staple crop and detect different types of diseases in cassava. By using the molecular mechanism and breeding program, disease-resistive crop varieties developed to tackle the disease. So diseases have particular symptoms. To save crop yield it becomes very important to predict and detect these diseases so that available methods and fungicides will be used to save the crop. Wang et al. [5]: This paper discusses the importance of prediction analysis and warning system by using IoT, cloud computing, and machine learning. It also shows that this system will help to get effective crops and it has a good future scope for research. Devraj and Jain [6]: This paper presents PulsExpert, an expert system that is a diagnostic tool for farmers to identify major diseases in pulse crops. This system is user friendly by making use of knowledge acquisitions and also become an interface for domain experts. Bao et al. [7]: This provides a database for crop disease and pests using GIS technologies. It allows flexible datasets to be managed and analyzed. This project will provide long-term information for the management of the disease. Many researchers are working for the integration of soft computing techniques and the agricultural situations, generally, techniques are used for different situations in agriculture. Here are a few of them:

Chlingaryan et al. [8] showed how yield prediction and nitrogen estimation for crops is beneficial in agriculture. Because of the advancement of machine learning and sensing technology, they are used for the prediction analysis and provide effective datasets for decision-making and estimations. Patrícioa and Riederb [9] provide a computer vision grading system for different crops which helps in analyzing samples and describing data. It motivates the use of artificial intelligence and machine learning to overcome the problems which are facing by the agricultural sector. Luo and Huang [10]: A GIS-based warning system was formed for crop diseases and pest, it predicts the disease by taking warning maps and forecast data collections, information, and

results which are viewable. This prediction system helps to control the disease. Liu [11]: A prediction system was made by using a neural network toolbox in MATLAB, it is about fruit tree diseases and pests by using the BP algorithm in prediction. Billah et al. [12]: This paper develops ANN (artificial neural network) to inspect diseases in the rice crop. Different stages of the disease in rice leaves are captured and those images are used for learning of ANN.

1.2 Fuzzy Logic

Fuzzy means Unclear or Imprecise. Fuzzy Logic is a method of reasoning, which resembles human reasoning. Before the use of fuzzy logic, traditional logic was used in which we have a crisp set of values either 1 or 0, either YES or NO. But in real-world problems, we have to face some situations in which we need the values either 1 or 0. So for that case, fuzzy logic used to get the fuzzy set of values between the crisp set of values. The fuzzy set of values contains the range of possibilities either YES or NO.

Fuzzy Inference System is the most important unit of Fuzzy logic having decision-making as a primary work. IF…THEN rules are applied to the inference system and one rule can have more than one type by using OR and AND function. MATLAB Fuzzy logic tool is the most efficient for applying fuzzification as it has all the required things inbuilt in it (Fig. 1).

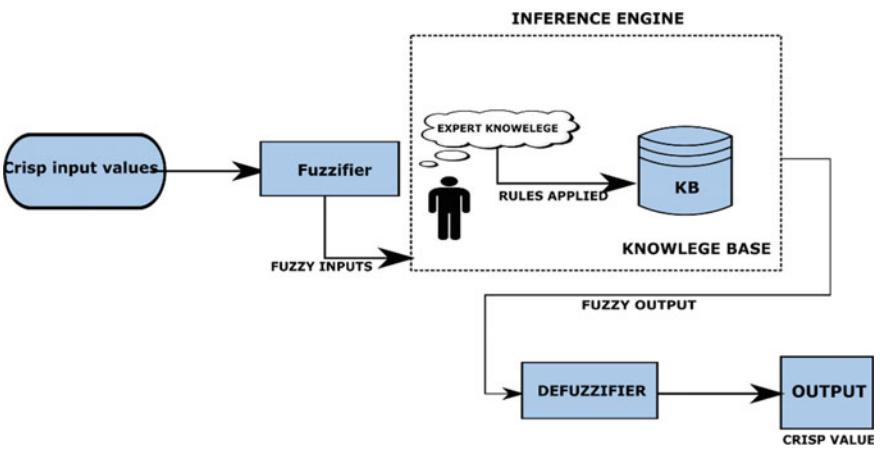


Fig. 1 Fuzzy inference system architecture

2 Fuzzy Logic in the Agricultural Sector

Fuzzy logic uses probability concept and has a decision-making capability because of which it is a very effective technique for rule-based soft computing in predicting and forecasting. Many scientists and researchers are using fuzzy logic in the farm field to make agriculture more effective and beneficial. Few of them are here:

Li and Yan [13]: A model was presented which uses fuzzy logic evaluation for emergy analysis used to check the health of agriculture land and return the status of health. It also uses reversion analysis for the improvement of land to maintain its health. Jawad and Choudhury [14]: This paper has proposed a system that will analyze the optimum crop cultivation based on Neuro-Fuzzy System. This system calculates the yield of crops by using the humidity, temperature, and rainfall values. Neamtollahi and Vafabakshi [15]: Here, the fuzzy logic system is used for making cropping patterns and the objectives of these patterns are reducing the amount of water and fertilizers for the crop and maximize the farmer's income. Rodriguez and Peche [16]: This paper uses Fuzzy logic as a quality assessment tool and used to design quality indexes of soil quality. These indexes check the dynamic quality of soil based on its characteristics. Roseline [17]: This paper proposed a few of the applications of fuzzy logic applied in the farm sector. It describes how the expert system works on a real-world scenario and helps farmers and it encourages to apply technology from the start of seed selection to the final stage of crop storage. Balanovskaya and Boretska [18]: Fuzzy inference system helps to increase the efficiency decision-making in agricultural enterprises. A quality management system is built. Different factors are defined which affects the quality of agricultural products. Marimin and Mushthofa [19]: It shows the use of the fuzzy system in the agroindustry engineering and technology for checking the soil and land suitability, climate prediction, pest and weed management, risk management, product quality assurance, as well as in the intelligent decision support system. Ingole and Katole [20] proposed the feature of the fuzzy system in predicting crop yield by remote sensing. It uses light and temperature sensing elements and mathematical modeling. Fuzzy logic use experts' knowledge to form rules and these rules will help to predict the level of the pathogen. Awoyelu and Adebisi [21], here this paper discusses how the fuzzy inference system can be used for predictive analysis for diagnosing the diseases in the cassava crop.

3 Discussion

AI techniques have vast impact on the research field. Different methods like classification, regression, image processing, rule-based techniques, etc., are very effective in solving different real-world situations. This paper discusses the roles of these techniques and especially Fuzzy logic techniques in the agricultural domain. This paper shows how fuzzy logic helps in different problems related to precision agriculture specially disease detection and prediction in the agricultural field will provide

benefits by decreasing the amount of crop loss and increase farmer's income by high yield.

4 Conclusion

Agricultural diseases are the major concerns for the farmers and all the agricultural scientists. Traditional methods to find the severity of diseases in crop plants were very time consuming, costly, and tiresome. The manual team for predicting the disease level is not effective and time efficient for agricultural problems. So from the past few years, there are computing techniques which provide the Digital Agricultural era. The expert system has a large dominance in solving different problems in agriculture fields like soil management, weather forecasting, pest management, disease management and many more. The prediction of disease severity will help the farmers to take the required measures to save the crop yield loss. Because less crop productivity creates a big problem in any nation, that is, poverty and starvation which lead to bad health conditions.

References

1. Y. Huang, Development of soft computing and applications in agriculture and biological engineering. Comput. Electron. Agric. (2010), www.elsevier.com/locate/compag, www.sciencedirect.com. Accessed 07 Jan 2019
2. M. Donatelli, R.D. Magarey, S. Bregaglio, Modeling the impacts of pests and diseases on agricultural systems. Agric. Syst. (2013), www.elsevier.com/locate/agsy, www.sciencedirect.com. Accessed 09 Jan 2019
3. M. Sharifa, M.A. Khana, Detection and classification of citrus diseases in agriculture based on optimized weighted segmentation and feature selection. Comput. Electron. Agric. (2018), www.elsevier.com/locate/compag. Accessed 12 Jan 2019
4. E.J. McCallum, R.B. Anjanappa, W. Gruissem, Tackling agriculturally relevant diseases in the staple crop cassava (Manihot esculenta). Sciencedirect, Elsevier (2017)
5. D. Wang, T. Chen, J. Dong, Research of the early warning analysis of crop diseases and insect pests. IFIP Int. Fed. Inf. Process. (2014), www.linkspringer.com. Accessed 16 Jan 2019
6. Devraj, R. Jain, PulsExpert: an expert system for the diagnosis and control of diseases in pulse crops. Expert. Syst. Appl. (Elsevier) (2011)
7. Y.-W. Bao, M.-X. Yuc, W. Wua, Design and implementation of database for a webGIS-based rice diseases and pests system. Procedia Environ. Sci., in *International Conference on Environmental Science and Information Application Technology ESIAT 2011 EI Comp* (Elsevier, 2011)
8. A. Chlingaryana, S. Sukkarieha, B. Whelanb, Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: a review. Comput. Electron. Agric. (Elsevier, Science Direct) **151** (2018)
9. D.I. Patrícioa, R. Riederb, Computer vision and artificial intelligence in precision agriculture for grain crops: a systematic review. Comput. Electron. Agric. (Elsevier, Sciencedirect.com) (2018)

10. J. Luo, W. Huang, The crop diseases and pest warning and prediction system, in *IFIP International Federation for Information*. Volume 294, Computer and Computing Technologies in Agriculture II, vol. 2, ed. by D. Li (Springer, Boston, 2009), pp. 937–945
11. G. Liu, H. Shen, X. Yang, Research on a prediction about fruit tree diseases and insects pest based neural network, https://link.springer.com/. Accessed 12 Jan 2019
12. M.M. Billah, M.P. Islam, M.G. Rahman, Identification of rice diseases using an artificial neural network. J. Bangladesh Soc. Agric. Sci. Technol. **4**(3 and 4), 189–194 (2007)
13. Q. Li, J. Yan, Assessing the health of agricultural land with energy analysis and fuzzy logic in the major grain-producing region. Catena (2012), www.elsevier.com/locate/catena. Accessed 15 Jan 2019
14. F. Jawad, T.U.R. Choudhury, Analysis of optimum crop cultivation using fuzzy system, in *IEEE ICIS 2016*, 26–29 June 2016, Okayama, Japan (2016)
15. E. Neamatollahi, J. Vafabakshi, Agricultural optimal cropping pattern determination based on fuzzy system. Fuzzy Inf. Eng. (2017), www.elsevier.com. Accessed 14 Jan 2019
16. E. Rodríguez, R. Peche, Dynamic quality index for agricultural soils based on fuzzy logic. Ecol. Indic. (2016), www.elsevier.com, www.sciencedirect.com. Accessed 15 Jan 2019
17. P. Roseline, A study of applications of fuzzy logic in various domains of agricultural sciences. Int. J. Comput. Appl. (0975–8887), in *International Conference on Current Trends in Advanced Computing (ICCTAC-2015)* (2015)
18. T.I. Balanovskaya, Z.P. Boretska, Application of fuzzy inference system to increase the efficiency of management decision-making in agricultural enterprises. Sci. J. Warsaw University of Life Sciences—SGGW Problems of World Agriculture vol. 14 (XXIX), number 4, (2014), pp. 15–24
19. M. Marimin, M. Mushthofa, Fuzzy logic systems and applications in agro-industrial engineering and technology, in *Second International Conference on Adaptive and Intelligence Agroindustry (2nd ICAIA)*, Bogor, 16–17 September 2013
20. K. Ingole, K. Katole, Crop prediction and detection using fuzzy logic in MATLAB. Int. J. Adv. Eng. Technol. IJAET (2013)
21. I.O. Awoyelu, R.O. Adebisi, A predictive fuzzy expert system for diagnosis of cassava plant diseases. Glob. J. Sci. Front. Res. C Biol. Sci. **15**(5) Version 1.0 the Year (2015)

Credit Card Fraud Detection Using Correlation-Based Feature Extraction and Ensemble of Learners



I. Sumaiya Thaseen and K. Lavanya

1 Introduction

Digitalization seen in the past few years has led to the tremendous use of credit cards for various applications. However, Credit cards are easily targeted for fraud since tricksters find money earning is easy in a short span of time, without the involvement of many risks. In addition, the crime is revealed only a few weeks after the transaction. Frauds can be classified into three categories, i.e., merchant related frauds, traditional card frauds and internet fraud. Fraud is identified by analyzing those fraudulent transactions. That is classifying the transactions into two classes: a genuine class and a fraudulent class. The major problem associated with fraud detection is the lack of real-world data for performing experimental results. This is due to the fact that financial data is sensitive, requiring high confidentiality. A fraud detection model can be developed by using machine learning techniques as uncertainty can be predicted in advance. The key idea in fraud detection is the provision of a computational model with a set of training data comprising of some attributes (e.g., obtained from the data of a sequence of financial transactions) that are intrinsic to the system in which the fraud detection is to be performed. After a learning process, the model is expected to make the correct identification of a transaction that has never been seen before as fraud or genuineness is based on the feature values of the transaction.

The structure of the article is as follows: Sect. 2 summarizes the findings from the literature related to applications of machine learning. The techniques deployed in the model are described in Sect. 3. The proposed model for credit card fraud identification is discussed in Sect. 4. The experimental results are illustrated in Sect. 5. The conclusion is given in the final section.

I. Sumaiya Thaseen (✉)
School of Information Technology and Engineering, Vellore Institute of Technology, Vellore, India
e-mail: sumaiyathaseen@gmail.com

K. Lavanya
School of Computing Science and Engineering, Vellore Institute of Technology, Vellore, India

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_2

2 Related Work

Credit card fraud detection has two unusual characteristics. The first one is the limited duration during which the decision of the acceptance or rejection has to be made. A large amount of credit card operations that have to be computed at a given time is the second one. Many of the supervised machine learning techniques are widely used for fraud identification. This is because the accuracy rate increases and the false positive rate decreases in the system.

A framework using logistic regression [1] with an incremental learning approach was utilized for credit card fraud detection to deal with unbalanced data. Credit card fraud detection [2] model was developed using Bayesian Networks (BN) and Artificial Neural Network (ANN). The authors have used a feature selection technique based on correlation for the identification of the primary attributes for classification. Different learning rates were deployed for ANN for the achievement of improved accuracy. However, the model has a false positive of 10%. Genetic Algorithm (GA) and Neural Network (NN) models have also been proposed for credit card fraud detection [3]. The decisions about network topology, hidden layers, the number of nodes required for designing the neural network model is determined by GA for credit card fraud detection. Feedforward backpropagation algorithm was also utilized for learning the model. A model using the association rule mining technique has been developed for the extraction of knowledge to enable getting the behavior pattern of unlawful transactions [4]. The authors used credit card information from prominent retail companies in Chile. GA and scatter search for fraud detection were utilized in another approach [5]. Among all the prevailing techniques, collective learning approaches are acknowledged as widespread and common due to their outstanding predictive performance on practical applications. A bagging classifier was used for the construction of the fraud detection model. Experimental analysis performed demonstrated the proof that bagging is better in comparison to other machine learning techniques [6]. SVM, Random Forest (RF), and Logistic Regression were compared for estimation of the accuracy and the results illustrated the performance of an overall higher precision and lesser recall by Random Forest [7]. This is the motivation for selecting random forest as one of the predictive classifiers in this article. Forest and Decision tree were combined to build an efficient online banking fraud detection [8] technique. Another fraud detection model [9] employs Long Short-Term Memory (LSTM) networks to identify transaction sequences. Feature aggregation strategies are also aggregated and compared with the RF. The results show that the sequential and nonsequential learning techniques benefit from manual feature aggregation approaches.

The credit card dataset contains many attributes which have no relevance for classification purpose. Hence, there is a critical need to identify the best discriminating features. The dataset is also highly unbalanced in nature as it contains a combination of fraudulent and genuine class transactions. This issue went unnoticed in the previous works. Thus, it is essential to consider the issues mentioned in developing a model for fraud identification in the credit card sector.

3 Background

3.1 Attribute Extraction

The process of determining the attributes that do not contribute or diminish the exactness of the perceptive model [10, 11] is known as attribute selection. Accurate prediction is achieved by eliminating overlapping or irrelevant features for the model. A feature in the subset is considered worthy if it is not much associated with other attributes of the class but has a high correlation with the class itself [12].

Correlation-Based Feature Selection
In this approach, the feature subset is obtained by evaluating the correlation between features based on the degree of duplication. Information theory based on entropy is used in CFS. Entropy is a measure of uncertainty which is given as follows:

$$H(z) = -\Sigma P(z_i) \log_2(P(z_i)) \tag{1}$$

The entropy of 'Z' after observing another variable values 'R' is defined as follows:

$$H(Z/R) = -\Sigma P(r_j) \Sigma P(z_i/r_j) \log_2(P(z_i/r_j)) \tag{2}$$

Here, the prior probability for the entire values of 'Z' is $P(z_i)$ and the posterior probability of 'Z' if 'R' is known is represented as $P(z_i/r_j)$. The decrease of entropy value in 'Z' reveals additional information relating to Z given by R which is the information gain as given in the equation below:

$$\text{IG}(Z/R) = H(Z) - H(Z/R) \tag{3}$$

A feature Z is correlated more to feature R than feature S if it satisfies the condition given below:

$$\text{IG}(R/Z) > IG(S/Z) \tag{4}$$

Symmetrical uncertainty (SU) is another measure which depicts the correlation between features as given below:

$$\text{SU(Z,}R) = 2\frac{\text{IG}(Z/R)}{H(Z) + H(R)} \tag{5}$$

SU normalizes the feature value in the range of [0, 1] wherein 1 shows the prediction of the value of others by the knowledge of either of the variables and 0 specifies the independence of X and Y. In this technique, the information bias of the gain toward features with more values is compensated.

3.2 Naive Bayes

One of the popular supervised machine learning algorithms is Naive Bayes, which utilizes the training data with identified class names for the prediction of the class labels of future instances. Bayes theorem is utilized in this technique. The membership probability for every class with regard to the various attributes in the dataset is calculated. The class with the maximum probability is defined as the foreseen class. Thus, the maximum posteriori probability is calculated as given below.

$$P(l|y) = \frac{P(y|l)P(l)}{P(y)} \tag{6}$$

Here $P(l|y)$ is the predicted class posterior probability given the feature value of the class. The prior probability of class is given by $P(l)$. $P(y|l)$ is the likelihood. The prior probability of the predicted class is denoted by $P(y)$.

3.3 KNN

The classification of an unknown data point is done on the basis of its nearest neighbor whose class is previously known. The inference of K is the basis for designating the number of closest neighbors to enable identification of the class label of the unidentified data sample. A majority of the closest neighbor class is predicted as the class label for the new test sample. The data samples are allocated with weights on the basis of the distance from the sample data. A K-nearest neighbor does the investigation of the pattern space for the 'm' training samples that are nearest to the unidentified sample. The Euclidean distance is utilized to determine the closeness between two points as given below:

$$\text{dist}(Z_1, Z_2) = \sqrt{\sum_{i-1}^{n} (z_{1i} - z_{2i})^2} \tag{7}$$

where, $Z_1 = (z_{11}, z_{12}, \ldots z_{1n})$ and $Z_2 = (z_{21}, z_{22}, \ldots z_{2n})$.

3.4 Random Forest

Many decision trees are constructed and the results are compiled into a final output in Random Forest. When a new sample requires classification, it is assigned to the various trees of the forest. A separate classification output for a class is determined by each tree. The maximum class label is selected from all the trees. This classifier is easy

to use for large and unbalanced datasets having many attributes. The computational speed of the classifier is also high. The output is accurate as the differences are minimized by selecting the maximum number of output for the final prediction.

4 Proposed Model

The proposed model for credit card fraud detection integrates correlation-based feature selection and ensemble of supervised classifiers. Figure 1 is the block diagram of the fraud detection model. Normalization is performed as the initial preprocessing step. In this phase, the attribute values are transformed in the range of [0, 1]. A correlation feature selection is used along with the best-fit search method. Highly ranked 12 attributes are selected from a total of 31 attributes based on the correlation value as given in Table 1. A threshold of 0.5 rank value is initialized and all

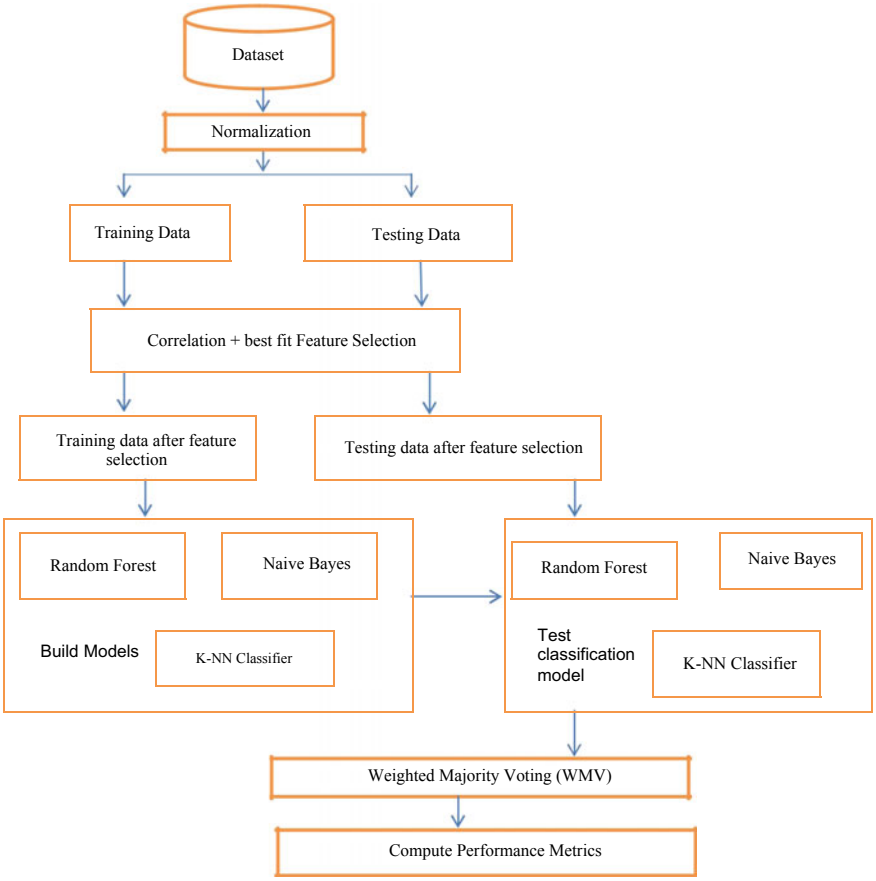


Fig. 1 Proposed fraud identification model

Table 1 Attributes selected by CFS subset evaluation + best-fit search

Attribute index	Correlation value
V1	0.9171
V2	0.854
V3	0.8217
V4	0.7895
V5	0.7572
V6	0.7249
V7	0.6927
V8	0.6604
V9	0.6282
V10	0.5959
V11	0.5637
V12	0.5314

attributes below the threshold are discarded from the subset (Pseudocode 1). Thus, the crucial attributes that play a pivotal role in fraud detection are identified. Table 1 shows the highly correlated features of the credit card dataset. As the attributes are confidential, the names of the attributes are not revealed. In the second stage, the subset is given as input to the supervised learners, namely, Naive Bayes, KNN, and Random Forest (Pseudocode 2). The primary advantage of using the collective layer of learners is improved performance. The subgroup is fed into the training phase for model building. The weights of each classifier are initialized by a validation set. WMV is deployed for integrating all the classifier weights. The determined class label of the test set is then compared with the target class present in the validation set. A learning factor β is utilized to deduct the weights, if there is an error in classifying the class label. The user can define the learning factor range from 0 to 1. This process is iterated for each record in the dataset.

Input : Credit Card Data Set (D= (x_1,…x_n)
Output: Optimal attribute set (OD)
Step 1: All the candidate features are examined by repeating steps 2 to 6.
Step 2: For each feature X_I in the candidate feature set, a CFS+ best fit search strategy is calculated.
Step 3: A feature X_i is chosen and included in the new feature set, if the rank is greater than the threshold of 0.5; OD = (OD , x_i)
Step 4: Else the next attribute is examined.
End

Pseudo Code 1: Retrieving best features using correlation-based feature selection with best-fit search strategy

Given: Classifier M_1,(NB) M_2, (KNN),M_3(Random Forest)
Input: Optimal attribute dataset OD
Output: Class label (L=*1…K)*
Step1: Initialize all the weights in D. $W_i=1/n$, where the total number of elements is represented by '*n*'.
Step 2: Fit the NB classifier to (x_t,y_t) for every sample data d_i , using weights w_i, Calculate the posterior probability and assign the class label for every sample as given in equation (6)
Step 4: Fit the KNN classifier to (x_t,y_t) using weights w_i. Calculate the maximum majority label using the distance measure as given in equation (7).
Step 5: Fit the random forest classifier using weights w_i. The class label is predicted by averaging all 'T' trees.
(i) For 1 to T do
(ii) Draw 'n' points D_l with replacement from D
(iii) Build full decision/regression tree on D_l
Each split only consider 'k' features, picked uniformly at random
new features for every split
(iv) Prune tree to minimize out-of-bag error.
(v) End For
(vi) Average all 'T' trees
(vii) End Procedure

Step 6: Predict the class label by Majority voting from results of steps (2), (3) and (4).
Step 7: Compute the performance metrics

Pseudo Code 2: Ensemble Classifier Prediction

5 Experimental Setup

5.1 Dataset

A public dataset of credit card transactions with fraudulent records has been used in the analysis. This dataset was deployed first by the authors [13] in their experiments and is available for download at http://www.ulb.ac.be/di/map/adalpozz/data/creditcard.Rdata. The dataset contains 31 attributes belonging to the numeric category. The attribute "Time" specifies the seconds from the first transaction and each of the other transactions present in the dataset. The attribute "Amount" denotes the transaction charge utilized for cost-sensitive learning. The attribute "class" is the resultant class label which is represented as "1" if it is a fraud and normal is denoted as "0". This dataset is highly unbalanced as frauds represent only 0.172% of all samples (492 frauds present in the total 284,807 transactions). The principal component analysis is used for transforming most of the features. The meanings of other features are not revealed for maintaining confidentiality. Each transaction is independent of others as the cardholder identifier is also not available in the records.

5.2 Training and Test Data

A tenfold CV is deployed to train and test the model and the process is repeated 10 times. The class proportion in the dataset was preserved similar across all the folds by utilizing a stratified CV. The resulting folds were unbalanced due to the imbalance in the original datasets. For each fold of CV, the random forest model was learned twice: one using all the samples and the other with the remaining observations after under-sampling. The training data contains 171,117 samples and test data contains 113,690 records. Both the models learned by the Random Forest are tested on the same testing set.

5.3 Discussion

Experiments were run using the *R* Environment. The primary metrics in the domain of fraud detection are "fraud catching rate" and "false alarm rate". The biased metrics are accuracy and error rate for datasets that are imbalanced. Thus in this paper, the four metrics evaluated are

$$\text{(i) Fraud Catching rate also True Positive rate} = \frac{\text{TP}}{(\text{TP} + \text{FN})} \tag{8}$$

$$\text{(ii) False Positive rate} = \frac{\text{TP}}{(\text{TP} + \text{FP})} \tag{9}$$

$$\text{(iii) Balanced Classification Rate (BCR)} = \frac{1}{2}(\text{TPR} + \text{TNR}) \tag{10}$$

$$\text{(iv) Matthews Correlation Coefficient (MCC)} = \frac{\text{TP} * \text{TN} - \text{FP} * \text{FN}}{\sqrt{(\text{TP} + \text{FN})(\text{TN} + \text{FP})(\text{TP} + \text{FP})(\text{TN} + \text{FN})}} \tag{11}$$

The terms TP, TN, FP, and FN are described as follows:

True Positive (TP) = Number of transactions identified as fake
True Negative (TN) = Number of transactions identified as legal
False Positive (FP) = Number of legal transactions identified as fraud
False Negatives (FN) = Number of fraud transactions identified

In order to perform a comparison with existing approaches, the other derived metrics that are calculated are precision, recall, and f-measure. The equations for obtaining the metrics are given below:

$$\text{Precision} = \frac{\text{TP}}{(\text{TP} + \text{FP})} \tag{12}$$

$$\text{Recall} = \frac{\text{TP}}{(\text{TP} + \text{FN})} \tag{13}$$

$$\text{F-measure} = \frac{2 * \text{Recall} * \text{Precision}}{\text{Recall} + \text{Precision}} \tag{14}$$

The precision measure shows the extent of reliability of the result of the classifier. The recall measure determines the effectiveness of the model in identifying the fake transaction. The harmonic mean of precision and recall is specified as F-measure. The confusion matrix of the test data is shown in Table 2. The true positive rate seen is 99.9% and false positive rate 99.63%. Table 3 shows the various performance metrics obtained using a random forest approach for the proposed credit card fraud identification model.

The proposed model is compared with other classifiers namely Naive Bayes, J48, ID3, Bayes Net, and NB Tree. The existing classifiers are compared with the proposed model and illustrated in Fig. 2. The precision of the proposed model is 99.96 which is the highest of all existing techniques. The recall of the existing classifiers with the proposed model is compared and shown in Fig. 3. The recall of the proposed model is 99.98 which is the highest among all existing techniques. The F-measure of the existing classifiers is compared with the proposed model and shown in Fig. 4. The F-measure of the proposed model is 99.96 which is the highest among all existing techniques. Thus the proposed model achieves better results in comparison to other classification techniques due to the deployment of CFS and ensemble layer of classifiers.

The proposed model can be implemented in other applications for classification as 99.9% of TPR is obtained deploying this model. Whenever there is a better performance in terms of an increase in accuracy and reduced false-positive rate, the technique can be utilized for other real-world problems also.

Table 2 Confusion matrix of the random forest model on credit card dataset

Fraud	Genuine
113,475 (TP)	14 (FN)
42 (FP)	159 (TN)

Table 3 Performance metrics of the credit card fraud identification using random forest approach

Metrics	Value
TPR	99.9
FPR	99.63
BCR	89.48
MCC	85

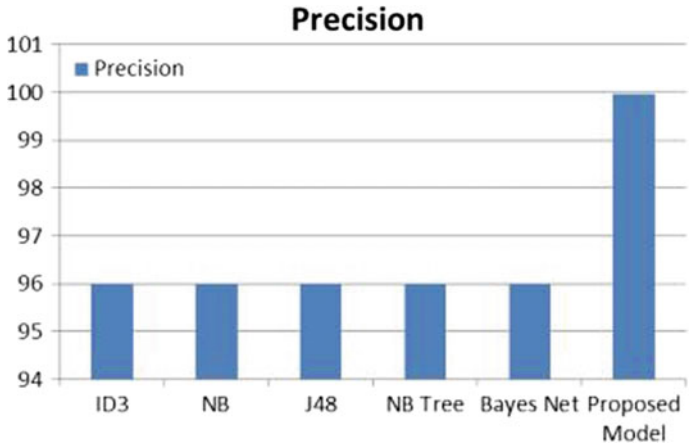


Fig. 2 Comparison of proposed model precision with other classifiers

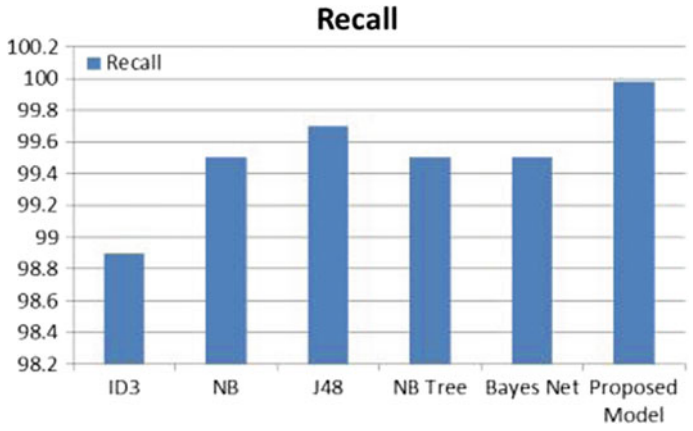


Fig. 3 Comparison of the proposed model's recall with other classifiers

6 Conclusion

Credit card fraud identification is recognized as a serious issue for financial organizations such as banks and credit card companies. Huge losses can be prevented with the expeditious detection of fraudulent transaction. The proposed technique is an integrated model of CFS and an ensemble layer of classifiers. The fraud detection model utilized the benefit of CFS (Correlation-based Feature Selection) such as improving the accuracy of the ensemble classifiers. The optimal attribute set retrieved from CFS is used by KNN, Naive Bayes and Random Forest. The final class label is obtained by a weighted majority voting technique. Experiments on real-time credit card fraud dataset reveals that the proposed approach outperforms the base classifiers such as

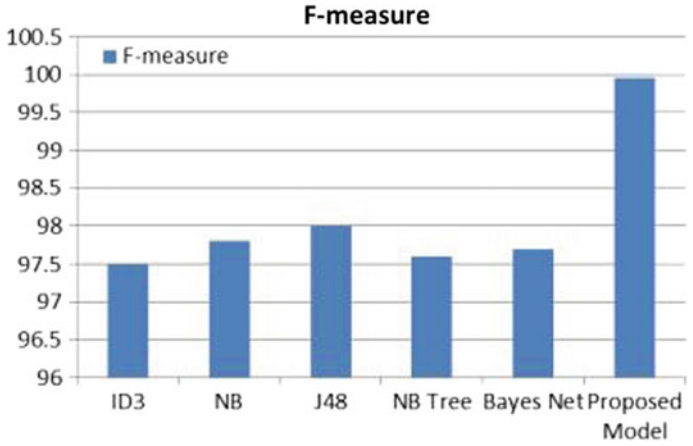


Fig. 4 Comparison of the proposed model' F-measure with other classifiers

Naive Bayes, Bayesian Network, Decision, and NB tree. The fraud detection rate of the proposed model is 99.63% and true positive rate is 99.9% which is superior in comparison to other base classifiers.

References

1. P. Kulkarni, R. Ade, Logistic regression learning model for handling concept drift with unbalanced data in credit card fraud detection system, in *Proceedings of the Second International Conference on Computer and Communication Technologies* (Springer, New Delhi, 2016), pp. 681–689
2. S. Maes, K. Tuyls, B. Vanschoenwinkel, B. Manderick, Credit card fraud detection using Bayesian and neural networks, in *Proceedings of the 1st International NAISO Congress on Neuro Fuzzy Technologies* (2002), pp. 261–270
3. R. Patidar, L. Sharma, Credit card fraud detection using neural network. Int. J. Soft Comput. Eng. (IJSCE) **1**(32–38) (2011)
4. L. Delamaire, H.A.H. Abdou, J. Pointon, Credit card fraud and detection techniques: a review. Banks Bank Syst. **4**(2), 57–68 (2009)
5. E. Duman, M.H. Ozcelik, Detecting credit card fraud by genetic algorithm and scatter search. Expert Syst. Appl. **38**(10), 13057–13063 (2011)
6. M. Zareapoor, P. Shamsolmoali, Application of credit card fraud detection: based on bagging ensemble classifier. Procedia Comput. Sci. **48**(2015), 679–685 (2015)
7. S. Bhattacharyya, S. Jha, K. Tharakunnel, J.C. Westland, Data mining for credit card fraud: a comparative study. Decis. Support Syst. **50**(3), 602–613 (2011)
8. W. Wei, J. Li, L. Cao, Y. Ou, J. Chen, Effective detection of sophisticated online banking fraud on extremely imbalanced data. World Wide Web **16**(4), 449–475 (2013)
9. J. Jurgovsky, M. Granitzer, K. Ziegler, S. Calabretto, P.E. Portier, L. He-Guelton, O. Caelen, Sequence classification for credit-card fraud detection. Expert Syst. Appl. **100**, 234–245 (2018)
10. D.M. Powers, Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation (2011)

11. J. Li, K. Cheng, S. Wang, F. Morstatter, R.P. Trevino, J. Tang, H. Liu, Feature selection: a data perspective. ACM Comput. Surv. (CSUR) **50**(6), 94 (2017)
12. M.A. Hall, Correlation-based feature selection of discrete and numeric class machine learning (2000)
13. A. Dal Pozzolo, O. Caelen, R.A. Johnson, G. Bontempi, Calibrating probability with undersampling for unbalanced classification, in *2015 IEEE Symposium Series on Computational Intelligence* (IEEE, 2015), pp. 159–166

A Model for Predicting Occurrence of Leaf Blast Disease in Rice Crop by Using Fuzzy Logic Techniques



Bhavna Chilwal and P. K. Mishra

1 Introduction

Agriculture is the process that is responsible for feeding the whole world. The food grain crops are the biggest feeder. But this important sector is facing so many factors which are responsible for yield loss. Rice crop is one of the most important crops among food grains. But there are many diseases that are affecting the productivity of this crop. Leaf blast is one of the major fungal diseases which has the capability to destroy 70% crop. So the detection of the occurrence of this disease could help farmers a lot to save their crops at the right time by applying the right treatments. The manual detection of diseases is time-consuming. Here we use the Fuzzy logic system as it is easy to understand, flexible in nature, and the results generated are satisfactory as compared to manual grading. This paper introduces a model that detect the level of disease and then predicts the occurrence of disease in rice crop leaves so that required measures and treatments could be applied to save the crop from productivity loss. The fuzzy logic system is used in two levels and the line of the equation is used to generate the dataset which is imported in ANFIS from the final model. MATLAB fuzzy logic toolbox and ANFIS are used and MS Excel is for finding the correlation and accuracy for the dataset.

B. Chilwal (✉) · P. K. Mishra
Department of Computer Engineering, GBPUAT, Pantnagar, India
e-mail: bhavnachilwal@gmail.com

P. K. Mishra
e-mail: pkmishra.cs@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_3

2 Literature Review

This section has discussed some works in which Fuzzy Logic has been used in agriculture for different situations like Pribira Kumar et al. [1]. This study develops a prototype which evaluates the severity of disease on rice crop by using computational intelligence and machine learning. This paper introduces K means segmentation with fuzzy logic to calculate the degree of the disease occurred in rice plants. The prototype has 86% accuracy. Neamtollahi and Vafabakshi [2], This paper uses fuzzy logic for getting the best cropping patterns in the farming sector. It is important to join economic principles to ecological principles to get an optimized model. The results showed that there is a need for changing the patterns and use the proposed ecological patterns for maximizing crop productivity. Awoyelu and Adebisi [3] discuss a model of the fuzzy expert system for predicting the severity of Cassava crop diseases. The disease's occurrence of diseases in the crop affects their yield and productivity. Here this paper proposed the fuzzy system for predicting diseases. The system was made by using the fuzzy tool available in MATLAB, 18 rules are applied for Mosaic disease, 27 rules for brown streak disease, and 25+ rules for bacterial blight for prediction and classification. This model will provide information about the possible disease. Ingole and Katole [4], This paper explores the applications of fuzzy in predicting crop, i.e., wheat productivity by using remote sensing RS. The fuzzy logic mimics the human reasoning ability and it contains rules by using multiple parameters for prediction. And by using different parameters it predicts the important conditions required for a particular crop.

3 Proposed Methodology

This paper shows the efficient use of fuzzy logic and ANFIS to detect the occurrence of leaf blast disease in rice crops. The flow graph in Fig. 1 shows the step to explain the proposed disease prediction model:

3.1 Input Parameters

We have to use a fuzzy inference system on two levels. The first level system has 3 input parameters which are the features of rice crop plants during the disease time. The parameters are taken from the Standard Evolution System, IRRI 2002 and thought as the correct measurements for leaf blast. The inputs for the first-level fuzzy system are the following:

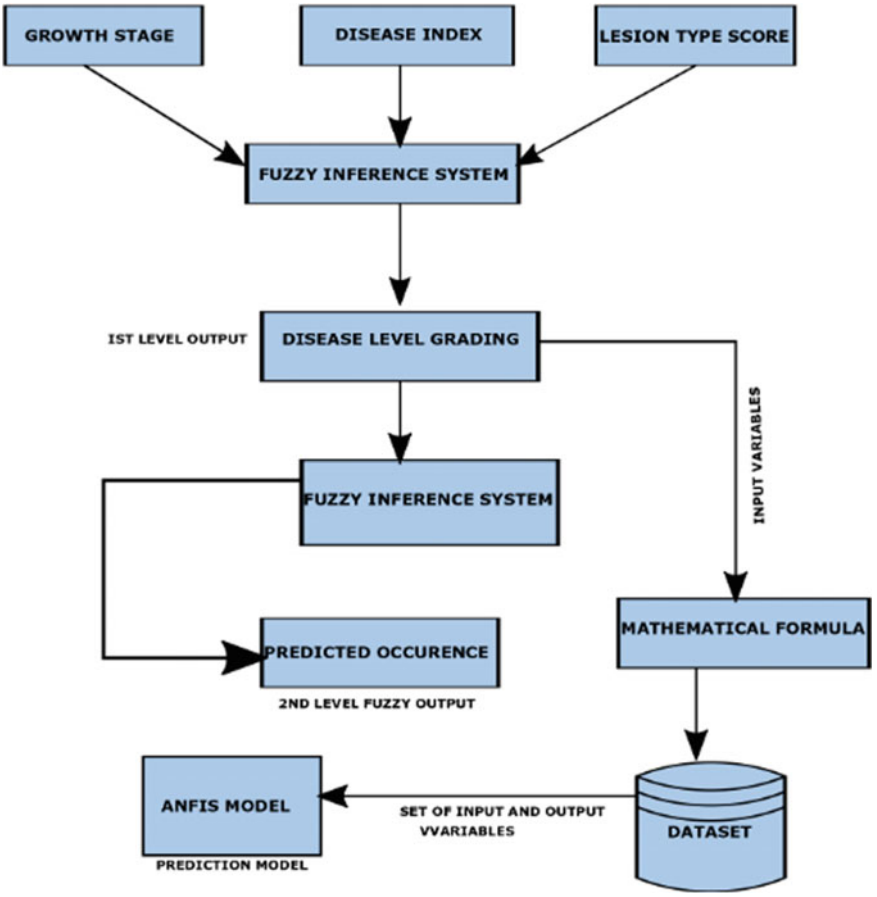


Fig. 1 The proposed disease prediction model

Growth Stage of Rice Plant (Standard Evaluation System of Rice (SES), IRRI, 2002)

Code	
1	Germination
2	Seedling
3	Tillering
4	Stem elongation
5	Booting
6	Heading
7	Milk stage

(continued)

(continued)

Code	
8	Dough stage
9	Mature grain

The growth stage parameter has nine codes for describing the stage of a plant according to IRRI, so for the fuzzy toolbox, we divided these 9 growth stages into three phases: Germinating Phase (1), Vegetative Phase (2–4), and Reproductive Phase (5–9).

Disease Index (Standard Evaluation System of Rice (SES), IRRI, 2002)

$$\text{PDI} = \frac{\text{Sum of individual rating}}{\text{Number of leaves assessed} \times \text{Maximum disease grade value}} \times 100$$

The leaf blast occurrence is calculated by the given formula above.
The scale for measuring the amount of disease percent in observed leaves is

Scale	(% infection)
0	No symptom observed
1	1–10%
3	11–30%
5	31–50%
7	51–70%
9	71–100%

Lesion Types (Standard Evaluation System of Rice (SES), IRRI, 2002)
Lesions are generally started from leaf tips or margins sometimes both and later extend to outer edges. The color of lesion changes with time during starting it is of pale green to grayish-green, then it turns to yellow and later gray (dead) with time. The scale for measuring the lesion type is given by IRRI is:

Scale range (0–9)
0 = No lesions
1 = Small brown specks of pinhead size without sporulating center
2 = Small roundish to slightly elongated, necrotic gray spots, about l–2 mm in diameter with a distinct brown margin and lesions are mostly found or lower leaves
3 = Lesions type is the same as in Scale 2, but a significant number of lesions are or the upper leaves
4 = Typical sporulating blast lesions, 3 mm or longer, infecting less than 2% of the leaf area

(continued)

(continued)

Scale range (0–9)
5 = Typical blast lesions infecting 2–10% of the leaf area
6 = Blast lesions infecting 11–25% leaf area
7 = Blast lesions infecting 26–50% leaf area
8 = Blast lesions infecting 51–75% leaf area
9 = More than 75% leaf area affected

3.2 Output

The output of the first level fuzzy system is the Level grading of disease occurrence. In this, the division of levels is in 3 and the output range is taken with membership function is [0–75]. In which

Level 1—[0.1–25] (output in this range has to be in Grade Level 1)
Level 2—[25.1–50] (output in this range has grading level 2)
Level 3—[50.1–75] (output in this range has grading level 3)

The rules are then generated which maps the inputs to the required output. The fuzzy system has a rule generator that takes IF…THEN statements using inputs and outputs, also have AND … OR functions to make different forms of one rule. Here for this model, only 15 rules were made for mapping the input to output.

The output of the first-level fuzzy system is then used as an input in second-level fuzzy inference system and output for the second level is the prediction percent of disease severity in a plant leaf.

The output range is taken as 0–100 (for percentage reference). The ranges provided for input and the output in second-level FIS and their relation with member function (trimf) is used in the mathematical formula for a straight line to prove the correlation between the obtained level grading (input) with the prediction percent of occurrence(output). The line of the equation for this model is

$$y = mx + b \tag{1}$$

y = Dependent variable, i.e., output (should be in range 0–100)
x = Independent variable, i.e., input (range 0–75).

4 Experimental Results

For result section to obtain the required equation for our model the following formula is used by using the input–output ranges:

$$(y-y1)/(y1-y2) = (x - x1)/(x1-x2)$$
$$(y1 \text{ should preferably be equal to } 0) \tag{2}$$

This formula is applied to the membership function of the input variable and membership function of output variable and after experimenting with the above step the required equation for this model is:

$$y = 1.199x + 0.3385 \tag{3}$$

After using this equation a dataset will be obtained in which every input value corresponds to every output value. To check the correlation between the two variables MS Excel Correlation function is used and the obtained correlation value is 99%, and to find the accuracy of this dataset the standard error is used which is found to be 0.17 and the accuracy formula is

$$\text{Accuracy} = (1 - \text{Standard Error})[1 \text{ is } 100\%] \tag{4}$$

So the accuracy obtained is 83%.

The dataset we get is then applied to the ANFIS function of the Matlab to obtain the model which corresponds to the grading level of diseases and their predicting disease occurrence value (Fig. 2 and Table 1).

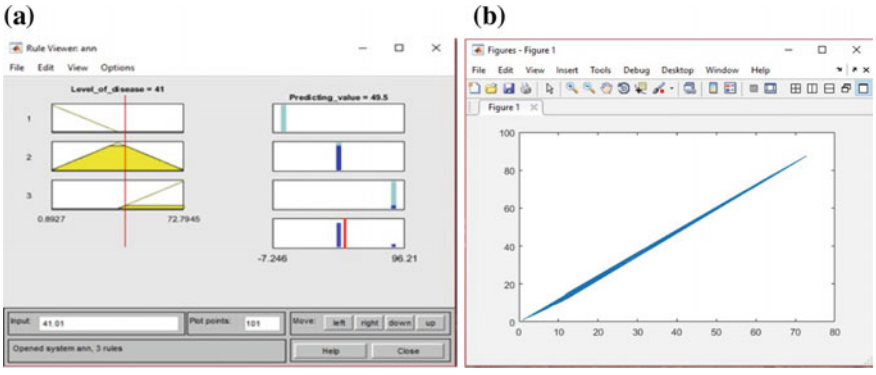


Fig. 2 **a** ANFIS model **b** graph plot

Table 1 Disease grading level and its predicted risk

Disease level	Range value [0–75]	Predicted risk	Percentage (%) Range [0–100]
No infection	0	No risk	0
Level 1	Between 1 and 25	Low risk	Between 1 and 30
Level 2	Between 25 and 50	Medium risk	Between 30 and 60
Level 3	Between 50 and 75	High risk	Between 60 and 100

5 Conclusions

This paper presents a model for measurement of leaf blast disease occurrence in rice crop plants and provides level grades of severity and corresponding predicting the percent of disease by using Fuzzy logic. The proposed model estimates the disease severity and by using mathematical formula the accuracy obtained is 83%. This model is computationally easy to understand and also easy to implement. This model is fast and accurate to predict disease risk in different stages of the plant. The future scope of this study applies more rules and can also be extended for another type of disease with different parameters or with a large dataset or by applying the different algorithms and mathematical model.

References

1. P.K. Sethy, B. Negi, Measurement of disease severity of rice crop using machine learning and computational intelligence, in *SpringerBriefs in Applied Sciences and Technology* (2018)
2. E. Neamatollahi, J. Vafabakshi, Agricultural optimal cropping pattern determination based on fuzzy system. Fuzzy Inf. Eng. (2017), www.elsevier.com. Accessed 14 Jan 2019
3. I.O. Awoyelu, R.O. Adebisi, A predictive fuzzy expert system for diagnosis of cassava plant diseases. Glob. J. Sci. Front. Res. C Biol. Sci. **15**(5) (2015) (Version 1.0 the Year)
4. K. Ingole, K. Katole, Crop prediction and detection using fuzzy logic in MATLAB. Int. J. Adv. Eng. Technol. IJAET (2013)

Improving the Effectiveness of Moving Target Defenses by Amplifying Randomization



Vaishali Kansal and Mayank Dave

1 Introduction

The constant nature of cyber systems gives the attacker enough time to plan attacks. To counter this threat, Moving Target Defense (MTD) changes the network constantly by dynamically shifting the attack surface area. By making random adaptations over time in the network configuration, MTD confuses the potential attacker thereby increasing complexity for attackers. The MTD forces the attacker to operate under uncertainty and unpredictability. It also attempts to develop resilient hardware which continues to operate while under attack.

MTD techniques can be applied at various levels of the system:

1. *Host level MTD*: Changing the host and OS level resources, naming and configuration.
2. *Network level MTD*: Changing the network topology, including IP-hopping and fake information about OS type and version.
3. *Application level MTD*: Changing the application environment like changing application type and routing them through various hosts, randomly arranging memory layout, altering the source code at every compilation, or changing programming languages to compile the source code.

Various MTD techniques are initiated so far to enhance security but effectiveness of these techniques is not well accessed. Effectiveness here describes the ability to enhance the security of the system by minimizing the efforts of the defender while maximizing the efforts of the attacker.

V. Kansal (✉)
Department of Computer Engineering, THDC Institute of Hydropower Engineering and Technology, New Tehri, India
e-mail: veshukansal@gmail.com

M. Dave
Department of Computer Engineering, National Institute of Technology, Kurukshetra, India
e-mail: mdave@nitkkr.ac.in

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_4

In this paper, an approach is presented that incorporates diversity at OS layer with shuffle-based MTD technique. It exhausts the attacker as it changes the attack surface by changing the vulnerability information with every shuffling.

2 Related Work

MTD has provided a promising way to network security. Jason et al. [1] productized moving target defense research technology, called Self-shielding Dynamic Network Architecture (SDNA) technology, currently known as Cryptonite NXT.

Rui et al. [2–4] presented a basic design schema of a MTD system. They have discussed three essential problems of MTD systems, which include the MTD Problem, the Adaptation Selection Problem, and the Timing Problem. They explained the effects of introducing randomization and diversification in the system with the help of simulation-based experiments.

MTDs have gained significant attention in the past few years and are attracting a growing number of researchers and system designers with the increasing adoption of several technologies such as program manipulation [5], compiler generated software diversity [6], Address Space Layout Randomization (ASLR) [7], Instruction Set Randomization (ISR) [8], IP address randomization [9, 10].

The authors in [11] proposed an approach that uses operating system level virtualization and portable checkpoint compilation for creating a virtual execution environment and for migrating a running application across different platforms while preserving the state of the application like open files, execution state, and network connections. Hong et al. [12] incorporated the MTD techniques Shuffle, Diversity, and Redundancy into the security model to assess their effectiveness. Jun et al. [13] compared the effectiveness of different MTDs using a three-layer model.

MTD has its potential applications in various areas. It has been used as a DDoS defense mechanism in [14, 15] where shuffling technique is adopted to isolate malicious users from innocent users. Also, the approach Early Detection and Isolation Policy(EDIP) proposed in [16, 17] is an efficient proactive approach that mitigates insider assisted DDoS attack using only shuffle MTD. But considering "greater the entropy of the system's configuration, results more effective MTD system", we have proposed an approach in this paper which integrates diversity at OS layer with shuffle technique to increase randomization.

3 Various MTD Techniques

Numerous MTD techniques have been proposed in various domains based on the kind of changes made to the system. Figure 1 shows the taxonomy of various MTD techniques discussed by researchers.

1. *Software-Based Diversification*:
   (a) *Software Manipulation* [5]: Switching between various implementations of the same function in a program and thus generating system variations.
   (b) *Compiler Generated Diversity* [6]: Applying compiler conversions to produce different variants of machine codes like redundancy elimination, loop distribution, etc.
2. *Run-time Based Diversification*:
   (a) *Address Space Layout Randomization (ASLR)* [7]: Randomize the location of objects in the memory.
   (b) *Instruction Set Randomization (ISR)* [8]: In this, during compilation, a randomization key is used to encode operators in machine instructions. Then just before execution, operators in randomized instructions are required to be correctly decoded.
3. *Communication Diversification*: Protect system against network-related attacks by changing network configuration parameters like IP addresses and communication protocols.
   (a) *IP Address Randomization (IPR)* [9, 10]: Changing IP addresses randomly and frequently
4. *Dynamic Platform Diversification*:
   (a) *Shuffle-Based* [12, 14, 15]: Shuffle rearranges system settings like topology rearrangement, migration, and virtual machine rotation.
   (b) *Diversity-Based* [12]: Diversity provides equivalent functions with different implementations, for e.g., OS level virtualization.
   (c) *Redundancy-Based* [12]: Redundancy provides multiple replicas of the network components like servers with diverse software to increase availability and diversity.

4 Proposed Work

4.1 Overview

We have proposed a model that consists of four interconnected components: application server, authentication server, layer of proxy servers, and end user. Figure 2 shows the proposed model.

The application server provides the online services. The authentication server performs various tasks like

– authenticate clients,
– assign proxy to clients,

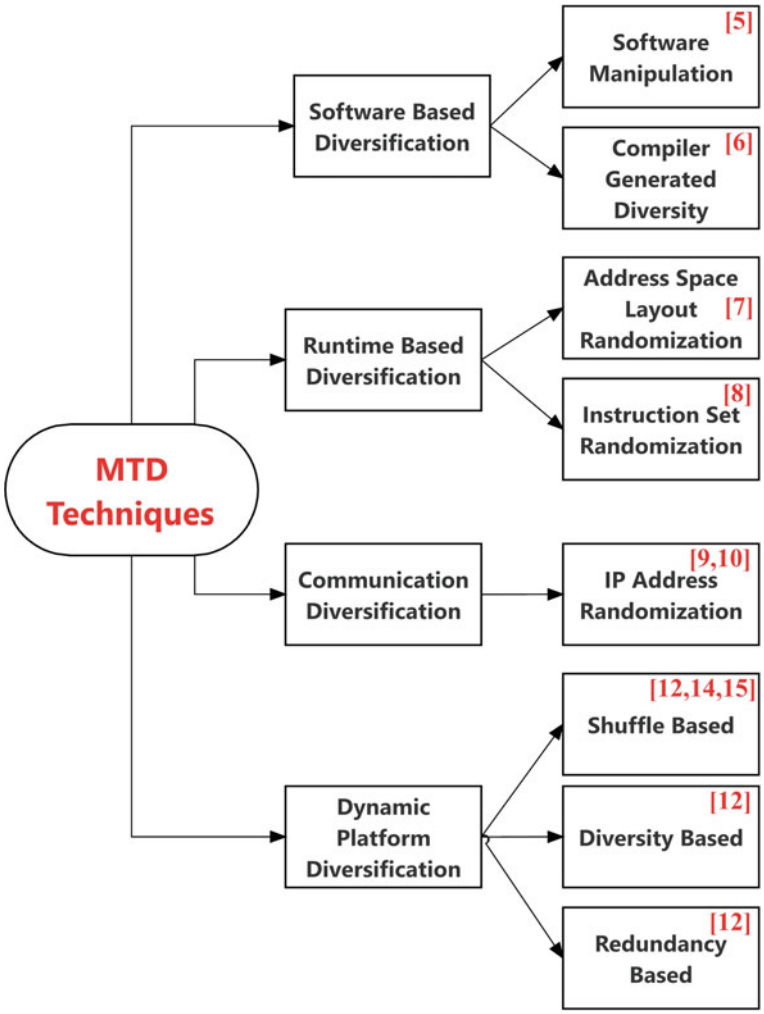


Fig. 1 Taxonomy

– detect proxies under attack,
– discover OS running on attacked proxies,
– instantiate new proxies running different OS, and
– coordinate shuffling of clients.

The layer of proxy servers is used to relay traffic between application server and clients. Only one proxy is assigned to a client but a proxy may have several clients connected to it. A client knows the IP address of assigned proxy only while other IP addresses are hidden. So, a malicious client can attack only on the assigned proxy.

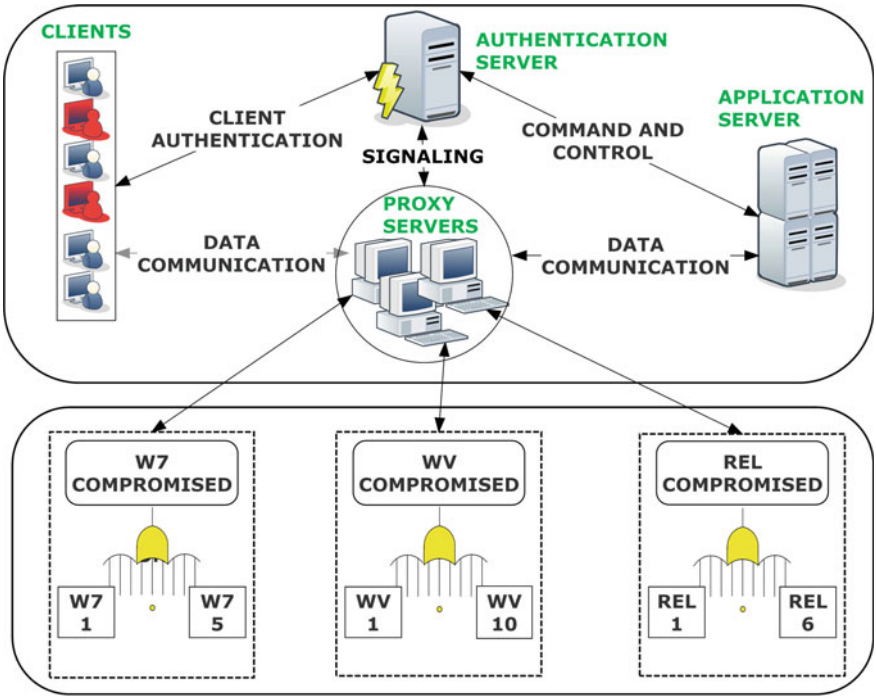


Fig. 2 Proposed model

We have considered that all proxy servers are running either Windows 7, Windows Vista, or Redhat Enterprise Linux associated with different vulnerabilities and the attacker can also exploit these OS vulnerabilities. There are five "Windows 7" vulnerabilities, ten "Windows Vista" vulnerabilities, and six "Redhat Enterprise Linux" vulnerabilities defined in [12]. We have modeled these OS vulnerabilities in our proposed model.

4.2 Client-to-Proxy Assignment

The proposed architecture allows client to access online services via a layer of proxy servers. Clients request access to authentication server, which in response authenticate clients using some cryptographic puzzles. For n number of clients and m number of available proxies, authentication server then runs Algorithm 1 to assign proxy to clients.

All proxy servers have a Dirty flag variable and a PC[ ] array. If Dirty flag for any proxy is set to "1" , then it indicates that it is under attack, otherwise not. PC[ ] array for a proxy contains the id of all the clients that are assigned to it. Authentication

Algorithm 1 Client-To Proxy Assignment

```
1: function INITIAL- ASSIGNMENT(P[1...m],C[1...n])
2:     for proxy i = 1 to m do
3:         DF[i] ← 0 ;
4:         PC_i[ ] ← 0 ;
5:     end for
6:     for client i = 1 to n do
7:         Assign any proxy k to client i ;
8:         Add i to PC_k[ ] ;
9:     end for
10: end function
```

server initializes Dirty flag and PC[ ] of all proxies to "0". After assignment of any proxy to a client, authentication server adds the client id into the PC[ ] of that proxy. Figure 3 shows the initial client-to-proxy assignment.

4.3 Proxy Server Under Attack

If any proxy server is under attack that is if one or more malicious client is connected to it, then that proxy server sets its Dirty Flag as "1". Illustration of proxy servers under attack because of the presence of malicious users in the system is shown in Fig. 4.

4.4 Attack Detection and Amplifying Randomization

Authentication server runs DETECT() function of Algorithm 2 at regular intervals to check the presence of malicious user. If it finds Dirty flag set to "1" for any proxy

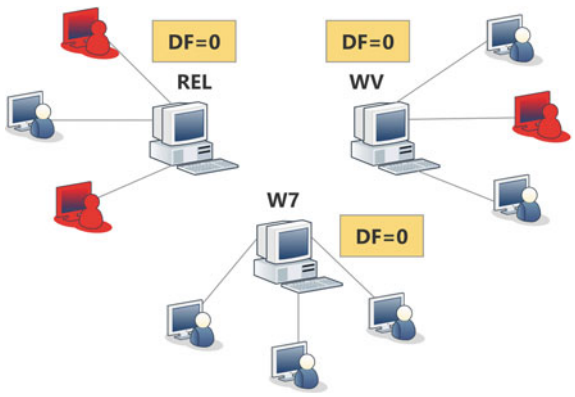


Fig. 3 Initial assignment

server, then it further calls SHUFFLE-WITH-DIVERSITY() function passing the attacked proxy server id and PC[ ] array of that attacked proxy as arguments in the function.

Algorithm 2 Attacker Detection and Adding Randomization

```
1: function DETECT
2:     for proxy i = 1 to m do
3:         if DF[i] == 1 then
4:             SHUFFLE-WITH-DIVERSITY(i, PC_i[ ])
5:         end if
6:     end for
7: end function
8: function SHUFFLE- WITH- DIVERSITY(i, PC_i[ ])
9:     Find OS running on proxy i ;
10:    Initialize new proxy p with different OS ;
11:    Assign PC_i[ ] clients to new initialized proxy p ;
12:    PC_p[ ] ← PC_i[ ] ;
13:    PC_i[ ] ← φ ;
14:    DF[i] ← φ ;
15:    Shutdown proxy i ;
16: end function
```

In SHUFFLE-WITH-DIVERSITY() function, it discovers the OS running on the attacked proxy and instantiate new proxy server running different OS. Then it shifts the clients which were assigned to attacked proxy to newly instantiated proxy and set PC[ ] array of new proxy equal to PC[ ] of attacked proxy. Further, it abandons the attacked proxy server and set its Dirty flag and PC[ ] array to ϕ indicating that it no longer serves the clients. Figure 5 shows how authentication server coordinates the shuffling of clients.

Therefore to deal with the OS vulnerabilities, we have used shuffle-based MTD technique with diversity-based MTD technique. We are shuffling the malicious clients

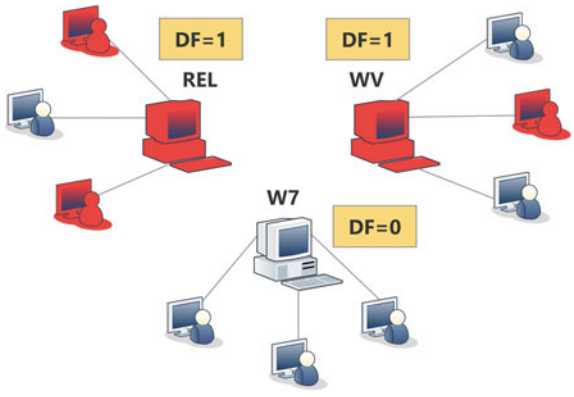


Fig. 4 Malicious user present

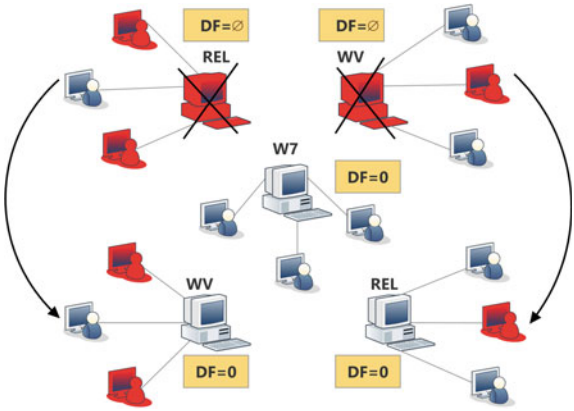


Fig. 5 Shuffling with diversity

Table 1 Comparison table

Factors/Approaches	Shuffle-based [12, 14, 15]	Diversity-based [12]	Shuffle-with-Diversity (SwD)
MTD technique	Only shuffle	Only diversity	Shuffle and diversity
Entropy	Lower than SwD	Lower than SwD	High
Attack cost	Lower than SwD	Lower than SwD	High
Attacker efforts	Requires less efforts of attacker	Requires less efforts of attacker	Requires more efforts of attacker
Complexity for attackers	Less complex than SwD	Less complex than SwD	More complex

among proxy servers running different OS. We are applying diversity on proxy servers at OS layer that will only change the vulnerability information providing the same functionality. This will result in change of attack surface that will force the attacker to use different exploits for different vulnerabilities with an assumption of nonoverlapping vulnerabilities. Thus, it increases uncertainty and complexity for attackers.

5 Security Analysis

Various approaches to enhance the security of the system have been developed but in most of the approaches, either shuffling is deployed or only diversity is deployed. Table 1 shows the comparison of the proposed approach Shuffle with Diversity (SwD) with other approaches.

Shuffle-Based: Deploying only shuffling technique results in shuffling of clients over different proxies irrespective of the vulnerability information.

Diversity-Based: Deploying only diversity technique results in equivalent functions with different implementations, for e.g., OS level virtualization without performing migration.

Shuffle-with-Diversity (SwD): The approach presented in this paper integrates shuffle MTD technique and diversity MTD technique. This approach has several advantages over other proposed approaches. Shuffling attacker again and again to new proxy server running different OS associated with different vulnerabilities increases complexity and attacker efforts. After every shuffle, attacker requires accurate information about the operating system running on the target proxy server to launch a successful attack. Thus, the previously compromised OS vulnerabilities provide no advantage to attacker increasing attack cost for attacker.

6 Conclusion

To enhance security in the cyber world, only a single MTD technique has been deployed so far, and the attacker who can break that defense will be successful. In order to overcome this, multiple techniques are required to be composed that force the attacker to incrementally break each of them. In this paper, we have proposed an integrated approach that combines shuffle-based MTD technique with diversity-based MTD technique applied at OS layer of proxy servers. This combination will result in improved effectiveness as it will amplify the value of entropy by dynamically re-diversifying each technique in an interleaved way.

References

1. L. Jason, J. Yackoski, N. Evancich, Moving target defense: a journey from idea to Product, in *Proceedings of the 2016 ACM Workshop on Moving Target Defense* (ACM, 2016), pp. 69–79
2. Z. Rui, S. Zhang, S.A. DeLoach, X. Ou, A. Singhal, Simulation-based approaches to studying effectiveness of moving-target network defense, in *National Symposium on Moving Target Research* (NIST, 2012), pp. 1–12
3. Z. Rui, S. Zhang, A. Bardas, S.A. DeLoach, X. Ou, A. Singhal, Investigating the application of moving target defenses to network security, in *2013 6th International Symposium Resilient Control Systems (ISRCS)* (IEEE, 2013), pp. 162–169
4. Z. Rui, S.A. DeLoach, X. Ou, Towards a theory of moving target defense, in *Proceedings of the First ACM Workshop on Moving Target Defense* (ACM, 2014), pp. 31–40
5. R. Martin, Manipulating program functionality to eliminate security vulnerabilities, in *Moving Target Defense* (Springer, New York, 2011), pp. 109–115
6. J. Todd, B. Salamat, A. Homescu, K. Manivannan, G. Wagner, A. Gal, S. Brunthaler, C. Wimmer, M. Franz, Compiler-generated software diversity, in *Moving Target Defense* (Springer, New York, 2011), pp. 77–98
7. S. Hovav, M. Page, B. Pfaff, E.-J. Goh, N. Modadugu, D. Boneh, On the effectiveness of address-space randomization, in *Proceedings of the 11th ACM Conference on Computer and Communications Security* (ACM, 2004), pp. 298–307

8. Kc. Gaurav, S., A.D. Keromytis, V. Prevelakis, Countering code-injection attacks with instruction-set randomization, in *Proceedings of the 10th ACM Conference on Computer and Communications Security* (ACM, 2003), pp. 272–280
9. A.-S. Ehab, Toward network configuration randomization for moving target defense, in *Moving Target Defense* (Springer, New York, 2011), pp. 153–159
10. Z. Jianjun, A.S. Namin, The impact of address changes and host diversity on the effectiveness of moving target defense strategy, in *2016 IEEE 40th Annual Computer Software and Applications Conference (COMPSAC)*, vol. 2 (IEEE, 2016) pp. 553–558
11. O. Hamed, A. Comella, E. Robinson, J. Haines, Creating a cyber moving target for critical infrastructure applications using platform diversity. Int. J. Crit. Infrastruct. Prot. **5**(1), 30–39 (2012). Elsevier
12. J.B. Hong, D.S. Kim, Assessing the effectiveness of moving target defenses using security models. IEEE Trans. Dependable Secure Comput. **13**(2), 163–177 (2016). IEEE
13. X. Jun, P. Guo, M. Zhao, R.F. Erbacher, M. Zhu, P. Liu, Comparing different moving target defense techniques, in *Proceedings of the First ACM Workshop on Moving Target Defense* (ACM, 2014), pp. 97–107
14. J. Quan, K. Sun, A. Stavrou, Motag: moving target defense against internet denial of service attacks, in *2013 22nd International Conference on Computer Communication and Networks (ICCCN)* (IEEE, 2013), pp. 1–9
15. W. Huangxin, Q. Jia, D. Fleck, W. Powell, F. Li, A. Stavrou, A moving target DDoS defense mechanism. Comput. Commun. **46**, 10–21 (2014). Elsevier
16. K. Vaishali, M. Dave, DDoS attack isolation using moving target defense, in *2017 International Conference on Computing, Communication and Automation (ICCCA)* (IEEE, 2017), pp. 511–514
17. K. Vaishali, M. Dave, Proactive DDoS Attack detection and isolation, in *2017 International Conference on Computer, Communications and Electronics (Comptelix 2017)* (IEEE, 2017), pp. 334–338

Remaining Life Assessment of Solid Insulation in Power Transformer Using Fuzzy Inference System (FIS)



Deepak Kanumuri, Veena Sharma and O. P. Rahi

1 Introduction

For solid insulation of power transformer which is in-service operated after its designed life of 25–30 years, it is essential to evaluate its status accurately; results obtained from this information is used in the prediction of remaining life and further, to arrange its maintenance scheme. Transformer insulation is majorly based on oil–paper insulation, which contains insulating oil, cellulose insulation paper and pressboards. Usually, degradation in solid insulation is more serious than that of transformer insulating oil because oil can be refurbished and replaced whereas paper insulation cannot be replaced. Hydrolytic, thermal and oxidative degradation are some forms of degrading mechanisms in cellulose. Cellulose is the basic element of paper, which is a glucose polymer where glucose units are linked to each other in the form of a chain structure. Cellulose is chemically represented as $[C_6\,H_{10}\,O_5]_n$, where 'n' represents the number of glucose units per polymeric chain, more specifically described as Degree of Polymerization (DP), an important factor of cellulose paper degradation. In some research works, it is observed that there are direct relations between the tensile strength of the insulating paper to the corresponding DP value. Generally, cellulose paper has a DP of 1200–200 when it is subjected to stress below which life in cellulose paper ends. During cellulose ageing, it is decomposed into different substances such as furanic compounds, acids, CO_2 and CO. Health monitoring and diagnostic approaches based on the ratio of carbon dioxide and carbon monoxide (i.e. CO_2/CO) gas values were investigated in [1]. Some researchers have given mathematical models which use 2FAL to assess the cellulose paper status in

D. Kanumuri · V. Sharma (✉) · O. P. Rahi
Department of Electrical Engineering, NIT Hamirpur, Hamirpur 177005, HP, India
e-mail: veena@nith.ac.in

D. Kanumuri
e-mail: deepakkanumuri95@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_5

a power transformer. In this work, the objective is to assess the status of Cellulose Insulation Paper (CIP) using fuzzy inference (FIS) model in which the concentrations of 2FAL, CO and CO_2 are considered as inputs with the value of DP as output and also estimate the DP with the help of mathematical models.

2 Calculation of DP Using Mathematical Models Formulated by Correlation Between 2FAL and DP of Paper Insulation

The collection of paper insulation sample from a transformer is quite difficult, especially when the transformer is still expected to continue its service, and it might lead to unit failure if it is not done properly with appropriate skill. Evaluation of paper insulation condition without exposing it to such a risk is desirable. It has been found that indirect tests could be done by analysing furanic compounds present in the oil, which are decomposed during the ageing process and through this the DP can be estimated. Although measuring of furanic compounds from an oil sample is relatively simple, its interpretation is very complex. Carbon oxide gases and moisture at low temperatures are the dominant ageing products. Furanic compounds at intermediate temperatures are dominant and at high temperatures are unstable [2]. Several researchers have been studying paper ageing and have attempted to correlate furans to the DP value [3–8]. Mathematical models relating to 2FAL and DP have been developed from laboratory values by these researchers.

Chengdong [3], Burton [4] and Vaurchex [5] were the first researchers to investigate the correlation between the log value of 2FAL and DP. De Pablo [6] and Pahlvanpour have illustrated five different equations correlating 2FAL concentration to DP. These equations were obtained considering different values given by laboratories.

Li and Song [7], this model was formed while determining the health index and considering their weights by qualitative analysis. In this method, it takes contents of CO and CO_2 gases and then quantifies them with piecewise linear function. The furfural content gives information about health index through these weights and functions, equation between 2FAL and DP was given.

Chaohui [8], here CO and CO_2 data in 57 transformers are normalized first and then they were used in the logical analysis as dependent variables. The regression equations formed have established the average DP that can estimate by furfural in the given equation (Tables 1 and 2).

Table 1 Mathematical model based on correlation between 2FAL and DP

S. No.	Author	Mathematical model	S. No.	Author	Mathematical model
(1)	Chengdong [3]	$DP = \frac{1.51 - \log_{10}(2FAL)}{0.0035}$	(4)	De Pablo [6]	$DP = \frac{7100}{8.88 + (2FAL)}$
(2)	Burton [4]	$DP = \frac{2.5 - \log_{10}(2FAL)}{0.005}$	(5)	Li and Song [7]	$DP = -121 * \ln(2FAL) + 458$
(3)	Vaurchex [5]	$DP = \frac{2.6 - \log_{10}(2FAL)}{0.0049}$	(6)	Chaohui [8]	$DP = 405.25 - 347.22 * \log_{10}(2FAL)$

Table 2 Estimated DP from mathematical models

S. No.	2FAL (ppm)	Chengdong [3]	Burton [4]	Vaurchex [5]	De Pablo [6]	Li and Song [7]	Chaohui [8]
1.	0.05	803.15	760.21	796.13	795.07	820.48	901.99
2.	0.06	780.53	744.37	779.97	794.18	798.42	874.50
3.	0.08	744.83	719.38	754.47	792.41	763.61	831.12
4.	0.09	730.22	709.15	744.03	791.53	749.36	813.36
5.	0.1	717.14	700	734.69	790.65	736.61	797.47
6.	0.2	631.13	639.80	673.26	781.94	652.74	692.95
7.	0.5	517.44	560.21	592.05	756.93	541.87	554.77
8.	1.72	364.13	452.90	482.55	669.81	392.378	368.47
9.	3	295.11	404.58	433.24	597.64	325.07	284.58
10.	3.85	264.15	382.9	411.13	557.74	294.88	246.97
11.	5	231.72	360.21	387.97	511.53	263.29	207.55
12.	6	209.10	344.37	371.81	477.15	241.20	180.06

3 Calculation of DP Using Fuzzy Logic System

Condition monitoring methods which are based on the ratio of CO_2 and CO gas values were discussed in [1]. Computing techniques like Artificial Neural Network (ANN), Fuzzy Logic (FL), Artificial Neural Fuzzy System (ANFIS), Wavelet Network and Support Vector Machine (SVM) were used to evaluate the health of insulating paper. These techniques use furaldehyde and carbon oxide concentrations for degradation evaluation in cellulose insulating paper.

3.1 Fuzzy Logic

The interpretation of transformer insulation health can be suitably analysed by fuzzy logic. The fuzzy-based logic helps to obtain an improved health evaluation which assures power transformer life maintainability, consistency and accessibility. The proposed scheme using Fuzzy Inference System (FIS) uses a graphical user interface (GUI) tool which was present and provided by MATLAB. Fuzzification of inputs and output is categorized into various sets whose ranges are varied from very low to high. A set of rules are used in fuzzy logic for analysing insulation condition in transformers and these rules play an important role in health diagnosis by using mapped fuzzified inputs and fuzzified output.

Types of membership functions. A fuzzy set is completely characterized by its MF (membership function). Types of membership functions [9] are triangular membership function, trapezoidal membership function, gauss membership function and gbell membership function. The triangular membership function is specified by three parameters $\{a, b, c\}$ and these parameters determine the x coordinates of the three corners underlying in triangular MF. Similarly, trapezoidal MF is specified by four parameters $\{a, b, c, d\}$ and these determine the x coordinates of the four corners in trapezoidal MF. Due to their simple formulas and computational efficiency, both triangular and trapezoidal MF has been frequently used, especially in real-time implementations. The rules for its parameterization are discussed in [9]. Since these MFs are straight line segments, in the corners it is not smooth so we introduce other membership functions. Gaussian and generalized bell (gbell) are the remaining MFs. In Gaussian, it is specified by two parameters $\{c, \sigma\}$ where c represents centre of MF and σ determines the MF width. In generalized bell MF, it is specified by three parameters $\{a, b, c\}$ where b is usually a positive parameter and this MF is the generalization of Cauchy distribution and is used in probability theory. The rules of these MFs are discussed in [9].

Fuzzy expert rules. The FIS model is built upon a set of rules in the form of linguistic variables. FIS model development is done by using 'IF-THEN' type of rules which are based on a set of linguistic variable rules on expert's knowledge, also known as implication [9]. FIS model maps space between inputs and outputs in an appropriate way; this will give mathematical strength to the design. The principle of cognitive uncertainty present in the fuzzy system gives an inference mechanism. The fuzzy expert rules formation is done by maintaining the correlation between carbon oxide gases, 2FAL and DP [10].

Defuzzification. This method brings out the desired value in the form of crisp output. It mostly applies the principle of centre of gravity in the fuzzy region on an account of weighted mean. The output which is desired is gained by using the formula discussed in [9]; this gives crisp output value of DP which is the output that predicts the condition of insulating paper. From the obtained value, condition of insulation can be determined as per some predefined ranges in MFs; these ranges and their corresponding conditions are determined in Table 3.

Table 3 Range of 2FAL, CO_2, CO and DP as stated in the IEEE standard C57.104^{TM}

2FAL (ppm)	CO_2 (ppm)	CO (ppm)	DP	Condition
0–0.1	0–2500	0–350	1200–700	Healthy transformer
0.1–1	2500–4000	350–570	700–450	Significant concern
1–10	4000–10000	570–1400	450–250	Exceedingly investigable
≥10	≥10000	≥1400	<250	End life

3.2 Scheme for Calculating DP Using Fuzzy Inference System (FIS)

In this work, 2FAL, CO and CO_2 concentrations are correlated to the values of DP and the FIS model is created. These concentrations are formed due to thermal degradation of cellulose paper and point out the process of ageing that decide the end life in paper insulation. In order to verify the developed FIS, data from [11] is used. This data set is utilized to evaluate the best MF used in the application of the FIS model. Ranges in the inputs and output DP values are fixed as per IEEE standard C57.104^{TM} [12] with respect to their linguistic variables.

Input and output attributes considered for FIS.
2FAL, CO_2 and CO concentrations data is fuzzified using different types of membership functions available. Four distinct linguistic variables are considered in the process of input fuzzification. Fuzzy rules are formed in such a manner that correlation exists between linguistic variables of input which were assigned to different input gas concentrations, to linguistic variables of output (i.e. DP), for evaluating the cellulose paper status shown in Fig. 1. Following are the membership functions assigned to inputs and output. And these membership functions ranges are considered and assigned to linguistic variables according to IEEE standard C57.104TM [12].

Membership function for input 2FAL. According to four linguistic variables, 2FAL is fuzzified into four different ranges; they are very low which is between 0 and 0.1 ppm, low which is between 0.1 and 1 ppm, moderate which is between 1 and 10 ppm, and high is for the value of 2FAL greater than 10 ppm.

Membership function for input CO_2. According to four linguistic variables, CO_2 is fuzzified into four different ranges; they are very low which is between 0 and 2500 ppm, low which is between 2500 and 4000 ppm, moderate which is between 4000 and 10,000 ppm, and high is for the value of CO_2 greater than 10,000 ppm. It plays a significant role in life evaluation of paper insulation.

Membership function for input CO. According to four linguistic variables, CO is fuzzified into four different ranges; they are very low which is between 0 and 350 ppm, low which is between 350 and 570 ppm, moderate which is between 570 and 1400 ppm, and high is for the value of CO greater than 1400 ppm. It is another important compound for evaluation and the ratio obtained from CO_2/CO obtained

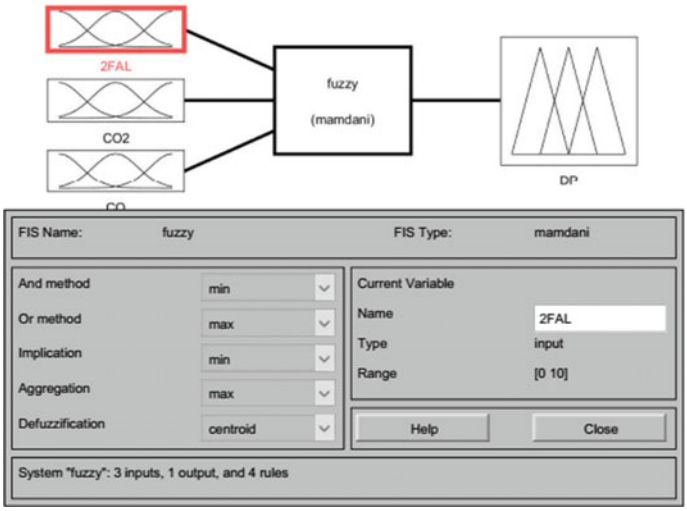


Fig. 1 FIS with inputs and output using different MFs

from DGA should be between 3 and 11, and when it is less than 3 it is serious for the health of paper insulation.

Membership function for output DP. According to four linguistic variables, DP is fuzzified into four different ranges; they are acceptable insulation (AI) which is in the range 1200–700, moderate deterioration which is in the range 700–450, extensive deterioration (ED) which is in the range 450–200, and urgent replacement (UR) is for the value of DP less than 250.

Expert rules. Depending on the FIS model proposed, four expert rules are formed to determine the DP value that gives the insulation health of the transformer. The combinations of furaldehyde with carbon gases decides the status of insulation in transformers based on the DP value. When all the three inputs show high, it indicates cellulose paper insulation end life. These below-listed expert fuzzy rules in Table 4 and Fig. 2 determines the way of designing in FIS.

Table 4 Expert fuzzy rules

Rule	2FAL	CO_2	CO	DP
1	Very low	Very low	Very low	Acceptable insulation (AI)
2	Low	Low	Low	Moderately deteriorated (MD)
3	Moderate	Moderate	Moderate	Extensively deteriorated (ED)
4	High	High	High	Urgent replacement (UR)

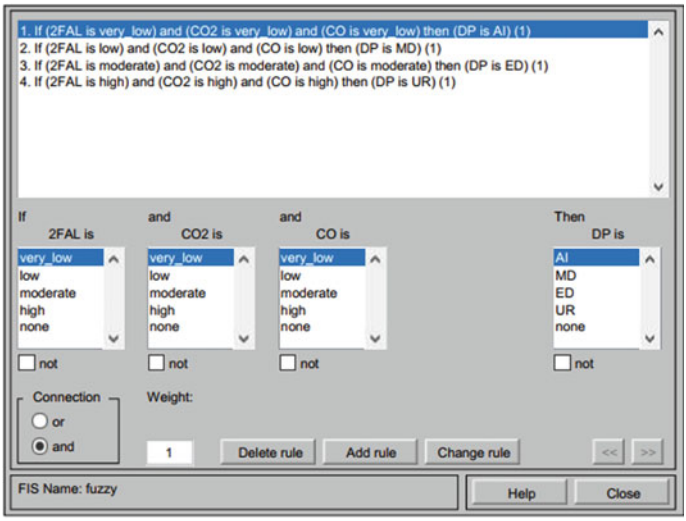


Fig. 2 Determining expert rules in FIS

4 Results and Discussions

Different MFs such as a generalized bell, Gaussian functions, trapezoidal and triangular, were utilized for the fuzzification of the inputs and for defuzzification of the output variables. As per IEEE standard ranges from the results obtained in Table 6, it is found that the best membership function is the trapezoidal membership function. The proposed FIS performance is evaluated using various membership function types for inputs and output. The DP value obtained from the FIS model gives output for different membership functions which is shown below in Table 5.

Table 5 shows the results obtained from developed fuzzy which determines DP for different types of MFs by considering particular combinations of inputs such that it covers all MFs considered as per fuzzy linguistic variables. The crisp output value of the DP is obtained and it predicts insulation paper condition. All the high values in inputs and output whose range is less than 250 should be urgently replaced. DP value which is associated with status of paper according to the IEEE standard is checked for deviation by comparing the results to experimental DP considered from [11], shown in Table 6.

Table 6 compares FIS given DP values to experimental values and the deviation. From this, it is observed that the deviation is less in trapezoidal than other membership functions and this is followed by triangular membership functions and output results give closer values to the trapezoidal function. Now with the obtained DP, the cellulose insulation paper is obtained. Table 2 shows the estimated DP by different mathematical models which were formed on the basis of correlation between 2FAL and DP. Each method gives different types of DP for the same value of 2FAL because

of the assumptions they considered for obtaining the equation. There is more deviation in the values of DP in mathematical equations to that of FIS. Anyway, both the methods are approximations and they cannot be reliable totally so we need to consider more parameters that affect the paper insulation. By knowing and considering more parameters, better models can be formed and life of paper can be determined accurately.

5 Conclusions

In the present work, the correlation with 2-furaldehyde and carbon oxide gases with DP is explored to assess the insulation paper status through fuzzy-based model and mathematical models. These models provide the status of insulation in which DP is used as a failure index. These models successfully predict DP values based on concentrations of 2FAL, CO_2 and CO gases in insulating oil. FIS model performance is tested using various fuzzy membership functions. Comparison of results which were acquired from the FIS model with the experimental DP data indicates there is a minimum deviation in trapezoidal function than other MFs.

Table 5 Results obtained for different MFs using FIS

S. No.	2FAL (ppm)	CO_2 (ppm)	CO (ppm)	DP				
				tria	trapez	Gaussian	Gauss2	gbell
1	0.05	1936	97	938	986	937	966	988
2	0.06	218	34	1035	1044	998	996	1042
3	0.08	1600	145	990	988	998	998	992
4	0.09	877	98	810	797	800	815	824
5	0.1	2500	350	428	462	427	426	456
6	0.2	2600	360	498	495	526	519	510
7	0.5	3089	405	636	606	661	617	634
8	1.72	4150	592	375	380	370	370	368
9	3	5315	662	362	362	382	375	378
10	3.85	4465	1062	272	280	275	279	265
11	5	5004	935	375	363	375	375	375
12	6	7800	1050	363	350	370	363	366

Table 6 Deviation in the DP obtained from FIS (fuzzy inference system) with respect to DP

S. No.	DP	Deviation				
		tria	trapez	Gaussian	Gauss 2	gbell
1	1055	117	69	118	89	67
2	1150	115	106	162	154	108
3	965	25	33	23	23	27
4	727	83	70	73	88	97
5	484	56	22	57	58	28
6	465	33	33	61	54	45
7	570	66	36	91	47	64
8	392	17	12	22	22	24
9	329	46	33	53	46	49
10	300	28	20	25	21	35
11	343	32	20	32	32	32
12	296	67	54	74	67	70

References

1. S. Mishra, H. Malik, A.K. Yadav, T. Mehto, Application of neurofuzzy scheme to investigate the winding insulation paper deterioration in oil-immersed power transformer. Int. J. Electr. Power Energy Syst. **53**(1), 256–271 (2013)
2. T.K. Saha, Review of modern diagnostic techniques for assessing insulation condition in aged transformers. IEEE Trans. Dielectr. Electr. Insul. **10**(5) (2003)
3. X. Chengdong, Monitoring paper insulation ageing by measuring furfural contents in oil, in *7th International Symposium on High Voltage Engineering* (1991), pp. 26–30
4. L. Cheim, D. Platts, T. Prevost, S. Xu, Furan analysis for liquid power transformers. Electr. Insul. Mag. (IEEE) **28**, 8–21 (2012)
5. R.D. Stebbins, D.S. Myers, A.B. Shkolnik, Furanic compounds in dielectric liquid samples: review and update of diagnostic interpretation and estimation of insulation ageing, in *7th International Conference on Properties and Application of Dielectric Materials* (2003), pp. 921–926
6. A. De Pablo, Furfural and ageing: how are they related, in *IEEE Colloquium on Insulating Liquids* (1999), pp. 5/1–5/4
7. E. Li, B. Song, Transformer health status evaluation model based on multi-feature factors, in *International Conference on Power System Technology* (2014), pp. 1417–1422
8. L. Chaohui, Z. Bide, Y. Yuchun, The aging diagnosis of solid insulation for oil-immersed power transformers and its remaining life prediction, in *Asia-Pacific Power and Energy Engineering Conference* (2010), pp. 1–3
9. J.-S.R. Jang, C.-T. Sun, E. Mizutani, Neuro fuzzy and soft computing a computational approach to learning and machine intelligence, Pearson
10. T.S. Ramu, M. Mandlik, Moisture aided degradation of oil impregnated paper insulation in power transformers. IEEE Trans Dielectr. Electr. Insul. **21**(1), 186–193 (2014)
11. Md. Manzar Nezami, S.A. Wani, S.A. Khan, Fuzzy approaches for residual life assessment of paper insulation, in *3rd International Conference on Condition Assessment Techniques in Electrical Systems,* 978-1-5386-3138-6/17/$31.00©2017 (IEEE, 2017)
12. IEEE C57.104-2008—IEEE guide for the interpretation of gases generated in oil-immersed transformers

Comparison of Fuzzy Logic Based MPPT of Grid-Connected Solar PV System with Different MPPT



Ashutosh Tiwari and Ashwani Kumar

1 Introduction

As there are various sources of electricity generation, one of the main sources is fossil fuels. Due to the depletion of fossil fuels, which is a major source of GHG emissions, electricity generation is shifting toward renewable energy sources [1]. Solar energy is the most widely used renewable energy source for electricity generation as it is free of pollution and requires less maintenance.

PV array is the connection of a number of modules in series or parallel. The panel is a connection of a number of solar cells in series or parallel. A solar cell is made up of a semiconductor through which sunlight passes. With the change in environmental conditions, output power generated always changes; so to track the maximum power, different MPPT methods are used irrespective of these conditions [2]. Various perturb steps have been discussed to increase efficiency of MPPT [3, 4]. In a grid-connected solar PV system, electricity generated by solar system can send the grid and can be taken by grid depending upon the requirement. For these, system battery storage requirements are not needed so power losses in storage are reduced.

In a solar photovoltaic system, MPPT is used for increasing the efficiency of the photovoltaic system. There are various MPPT techniques such as perturb and observe, incremental conductance, fuzzy logic based, and neural network MPPT methods [5]. Comparison of different MPPTs is discussed in [6]. For variable-step incremental conductance, MPPT is used, but it is more complex as compared to P&O MPPT, and for varying climate conditions fuzzy logic based MPPT exhibits good performance as compared to others but it is more expensive due to its design complexity [7].

A. Tiwari (✉) · A. Kumar
National Institute of Technology, Kurukshetra, Kurukshetra 136119, Haryana, India
e-mail: asutosh.tiwari0@gmail.com

A. Kumar
e-mail: ashwaks@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_6

Fuzzy logic based controller for standalone and grid-connected system is discussed [8]. The perturb and observe method is widely used because it is cheap and can be easily implemented, but it oscillates around the maximum power point so it causes power losses. On comparing P&O and INC, INC exhibits better performance but it is more complex [3].

In this paper, comparison of MPPTs is discussed for a 150 kW grid-connected solar system which has been designed with the use of MATLAB/Simulink. The proposed system consists of a solar photovoltaic array, MPPT for extracting maximum power from the solar system, boost converter for regulating and boosting the solar system output, voltage source inverter for inverting the DC output into AC for supplying to the grid, and an LC filter for filtering current harmonics from the output of inverter current. Section 2 presents the detailed structure of the proposed system, Sect. 3 describes the simulation and experimental result of the model and finally, Sect. 4 describes the conclusion.

2 Methodology

2.1 Modeling of PV Array

A photovoltaic array is the connection of a number of modules or panels connected in series/parallel. A PV module can be defined as the number of series and parallel connections of solar cell which in turn is a p–n junction that converts sunlight into electrical energy. If the intensity of solar rays is higher, current from the solar cell is more. Figure 1 shows the circuit of a solar cell.

Now from the figure,

$$I = I_{ph} - I_d - I_p \tag{1}$$

For a solar cell equation,

$$I_d = I_0\left[e^{\frac{V+I_p}{nV_T}} - 1\right] \tag{2}$$

$$I_p = \frac{V + IR_s}{R_p} \tag{3}$$

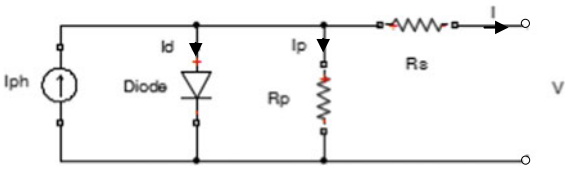


Fig. 1 Circuit of solar cell

Equation (1) can now be written as

$$I = I_L - I_0\left[e^{\frac{V+I_{sh}}{nV_T}} - 1\right] - \frac{V + IR_s}{R_{sh}} \tag{4}$$

where

I_0	Reverse saturation current,
R_s and R_p	Series and parallel resistance of solar cell, respectively,
n	Diode ideality factor,
V_T	Thermal voltage,

and V_T is given by

$$V_T = \frac{KT}{q} \tag{5}$$

K = Boltzmann Constant = $1.38 \times 10 - 23\,\text{J/K}$,

$q = 1.6 \times 10^{-19}$ Coulomb

A P–V and I–V characteristic of the PV system is shown in Fig. 2 at 25 °C temperature.

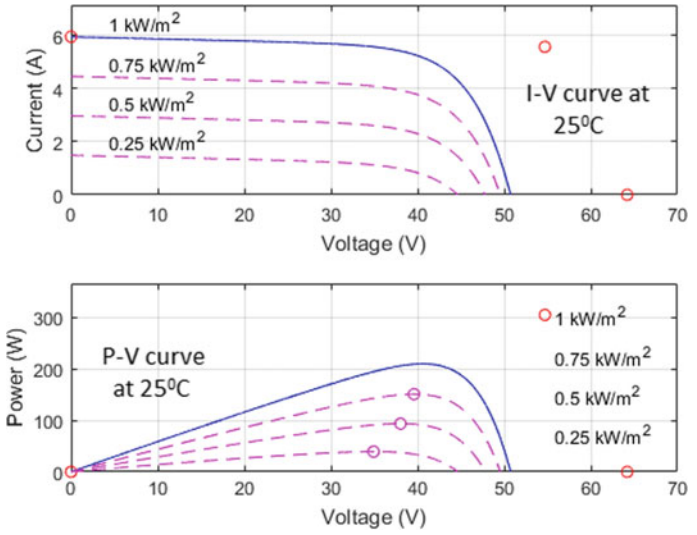


Fig. 2 P–V and I–V characteristics at 25 °C temperature

2.2 Maximum Power Point Tracking (MPPT)

For varying environmental conditions, it is necessary to use MPPT. There are different types of MPPT methods that are available as fuzzy logic based, neural network based, perturb and observe based, INC, etc. Among all MPPTs, fuzzy logic based MPPT is the best; it is more efficient than others but its drawback is that it is more expensive than the others [9]. Perturb and observe is mostly used because of its easiness and INC type is better in performance than P&O [10].

Perturb and Observe MPPT: It is one of the Hill climbing methods. Since it is very easy to implement and maintain, it is widely used. This MPPT tracks the maximum power based on $\frac{\Delta P}{\Delta V}$ in the P–V characteristics [6]. If $\frac{\Delta P}{\Delta V} > 0$, then the maximum power point is at its right and if $\frac{\Delta P}{\Delta V} < 0$, then the maximum power point is at its left. Flowchart of the MPPT algorithm and Simulink diagram is shown in Fig. 3 and Fig. 4, respectively.

Incremental Conductance MPPT: It is an improvement of the P&O MPPT algorithm used for variable steps. This works on the I–V characteristics of the PV. It is also easier in implementation and simple as compared to fuzzy. INC determines whether the MPP is reached or not by checking the condition [10]. If MPP is obtained, then it stops the perturbing; if not obtained, it again checks the condition as shown in the

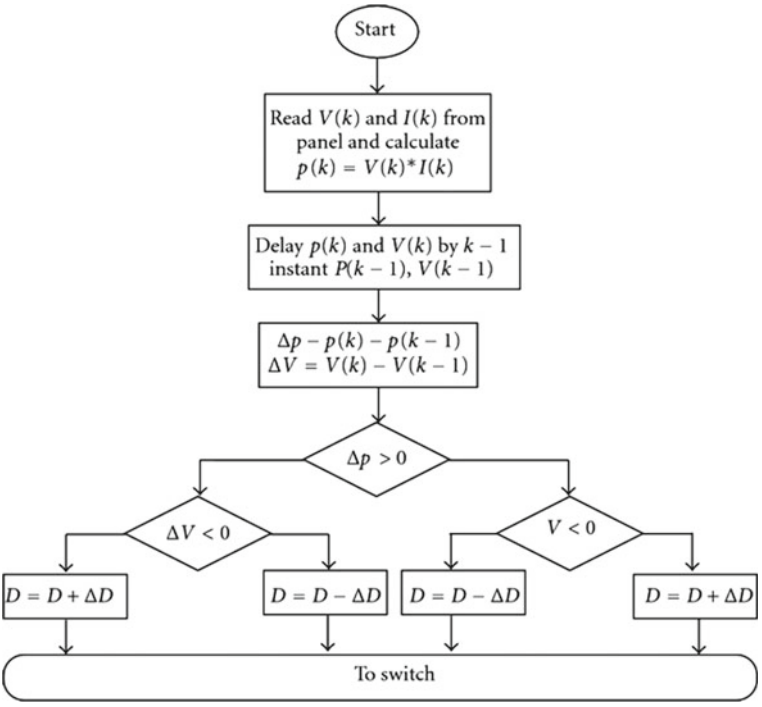


Fig. 3 Flowchart of P&O algorithm

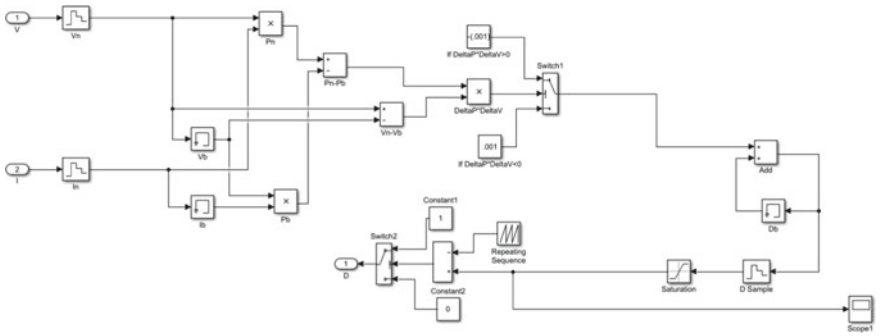


Fig. 4 Simulink diagram of P&O algorithm

algorithm of INC. MPP is at the right, if $\frac{\Delta P}{\Delta V} < 0$, i.e., $-I/V$ and if positive, then MPP is at the left. Flowchart and Simulink circuit of INC is given in Fig. 5 and Fig. 6, respectively.

Fuzzy Logic based MPPT: As the variation of temperature and irradiance response of perturb and observe based and INC-based MPPT methods are slow, to improve

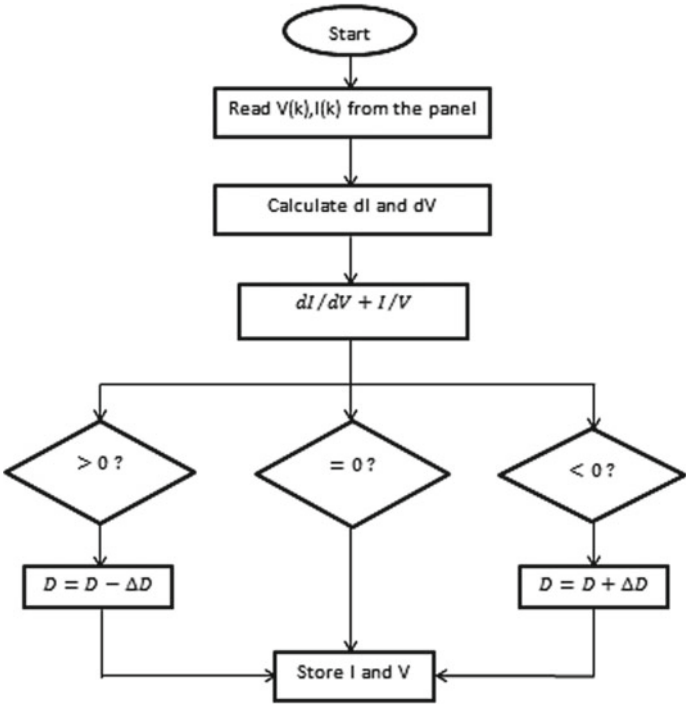


Fig. 5 Flowchart of INC algorithm

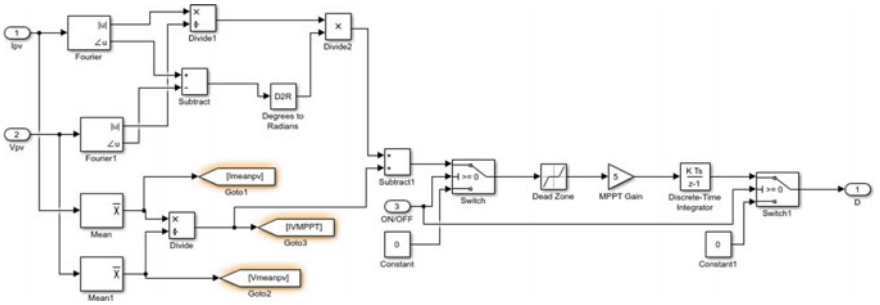


Fig. 6 Simulink diagram of INC algorithm

these problems fuzzy logic based methods are used. This method has fast response as compared to others [11]. It also increases the stability of the system. There is no need for mathematical modeling in this case. Figures 7 and 8 are the fuzzy logic designer and membership function of fuzzy systems.

There are three stages of a fuzzy logic controller: fuzzification, logic building, and defuzzification. In fuzzification E, i.e., $\frac{\Delta P}{\Delta V}$ is coded into a linguistic variable through membership function as shown [6]. There are two inputs E and CE and the duty cycle, i.e., D is the output which is a numeric variable. E and CE can be written as

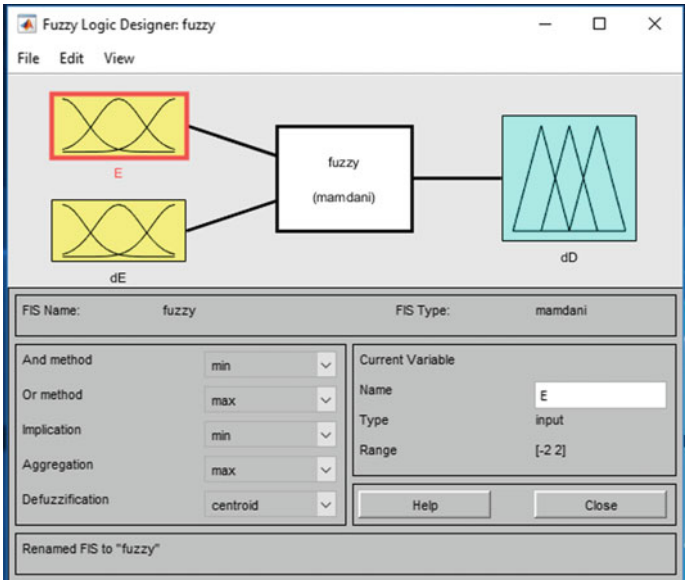


Fig. 7 Fuzzy logic designer

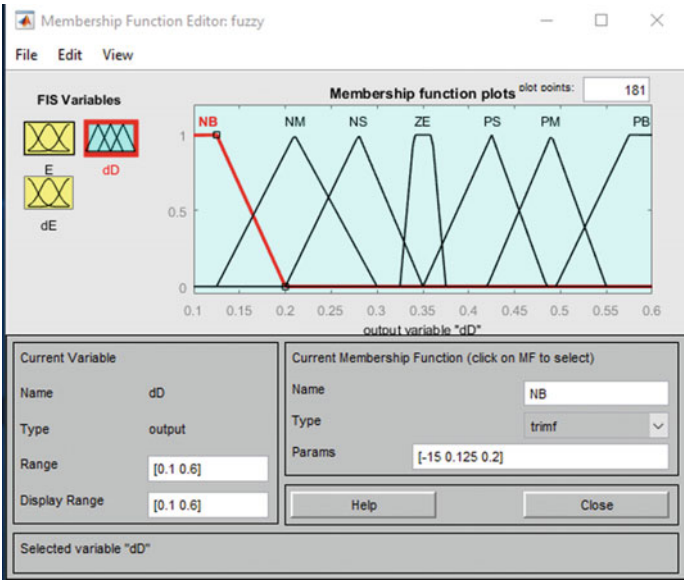


Fig. 8 Membership function of fuzzy

$$E = \frac{P[k] - P[k-1]}{V[k] - V[k-1]} \tag{6}$$

$$CE = V[k] - V[k-1]$$

Simulink circuit of the fuzzy logic based MPPT is shown in Fig. 9.

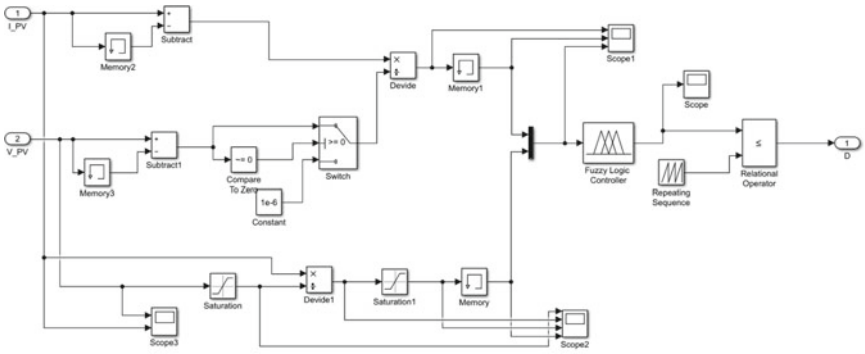


Fig. 9 Simulink diagram of fuzzy logic based MPPT

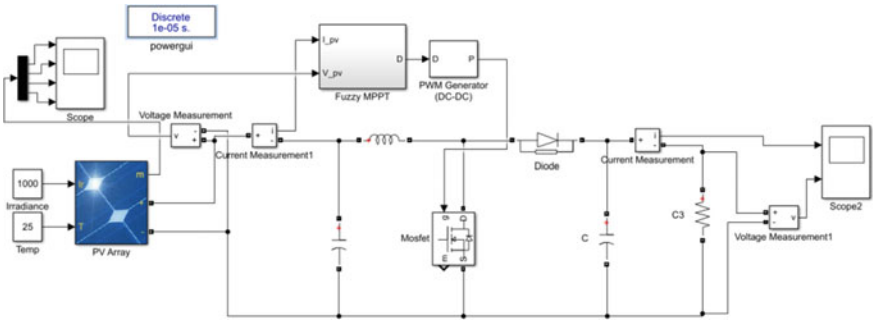


Fig. 10 Simulation model of PV array with boost converter

2.3 Boost Converter

Boost converter is a DC converter that is used for increasing the voltage level. With this converter, $V\text{out} > V\text{in}$ and $I\text{out} < I\text{in}$ [10]. Boost converter consists of IGBT and diode as shown in the figure. It also used to regulate the voltage of PV system. Adjusting the duty cycle of MPPT for different environmental conditions can get maximum power [9]. Simulink diagram of boost converter is shown in Fig. 10.

2.4 Three-Level Inverter

Three-phase VSC is used for converting the DC level into AC for supplying the power generated by the solar array to the grid. Simulink block of three-phase VSC is shown in Fig. 11. Two capacitors are used at the output of the boost converter [12]. These capacitors provide a neutral point for the inverter. Since IGBT is having fast switching speed, it is used for high power levels. For controlling the inverter, HCC is used. HCC (Hysteresis current controller) provides the gate signal for voltage source inverter [13]. Hysteresis current controller provides PWM (pulse width modulated signal for) signal to the inverter [14].

3 Simulation and Result

A complete Simulink model of this paper is shown in Fig. 15. Different MPPTs are used in this circuit to compare their performance. PV array power and load voltage of the boost converter for each MPPT as P&O, INC, fuzzy are shown in Fig. 12, respectively. By using 71 parallel strings and 7 series strings for PV array with each string consisting of 96 panels. Grid voltage and current for each MPPT are shown in Fig. 13 and power supplied to the grid for each MPPT is shown in Fig. 14. On

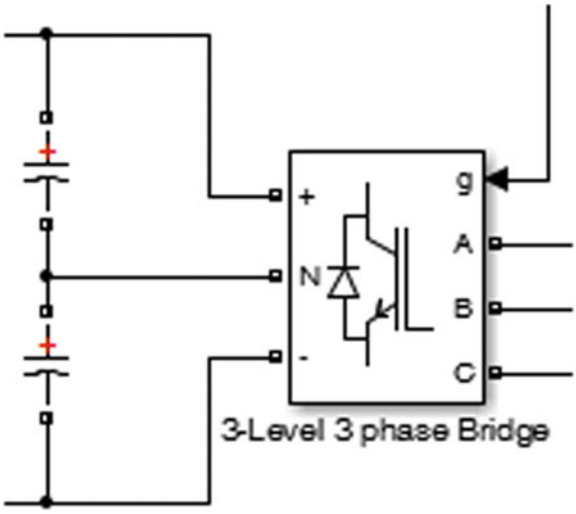


Fig. 11 Simulink block of three-phase three-level inverter

comparison of all these MPPTs, it is found that fuzzy MPPT is more efficient as shown.

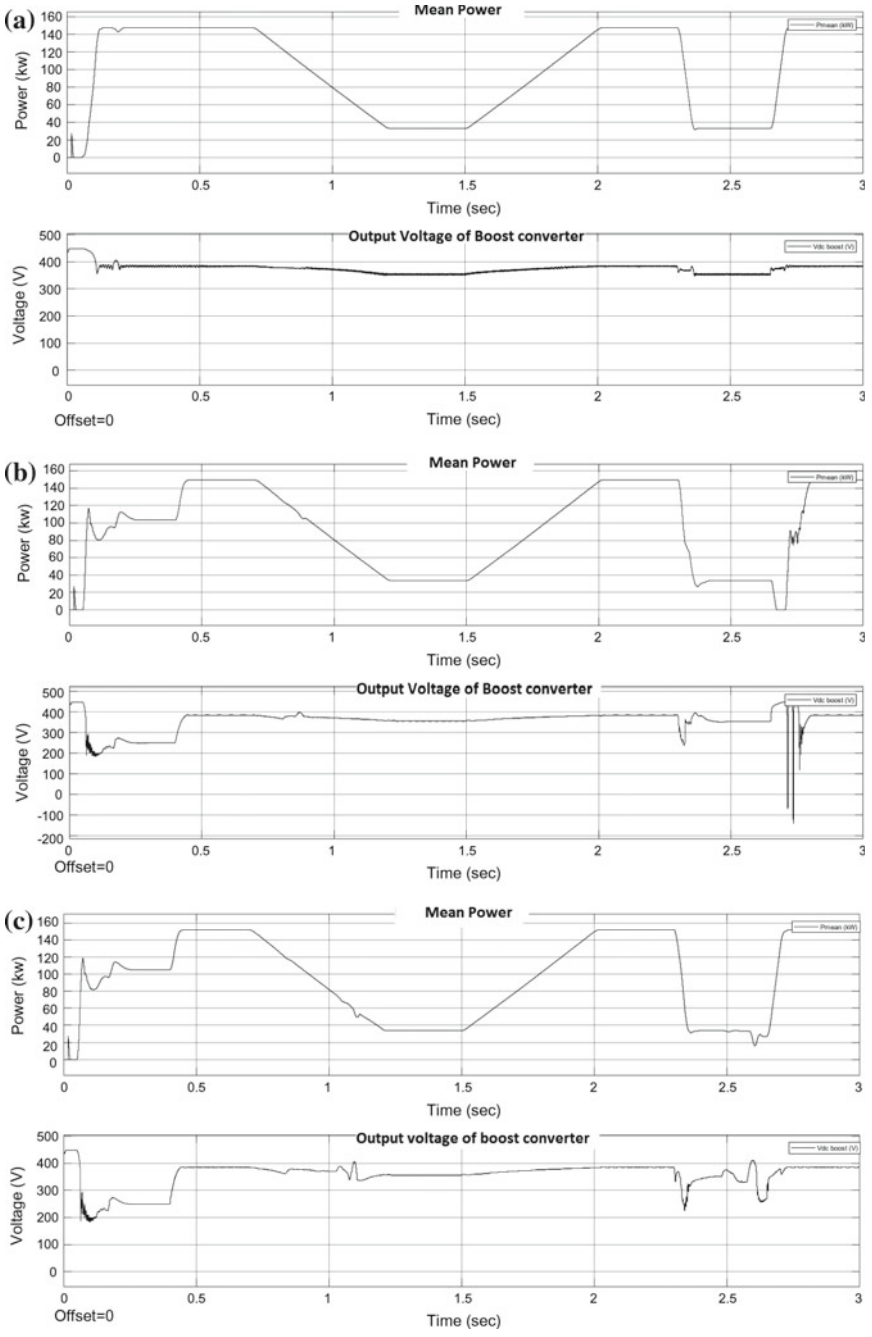


Fig. 12 Waveforms of mean output power of PV array and load voltage of boost converter for **a** P&O MPPT **b** INC MPPT **c** Fuzzy logic based MPPT

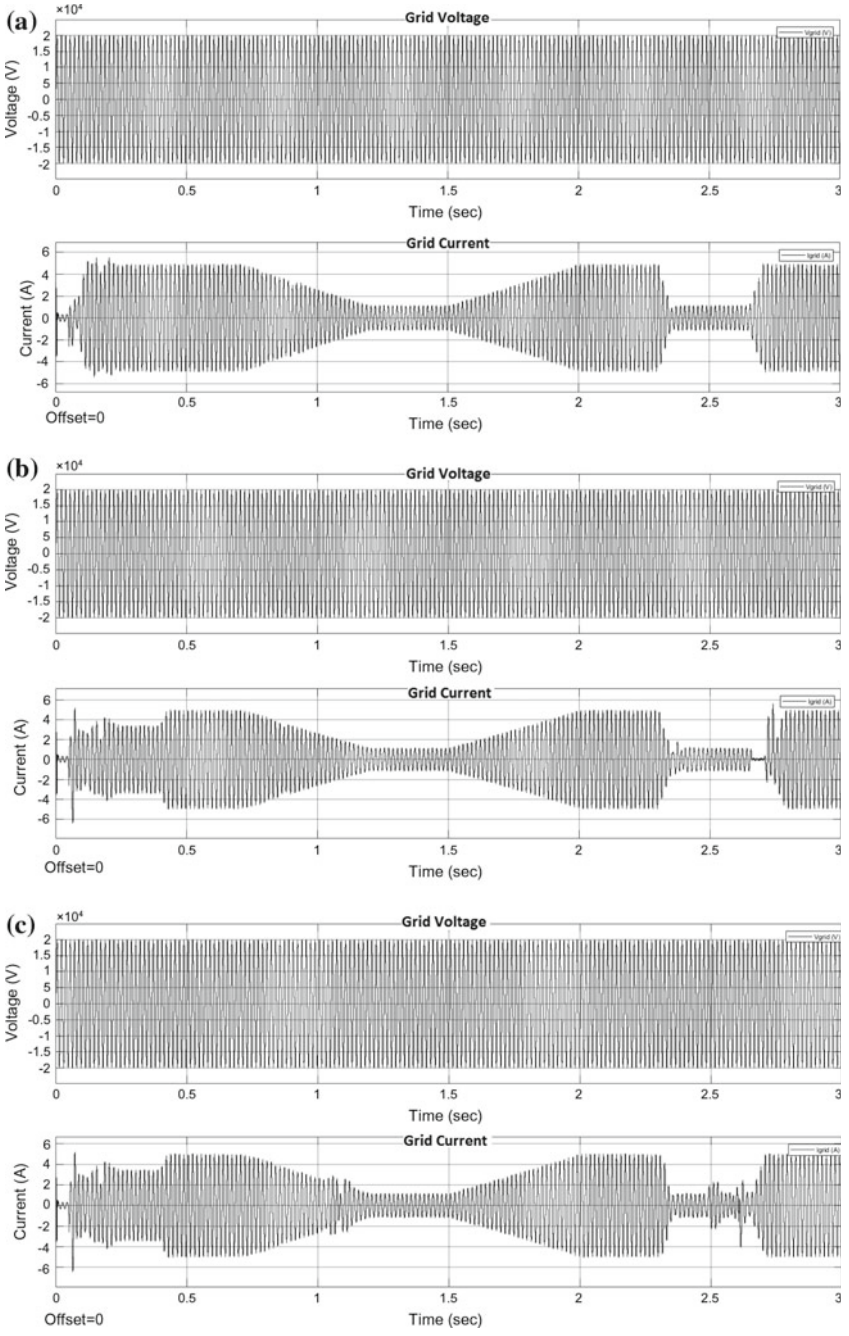


Fig. 13 Waveforms of grid voltage and current for **a** P&O MPPT **b** INC MPPT **c** Fuzzy logic based MPPT

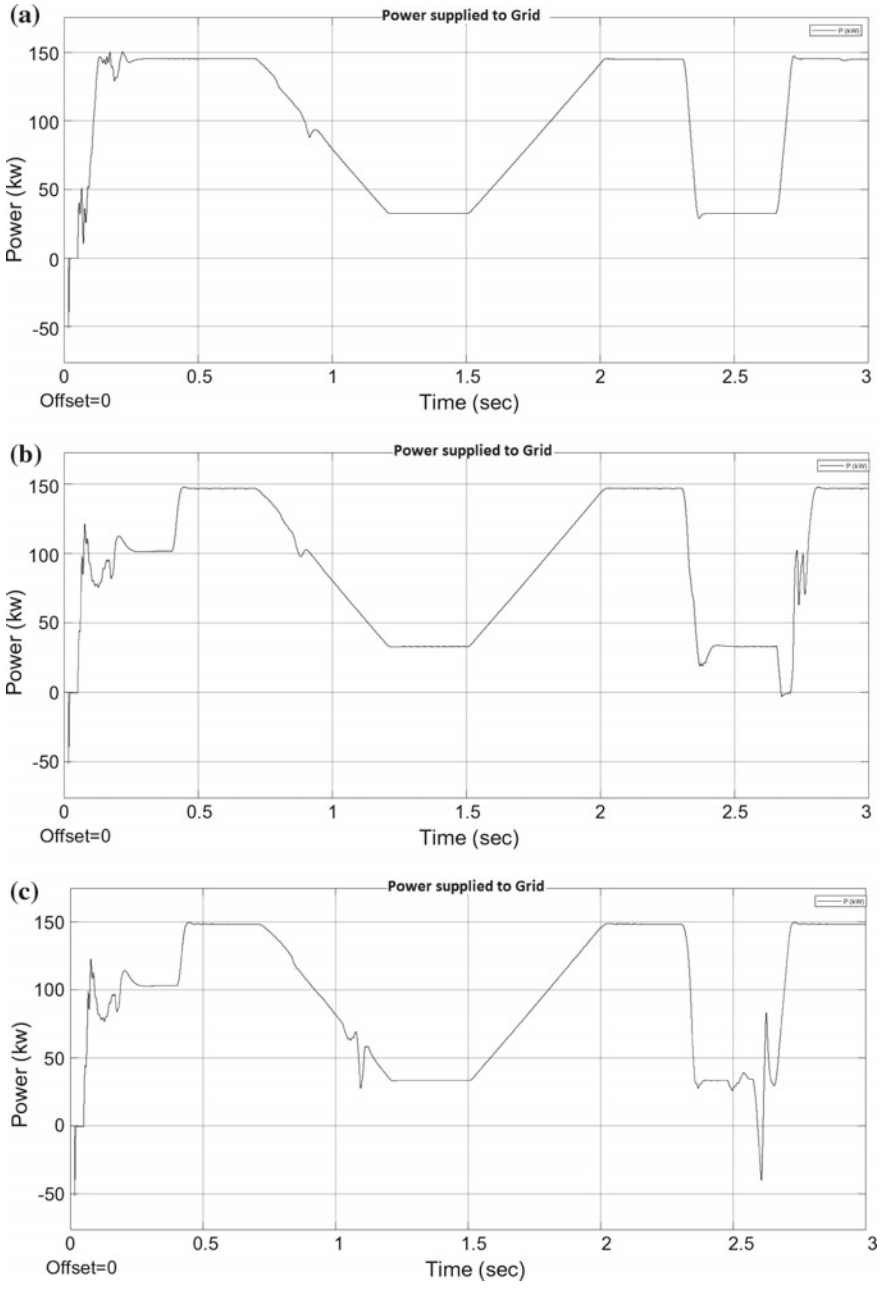


Fig. 14 Waveforms of power at bus connected to grid for **a** P&O MPPT **b** INC MPPT **c** Fuzzy logic based MPPT

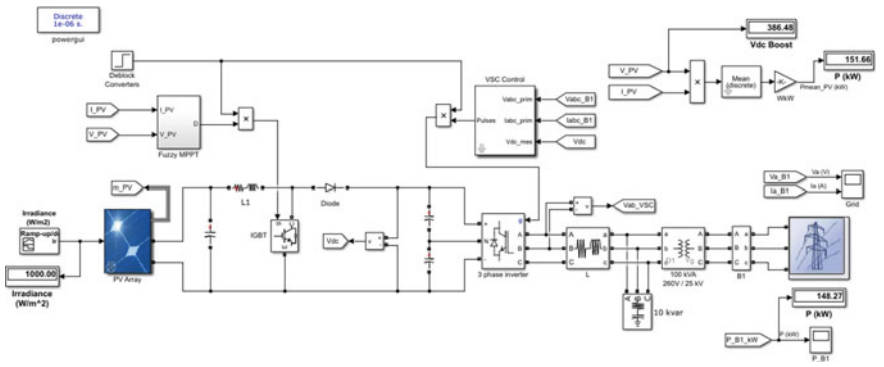


Fig. 15 Complete Simulink model

4 Conclusion

As the waveforms shown above for different MPPTs for PV array power and load voltage of DC–DC boost converter is better of fuzzy MPPT than incremental and perturb and observe type. Also, the output power at the bus level for grid connection is better for fuzzy than others. Fuzzy-based MPPT is more efficient than others as shown in the figure and also in Table 1. From the table, it is clear that speed of response is better in fuzzy and moderate for INC type. P&O type is slow but less complex to design and easy in maintenance. Overall, fuzzy-based MPPT is more efficient. Hysteresis current controller is used for controlling the pulse for voltage source inverter. This type of controller is easy to implement and has better response in speed.

Table 1 Comparison of MPPTs

MPPT method	Speed	Reliable	Complex	PV power (kW)	Load voltage of boost converter (V)	Power at bus (kW)
P&O	Slow	Less	Less	147	375	145
INC	Vary	Moderate	Moderate	149	379	146
Fuzzy	Fast	Moderate	Moderate	151	386	148

References

1. M.E. Ropp, S. Gonzalez, Development of a MATLAB/Simulink model of a single-phase grid-connected photovoltaic system. IEEE Trans. Energy Convers. **24**(1), 195–202 (2009)
2. M. Bharathkumar, H.V. Byregowda, Performance evaluation of 5 MW grid connected solar photovoltaic power plant established in Karnataka. Int. J. Innov. Res. Sci. Eng. Technol. **3**(6) (2014)
3. A.K. Abdelsalam, A.M. Massoud, S. Ahmed, P.N. Enjeti, High-performance adaptive perturb and observe MPPT technique for photovoltaic-based microgrids. IEEE Trans. Power Electron. **26**(4), 1010–1021 (2011). https://doi.org/10.1109/TPEL.2011.2106221
4. N. Femia, G. Petrone, G. Spagnuolo, M. Vitelli, Optimization of perturb and observe maximum power point tracking method. IEEE Trans. Power Electron. **20**(4), 963–973 (2005). https://doi.org/10.1109/TPEL.2005.850975
5. E. Roman, et al., Intelligent PV module for grid-connected PV systems. IEEE Trans. Ind. Electron. **53**(4), 1066–1073 (2006)
6. D.P. Hohm, M.E. Ropp, Comparative study of maximum power point tracking algorithms. Prog. Photovolt. Res. Appl. **11**, 47–62 (2002). https://doi.org/10.1002/pip.459
7. Z. Ahmad, S.N. Singh, Modeling and control of grid connected photovoltaic system—a review. Int. J. Emerg. Technol. Adv. Eng. **3**(3), 2250–2459 (2013)
8. B. Krishna Naick, T.K. Chatterjee, K. Chatterjee, Fuzzy logic controller based PV system connected in standalone and grid connected mode of operation with variation of load. Int. J. Renew. Energy Res. **7**(1) (2017)
9. B. Subudhi, R. Pradhan, A comparative study on MPPT techniques for PV power systems. IEEE Trans. Sustain. Energy **4**(1) (2013)
10. J. Atiq, P.K. Soori, Modelling of a grid connected solar PV system using MATLAB/Simulink. Int. J. Simul. Syst. Sci. Technol. **17**(41) (2016)
11. N. Rav, Performance study of incremental inductance algorithm for PV applications. Int. J. Sci. Eng. Technol. **5** (2016)
12. J. Kasera, V. Kumar, R.R. Joshi, J.K. Maherchandani, Design of grid connected photovoltaic system employing incremental conductance MPPT algorithm. J. Electr. Eng. **12**, 172–177 (2012)
13. F. Liu, S. Duan, F. Liu, B. Liu, Y. Kang, A variable step size INC MPPT method for PV systems. IEEE Trans. Industr. Electron. **55**(7), 2622–2628 (2008)
14. A.S. Swathy, R. Archana, MPPT using modified incremental inductance for solar PV system. Int. J. Eng. Innov. Technol. **3**(2) 2013

Comparative Study of Different Classification Models on Benchmark Dataset of Handwritten Meitei Mayek Characters



Deena Hijam and Sarat Saharia

1 Introduction

Optical Character Recognition (OCR) is the technique by which computers can recognize printed or handwritten text which can be used within its processing applications. Handwritten Character Recognition (HCR) is a subfield of OCR where the text concerned are handwritten ones. For decades, this research area has been explored by many researchers. A significant amount of work is done for major scripts such as Latin, Chinese, Arabic, and Indic scripts such as Bangla and Devnagari. However, the number of scripts available worldwide is huge. Since there are different issues and challenges pertaining to different scripts, it is required that evaluation of recognition systems be carried out on different scripts. In order to get an idea of results of different classification models, this paper attempts to provide a comparative study of seven popular classification models (available in scikit-learn) on Meitei Mayek handwritten character dataset. A CNN model is also proposed which gives state-of-the-art results on the concerned dataset.

Some works have been reported in HCR of Meitei Mayek. However, the datasets used in the said works are not publicly available and hence benchmark results cannot be obtained for future research. Earlier works have used features such as probabilistic and fuzzy features [1], neural networks [2, 3], Gabor filters [4], projection histograms, Histogram Oriented Gradient (HOG) and Background Directional Distribution (BDD) [5], hybrid point features [6], HOG features [7], and features based on Local Binary Pattern [8]. Classifiers used are SVM, KNN, and Neural Networks.

The rest of the paper is organized as follows: Sect. 2 briefs about Meitei Mayek script. Section 3 talks about classification models used in the present work. It also discusses three architectures of CNN proposed for recognition of characters in the

D. Hijam (✉) · S. Saharia
Tezpur University, Napaam 784028, Assam, India
e-mail: deenahi@tezu.ernet.in

S. Saharia
e-mail: sarat@tezu.ernet.in

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_7

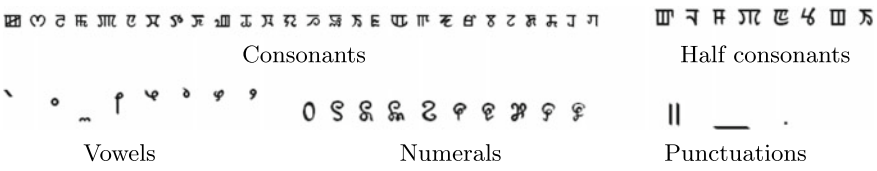


Fig. 1 Meitei Mayek character set

dataset. Section 4 gives comparative study of the different models. Section 5 gives future research direction and conclusion.

2 Meitei Mayek Script

Meitei Mayek script is used to write Manipuri language. It is a Tibeto-Burman language and one of the 22 official Indian languages as per the Eighth Schedule of the Constitution of India. Manipur is one of the few states of Northeastern India with a script of its own. The script has been ignored for centuries when it was replaced by Bengla in the early eighteenth century. The script has gained its due recognition when it got revived in 1980 and it was introduced as part of academic curriculum during 2005–2006 session in schools across Manipur [9].

Meitei Mayek character set consists of 56 letters with 27 Iyek Ipee(consonants), 8 Lonsum Iyek (final consonants), 8 Cheitap Iyek (vowels), 3 Khudam Iyek (punctuation marks), and 10 Cheising Iyek (numerals) (Fig. 1). The script is written from left to right. There is no concept of uppercase and lowercase letters in Meitei Mayek. Presence of similar characters can be seen in the dataset.

3 Classification Models

For our work, we have used the classification models provided in the scikit-learn library. It is an open-source machine learning library developed for the Python programming language [10]. It provides various supervised and unsupervised learning algorithms among others. It can be used in both commercial and academic settings as it is licensed under the simplified BSD license. A number of groups of models are listed under scikit-learn but our interest at this point is the supervised classification model which includes models such as decision trees, SVM and neural networks and model selection modules to calculate accuracy and confusion matrix. The models considered for our work are based upon the work reported in [11] where they have also selected similar set of models to develop a benchmark on their dataset.

Table 1 Parameters used for learning algorithm

Name	Value
Optimization algorithm	Mini-batch stochastic gradient descent
Learning rate	0.01
Decay rate	1e−7
Mini-batch size	32
Dropout	0.3
Number of epochs	50

3.1 Convolutional Neural Network Model

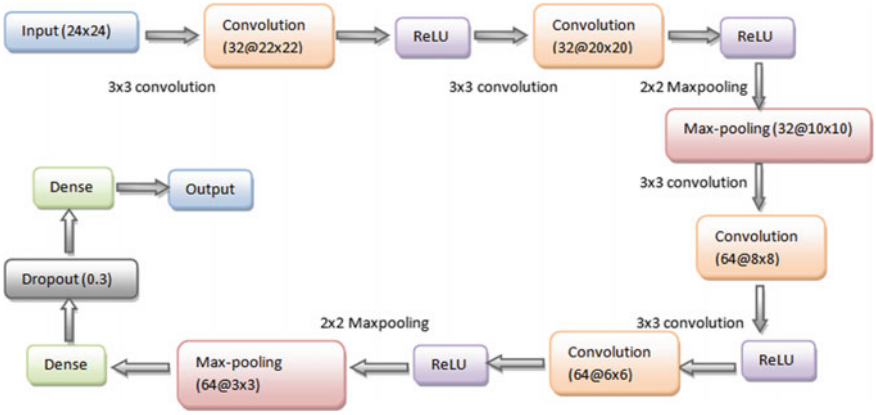
Three architectures of CNN have been studied and are shown in Fig. 2. First architecture has four convolutional layers, two max-pooling layers, two fully connected layers, one dropout layer, input, and output layers. Second and third architectures have two batch normalization layers in addition to the layers of first architecture.

The first two convolutional layers have 32 kernel masks of size 3×3 resulting in 32 feature maps. The max-pooling layer has window size of 2×2 pixels with stride 2. Each max-pooling layer produces feature maps of size half of that of its input. The last two convolutional layers have 64 convolution kernel masks of size 3×3 outputting 64 feature maps. Each convolutional layer has ReLU activation function ($f(x) = max(0, x)$). There are two fully connected layers with 1024 and 512 nodes with a dropout of 0.3 between them. These two layers use sigmoid ($f(x) = tanh(x)$) activation function. The last layer, i.e., the output layer has softmax function with cross-entropy loss function.

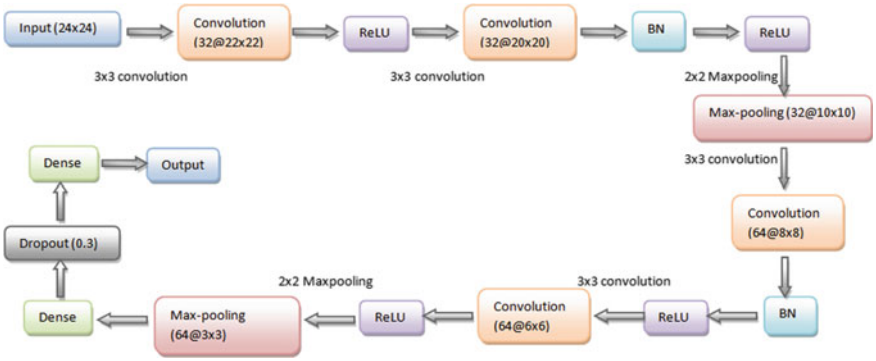
Batch Normalization Batch normalization is employed after every convolutional layer in our model. It is a technique to normalize the inputs of each layer (normally internal layers) in order to reduce the internal covariate shift problem [12]. During training Deep Neural Networks, the distribution of inputs to each layer changes due to the change in previous layer's parameters and hence it becomes necessary for each layer to learn the new distribution during each training step. This problem is known as internal covariate shift problem and it slows down the training phase of Neural Networks.

Dropout Dropout is used between the last two fully connected layers in the present model. It is a regularization method to fight overfitting of data in Neural Networks [13]. It is a process of dropping certain units including its incoming and outgoing connection during training so that the network does not become too biased to the patterns that it sees during training.

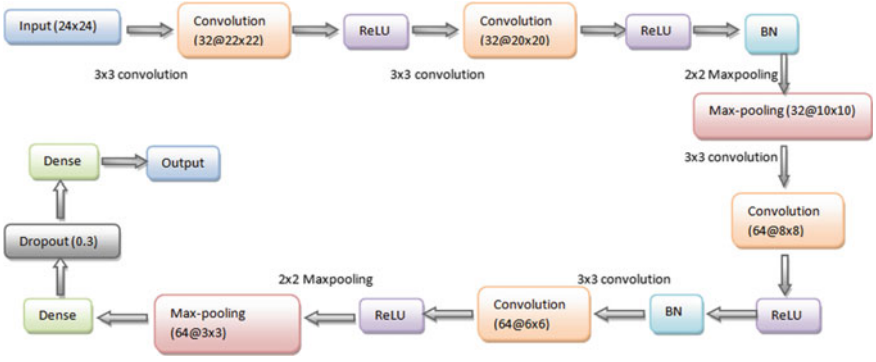
Parameters Used: Mini-batch stochastic gradient descent optimization algorithm is used as the learning algorithm. The parameters used for this algorithm are listed in Table 1.



(a) CNN architecture 1 - CNN+Dropout



(b) CNN architecture 2 - CNN+Dropout+BatchNormBeforeActivation



(c) CNN architecture 3 - CNN+Dropout+BatchNormAfterActivation

Fig. 2 Three CNN architectures studied in the present work. BN is a batch normalization layer

4 Experimental Results

4.1 Dataset

The dataset used in the present work is the MMHC dataset available at http://agnigarh.tezu.ernet.in/~sarat/resources.html [14]. To the best of our knowledge, this is the only publicly available dataset of the concerned script. It has a total of 60,285 character images with 54,239 and 6,046 images in training and testing sets, respectively. There are 37 classes (27 consonants and 10 numerals). The images are grayscale, size-normalized to 24 × 24 pixels and saved in TIFF format.

As a preprocessing step, images are normalized to pixel intensity range [0,1] for faster computation.

4.2 Experimental Results and Discussion

Each classification algorithm is run three times with the specified parameters and the average test accuracies achieved are reported as shown in Table 2. The best model of each algorithm is chosen and a classwise accuracy of the same is shown in Fig. 3.

The above results show that classification rates of certain characters are low compared to other characters. This is due to structural similarities between characters which makes the task of classification challenging. The top six characters in terms of highest misclassification rates by all eight models are shown in Fig. 4.

We have also analyzed the effect of using batch normalization before and after the activation function in CNN and it is observed that using batch normalization after activation gives a faster convergence and highest test accuracy of 98.11% as compared to other two architectures. The test losses and test accuracies achieved by three architectures are shown in Fig. 5. CNN and dropout (architecture 1) achieve highest test accuracy of 97.40%, CNN with dropout and batch normalization before activation function (architecture 2) and CNN with dropout and batch normalization after activation function (architecture 3) achieve highest test accuracies of 97.95% and 98.11%, respectively.

Table 2 Test accuracies achieved using different classification models

Classifier	Parameters	Test accuracy (%)
KNeighbor classifier	n_neighbors = 1, weights = 'uniform', p = 1, algorithm = 'auto'	84.44
	n_neighbors = 3, weights = 'uniform', p = 1, algorithm = 'auto'	83.76
	n_neighbors = 5, weights = 'uniform', p = 1, algorithm = 'auto'	83.68
	n_neighbors = 5, weights = 'distance', p = 2, algorithm = 'auto'	82.60
	n_neighbors = 7, weights = 'distance', p = 2, algorithm = 'auto'	81.28
	n_neighbors = 5, weights = 'distance', p = 1, algorithm = 'auto'	84.25
	n_neighbors = 1, weights = 'distance', p = 1, algorithm = 'auto'	84.44
	n_neighbors = 3, weights = 'distance', p = 1, algorithm = 'auto'	83.76
	n_neighbors = 7, weights = 'distance', p = 1, algorithm = 'auto'	83.50
	n_neighbors = 9, weights = 'distance', p = 1, algorithm = 'auto'	82.92
	n_neighbors = 1, weights = 'distance', p = 2, algorithm = 'auto'	82.78
	n_neighbors = 3, weights = 'distance', p = 2, algorithm = 'auto'	82.93
Decision tree classifier	criterion = 'entropy', split = 'best', max_depth = 200, min_samples_leaf = 1	62.73
	criterion = 'gini', split = 'best', max_depth = 200, min_samples_leaf = 1	58.50
	criterion = 'entropy', split = 'best', max_depth = 50, min_samples_leaf = 1	62.81
	criterion = 'gini', split = 'best', max_depth = 50, min_samples_leaf = 1	58.96
	criterion = 'entropy', split = 'best', max_depth = 100, min_samples_leaf = 5	62.77
	criterion = 'gini', split = 'best', max_depth = 100, min_samples_leaf = 5	59.23
	criterion = 'entropy', split = 'random', max_depth = 100, min_samples_leaf = 1	59.99
	criterion = 'gini', split = 'random', max_depth = 100, min_samples_leaf = 1	58.35
	criterion = 'entropy', split = 'random', max_depth = 50, min_samples_leaf = 1	59.72
	criterion = 'gini', split = 'random', max_depth = 50, min_samples_leaf = 1	58.15
	criterion = 'entropy', split = 'best', max_depth = 200, min_samples_leaf = 1	59.01

(continued)

Table 2 (continued)

Classifier	Parameters	Test accuracy (%)
	criterion = 'gini', split = 'random', max_depth = 200, min_samples_leaf = 5	57.55
Support vector classifier	kernel = 'linear', c = 1	82.93
	kernel = 'linear', c = 10	81.46
	kernel = 'linear', c = 100	80.89
	kernel = 'RBF', c = 1	81.56
	kernel = 'RBF', c = 10	86.33
	kernel = 'RBF', c = 100	87.51
	kernel = 'sigmoid', c = 1	74.32
	kernel = 'sigmoid', c = 10	81.49
	kernel = 'sigmoid', c = 100	79.76
	kernel = 'poly', c = 1	81.47
	kernel = 'poly', c = 10	86.33
	kernel = 'poly', c = 100	87.17
Random forest classifier	criterion = 'gini', max_depth = 50, n_estimates = 100	88.07
	criterion = 'gini', max_depth = 100, n_estimates = 100	88.07
	criterion = 'gini', max_depth = 200, n_estimates = 100	88.07
	criterion = 'entropy', max_depth = 50, n_estimates = 100	88.09
	criterion = 'entropy', max_depth = 100, n_estimates = 100	88.08
	criterion = 'entropy', max_depth = 200, n_estimates = 100	88.41
	criterion = 'entropy', max_depth = 50, n_estimates = 50	86.94
	criterion = 'entropy', max_depth = 100, n_estimates = 50	86.94
	criterion = 'entropy', max_depth = 50, n_estimates = 20	82.89
Classifier	Parameters	Test accuracy (%)
MLP classifier	activation = 'relu', batch_size = 'auto', hidden_layer_sizes(100,), solver = 'sgd'	82.61
	activation = 'relu', batch_size = 'auto', hidden_layer_sizes (100), solver = 'lbfgs'	76.86
	activation = 'relu', batch_size = 'auto', hidden_layer_sizes (100), solver = 'adam'	83.98
	activation = 'relu', batch_size = 'auto', hidden_layer_sizes (50, 50), solver = 'adam'	85.22
	activation = 'relu', batch_size = 'auto', hidden_layer_sizes (50, 50), solver = 'sgd'	82.61

(continued)

Table 2 (continued)

Classifier	Parameters	Test accuracy (%)
	activation = 'relu', batch_size = 'auto', hidden_layer_sizes (50, 50, 50), solver = 'adam'	86.38
	activation = 'relu', batch_size = 'auto', hidden_layer_sizes (100, 50, 50), solver = 'adam'	87.41
	activation = 'relu', batch_size = 'auto', hidden_layer_sizes (100, 100, 100), solver = 'adam'	89.55
Linear SVC	penalty = '12', loss = 'squared_hinge', c = 1.0, multi_class = 'ovr', max_iter = 1000	73.51
	penalty = '12', loss = 'squared_hinge', c = 10, multi_class = 'ovr', max_iter = 10000	71.11
	penalty = '12', loss = 'squared_hinge', c = 1.0, multi_class = 'ovr', max_iter = 5000	75.80
	penalty = '12', loss = 'hinge', c = 1.0, multi_class = 'ovr', max_iter = 5000	72.86
	penalty = '12', loss = 'squared_hinge', c = 1.0, multi_class = 'crammer_singer', max_iter = 5000	76.94
	penalty = '12', loss = 'squared_hinge', c = 10.0, multi_class = 'crammer_singer', max_iter = 5000	75.26
	penalty = '12', loss = 'squared_hinge', c = 0.1, multi_class = 'crammer_singer', max_iter = 5000	79.32
GaussianNB	priors = None, var_smoothing = 1e–09	62.53
Convolutional neural network	CNN + Dropout	97.40
	CNN + Dropout + Batch Norm Before Activation	97.95
	CNN + Dropout + Batch Norm After Activation	98.11
	(As described in Sect. 3.1)	

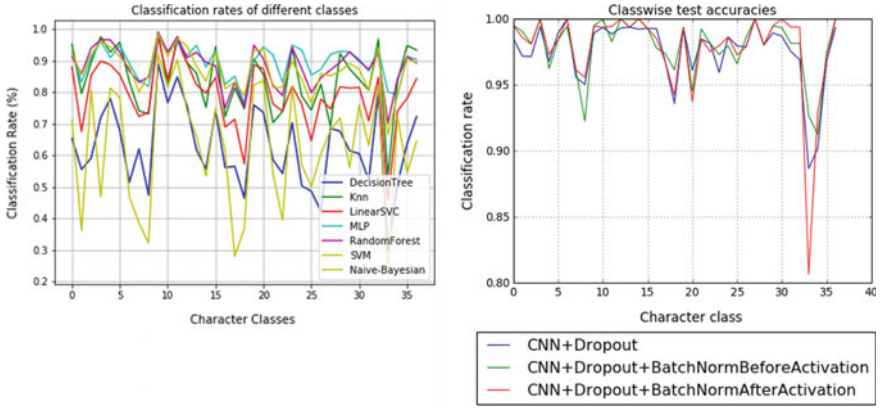
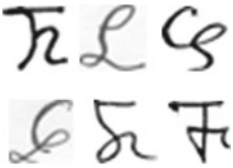


Fig. 3 Test accuracies of each class for different classification models

Fig. 4 Top six misclassified characters by seven models



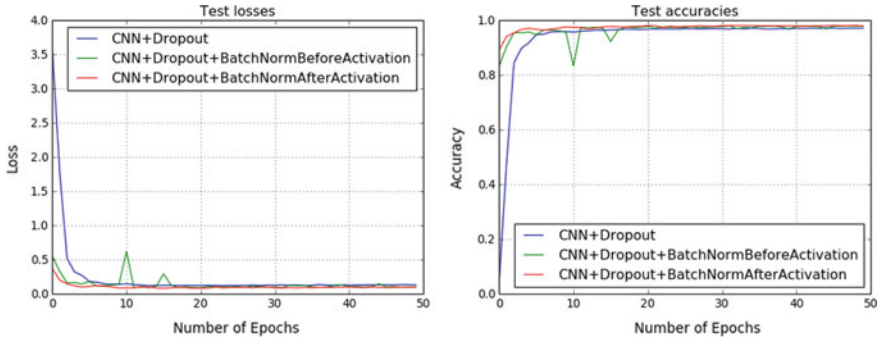


Fig. 5 Test losses and test accuracies achieved by three CNN models

Performance evaluation of the models is carried out in terms of test accuracies and weighted average precision, recall, and f1-score. Since the number of samples in each class vary slightly, the weighted evaluation measures have been considered. Table 3 shows results of the same.

Experimental results show that out of the seven classification algorithms, MLP gives highest accuracy of 89.55% followed by Random Forest and SVM with test accuracies of 88.41% and 87.51%, respectively. Lowest test accuracy is shown by Gaussian Naive Bayes algorithm with an accuracy of 62.53%. This gives an idea of how complex the dataset is as compared to other handwritten datasets such as MNIST for which the mentioned classification models could achieve better test accuracies [11].

Table 3 Values of evaluation measures for seven classification models

Classifier	Weighted avg precision	Weighted avg recall	Weighted avg f1-score	Test accuracy (%)
Decision tree	0.63	0.63	0.63	62.81
GaussianNB	0.63	0.63	0.63	62.53
KNeighbor classifier	0.86	0.85	0.85	84.76
Linear SVC	0.79	0.79	0.79	79.32
MLP	0.89	0.89	0.89	89.55
Random forest	0.88	0.88	0.88	88.41
SVM	0.88	0.88	0.88	87.51
CNN (Architecture 3)	0.98	0.98	0.98	98.11

5 Conclusion and Future Work

This paper presents a comparative study of seven popular classification models on the dataset of Meitei Mayek handwritten characters. A CNN model is also proposed which gives state-of-the-art test accuracy of 98.11%. A classwise analysis of the test accuracy shows that presence of certain confusing characters makes it challenging for classification models to achieve good accuracy. As part of future work, performance of classifiers using different hand-crafted features can be studied for the concerned dataset.

References

1. T. Thokchom, P. Bansal, R. Vig, S. Bawa, Recognition of handwritten character of manipuri script. JCP **5**(10), 1570–1574 (2010)
2. R. Laishram, A.U. Singh, N.C. Singh, A.S. Singh, H. James, Simulation and modeling of handwritten meitei mayek digits using neural network approach, in *Proceedings of the International Conference on Advances in Electronics, Electrical and Computer Science Engineering-EEC* (2012), pp. 355–358
3. R. Laishram, P.B. Singh, T.S.D Singh, S. Anilkumar, A.U. Singh, A neural network based handwritten meitei mayek alphabet optical character recognition system, in *2014 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC)* (IEEE, 2014), pp. 1–5
4. K.A. Maring, R. Dhir, Recognition of cheising iyek/eeyek-manipuri digits using support vector machines. Ijcsit **1**(2) (2014)
5. C.J. Kumar, S.K. Kalita, Recognition of handwritten numerals of manipuri script. Int. J. Comput. Appl. **84**(17) (2013)
6. C.J. Kumar, S.K. Kalita, Point feature based recognition of handwritten meetei mayek script, in *Advances in Electronics, Communication and Computing* (Springer, Berlin, 2018), pp. 431–439

7. K. Nongmeikapam, I. Manipur, I.W.K. Kumar, M.P. Singh, Exploring an Efficient Handwritten Manipuri Meetei-mayek Character Recognition using Gradient Feature Extractor and Cosine Distance Based Multiclass k-nearest Neighbor Classifier
8. S. Inunganbi, P. Choudhary, Recognition of Handwritten Meitei Mayek Script Based on Texture Feature
9. N. Kshetrimayum, A comparative Study of Meetei Mayek: From the Inscribed Letterform to the Digital Typeface. Unpublished Masters Dissertation. University of Reading. Reading, UK (2010)
10. F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, et al, Scikit-learn: machine learning in python. J. Mach. Learn. Res. **12**, 2825–2830 (2011)
11. H. Xiao, K. Rasul, R. Vollgraf, Fashion-Mnist: A Novel Image Dataset for Benchmarking Machine Learning Algorithms (2017). arXiv:1708.07747
12. S. Ioffe, C. Szegedy, Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift (2015). arXiv:1502.03167
13. N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, R. Salakhutdinov, Dropout: a simple way to prevent neural networks from overfitting. J. Mach. Learn. Res. **15**(1), 1929–1958 (2014)
14. D. Hijam, S. Saharia, Convolutional neural network based meitei mayek handwritten character recognition, in *International Conference on Intelligent Human Computer Interaction* (Springer, Berlin, 2018), pp. 207–219

A Literature Review on Energy-Efficient Routing Protocols for Heterogeneous WSNs



Isha Pant and S. K. Verma

1 Introduction

Wireless Sensor Networking (WSNs) domain has engrossed much consideration in the research field over the last few years, driven by affluence of practical and theoretical challenges and a rising number of practical civilian applications. The development of sensor networking mainly stands on the phenomenon of one-time deployment with multiple growing applications. Moreover, the expenditure of deploying thousands or even more figures of sensor nodes over a widespread geographical area is huge. The wireless sensor networks field consists of spatially dispersed self-directed schemes using mainly sensors to monitor environmental or physical conditions. All the respective nodes send their sensed value to the sink. The Lifetime of the network, Energy, and Stability are the main prominent parameters of sensor networks. The basic components of WSN field are an interconnecting network, assortment of dispersed or localized sensors, a sink for gathering data sensed by the sensors and a firm number of assets to handle data association at the sink [1]. WSNs comprises random type nodes or fixed type of nodes which develop themselves in the sensor field following a proper deployment or arbitrary placement mechanism. So, designing of solicitations and protocols for such type of networks needs to be energy efficient in order to lengthen the network lifetime because replacing the embedded sensors once installed tends to be a very difficult process. To overcome this situation, a methodology clustering came into light.

The rest part of the paper is systematized in the following sections: Sect. 2 deals with the process of clustering; Sect. 3 says about the literature-related work done.

I. Pant (✉) · S. K. Verma
Department of Computer Science, GB Pant Institute of Engineering & Technology, Pauri Garhwal, India
e-mail: kkavyapant97@gmail.com

S. K. Verma
e-mail: skverma.gbpec@rediffmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_8

Section 4 gives the findings from the literature, Sect. 5 contains the summary followed by the conclusion of the paper.

2 Clustering Process in Heterogeneous Wireless Sensor Networks

Prolonging network lifetime is the key aspect in wireless sensor networking and one such methodology for lengthening the lifetime of WSN is clustering. WSN consists of a small type of sensor nodes where nodes are assembled into clusters. Grouping the same capability or same functionality related elements together, or arranging them for a specific task, for the vital outcome is termed as cluster. Clustering is a mechanism through which the sensor network is distributed into clusters with a central controller for each, usually said to be the cluster head (CH). It performs all the work from collecting the sensed data value from the respective sensor nodes to transmitting that data value to the sink or the base station [2]. As a result, cluster head (CH) drops its energy very fast.

The clustering process constitutes three main stages. These stages are common for the clustering structure schema, as cluster head selection phase, cluster's setup phase, and the steady phase. The first stage of the process is the cluster head selection process; here, the desired (CH) is selected from the pool of available sensor network nodes. The (CH) can be fixed or it can be elected based on some sort of strategy, which is totally governed by the system working over the network [3]. After electing the cluster head, the next step is that it intimates all other cluster member's nodes which come under the coverage of communication radius. Intimation is all the way through advertising its identifiers based on id, the available energy level, its nearness with sink or all other cluster members. The second stage is the Cluster Setup Phase; here, the cluster member in response replies to the cluster head (CH) by updating appropriate information in the routing table such as receiving the hop count, relative distance criteria, strength of the signal, with the respective cluster head (CH). It is called the cluster setup phase. The last stage of the clustering process is the Steady Phase; here, all the cluster members send out their respective collected data to the CH node. The data placed on the unreachable accessing place is sensed by the sensor network. The basic working mechanism of CH is that it masses the sensed data and thereafter, sends it out to the sink.

Figure 1 defines the creation of various clusters in the WSNs field. These clusters contain various types of sensor nodes and each node transmits the sensed data to the desired cluster head which is held responsible for forwarding this form of sensed data to the sink node.

Homogeneous and heterogeneous are two main forms of clustering in WSN.

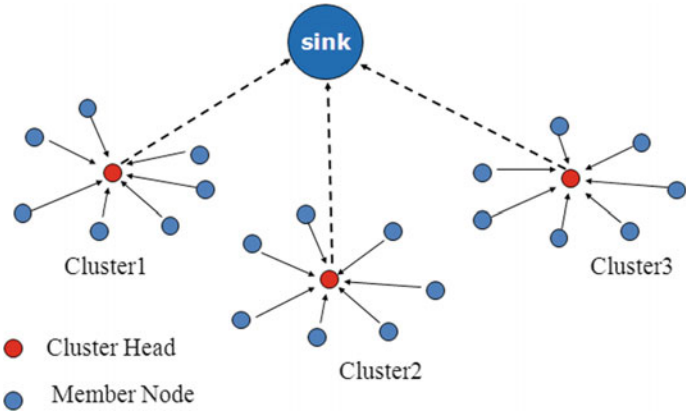


Fig. 1 WSNs cluster formation

2.1 Homogeneous Clustering Approach

In the homogeneous form of clustering, every node is considered equal in distribution of energy. The wireless sensors' life crucially depends on the lifetime of the battery. There are many energy-saving methodologies that are mainly introduced to minimalize the energy consumption and enhance the nodes' stability mechanism. There are many protocols for homogeneous WSNs type and heterogeneous WSNs type. The LEACH protocol is considered as one of the oldest homogeneous reactive communication WSN protocol based on the clustering mechanism given by Heinzelman [4]. In LEACH, the energy burden is spread throughout the network by using a suitable randomization type of practice. Clusters are created basically when the sensor nodes are combined, thereafter from designed cluster, one such node is called the CH. The CH wraps the data being gathered from the nodes and sends that particular sensed data to the desired (BS) [5]. CH election criteria are mainly grounded on the energy left at the respective node type. A sensor node randomly gets selected as the CH node having a certain type of probability at a given time. All sensor nodes in a homogeneous type of network hold similar expanse of energy; therefore, the LEACH protocol works in a homogeneous network, also it is not capable to work upon the heterogeneous network. Other protocols in the homogenous category are PEGASIS [3] and HEED protocol [3]. These algorithms will give poor performance in the heterogeneous type of network.

2.2 Heterogeneous Clustering Approach

In the heterogeneous WSNs type, the nodes are sited with different initial energies. In the heterogeneous clustering approach, different types of nodes are considered consuming distinct energy. The "Stable Election Protocol" (SEP), "Distributed Energy Efficient Clustering" (DEEC) [3], "Threshold Sensitive Stable Election Protocol" (TSEP), "Enhance threshold sensitive stable election protocol" (ETSSEP), "Dual Cluster Head Routing Protocol" (DCHRP), and "DCHRP with level four Heterogeneity" (DCHRP4) are the various protocols from the category of heterogeneous WSNs.

3 Literature Survey

Heterogeneous forms of protocols for wireless sensor networks are classified based on the performance metrics, namely the energy efficiency phenomenon, no. of heterogeneity levels, stability criteria, and CH cluster head selection [3]. Effective energy management is used for choosing the proper cluster head effectively. "Stable Election Protocol" (SEP), "Distributed Energy Efficient Clustering" (DEEC), "Threshold Sensitive Stable Election Protocol" (TSEP), "Enhance threshold sensitive stable election protocol" (ETSSEP), "Dual Cluster Head Routing Protocol" (DCHRP), and "DCHRP with four level of Heterogeneity" (DCHRP4) perform the desired election process by sorting the sensor nodes into different kinds [2].

3.1 The (SEP) Stable Election Protocol for Heterogeneous Networks

It is a heterogeneous wireless network protocol given by Smaragdakis in 2004. The protocol provides level two of heterogeneity containing some percentage of advance nodes (m) and normal nodes [3, 4]. The normal type nodes contain less aggregate of energy as equated to advanced node. Also, the advance node descends (α) extra energy than the general normal node type. The SEP is considered dynamic in nature, i.e., there is no such need to take up prior dissemination of the different energy stages in the sensor nodes.

In the SEP protocol, the sensor nodes are scattered uniformly, also the size of the network field is predefined, i.e., fixed, and the sink coordinates are known [4]. Nodes choose themselves as desired CHs based on lasting energy left in each sensor node. SEP uses the phenomenon of "Weighted election probability" for selecting the desired CH. The advantage of the protocol is that at every election round, the SEP protocol does not entail any inclusive energy knowledge. The downside of the SEP protocol methodology is that the election of CHs midst the two types of sensor

nodes is actually not of dynamic nature [3], which results in sensor nodes away from powerful nodes eventually dying first.

An enhancement over the SEP protocol was "Enhanced Stable Election Protocol" given by Femi Aderohunmu in 2009 [3]. It involves three forms of nodes referred to as the three levels in cluster development in a two-level order network as compared to the general SEP protocol. In 2012, a further enhancement over SEP was given by M. Islam namely the Extended SEP. The procedure to become a CH in ESEP is mainly the same as in the SEP protocol procedure by creating a random number and then relating it with the threshold value. In ESEP, the modest nodes or the intermediate type of nodes are elected in two conducts either by the energy level threshold between normal nodes and the advanced form of nodes or by taking relative space of advance nodes to the simple normal nodes. Another variant of the SEP family is Hierarchical-based SEP protocol (HSEP) proposed after ESEP. The growing distance formed between the (CH) and the sink or the base station (BS) outcomes in increasing transmission energy because most of the energy gets consumed in the transmission process. The HSEP protocol is proposed aiming to reduce the amount of transmission energy between the respective CH and BS. It brings into attention, the clustering order which lowers the energy and hence the transmission cost. In this type of clustering, two types of cluster heads are used, namely primary CHs and secondary CHs. The process of choosing the primary type of CH is exactly the same as in SEP, i.e., generating a random type of number ranging between value 0 and value 1 and then comparing it with the threshold value. The secondary CHs are selected from the primary CHs and they are chosen on the foundation of probability aspects from those types of nodes that had once previously become the primary CHs.

3.2 The (DEEC) Distributed Energy Efficient Clustering Protocol

The DEEC protocol, mainly based on the working of two stages of energy, is the same as the SEP protocol but rather it provides a better stability period. Here, cluster heads election is done by the possibility created on ratio basis among the energy remaining of the node in the sensor network and the average form of energy of the WSNs. DEEC constitutes of three forms of sensor nodes, mainly the normal nature nodes, the advanced nature nodes, and the super nature nodes having diverse energy levels [3]. The nodes having high residual amount of energy will likely have higher chances to be the next CH than the sensor nodes having less level of energy. That means, the advance nodes have greater possibilities to be selected as (CHs) than the normal type of nodes. The main advantages of the DEEC protocol are, like SEP, in here global information of energy is not required. Unlike the SEP protocol, the DEEC protocol performs better in heterogeneous multilevel wireless network [6].

It also constitutes a limitation, i.e., advanced nodes always get punished in the DEEC protocol, mainly when their left residual energy gets reduced and comes in

the normal node range. In this position, the advanced type of nodes die rapidly as compared to the other form of nodes. Various variants of the DEEC protocol are Developed DEEC (DDEEC) which is said to be the advanced version of DEEC proposed by Brahim Elbhiri in 2010; this protocol resolves the DEEC protocol's penalizing effect. DDEEC balances the election criteria of the CH, i.e., the cluster heads by following nodes' remaining energy. Therefore, advance nodes are possibly the ones to become CHs during the initial epoch but as soon as their energy functioning depletes, they become comparable to the normal type of sensor nodes. Enhanced DEEC (EDEEC) is the extended version of DEEC with normal nodes; the advance nodes and super form are classified on the basis of node energy proposed by Parul et al. in 2010. Another class of the DEEC protocol is Threshold-DEEC (TDEEC) [3] protocol having three types of different energy nodes with modified probability function. It uses the same selection criteria for the CHs as the DEEC protocol does. The EDDEEC protocol [7, 8] is the merged and more generalized improved EDEEC and DDEEC protocol versions.

3.3 The (TSEP) Threshold Sensitive Stable Election Protocol

The TSEP protocol is an enhancement of the base SEP protocol introduced by Kashaf et al. It comprises two features: first, "Level three heterogeneity" and second, "a reactive routing mechanism". TSEP is termed as a "reactive protocol" as here the transmission devours much energy than sensing; it occurs only when some specific threshold criteria are reached. It contains level three heterogeneity of the nodes comprising advance nodes, the intermediate nodes, and normal nodes [9]. The advance type of node possesses more energy than all further forms of nodes, the intermediate nodes comprises energy in between that of the normal nodes and the respective advance nodes while all other remaining nodes are termed as the normal nodes. The CH selection occurs on the basis of threshold criteria, namely the Hard-Threshold type and Soft-Threshold type.

The Hard-Threshold (HT). It is said to be the absolute form of sensed attribute past which nodes will start their transmission of the info to the cluster heads. As rapidly as the sensed data by the node becomes equal or greater to the desired threshold value, the node turns on its transmitter automatically and sends that sensed info to preferred CH.

The Soft-Threshold (ST). It is the minutest form of sensed type of value by which the respective nodes mainly switch on their operational transmitter automatically and start transmitting. So, the stability and lifetime criterion of the network firmly increases in the case of TSEP.

3.4 The (ETSSEP) Enhance Threshold Sensitive Stable Election Protocol

The ETSSEP protocol was given for the heterogeneous type of sensor networks by adding additional criteria of parameters by Shekhar [2]. Based on a cluster type reactive protocol, it constitutes of three forms of heterogeneity phenomenon, namely normal nodes, advance nodes, and intermediate nodes. The normal nodes are α times much less as in comparison to the advance nodes and β times as in comparison to the intermediate nodes. Also here, $\beta = \alpha/2$. The CH election is centered on the probability function and this function is dependent on the left or residual and usual network energy. Here in ETSSEP, the threshold value is planned mainly on the ratio of left residual and average value of energy and also on optimal no. of clusters at each round [10, 11]. Afterward, the calculated or generated value of threshold is adjusted for electing the desired CH. So, only the node carrying the highest energy left among other nodes will be considered as a cluster head. The concept of the discussed protocol comprises one cluster head selection phenomenon.

3.5 The (DCHRP) Dual Cluster Head Routing Protocol

An enhancement over ETSSEP is the Dual Cluster Head Routing DCHRP Protocol with three levels of heterogeneity; the main objective here is to lessen the election no. of CHs so that wastage of energy can be lessened. The dual cluster protocol implements the same concept as that of the ETSSEP protocol in terms of heterogeneity of sensor nodes. However, in the former, i.e., the ETSSEP protocol, [2] it is built on the phenomenon of selection of one cluster head for data aggregation, and the dual cluster head DCHRP protocol customs the conception of dual CHs for enhancing the life span of cluster heads resulting in the enhancement of the lifetime of the desired system [1].

3.6 The (DCHRP4) Dual Cluster Head Routing Protocol with Four-Level Heterogeneity

The DCHRP protocol comprises super nodes that describe WSN consisting mainly of four forms of sensor node types varying in energy levels. The nodes are divided as super nodes, the advance nodes, the intermediate form of nodes, and the normal type of nodes, respectively [1]. As the CH cluster head is having more work, the energy loss is imminent. Super, advanced, intermediate, and normal type of nodes are consigned with energy as μ, γ, β, and α, which are gauged using the Eqs. (1)–(4). The Normal nodes comprise μ less energy value than the super nodes and it is presumed that $\gamma = \alpha/4$. Also, the protocol ruptures the charge into two general

parts, thereby reducing the threat of the cluster head losing its energy too quickly. The introduction of super type of node in this protocol constitutes an additional bit of energy than the previously presented protocols. Energy for nodes can be calculated by formulas:

The Normal Nodes are given by

$$\text{Enrm} = \text{Eo} * (1 + \alpha) \tag{1}$$

The Intermediate Nodes are given by

$$\text{Eint} = \text{Eo} * (1 + \beta) \tag{2}$$

The Advanced Node given by

$$\text{Eadv} = \text{Eo} * (1 + \gamma) \tag{3}$$

The Super Node given by

$$\text{Esup} = \text{Eo} * (1 + \mu) \tag{4}$$

Inequalities of energy is given as Enrm < Eint < Eadv < Esup. Total amount of energy calculation is done as follows:

$$\begin{aligned} \text{Etot} &= \text{Eo}^* (1 + \alpha) + \text{Eo}^* (1 + \beta) + \text{Eo}^* (1 + \gamma) + \text{Eo}^* (1 + \mu) \\ &= \text{Eo}(4 + \alpha + \beta + \gamma + \mu) \end{aligned} \tag{5}$$

Equation (5) shows the nodes with four-level heterogeneity.

To find out which node must be the CH, first the maximum amount of node energy is established. For instance, let us consider the super form of nodes. From the assortment of energies of the super nodes, take the node with maximum function among all nodes. At the initial phase, all the nodes which are super type would be consuming the same amount of initial energies, so the first node among all would get elected from the super nodes. Suppose, super nodes = {0.85, 1, 0.95, 0.7, 08, 0.75, 0.82}; in the sequence, the node number 2 is having the max energy [1]; therefore, node 2 would be elected to figure out the probability. The equivalent procedure is continued for the advance nodes, intermediate nodes, and the normal type of nodes. From all the above, the DCHRP4 protocol with level four Heterogeneity is considered as the best one to select the leader or cluster head to date, as it gives an upgradation on the lifetime of the network up to an assured degree (Table 1).

Table 1 A comparison on WSN heterogeneous protocols

Performance criteria	SEP	TSEP	DEEC	ETSSEP	DCHRP	DCHRP4
Proposed year	2004	2012	2012	2015	2017	2018
Heterogeneity level	Two	Three	Multilevel	Three	Three	Four
Stability	High	More than SEP	High	Higher than SEP and TSEP	Higher than ETSSEP	More than ETSSEP and DCHRP
Cluster-head selection	Based on initial energy and the residual energy	Based on threshold	Based on initial, residual, and average node energy	Based on initial, residual, and average node energy	Based on maximum probability and node's energy	Based on threshold distance, energy, and probability of nodes
Energy efficiency	Good	Good	Good	Very good	Very good	Best

4 Findings from the Reviewed Literature

Selecting the cluster head (CH) at the center offers the best connectivity in the network. Researchers mainly prefer the random method of deployment for the sensor node placement; this causes a much less stable type of network than uniformly deployed network.

5 Summary of the Reviewed Literature

The Heterogeneous WSN protocols have the capability of managing the clusters in the sensor network field and their associated cluster member nodes. They are basically considered better in balancing the energy intake of the sensor type nodes in the whole sensor network [5]. Moreover, multi-hop path among cluster heads (CHs) to the sink is an important concern in saving energy during the transmission of the data. It is acknowledged that more precise the selection of the cluster head, better will be the cluster connectivity hence resulting in enhanced performance measures.

Hence at the end of this literature review, we say that the DCHRP4 protocol among all contributes all the way for better stability than the others [12], providing a four-level heterogeneity mechanism. Here, the concept of two main Cluster Heads, i.e., dual CH for enhancing the life span of CH is used. The other enhanced form of the protocol, ETSSEP, is mainly based on the selection of one cluster head phenomenon for data aggregation. The DEEC protocol mainly focuses on totally targeting the heterogeneity and average energy existing in the network, resulting in enhancement of life.

According to some of the authors, selected CH at the center gives the best node connectivity and better coverage results. Thus resulting in more stability and prolonged lifetime of WSN.

6 Conclusion

This paper delivers a general review of some of the latest and impactful researches being made in the field of WSNs mainly focusing on the criteria of enhancing the performance aspect in heterogeneous WSNs through effective WSN routing algorithm. The following paper discusses a variety of WSN routing algorithms of heterogeneous WSNs covering a wide analysis of work. The paper found that these discussed algorithms of heterogeneous WSN states can be exploited beneficially effectively for getting improved WSNs' performances, e.g., extending the desired network lifetime criteria, making improvements in reliable data delivery mechanism, and also making reduction in data transmission latency. Although most work taken up here is centered on the findings from different researches, the concepts specified in the paper are noteworthy in the development of improved routing algorithms for application-precise heterogeneous WSN-type scenarios. In the paper, various WSN heterogeneous forms of protocols for getting efficient energy are discussed. The energy efficiency criteria are the main points of discussion in any wireless networks. All the protocols discussed here overcome the energy consumption problem. The DCHRP4 with level four heterogeneity among all energy-efficient protocols has been provide better efficiency to date.

References

1. S.K. Verma, Y. Istwal, Dual cluster head routing protocol in WSN, in *ICCNT-2017* (2017)
2. S. Kumar, S.K. Verma, A. Kumar, ETSSEP: enhanced threshold sensitive stable election protocol for heterogenous WSN. Wirel. Pers. Commun. (2015)
3. F. Aderohunmu, J. Deng, An enhanced stable election protocol (SEP) for clustered heterogeneous WSN. Discussion Paper Series, October 2009 series, University of Otago, January 2009
4. S. Fahmy, O. Younis, HEED: a hybrid, energy-efficient, distributed clustering approach for ad-hoc sensor networks. IEEE Trans. Mobile Comput. 366–379 (2004)
5. G. Smaragdakis, I. Matta, A. Bestavros, SEP: a stable election protocol for clustered heterogeneous wireless sensor networks, in *Second International Workshop on Sensor and Actor Network Protocols and Applications (SANPA 2004)*, Boston (2004)
6. S. Bandyopadhyay, E.J. Coyle, An energy efficient hierarchical clustering algorithm for wireless sensor networks, in *Proceedings of INFOCOM*, April 2003
7. A. Kashaf, N. Javaid, I. Khan, Z. Khan, TSEP: threshold-sensitive stable election protocol for WSNs, in *10th International Conference on Frontiers of Information Technology (FIT)*, vol. 164 (2012), pp. 17–19

8. M. Liu, E. Duarte-Melo, Analysis of energy consumption and lifetime of heterogeneous wireless sensor networks, in *Proceeding of Global Telecommunications Conference (GLOBECOM)* (IEEE, 2002), pp. 21–25
9. H. Mohammad, M. Afsar, N. Tayarani, Clustering in sensor network: a literature survey. J. Netw. Comput. Appl. (2014)
10. C. Rosenberg, V. Mhatre, Homogeneous versus heterogeneous clustered sensor n/w: a comparative study, in *Proceedings of IEEE International Conference on Communications (ICC)* (2004)
11. W. Heinzelman, H. Balakrishnan, A. Chandrakasan, An application-specific protocol architecture for wireless microsensor networks. IEEE Trans. Wirel. Commun. **1**(4) (2002)
12. Y. Istwal, S.K. Verma, Dual cluster head routing protocol with super node in WSN. Wirel. Pers. Commun. (2018)

Location-Based Proactive Handoff Mechanism in Mobile Ad Hoc Network



D. Kalyani, Somula Ramasubbareddy, K. Govinda and V. Kumar

1 Introduction

In telecom, the term handoff alludes to the way toward exchanging ongoing call or session from one base station to another base station channel without the loss of information. Hard handover is a sort in which the divert in the source cell is discharged and at exactly that point, the direct in the objective cell is locked in. Thus, the connection to the source is broken before or "as" the connection to the target is made; for this reason, such handovers are referred to as *break-before-make*. Hard handovers are planned to be immediate, keeping in mind the end goal to limit the disturbance to the call. Soft handover is another type in which the divert in the source cell is held and utilized for some time in parallel with the direct in the target cell. For this situation, the association with the target is built up before the association with the source is broken; thus this handover is called make-before-break. The interval and amid are two associations which are utilized as a part of parallel, might be brief or substantial. Transmission Control Language is a high-level programming used for general purpose and is powerful; it is an interpreted and dynamic programming language. It supports procedural, functional, and object-oriented style of programming. It is commonly used in embedded applications along with C application programs for scripts, graphical user interface, and operating system applications and can be used for small- to large-scale applications.

D. Kalyani · S. Ramasubbareddy (✉)
Department of Information Technology, VNRVJIET, Hyderabad, Telangana, India
e-mail: svramasubbareddy1219@gmail.com

K. Govinda
SCOPE, VIT University, Vellore, Tamil Nadu, India

V. Kumar
Department of CSE, DRKIST, Hyderabad, Telangana, India

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_9

The aim of this work is to propose a method that will significantly improve handoff latency during handovers in a MANET which occurs due to the movement of nodes and inability to provide long-range coverage and hence transmission errors should be minimized and continuous connection with the Zone Controller is necessary. This calls for a revision in our existing systems and the work aims to do the same. The proposed architecture of WiMAX will follow the normal trend followed in MANETs which uses I telecommunication that has been validated in the market and adheres to the standards of wireless communication in open-system architecture The fact that a location-based handoff system is being used will ensure that data transmission does not falter when switching from one AP to another. Handoff latency will further be enhanced as it only occurs at predetermined points. Usage of control signal prevents continuous probing for subsequent candidate APs and also reduces the load on the ZC.

2 Literature Survey

MIMO-aided handover used in CBTC framework will minimize the handoff latency. In a handoff mechanism, the mobile station communicates with AP and the nearest access point in one base station to another base station access point without losing the information during communication. Handoff execution and transmission delay of the train and control information is analyzed in this scheme compared to traditional ones. The highest handoff latency is found to be 100 ms which makes this more efficient and reliable than other handoff techniques [1].

In this paper, the author assumed that the vehicle can move from one base station to another base station attached to fixed infrastructure. The handoff takes place from one access point to another access point in a fixed infrastructure for continuous communication between mobile nodes without information loss. The mobile node will move at a constant rate between the base stations and the handoff is also carried out at a constant rate without losing information. This scheme used the RFID tags which are used to reduce the handoff latency between the access points in fixed infrastructure. This is also helps in determining the upcoming AP and routing the data in an efficient way; handover latency is found to be 50 ms, hence this approach is $n =$ more efficient than other handoff mechanisms [2].

In this paper, they have explained about ERTMS (European Rail Traffic Management System). The communication technology that supports ERTMS is GSM-R, which supports high-mobility scenarios (max 500 km/h). Nevertheless, this technology offers the same data communication characteristics as the second generation mobile GSM. In heavy load situations such as complex yards or busy junctions, communication flows have presented bottlenecks for new high-priority connections. At the same time, in off railway environments, there are many standardization groups, IEEE and IETF, that focused on mobilizing the Internet by standardizing new access technologies and protocols. Some of the research working groups are related to technologies such as 802.11p, 802.16e or 802.20. In this paper, they have presented the

performance results obtained when deploying an ETCS application on an 802.16e WiMAX-based telecom architecture. The same QoS KPI demanded to GSM-R in EIRENE specification have been evaluated [3–5].

The main shortcomings of the above systems are the unreliability of the transmitted data in the network. The handoffs proposed either do not provide enough sufficiently low handoff mechanisms or do not enable reliable data transmission and could be prone to errors. In a CBTC, errors in data transmission cannot be tolerated as a train's safety is in question.

The above survey indicates that the proposal of a new scheme might be required which could eventually upgrade the existing methods and provide the CBTC networks with a more reliable and fast method of communication. The bandwidth requirement of the network is also a primary problem that the survey has indicated which the current methods fail to notice, thus providing a method which either overshoots the bandwidth requirements or fails to comply with the required speeds or amount of data.

3 Proposed Method

In a communication-based mobile node, one of the most important aspects is ensuring that a mobile station is connected to the Zone Controller to ensure the continuous and precise transmission of the node's location, speed, and other relevant information. Due to the continuous motion of the node, handoffs between different access points (AP) are necessary to remain connected. This system will minimize handoff latency and provide better connections. The architecture has been chosen keeping in mind the bandwidth requirements of the system and also the required network coverage area. The handoff mechanism in the proposed scheme is a *make-before-break* mechanism where the connection to a new AP is first made before an existing connection is terminated. The location-based handoff mechanism uses the concept of placing AP at the predetermined locations (end points of areas covered by each AP). Each mobile node is equipped with a radio transceiver and a WiMAX router. The radio transceiver contains channel information about the subsequent AP. Handoff is triggered at the location where the radio transceiver is present and a connection is made to the new AP based on the channel information obtained. Once the connection is successfully made, information about the node's location, speed, etc., are transmitted to the zone control and relevant braking information can be relayed back. The entry in the table is removed once the node leaves its coverage area. The proposed handoff method provides better handoff latency and error-free transmission in a system where precise information transmission is necessary.

4 Architecture

In this scheme, every node is furnished with a WiMAX router and a transceiver. The WiMAX routers are utilized to interface with the access points. It is productive in nature to store information about channel recurrence of the forthcoming access point. A group of APs forms a subnet. Each subnet is checked by a designated access point called an access router (AR). The starting router connected to a local server (LS) keeps track of local nodes within the subnet like node data and its relationship with it, the local server is associated with the global server (GS). The global server keeps track of all the nodes in the system and which node falls under which subnet and communication via different access points from the source to destination through handoff mechanism without losing the data. The local server keeps track of the node information such as MAC address of the node, Care of address of the node, and IPv6 address of the access point. The global server maintains node location table containing MAC address of the node, CoA of the node, and IPv6 address of the access router. Each access point gets access to node and other access points through a router; there is no redundancy of CoA as the CoA is doled out by the local server at the access router, likewise keeping an address table which is maintained to dole out CoA to nodes. This is a simple mechanism to keep track of all the nodes in the access router (Tables 1, 2, and 3).

In Fig. 1, the transaction of data during the handoff mechanism is shown. After scanning and obtaining the information of the next AP using the WiMAX router, the node sends REQ_IP to the corresponding access point. The access point at that point includes its IP address with a message and advances it to AR. LS related to the AR checks the accessibility of the MAC address of the node in the node interface table. In the event that it is available, the LS refreshes the table with the relating AP address in the table.

If it is not present in the table, at that point the LS includes the Care Of Address (COA) to the node in the table and updates the MAC address of the relating AP. At that point, AR sends an REP_IP to the node through AP. The AR sends the LOC_INFO to the GS through the zone controller. The GS checks for the availability of the node's location in the node location table. If present, then it updates CoA field and sends the old Care of Address to the old AR as an HO_CONFIRM message to advise that the old CoA is accessible for use.

In this plan, each of the hubs is outfitted with a WiMAX router. The WiMAX router is utilized to be associated with access points. They contain data on the AP

Table 1 Node link table

Media access control address of the node	CoA of node	IPv6 address of AP

Table 2 Node location table

Media access control address of node	CoA of node	IPv6 address of AR

Table 3 Address table of access router

Node ID	AB
2000:F000:0001	0
2000:F000:0001	1
2000:F000:0001	1

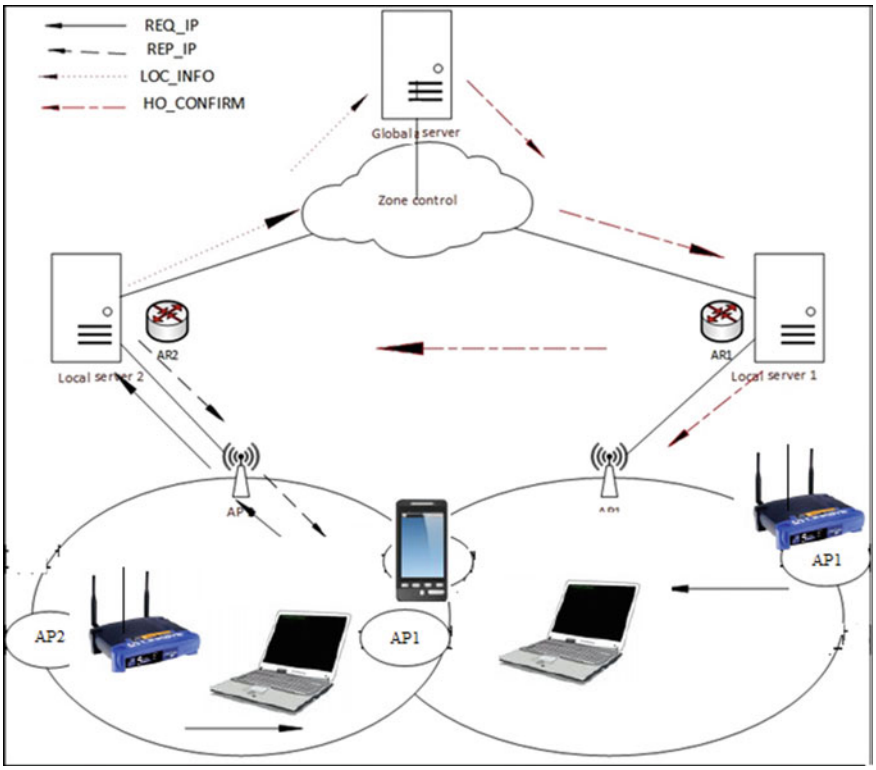


Fig. 1 Proposed handoff mechanism

that discloses the moving node in a specific direction. It is more proficient to store the data about path recurrence of the upcoming AP as opposed to broadcasting data tirelessly. All the access points in a certain range form a subnet and this access point are associated with the local server which in turn is connected to the global server via the Internet. The global server is responsible for keeping track of all the mobile nodes in the system as well as the access points and access router; the LS maintains node interface table containing MAC of the nodes, CoA of the node and IP address of the AP. Similarity, the global server also maintains the same data differentiated from AP, AR, and the node. The IP address of the AP indicates the rolled out node's IP address and propagates the same information to LS and AR,which is very simple to keep track of all the entities in the AR.

5 Result Analysis

The work was simulated using the Network Simulator 2 software and the results obtained through the simulation are shown in Figs. 2, 3, 4, and 5.

Figure 6 demonstrates the latency improvement after using the proposed plot. The difference in handoff latency for traditional methods has always been above 50 ms. The proposed system, however, provides the mobile node with a difference in handoff latency which is almost negligible. This is due to the very low Frame Error Rate. Therefore, the implementation of the scheme could prove very helpful.

Figure 7 compares the error-free periods between various handoff methods. For the traditional handoff schemes, the mean number of handoffs is identified with hysteresis and the receiving power threshold. The quantity of handoffs diminishes with an expansion in hysteresis. The number increases when the handoff threshold increases. The lower the threshold, the fewer the number of handoffs. The number of handoffs in our proposed scheme is fixed at 1 because the handoff occurs only at a predetermined location. In this way, the time interim between two back-to-back handoffs is the time that the node takes to traverse between two consecutive APs.

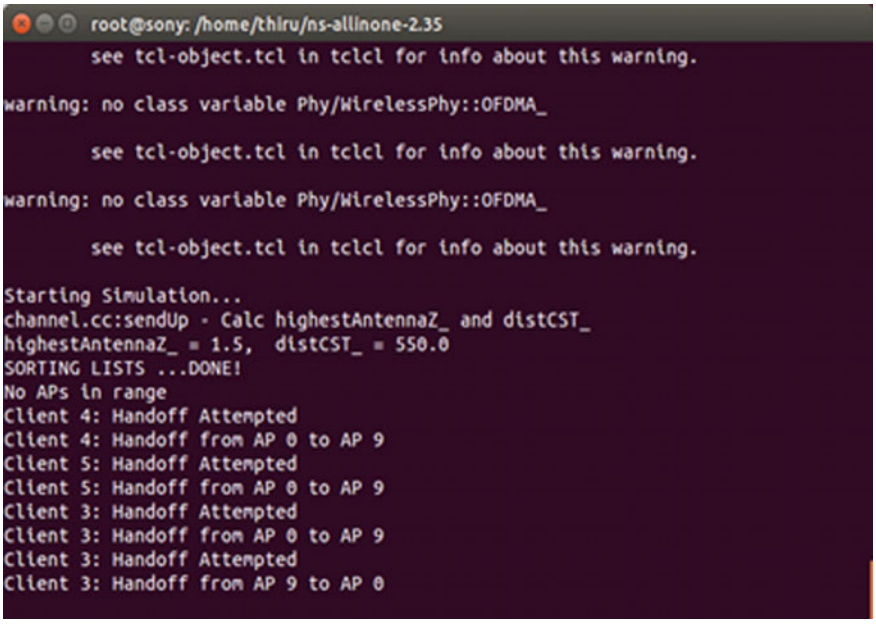


Fig. 2 Starting simulation

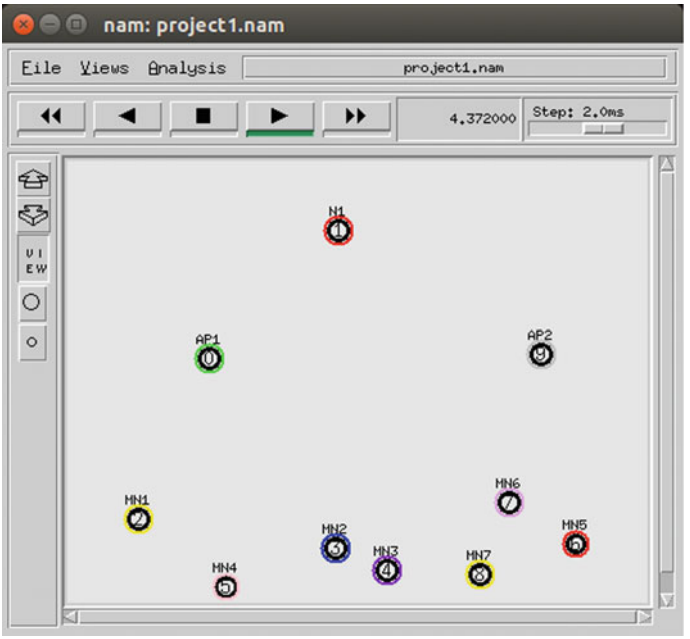


Fig. 3 AP and mobile nodes in position

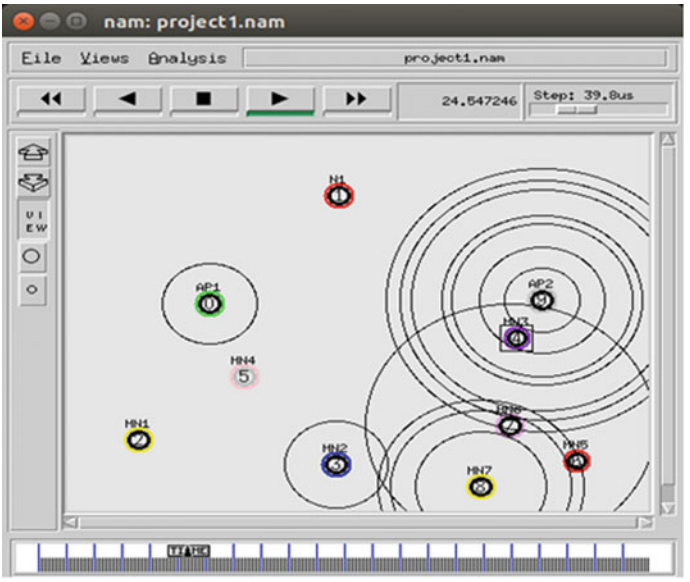


Fig. 4 Mobile node 6 during handoff

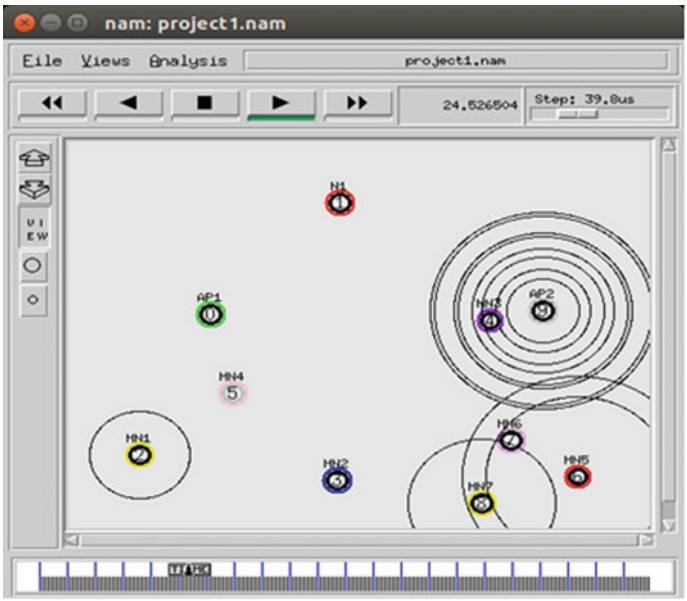


Fig. 5 Mobile node 6 after the handoff

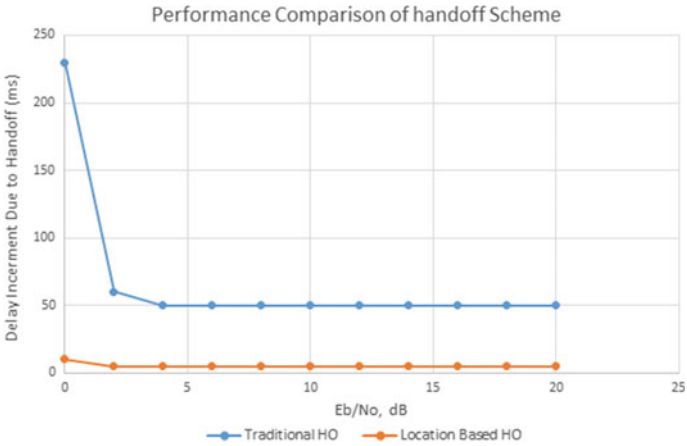


Fig. 6 Comparison of handoff latency between the traditional handoff schemes with the proposed one

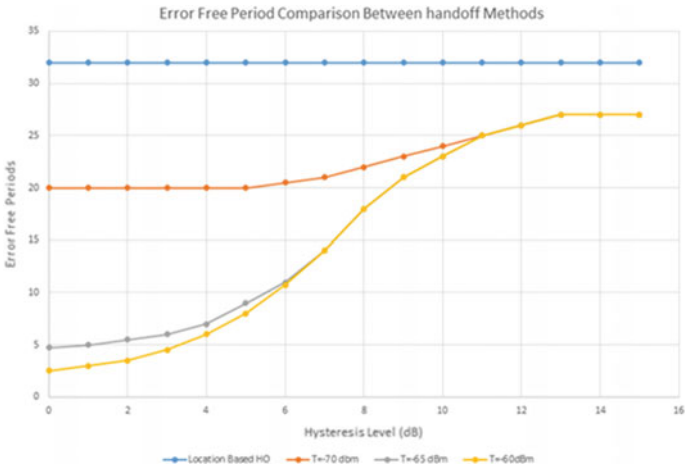


Fig. 7 Comparison of error-free periods between various handsoff methods

6 Conclusion

It can be concluded from our above work that the introduction of a location-based handoff system in the existing mobile node scenario could prove to be very useful in improving the standards of communication and node safety monitoring. The results prove that the new system performs significantly better than the existing methods and provides lower handoff latency and better FER (Frame Error Rates). The usage of AP at predetermined locations reduces the need for continuous broadcasting of messages to the Zone Controller which could increase the data traffic in the network and lead to a reduction in the QoS of the network. The assignment of addresses by the routers is also unique due to the presence of tables, which remove the worry of assignment of the same address to more than one node.

References

1. H. Jiang, V.C.M. Leung, C. Gao, T. Tang, MIMO-assisted handoff scheme for communication based train control system. IEEE Trans. Veh. Technol. **64**(4) (2015)
2. IEEE802.16e, IEEE Standard for Local and Metropolitan Area Networks. Part 16: Air Interface for Fixed and Mobile Broadband Wireless Access Systems Amendment 2: Physical and Medium Access Control Layers for Combined Fixed and Mobile Operation in Licensed Brands and Corrigendum 1 (2005)
3. X. Wang, H. Qian, A mobility handover scheme for IPv6-based vehicular ad hoc networks. Wirel. Pers. Commun. **70**(4), 1841–1857 (2013)

4. I. Ramani, S. Savage, Syncsan: practical fast handoff for 802.11 infrastructure networks, in *Proceedings of the IEEE 24th Annual Joint Conference, INFOCOM*, Istanbul, Turkey (2005), pp. 675–684
5. K. Ramachandran, S. Rangarajan, J. Lin, Make before break mac layer handoff in 802.11 wireless networks, in *Proceedings of the IEEE International Conference on Communication*, Istanbul, Turkey (2006), pp. 4818–4823

PSO-Based Improved DV-Hop Localization Algorithm in Wireless Sensor Networks



Gaurav Sharma and Arunacharam Rajesh

1 Introduction

Wireless Sensor Network (WSN) is an emerging technology and has attracted worldwide interest in research and industries, because of having a large number of applications in various areas such as surveillance, habitat and environmental monitoring, military applications, health care, structural monitoring, and disaster management [1, 2]. Among all these, many applications depend on the accurate locations of the nodes. Due to this, localization has become a crucial issue in wireless sensor networks.

Finding the geographical position of the nodes is known as node localization. Generally, node localization is of two types (i) Relative Location and (ii) Absolute Location. The relative location of the nodes is the location of the nodes with respect to any other object or node, whereas absolute location is the exact coordinates of the nodes. In this paper, absolute location of the nodes has been calculated using particle swarm optimization. Localization algorithms are basically divided into two categories: range based and range free [3, 4]. The range-free schemes use only connectivity information between the nodes. But range-free based algorithms show poor localization accuracy. In range-free algorithms, Distance Vector Hop (DV-Hop) algorithm has attracted the interests of researchers due to its simplicity and cost-effectiveness. Also, in range-free localization algorithms, there is no need for extra hardware. In this algorithm, the distance between the nodes is calculated using hop count and hop size of the nodes. To perform the effective node localization, only a

G. Sharma (✉) · A. Rajesh
Faculty in ECE Department, CVR College of Engineering, Hyderabad 501510, India
e-mail: ergaurav209@yahoo.co.in

A. Rajesh
e-mail: drrajesha86@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_10

small amount of location-aware nodes are deployed called anchor nodes, along with location-unaware nodes known as target nodes.

In this paper, a Particle Swarm Optimization (PSO) based Improved Distance Vector (PSO-IDV) hop algorithm which is a range-free distributed scheme has been proposed. PSO [5] is used to improve the localization accuracy of the DV-Hop algorithm.

2 Related Works

In the literature, a number of range-free localization algorithms have been proposed during the last two decades such as Centroid [6, 7], Distance Vector Hop (DV-Hop) [6, 8], Approximate Point In Triangle (APIT) [9], Convex Position Estimation (CPE) [10], Multidimensional Scaling (MDS) [11], and many more. But among all these range-free algorithms, DV-Hop has attracted more concentration of researchers because of its simplification, stability, feasibility, and fewer requirements of hardware. However, DV-Hop has poor localization accuracy and needs improvement in it. Therefore, an improved DV-Hop algorithm based on PSO is presented in the coming sections of this paper.

DV-Hop localization algorithm was first proposed by Niculescu and Nath [8]. Although it has poor localization accuracy, it is most widely used for large-scale applications.

Dengyi and Feng [12] proposed an improved DV-Hop algorithm for WSNs. In this, localization accuracy is improved by two methods throughout the whole process. First, the normal node refines the average distance per hop by taking the mean of anchor node's hop size. Second, the anchor nodes also refine the average distance per hop with the angle information between normal and anchor nodes.

Chen and Zhang [13] proposed an improved DV-Hop algorithm based on Particle Swarm Optimization (PSO). There are mainly four steps to describe the algorithm. Distance is calculated by normal nodes and position is estimated using 2D hyperbolic location algorithm [14]. Finally in the algorithm, PSO is used to improve the accuracy of the algorithm. It is very difficult to place the anchor nodes at the boundary of the inaccessible sensing field.

Peng and Li [15] proposed a genetic algorithm (GA) based DV-Hop algorithm. The complete algorithm is quite similar to the traditional DV-Hop algorithm. But in this method, localization accuracy is improved by using the genetic algorithm (GA). In GA, it is also difficult to find the algorithm-controlling parameters (crossover, selection, and mutation) to find the optimal solution.

From the above review, it is concluded that in WSN, localization is the optimization problem and its overall estimation error needs to be minimized. This encourages us to propose PSO-IDV.

3 Brief Overview of PSO

Particle Swarm Optimization (PSO) [5] is a computational method that optimizes a problem by iteratively trying to improve a candidate solution with regard to a given measure of quality. It solves a problem by having a population of candidate solutions, here dubbed particles, and moving these particles around in the search-space according to simple mathematical formulae over the particle's position and velocity. Each particle's movement is influenced by its local best known position, but is also guided toward the best known positions in the search-space, which are updated as better positions and are found by other particles. This is expected to move the swarm toward the best solutions.

PSO is originally attributed to Kennedy, Eberhart, and Shi and was first intended for simulating social behavior, as a stylized representation of the movement of organisms in a bird flock or fish school. The algorithm was simplified and it was observed to be performing optimization. The book by Kennedy and Eberhart [5] describes many philosophical aspects of PSO and swarm intelligence. An extensive survey of PSO applications is made by Poli. Recently, a comprehensive review of theoretical and experimental works on PSO has been published by Bonyadi and Michalewicz.

PSO is a metaheuristic as it makes a few or no assumptions about the problem being optimized and can search very large spaces of candidate solutions. However, metaheuristics such as PSO do not guarantee an optimal solution that is ever found [16]. Also, PSO does not use the gradient of the problem being optimized, which means PSO does not require that the optimization problem be differentiable as is required by classic optimization methods such as gradient descent and quasi-Newton methods.

A basic variant of the PSO algorithm works by having a population (called a swarm) of candidate solutions (called particles). These particles are moved around in the search-space according to a few simple formulae [16, 17]. The movements of the particles are guided by their own best known position in the search-space as well as the entire swarm's best known position. When improved positions are discovered, these will then come to guide the movements of the swarm [5]. The process is repeated and by doing so it is hoped, but not guaranteed, that a satisfactory solution will eventually be discovered.

Formally, let f: $Rn \rightarrow R$ be the cost function which must be minimized. The function takes a candidate solution as an argument in the form of a vector of real numbers and produces a real number as output which indicates the objective function value of the given candidate solution. The gradient of f is not known. The goal is to find a solution a for which $f(a) \leq f(b)$ for all b in the search-space, which would mean a is the global minimum. Maximization can be performed by considering the function $h = -f$ instead.

Let S be the number of particles in the swarm, each having a position $xi \cup Rn$ in the search-space and a velocity $vi \cup Rn$. Let pi be the best known position of particle i and let g be the best known position of the entire swarm.

The choice of PSO parameters can have a large impact on optimization performance. Selecting PSO parameters that yield good performance has therefore been the subject of many researches. The PSO parameters can also be tuned by using another overlaying optimizer, a concept known as meta-optimization, or even fine-tuned during the optimization, e.g., by means of fuzzy logic. Parameters have also been tuned for various optimization scenarios.

In relation to PSO, the word convergence typically refers to two different definitions:

- The convergence of the sequence of solutions (stability analysis, converging) in which all particles have converged to a point in the search-space, which may or may not be the optimum,
- Convergence to a local optimum where all personal bests p or, alternatively, the swarm's best known position g, approaches a local optimum of the problem, regardless of how the swarm behaves.

Convergence of the sequence of solutions has been investigated for PSO. These analyses have resulted in guidelines for selecting PSO parameters that are believed to cause convergence to a point and prevent divergence of the swarm's particles (particles do not move unboundedly and will converge somewhere). However, the analyses were criticized by Pedersen for being oversimplified as they assume the swarm has only one particle, that it does not use stochastic variables, and that the points of attraction [5], that is, the particle's best known position p and the swarm's best known position g remain constant throughout the optimization process. However, it was shown that these simplifications do not affect the boundaries found by these studies for parameters where the swarm is convergent.

Convergence to a local optimum has been analyzed for PSO. It has been proven that PSO needs some modification to guarantee finding a local optimum.

This means that determining convergence capabilities of different PSO algorithms and parameters therefore still depends on empirical results. One attempt at addressing this issue is the development of an "orthogonal learning" strategy for improved use of the information already existing in the relationship between p and g, so as to form a leading converging exemplar and to be effective with any PSO topology. This aims to improve the performance of PSO overall, including faster global convergence, higher solution quality, and stronger robustness. However, such studies do not provide theoretical evidence to actually prove their claims.

Consider a search space is d-dimensional and ith particle in the swarm which can be represented as $Xi = [xi1; xi2; : : :; xid]$ and its velocity can be represented by another d-dimensional vector $Vi = [vi1; vi2; : : :; vid]$. Let the best position ever visited in the past by the ith particle be denoted by $Pi = [pi1; pi2; : : :; pid]$. Many a time, the whole swarm is subdivided into smaller groups and each group/sub-swarm has its own local best particle denoted by $Pl = [pl1; pl2; : : :; pld]$, and an overall best particle, denoted as $Pg = [pg1; pg2; : : :; pgd]$, where subscripts l and g are particle indices. The particle iterates in every unit time according to (1) and (2):

$$v_{id} = wv_{id} + c_1 r_1 (p_{id} - x_{id})$$

$$+ c_2 r_2 \left(p_{gd} - x_{id}\right) + c_3 r_3 (p_{ld} - x_{id}) \quad (1)$$

$$x_{id} = x_{id} + v_{id} \quad (2)$$

The parameters w, c_1, c_2, and c_3 are termed as inertia weight, cognitive, social, and neighborhood learning parameters, respectively.

4 PSO-Based DV-Hop Algorithm

The proposed algorithm, i.e., PSO-based IDV-Hop algorithm for node localization in WSN is presented in this section. It is well-known that the DV-Hop algorithm is based on Distance Vector (DV) routing [14, 18].

A distance vector routing protocol requires that a router inform its neighbors of the topology changes periodically. Compared to link-state protocols, which require a router to inform all the nodes in a network of topology changes, distance-vector routing protocols have less computational complexity and message overhead.

The term distance vector refers to the fact that the protocol manipulates vectors (arrays) of distances to other nodes in the network. The Distance Vector algorithm was the original ARPANET routing algorithm and was also used on the Internet under the name of RIP (Routing Information Protocol).

Distance between the nodes is estimated by the hop count value multiplied by the hop size of the anchor. In the proposed algorithm, the beginning two steps are the same as the original DV-Hop algorithm. In step 3 of the algorithm, the calculated distance is modified. Consider D_{ij} is the distance between two anchors i and j, calculated according to (3), as follows:

$$D_{ij} = \text{HopSize}_i \times \text{hop}_{ij}, \text{ where } i \neq j \quad (3)$$

The actual distance between two anchor nodes can be calculated by their coordinates as follows:

$$D'_{ij} = \sqrt{\left(x_i - x_j\right)^2 + \left(y_i - y_j\right)^2} \quad (4)$$

where (x_i, y_i) and (x_j, y_j) are the coordinates of anchors i and j, respectively. Now from (3) and (4), the error between the estimated distance and actual distance: d^e_{ij} can be calculated by the following relationship [17].

$$d^e_{ij} = \left|D_{ij} - D'_{ij}\right| \quad (5)$$

Now, this distance error value is used to modify the average distance per hop value of the anchor node by adding the correction factor ψ to the previous hop size value

of the particular anchor. This correction factor is used to refine the average distance per hop deviation of the anchor node, which can be calculated as follows:

$$\psi_i = \frac{\sum_{i \neq j} d'_{ij}}{\sum_{i \neq j} h_{ij}} \tag{6}$$

Now finally, the specific distance between the unknown node and anchor node can be calculated as (7).

$$d'_{ij} = \left(\text{HopSize}_i + \psi_i\right) \times \text{hop}_{ij} \tag{7}$$

Still, the modified distance value calculated from (7) is not so much accurate because it is an estimated value; there must be error. Therefore, further to reduce the localization errors, an efficient optimization technique, i.e., PSO is used. So the localization problem can be formulated in terms of objective function as

$$f(x_u, y_u) = \text{Min}\left(\sum_{\substack{i=1,2\ldots k \\ u=k+1\ldots n}} \left| \sqrt{(x_u - x_i)^2 + (y_u - y_i)^2} - d_{iu} \right| \right) \tag{8}$$

where (x_i, y_i) is the coordinates of anchor nodes, $i = 1, 2\ldots k$. (x_u, y_u) is the estimated coordinates of unknown nodes, $u = k + 1, k + 2\ldots n$ and d_{iu} is the distance between anchor nodes and unknown nodes, calculated according to formula (6). There are total N sensor nodes in the network, i.e., $N = k + n$.

In Sect. 5, the simulation parameters, experimental results, and comparisons of the proposed algorithm to some existing algorithms are presented.

5 Results and Discussions

In our simulations, 100 sensor nodes (anchor nodes (k), and unknown nodes (n)) are evenly distributed randomly in two-dimensional fixed square area of 100 × 100 m^2 as shown in Fig. 1, where red pentacles represent the anchor nodes and black dots represent unknown nodes. Each node has an equal communication radius which is set to 25 m. Simulation parameters are shown in Table 1.

Localization error (LE) and mean localization error of the nodes have been calculated with different parameters under the same scenario to compare the performance of our proposed algorithms with other existing algorithms as follows:

$$\text{LE} = \sqrt{(x_u - x_a)^2 + (y_u - y_a)^2} \tag{9}$$

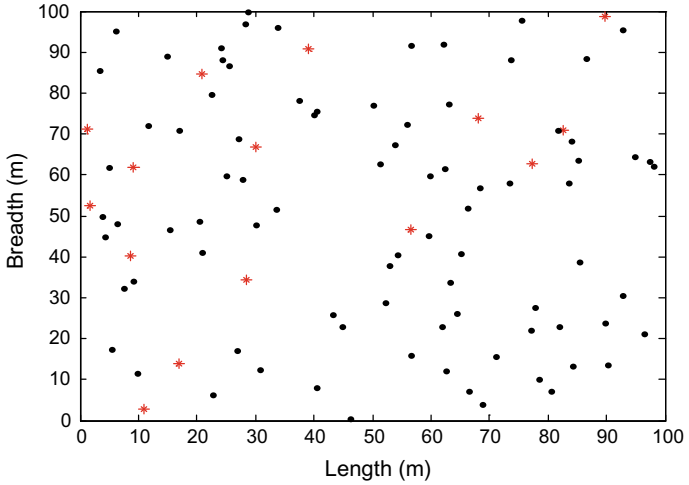


Fig. 1 Nodes distribution

Table 1 Simulation parameters

Parameters	Value
Total number of nodes	100
Area	100×100 m^2
Communication range	25 m
Number of anchor nodes	20%
Population	50
Generations	500
w	0.01
$c1$, $c2$, and $c3$	2.0

where (x_u, y_u) is the estimated coordinate and (x_a, y_a) is the actual coordinate of unknown node.

Mean localization error (MLE) can be calculated as the ratio of total localization error to the number of unknown nodes (n) and can be expressed as follows:

$$\text{MLE} = \frac{\sum_{a,u=1}^{n} \sqrt{(x_u - x_a)^2 + (y_u - y_a)^2}}{n} \tag{10}$$

Figure 2 shows the estimated localization mean error of these algorithms with a different number of anchor nodes deployed in the sensing field. It can be seen from the results that our algorithm performs better than the existing three algorithms of [8, 15], and [5, 13, 19]. It is observed that when the number of anchor nodes is more in the sensing field, MLE of the nodes becomes less.

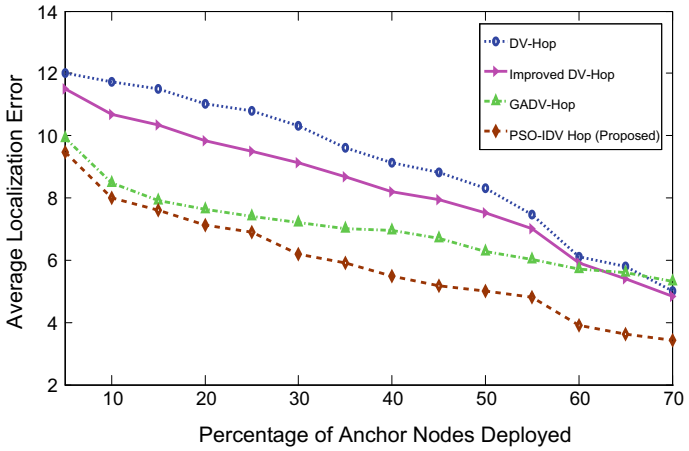


Fig. 2 MLE with different number of anchor nodes

The result data of each algorithm is taken as the average of 100 independent simulation experiments.

Figure 3 shows the MLE with a different number of sensor nodes. Sensor nodes are taken from 50 to 400 in the sensing field in which 20% of the nodes are assumed as anchors nodes. It is observed from the figure that MLE decreases as number of sensor nodes increases. It is due to the reasons that when the number of sensor nodes is more in the sensing field, the average hop size of the anchor nodes becomes more accurate and as a result localization accuracy is improved.

It can be seen from the simulation results of Figs. 2 and 3, the proposed algorithm, viz., IDV-Hop using PSO outperforms the existing algorithms.

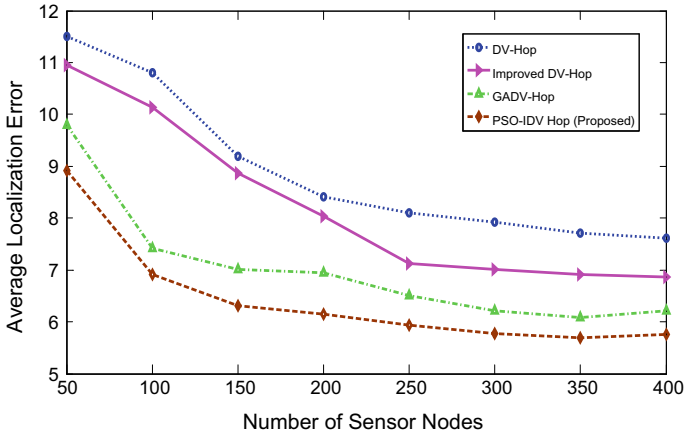


Fig. 3 MLE with different number of sensor nodes

6 Conclusions and Future Scope

Finding the accurate location of the sensor nodes in WSN is considered as the most critical issue because many WSN applications depend on the accurate location of the sensor nodes. Range-free localization algorithms have many advantages over range-based localization algorithms. But range-free algorithms show poor localization accuracy. To improve the localization accuracy of the range-free algorithms, PSO-IDV Hop technique for WSNs has been proposed. In the proposed algorithm, average hop size of an anchor is modified using correction factor and distance between the unknown nodes and anchors is calculated based on modified hop size. For further improvement in localization accuracy, an efficient optimization technique, viz., PSO has been used in this paper. Simulation results show that the proposed algorithm has better accuracy than DV-Hop, GADV-Hop, and DV-Hop using PSO algorithms.

References

1. I.F. Akyildiz, W. Su, Y. Sankarasubramaniam, E.E. Cayirci, A survey on sensor networks. IEEE Commun. Mag. **40**(8), 102–105 (2002)
2. A. Boukerchie, H.A.B.F. Oliveria, E.F. Nakamura, A.A.F. Loureiro, Localization systems for wireless sensor networks. IEEE Wirel. Commun. **14**(6), 6–12 (2007)
3. B.H. Wellenhof, H. Lichtenegger, J. Collins, *Global Positioning System: Theory and Practice* (Springer, 1997)
4. J. Zheng, C. Wu, H. Chu, Y. Xu, An improved RSSI measurement in wireless sensor networks. Procedia Eng. (Elsevier) **15**, 876–880 (2011)
5. J. Kennedy, R. Eberhart, Particle swarm optimization, in *Proceedings of the IEEE International Conference on Neural Networks*, vol. 4, pp. 1942–1948 (1995)
6. D. Niculescu, B. Nath, Ad Hoc positioning system (APS), in *Proceedings of the Global Telecommunication Conference (Globecom '01),* vol. 1, pp. 2926–2931, Nov. 2001
7. N. Bulusu, J. Heidemann, D. Estrin, GPS-less low cost outdoor localization for very small devices. IEEE Pers. Commun. Mag. **7**(5), 28–34 (2000)
8. D. Niculescu, B. Nath, DV based positioning in ad hoc networks. J. Telecommun. Syst. **22**(1–4), 267–280 (2000)
9. T. He, C. Huang, B.M. Blum, J.A. Stankovic, T. Abdelzher, Range-free localization schemes for large scale sensor networks, in *Proceedings of the ACM MobiCom '03*, pp. 81–95, Sept. 2003
10. L. Doherty, K.S.J. Pister, L.E. Ghaoui, Convex position estimation in wireless sensor networks, in *Proceedings of the IEEE INFOCOM '01*, vol. 3, pp. 1655–1663, Apr. 2001
11. Y. Shang, W. Ruml, Y. Zhang, M. Fromherz, Localization from mere connectivity, in *Proceedings of the ACM MobiHoc*, Annapolis, MD, June 2003, pp. 201–212 (2003)
12. Z. Dengyi, L. Feng, Improvement of DV-Hop localization algorithms in wireless sensor networks, in *International Symposium on Instrumentation & Measurement, Sensor Network and Automation (IMSNA)*, pp. 567–569 (2012)
13. X. Chen, B. Zhang, Improved DV-Hop node localization algorithm in wireless sensor networks. Int. J. Distrib. Sens. Netw. (2012). https://doi.org/10.1155/2012/213980
14. G. Sharma, A. Kumar, Fuzzy logic based 3D localization n wireless sensor networks using invasive weed and bacterial foraging optimization. Telecommun. Syst. **67**(2), 149–162 (2017)
15. B. Peng, L. Li, An improved localization algorithm based on genetic algorithm in wireless sensor networks. Cogn. Neurodyn. (Springer) **9**, 249–256 (2015)

16. G. Sharma, A. Kumar, Dynamic range normal bisector localization algorithm for wireless sensor networks. Wirel. Pers. Commun. **9**(3), 4529–4549 (2017)
17. G. Sharma, A. Kumar, Improved DV-Hop localization algorithm using teaching learning based optimization for wireless sensor networks. Telecommun. Syst. **67**(2), 163–178 (2017)
18. G. Sharma, A. Kumar, Modified energy-efficient range-free localization using teaching–learning-based optimization for wireless sensor networks. IETE J. Res. **64**(1), 124–138 (2017)
19. G. Sharma, A. Kumar, Improved range-free localization for three-dimensional wireless sensor networks using genetic algorithm. Comput. Electr. Eng. **72**(11), 808–827 (2018)

Countermeasures Against Variants of Wormhole in Wireless Sensor Networks: A Review



Manish Patel, Akshai Aggarwal and Nirbhay Chaubey

1 Introduction

Security is very important for resource-constrained wireless sensor networks due to their fundamental nature. The possible attacks in wireless sensor networks include sinkhole, wormhole, sybil, selective forwarding, black hole, etc. Wormhole is a gateway to many more attacks. To launch the attack is easy, but to detect it is very hard. For launching the attack, it is not required to know the cryptographic material or protocols used in the network. A malicious node attracts the traffic from one location and tunnels to another location and disturbs the whole routing process [1–3].

As shown in Fig. 1, location of malicious node M1 is far away from location of the malicious node M2. First, malicious node receives traffic from one part and tunnels it to the second malicious node which replies the traffic to another part of the network. Therefore, routing process in the network is disturbed. As shown in Fig. 1, node Y and Z become one-hop neighbors of node W and vice versa. There is no need to know the cryptographic mechanism used in the network and protocols or services offered in the network for launching the attack. The packets that pass through the wormhole can propagate faster compared to the normal path.

Section 2 discusses sinkhole-based wormhole attack and its countermeasures. Section 3 discusses the denial of service-based wormhole attack and its countermeasures. Section 4 discusses black hole-based wormhole attack and its countermeasures. Finally, conclusion is presented in Sect. 5.

M. Patel (✉)
Smt. S R Patel Engineering College, Unjha, Gujarat, India
e-mail: it43manish@gmail.com

A. Aggarwal · N. Chaubey
Gujarat Technological University, Ahmedabad, Gujarat, India

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_11

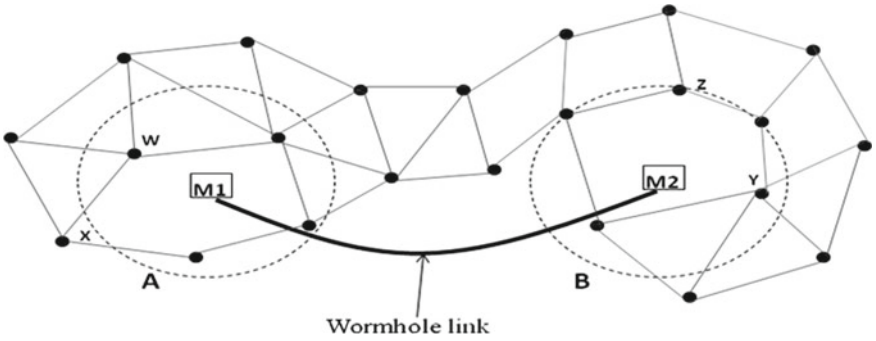


Fig. 1 Wormhole attack

2 Sinkhole-Based Wormhole and Its Countermeasures

In sinkhole-based wormhole attack, an attacker attracts the traffic toward it and then selectively forward the packets. One malicious node is located near to the destination and the second malicious node is located near to the source. The malicious node captures the route reply packet which is sent by the destination and tunnels it to second malicious node. In this way, path is established via malicious node.

In [4], the authors have proposed Received Signal Strength Indicator (RSSI) based approach for the detection of sinkhole and selective forwarding attack. Some extra monitor nodes are added and the RSSI value from the extra monitor nodes is used to determine the position of the sensor nodes. The extra monitor nodes are used to monitor the traffic. Sinkhole attack is detected using RSSI value and selective forwarding attack is detected by monitoring the traffic.

In [5], the authors have proposed sinkhole attack detection method by analyzing the consistency of data. Using data consistency, a list of suspected nodes is found and then, an intruder is identified by analyzing network flow information. It can effectively deal with multiple malicious nodes.

The authors have presented the vulnerabilities of Mintroute protocol and the detection method for sinkhole attack in [6]. The nodes send their local decision regarding the attack to the sink node. The neighbor nodes broadcast the route update messages and it is promiscuously listened by every node. Sink node receives the suspected node list if any rule is violated and the decision is taken based on the received alarms from the sensor nodes. The sink node checks the common suspected nodes in the lists sent by the sensor nodes.

Algorithm proposed in [7] is based on network flow graph. Initially, the suspected nodes are found and then, the intruder is located from the attacked area. The base station obtains more than one tree for network flow information. Using the depth-first search method, authors have calculated number of nodes in different trees. The intruder attracts more traffic; therefore, it is located at the root of the biggest tree.

3 Denial of Service-Based Wormhole and Its Countermeasures

The first malicious node tunnels route request packet to the second malicious node. The second malicious node broadcasts to its neighboring nodes and through the neighboring nodes, it reaches the destination. The neighboring nodes also receive the route request packet through the legitimate path. It is a duplicate packet received by neighboring nodes and will be dropped. Therefore, it cannot reach the destination. When the neighboring nodes receive route reply packet sent by the destination, the neighboring nodes cannot forward the route reply packet because they will not have the reverse route.

Authors have proposed a novel Message Observation Mechanism (MoM) for detecting the DoS attack in [8]. For isolating malicious nodes, it uses rekey and reroute countermeasures. Energy consumption is also reduced. Message observation mechanism consists of two types of lists: Normal Message List (NML) and Abnormal Message List (AML). MoM is deployed in a cluster head.

In [9], authors have proposed an approach for preventing denial of service attacks. It is operated in two phases: (1) Control node election and (2) Detection and blocking of malicious node. Every cluster head node monitors traffic in its cluster. If any node transmits number of messages higher than the threshold, then it is considered as a malicious node. After detecting the malicious node, all messages transmitted by the malicious node will be blocked.

The authors have proposed a localized clustering scheme for detecting attacks using traffic monitoring proxies on some nodes in WSNs in [10]. The method consists of two modules: (1) Cluster formation and the cluster head selection and (2) Session key establishment. The security is achieved using session key establishment.

The authors have proposed a hybrid approach using table and swarm-based protection for denial of service attacks in WSNs in [11]. Cluster head periodically checks that each node has sufficient trust value. If yes, then the routing path is formed. If no, then new routing path is calculated based on swarm-based defense. The variation in channel behavior is also identified. The faulty channel is mitigated using swarm-based approach.

For identifying compromised nodes, authors have to use recursive clustering-based approach in [12]. The recursive clustering process is done until the desired granularity is obtained. This approach is called k-clustering. The detection approach is based on building multicast tree.

In [13], the authors have explored the flooding-based denial of service attack and assess the lifetime of the network under the attacking scenario. The authors have presented multilevel analysis of the distributed denial of service attack. The role of the attacker is to generate a flood command. Authors have used protection modeling language.

Authors have presented DDoS attack prevention mechanism using Hidden Semi-Markov Model [14]. The system collects packets every second. For captured packets, selected features are calculated. These features are used to classify the incoming attack packets. Packet classification algorithm is applied and is estimated using Hidden Semi-Markov Model.

4 Black hole-Based Wormhole and Its Countermeasures

Source node broadcast the route request packet for establishing the path to the destination. This packet is immediately captured by the malicious node M1and through the tunnel; it is forwarded to the malicious node M2. Malicious node M2 sends it to the destination. The destination node sends route reply packet. Source node receives this route reply packet via the tunnel and path is created between source and destination via the tunnel. Data packets sent by the source node will be dropped without forwarding to the destination by the malicious node. It forms a black hole attack.

In indirect black hole attack, malicious node captures route reply packet and forward it to the target node T. The source node receives it from the target node. The source and other neighboring nodes consider the target node as the one-hop neighbor. The target node does not have a complete route toward the destination. Therefore, packet dropping occurs.

The authors have proposed Unmanned Aerial Vehicles (UAVs) based black hole attack detection method in [15]. Sequential Probability Ratio Test method is used to decide whether a node is black hole or not. UAV visits each node during network traversing. When messages are not received from a node, it is considered as a black hole node. If threshold is set to low, then genuine nodes are treated as black hole nodes. If threshold is set to high, then black hole nodes are treated as genuine nodes.

The authors have reviewed several existing black hole detection methods in [1]. In [2], each node observes the behavior of its neighbors. Every node overhears packets transmitted by its neighbors and identifies the suspicious nodes based on the behavior. The node detects the misbehavior of its neighbor. If the misbehavior entries exceed the threshold, then the node is considered as malicious node. All the suspicious nodes are verified during the next stage. The verification messages are sent to the root via an alternative path.

In [3], authors have presented black hole attack detection in cluster-based wireless sensor networks. Cluster head stores all node's IDs in its table. It starts timer and sensor nodes must send the data within that time. The malicious node will not forward the packets. Therefore, it is detected by the cluster head. If the cluster head becomes malicious, then it is detected by the base station.

In [16], authors have simulated black hole and selective forwarding attack. For detecting the attack, the base station monitors all the sensor nodes as it has high resources compared to other sensor nodes. The detection method is also energy efficient as there is no extra burden on sensor nodes to detect the attack.

Authors have proposed a black hole attack detection method based on data mining [17]. Network dataset is extracted. The ability of every node in terms of receive, send, forward, and drop is analyzed as high and low by keeping the threshold value. The main idea is to verify the node's behavior. Receiving maximum packets but not forwarding them, then that node will be treated as a black hole node. Nodes with maximum packet receiving ratio and zero forwarding ratios are treated as black hole nodes.

Authors have proposed an authentication mechanism for detecting black hole attack in wireless sensor networks [18]. After cluster formation, the cluster head node is elected. The responsibility of the cluster head is to detect the malicious node in its cluster. The cluster head maintains a table to include IDs of sensor nodes and sends the authentication packet to each of the sensor nodes. The legitimate nodes only send the reply message.

Authors have proposed a lightweight security scheme for the detection of selective forwarding attack in [19]. After obtaining responses from intermediate nodes, it uses multi-hop acknowledgment for launching alarms. Any packet loss is detected by an intermediate node with low communication overhead and good detection accuracy. In the case of poor radio conditions, detection accuracy is guaranteed.

In [20], authors have proposed support vector machines based centralized detection method. It is a part of machine learning algorithms. A one-class support vector machine is used for data pattern classification. The problems of curse of dimensionality and overfitting are avoided by SVMs. Without depleting the node's energy, high detection accuracy is achieved.

The authors have proposed checkpoint-based multi-hop acknowledgment method for the detection of selective forwarding attack [21]. The checkpoint nodes are used to find the area where selective forwarding attack is launched. For selecting checkpoint nodes, a fuzzy rule system is used. It considers the count of suspected nodes and the estimated distance from BS.

The authors have proposed a lightweight detection method based on neighbor information [22]. Each node maintains two-hop neighbor lists. Each node maintains a malicious counter. The node monitors the behavior of its neighbors. It monitors that the neighbor node forward the packet toward the sink node. If not, then counter value is incremented. When the counter value exceeds the threshold, the node is declared as a malicious node. Overhearing is used for reducing the transmission of alert packets and thereby consumes less energy.

In [23], IDS is proposed for the detection of selective forwarding attack in mobility-based WSNs. Each node collects information from all its neighbors and stores the number of packets sent and received into the table. During the data analysis step, sink node calculates the number of dropped packets and the probability of dropped packet. Sequential Probability Ratio Test is used during decision step. Finally, a compromised node is eliminated.

A fog computing-based system is proposed for the detection of selective forwarding attack in mobile wireless sensor networks [24]. The intrusion detection system in a fog server collects the information regarding received and forwarded packets of mobile monitor nodes. Watchdog is used to receive the information and maintains a

monitoring table for all nodes. The received information is analyzed and malicious node is detected using voting.

The authors have proposed a network area monitoring based approach for attack detection in [25]. An IDS is installed in each sensor node. It continuously monitors the sent packets and the overheard packets. A time attribute is assigned to each packet for detecting delay attack.

In [26], authors have proposed an adaptive and channel aware detection approach for selective forwarding attack. The adaptive detection threshold channel aware reputation system evaluates the behaviors of sensor nodes and accurately detects the compromised node. The authors have also proposed an attack tolerant data forwarding approach. Using this approach, the data delivery ratio is significantly improved.

For detecting the forwarding misbehavior, the authors have proposed a hop-by-hop cooperative detection approach in [27]. Here, each node overhears the packets and make records of forwarding operations. If the node has overheard a packet, the flag value is 1, otherwise, it is 0. Forwarding probability is quickly reduced for those malicious nodes which are detected multiple times.

Authors have proposed a multipath routing scheme for defense against selective forwarding attack in [28]. If packet drops occur, then the source resends the packets on the different alternate routes. The proposed algorithm consists of two phases: (1) Network construction and (2) Multipath routing. There exist many paths from source node to the base station. Using minimum number of hop count, packets are transmitted between source and the sink node. If the sender node cannot overhear the transmission of packet from the destination, then the sender node selects different alternate routes.

In [29], the authors have proposed detecting selective forwarding attack scheme using watermarking in WSNs. A trust value of each node is calculated for selecting the path for message forwarding. Watermark technology is applied for the detection of malicious nodes. Watermarking technique is used to calculate the packet loss rate. If the detected packet loss rate is higher than the normal rate, then the base station detects the misbehaving node hop-by-hop.

For detecting a selective forwarding attack, the authors have proposed a game theory model in [30]. Malicious nodes are detected using Zero-Sum game theory and selective node acknowledgment. The existing system is treated as a game. Two players, intruder and the detection system are involved in the game theory model. Multi-hop acknowledgment-based detection method is used for attack detection. Intermediate nodes are selected along the path for detecting malicious nodes.

Authors have presented the intrusion detection method to detect black hole and selective forwarding attacks using local information [31]. For detecting malicious behavior, nodes monitor their neighbors. Watchdogs are used to analyze communication links. For each packet that a node sends to its neighbors, watchdog temporarily buffers the packet and waits to examine that the neighbor node forward the packet or not. If neighbor node is not forwarding the packet, then the counter value is incremented.

In [32], authors have presented a selective forwarding attack detection method in WSNs using binary search. When the number of dropped packets is more than

the threshold value, then the cluster head raises an alarm message. For detecting the compromised node, hello packets and control packets are exchanged on a suspicious path. Multiple compromised nodes are also detected.

In [33], authors have proposed a selective forwarding attack detection method in heterogeneous sensor networks using a sequential probability ratio test. H-sensors and L-sensors are deployed in the networks. Authors have proposed a method that can report successful forwarding packets and drop packets to an H-sensor node. After receiving the report, H-sensor will execute a test to examine whether L-sensor is malicious or not.

In [34], authors have proposed a cumulative acknowledgment-based detection method for detecting selective forwarding attack. The proposed approach consists of three steps: (1) Construction of topology and selecting route, (2) Data transmission, and (3) Process of detection. The attack is detected using multi-hop acknowledgments. Some nodes are selected as checkpoint nodes along the forwarding route to send acknowledgments after receiving each packet. Malicious nodes are detected after receiving the acknowledgments.

The authors have proposed selective forwarding attack detection method based on sequential mesh test [35]. After receiving the packet drop report, the cluster head node detects the packet dropping node using sequential mesh test. This test extracts a small quantity of samples for running the test. It requires less computation and communication power with shorter detection time.

5 Conclusion

Sensor networks pose unique challenges. Therefore, traditional security algorithms cannot be applied to wireless sensor networks. There is enormous research potential in the field of security in WSNs. We have discussed variants of wormhole attacks and their countermeasures. Future work includes analysis and detection of variants of wormhole attacks (Denial of service, Sinkhole, Black hole, Selective Forwarding, etc.).

References

1. B.K. Mishra, M.C. Nikam, P. Lakkadwala, Security against black hole attack in wireless sensor network–a review, in *4th IEEE International Conference on Communication Systems and Network Technologies* (2014)
2. F. Ahmed, Y.-B. Ko, Mitigation of black hole attacks in routing protocol for low power and lossy networks. J. Secur. Commun. Netw. (2016) Wiley, New York
3. P. Dewal, G.S. Narula, V. Jain, Detection and prevention of black hole attacks in cluster based wireless sensor networks, in *3rd IEEE International Conference on Computing for Sustainable Global Development* (2016)

4. C. Tumrongwittayapak, R. Varakulsiripunth, Detecting sinkhole attack and selective forwarding attack in wireless sensor networks, in *7th International Conference on Information, Communications and Signal Processing (ICICS)* (2009), pp. 1–5
5. S.A. Salehi, M.A. Razzaque, P. Naraei, A. Farrokhtala, Detection of sinkhole attack in wireless sensor networks, in *IEEE International Conference on Space Science and Communication* (2013), pp. 361–365
6. M.A. Rassam, A. Zainal, M.A. Maarof, M. Al-Shaboti, A sinkhole attack detection scheme in mintroute wireless sensor networks, in *International Symposium on Telecommunication Technologies* (2012), pp. 71–75
7. C.H.E. Ngai, J. Liu, M.R. Lyu, On the intruder detection for sinkhole attack in wireless sensor networks, in *IEEE International Conference on Communications*, vol. 8 (2006), pp. 3383–3389
8. Y. Zhang, X. Li, Y. Liu, The detection and defence of DoS attack for wireless sensor network. J. China Univer. Posts Telecommun. **19**, 52–56 (2012). Elsevier
9. D. Mansouri, L. Mokddad, J. Ben-othman, M. Ioualalen, Preventing denial of service attacks in wireless sensor networks. in *IEEE Mobile and Wireless Networking Symposium* (2015)
10. P.P. Joby, P. Sengottuvelan, A localised clustering scheme to detect attacks in wireless sensor network. Int. J. Electron. Secur. Digit. Forensics **7**(3) (2015)
11. M. Gunasekaran, S. Periakaruppan, A hybrid protection approaches for denial of service (DoS) attacks in wireless sensor networks. Int. J. Electron. Taylor Francis 1362–3060 (2017). ISSN: 0020–7217 (Print)
12. S. Fouchal, D. Mansouri, L. Mokdad, M. Iouallalen, Recursive-clustering-based approach for denial of service (DoS) attacks in wireless sensors networks. Int. J. Commun Syst. **28**, 309–324 (2015)
13. K. Mazur, B. Ksiezopolski, R. Nielek, Multilevel modeling of distributed denial of service attacks in wireless sensor networks. J. Sens. **2016.** Article ID 5017248, 13
14. K.K. Oo, K.Z. Ye, H. Tun, K.Z. Lin, E.M. Portnov, Enhancement of preventing application layer based on DDOS attacks by using hidden semi-markov model, in *9th International Conference on Genetic and Evolutionary Computing* (2015)
15. M. Motamedi, N. Yazdani, Detection of black hole attack in wireless sensor network using UAV, in *7th IEEE International Conference on Information and Knowledge Technology* (2015)
16. M. Tripathi, M.S. Gaur, V. Laxmi, P. Sharma, Detection and countermeasure of node misbehaviour in clustered wireless sensor network. ISRN Sens. Netw. **2013**. Article ID 843626, 9
17. G. Kaur, M. Singh, Detection of black hole in wireless sensor network based on data mining, in *5th IEEE International Conference Confluence The Next Generation Information Technology Summit* (2014)
18. M. Wazid, A. Katal, R. Singh, Detection and prevention mechanism for blackhole attack in wireless sensor network, in *International Conference on Communication and Signal Processing* (India, 2013)
19. B. Yu, B. Xiao, Detecting selective forwarding attacks in wireless sensor networks, in *Proceedings 20th IEEE International Parallel and Distributed Processing Symposium* (2006)
20. S. Kaplantzis, A. Shilton, N. Mani, Y.A. Sekercioglu, Detecting selective forwarding attacks in wireless sensor networks using support vector machines, in *3rd International Conference on Intelligent Sensors, Sensor Networks and Information* (2007), pp. 335–340
21. S.J. Lee, I.G. Chun, W.T. Kim, S.M. Park, Control method for the number of checkpoint nodes for detecting selective forwarding attacks in wireless sensor networks, in *International Conference on Information and Communication Technology Convergence (ICTC)* (2010), pp. 537–538
22. T.H. Hai, E.-N. Huh, Detecting selective forwarding attacks in wireless sensor networks using two-hops neighbor knowledge, in *Seventh IEEE International Symposium on Network Computing and Applications* (2008), pp. 325–331
23. F. Gara, L.B. Saad, R.B. Ayed, An intrusion detection system for selective forwarding attack in IPv6-based mobile WSNs, in *13th International Wireless Communications and Mobile Computing Conference (IWCMC)* (2017), pp. 276–281

24. Q. Yaseen, F. AlBalas, Y. Jararweh, M. Al-Ayyoub, A fog computing based system for selective forwarding detection in mobile wireless sensor networks, in *IEEE 1st International Workshops on Foundations and Applications of Self Systems* (2016), pp. 256–262
25. M. Stehlik, V. Matyas, A. Stetsko, Towards better selective forwarding and delay attacks detection in wireless sensor networks, in *IEEE 13th International Conference on Networking, Sensing, and Control (ICNSC)* (2016), pp. 1–6
26. J. Ren, Y. Zhang, K. Zhang, X. Shen, Adaptive and channel-aware detection of selective forwarding attacks in wireless sensor networks. IEEE Trans. Wirel. Commun. **15**(5), 3718–3731 (2016)
27. S. Lim, L. Huie, Hop-by-Hop cooperative detection of selective forwarding attacks in energy harvesting wireless sensor networks, in *International Conference on Computing, Networking and Communications (ICNC)* (2015), pp. 315–319
28. P.C. Geethu, A.R. Mohammed, Defense mechanism against selective forwarding attack in wireless sensor networks, in *Fourth International Conference on Computing, Communications and Networking Technologies* (2013), pp. 1– 4
29. D.-Y. Zhang, C. Xu, L. Siyuan, Detecting selective forwarding attacks in WSNs using watermark, in *International Conference onWireless Communications and Signal Processing* (2011), pp. 1–4
30. Y.B. Reddy, S. Srivathsan, Game theory model for selective forward attacks in wireless sensor networks, in *17th Mediterranean Conference on Control and Automation* (2009), pp. 458–463
31. M. Tiwari, K.V. Arya, R. Choudhari, K.S. Choudhary, Designing intrusion detection to detect black hole and selective forwarding attack in wsn based on local information, in *Fourth International Conference on Computer Sciences and Convergence Information Technology* (2009), pp. 824–828
32. S. Mukherjee, M. Chattopadhyay, S. Chattopadhyay, P. Bose, A. Bakshi, Detection of selective forwarding attack in wireless ad hoc networks using binary search, in *Third International Conference on Emerging Applications of Information Technology* (2012), pp. 382–386
33. J. Brown, X. Du, Detection of selective forwarding attacks in heterogeneous sensor networks, in *IEEE International Conference on Communications* (2008), pp. 1583–1587
34. Y.K. Kim, H. Lee, K. Cho, D.H. Lee, CADE: Cumulative acknowledgement based detection of selective forwarding attacks in wireless sensor networks, in *Third International Conference on Convergence and Hybrid Information Technology*, vol. 2 (2008), pp. 416–422
35. G. Li, X. Liu, C. Wang, A sequential mesh test based selective forwarding attack detection scheme in wireless sensor networks, in *International Conference on Networking, Sensing and Control* (2010), pp. 554–558

Adaptive Backup Power Management in Ad Hoc Wireless Network



Ganesh Gupta, Vivek Jaglan and Ashok K. Raghav

1 Introduction

Due to the architecture of mobile ad hoc networks, power control is one of the difficult problems. Here, one node at a time can act as source, receiver, and router, which consumes extra power for operating themselves. Ad hoc network was first introduced for facilitating military services. The source node can transmit the message to destination node if both nodes are within communication range otherwise the source node can send message through intermediate nodes to the destination network [1]. Wireless nodes are powered by battery having limited capacity of energy. The network can adopt one solution by controlling battery power consumption to prolong their life. Remaining battery power is one of the main facts for determining the actual life of battery and node. The consumption of energy can be also minimized with help of utility software and compact hardware handling techniques in wireless Ad hoc network. The overall performance of network could be greatly influenced by efficient utilization of battery power. Power controlling techniques are one of the serious issues of mobile Ad hoc network [2].

G. Gupta (✉) · V. Jaglan · A. K. Raghav
School of Engineering and Technology, Amity University, Haryana Gurugram, India
e-mail: ganeshgupta81@gmail.com

V. Jaglan
e-mail: vjaglan@ggn.amity.edu

A. K. Raghav
e-mail: akraghav@ggn.amity.edu

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_12

2 Related Work

The available literature is mainly focused on different power controlling algorithms and various distributed power control scheme which is also concerned with scheduling of power, software, and hardware effect on power consumption.

2.1 Power Control Algorithms

The noncooperative power control with self-incentive property in Ad hoc network was proposed by Long et al. Here, the power level has been controlled after considering QoS requirement and the energy-efficient property. The concept of power is not matching with the current situation where security also matters [3]. Zheng et al. have told about combination of power control, link schedule, and rate control. Here, they considered the practical aspects of link scheduling interference; however, transport layer controls the congestion of data. The suggested algorithm only works for small and less complex network [4]. The heuristic algorithm suggested by Song Guo et al. handles the larger network with minimum transmission energy which is ineffective in maintaining the node power level for link scheduling [5].

2.2 Gap

It has been observed from the above section literature survey that determination of individual node power level is a difficult task and not possible at all; however, estimated power could be decided using reinforcement learning algorithm. The results obtained by the algorithm were not significant due to not proper encoding of interfacing scheme.

2.3 Different Power Control Scheme and Power Scheduling

The performance of Ad hoc network can improve with an optimal power consumption of every node which is maintained by power control schemes and by scheduling. Yan Chen et al. have measured a different power allocation for maintaining power efficiently on a given data rate between the source node and intermediate nodes of the given network. Due to the congestion problem, these networks are not considered to be much reliable [6]. V.P Singh et al. have shown the significant power control improvement in neural base energy consumption and throughput [7]. H. Pimentel et al. have focused on creative behaviors of ad hoc networks, where the introduction of extra routes increases the performance of the network as well as its energy-saving

due to solving the congestion issue in the network [8]. Lee et al. have searched power scheduling and controlling scheme for the communication. The network node in this scheme consumes more energy here [9]. Acharya et al. have given a scheme for random selection of nodes, where it consumes less energy in transmission of data and network never fails down. The performance of this scheme degraded with high data transfer rates [10].

2.4 Gap

It is analyzed that the power control scheme and its scheduling can improve the energy efficiency; however, these schemes have several limitations in terms of performance and data transfer speed. The efficiency of network will also be affected due to congestion in routing path of the network.

3 Proposed Power Awareness Scheme

As shown in Fig. 1, the given network topology has source node S, destination node D, and rest nodes are intermediate node for forwarding outgoing and receiving incoming packets. If we assume current battery power of node E, F, and G is enough for receiving and forwarding data packet which can sustain this route for a long lifetime, then it will follow it.

Let if node F has (in Fig. 2) insufficient battery after time t so that it can't participate in the network, then the neighbor node E will first inform to the rest of adjacent nodes and it will try to discover new route, which will be less energy-consuming with longtime sustainability.

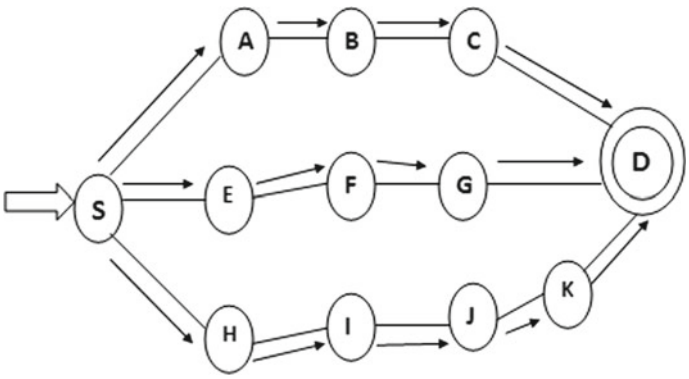


Fig. 1 Network topology [11]

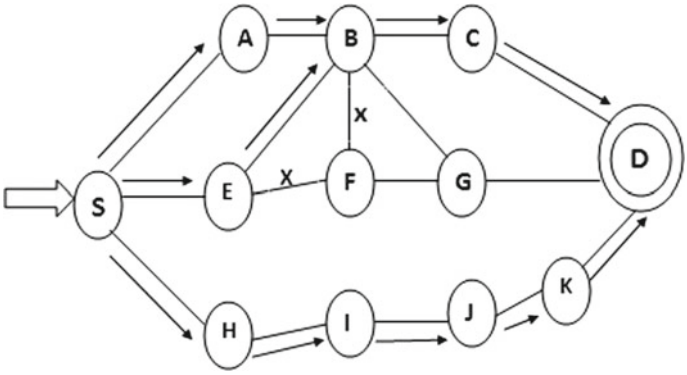


Fig. 2 Efficient route selection

After finding out the appropriate adaptive energy backup route, node E will find a route toward the destination and it will reroute the packet through node B.

Hence after some time, we can notice that energy-draining through node B and C is quicker than rest routes. As both nodes are forwarding more packet toward the destination. Certainly, before reaching threshold level of energy, the network should decide the alternative power-saving route (Fig. 3).

When the links break inside topology, an adaptive backup route request is generated, and the process of handshaking backup route reply takes place. It provides new path information based on current battery power as well as their neighbors signal range. The negotiations are based on alternate path longtime backup battery life. This adaptive scheme will also enhance the throughput of network.

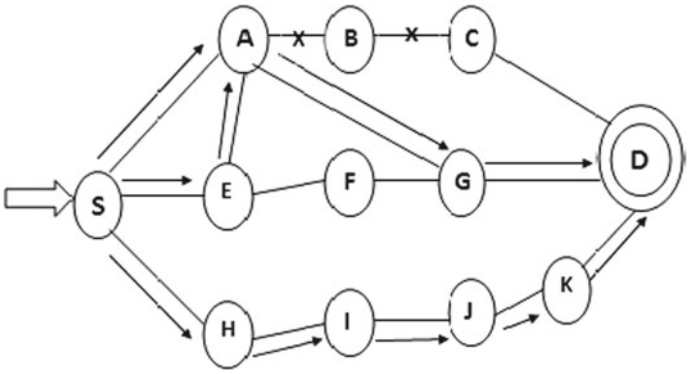


Fig. 3 Reroute selection

4 Simulation and Results

Assuming the following four linguistic input parameters.

1. Hop Count
2. Throughput of Network
3. Energy Remain
4. Number of packets drop

Here Route lifetime (RL) is considered as an output parameter (Table 1).

The variables are further divided into seven linguistic variables for both input and output parameters like very low, low to high, and very high. One can decide a most suitable route in terms of maximum RL and rest route as options for alternating routes.

$$\mathrm{RL} = \mathrm{C1} + \mathrm{C2} + \mathrm{C3} + \mathrm{C4}/4 \tag{1}$$

where C1, C2, C3……are components of fuzzy suitability index (Tables 2, 3, 4, 5).

Table 1 Linguistic hop count variable

Linguistic value	Notation	Range
Very low	HCVL	[HCVLa, HCVLb]
Low	HCL	[HCLa, HCLb]
Less low	HCLL	[HCLLa, HCLLb]
Medium	HCM	[HCMa, HCMb]
Less high	HCLH	[HCLHa, HCLHb]
High	HCH	[HCHa, HCHb]
Very high	HCVH	[HCVHa, HCVHb]

Table 2 Linguistic throughput variable

Linguistic value	Notation	Range
Very low	TNVL	[TNVLa, TNVLb]
Low	TNL	[TNLa,TNLb]
Less low	TNLL	[TNLLa, TNLLb]
Medium	TNM	[TNMa, TNMb]
Less high	TNLH	[TNLHa,TNLHb]
High	TNH	[TNHa, TNHb]
Very high	TNVH	[TNVHa, TNVHb]

Table 3 Linguistic energy remain variable

Linguistic value	Notation	Range
Very low	ERVL	[ERVLa, ERVLb]
Low	ERL	[ERLa, ERLb]
Less low	ERLL	[ERLLa, ERLLb]
Medium	ERM	[ERMa, ERMb]
Less high	ERLH	[ERLHa, ERLHb]
High	ERH	[ERHa, ERHb]
Very high	ERVH	[ERVHa, ERVHb]

Table 4 Linguistic packet drop variable

Linguistic value	Notation	Range
Very low	PDVL	[PDVLa, PDVLb]
Low	PDL	[PDLa, PDLb]
Less low	PDLL	[PDLLa, PDLLb]
Medium	PDM	[PDMa, PDMb]
Less high	PDLH	[PDLHa, PDLHb]
High	PDH	[PDHa, PDHb]
Very high	PDVH	[PDVHa, PDVHb]

Table 5 Linguistic route lifetime variable

Linguistic value	Notation	Range
Very low	RLVL	[RLVLa, RLVLb]
Low	RLL	[RLLa, RLLb]
Less low	RLLL	[RLLLa, RLLLb]
Medium	RLM	[RLMa, RLMb]
Less high	RLLH	[RLLHa, RLLHb]
High	RLH	[RLHa, RLHb]
Very high	RLVH	[RLVHa, RLVHb]

4.1 Evaluation of Different Case Statements

Case1: Inference Rules between Hop Count and Throughput of Network
R1: When Hop Count is HCVH, then Throughput of Network is TNVL
R2: When Hop Count is HCH then Throughput of Network is TNL
R3: When Hop Count is HCLH, then Throughput of Network is TNLL
R4: When Hop Count is HCM, then Throughput of Network is TNM
R5: When Hop Count is HCLL, then Throughput of Network is TNLH

R6: When Hop Count is HCL, then Throughput of Network is TNH
R7: When Hop Count is HCVL, then Throughput of Network is TNVH

Case2: Inference Rules between Energy Remain and Number of Packets Drop
R1: When Energy Remain is ERVH, then Number of Packets Drop is PDVL
R2: When Energy Remain is ERH, then Number of Packets Drop is PDL
R3: When Energy Remain is ERLH, then Number of Packets Drop is PDLL
R4: When Energy Remain is ERM, then Number of Packets Drop is PDM
R5: When Energy Remain is ERLL, then Number of Packets Drop is PDLH
R6: When Energy Remain is ERL, then Number of Packets Drop is PDH
R7: When Energy Remain is ERVL, then Number of Packets Drop is PDVH

Case3:: Inference Rules between Hop Count and Route Lifetime
R1: When Hop Count is HCVH, then Route Lifetime is RLVH
R2: When Hop Count is HCH, then Route Lifetime is RLCH
R3: When Hop Count is HCLH, then Route Lifetime is RLLH
R4: When Hop Count is HCM, then Route Lifetime is RLM
R5: When Hop Count is HCLL, then Route Lifetime is RLLL
R6: When Hop Count is HCL, then Route Lifetime is RLL
R7: When Hop Count is HCVL, then Route Lifetime is RLVL

Case4: Inference Rules between Packet Drop and Route Lifetime
R1: When Packet Drop is PDVH, then Route Lifetime is RLVL
R2: When Packet Drop is PDH, then Route Lifetime is RLL
R3: When Packet Drop is PDLH, then Route Lifetime is RLLL
R4: When Packet Drop is PDM, then Route Lifetime is RLM
R5: When Packet Drop is PDLL, then Route Lifetime is RLLH
R6: When Packet Drop is PDL, then Route Lifetime is RLH
R7: When Packet Drop is PDVL, then Route Lifetime is RLVH

Function for Different Possible Route Lifetime
As the declaration is required for different route lifetime, Table 6 shows these possible linguistic values (Figs. 4, 5, 6, 7 and 8).

Table 6 Route membership function

Linguistic value	Value of range
Very low	(0,0,0,0.16),
Low	(0,0.16,0.16,0.32)
Less low	(0.16,0.32,0.32,0.48)
Medium	(0.32,0.48,0.48,0.68)
Less high	(0.48,0.64,0.64,0.80)
High	(0.64,0.80, 0.80.0.96)
Very high	(0.80, 0.80,1,1)

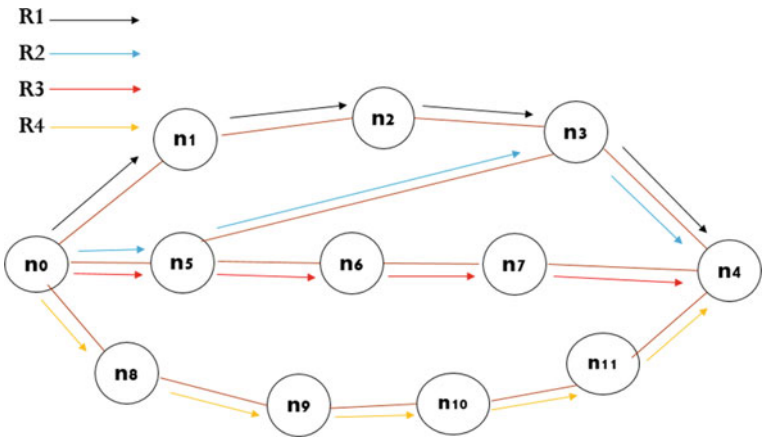


Fig. 4 Route selection [12]

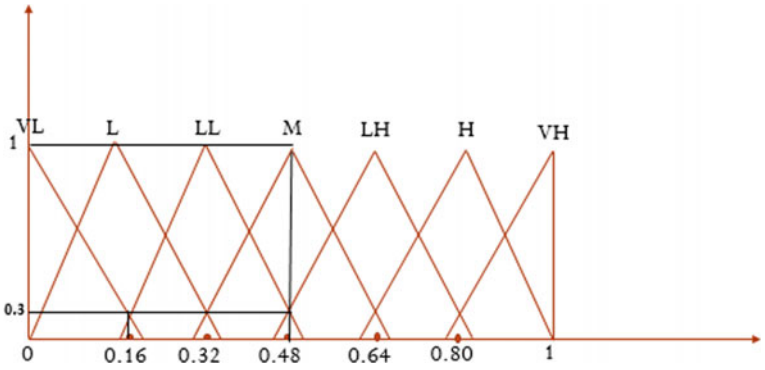


Fig. 5 Value of truth for route 1

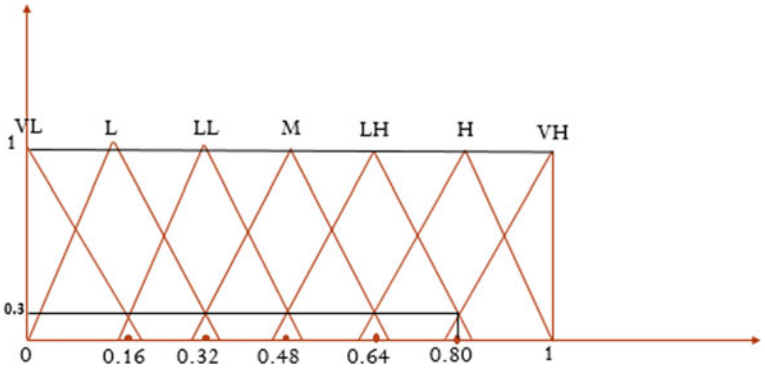


Fig. 6 Value of truth for route 2

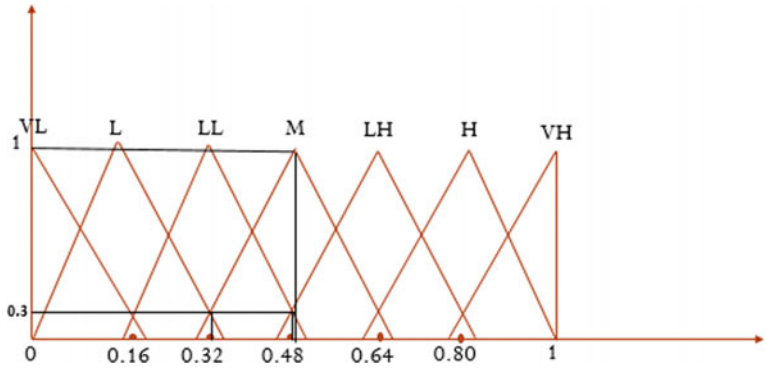


Fig. 7 Value of truth for route 3

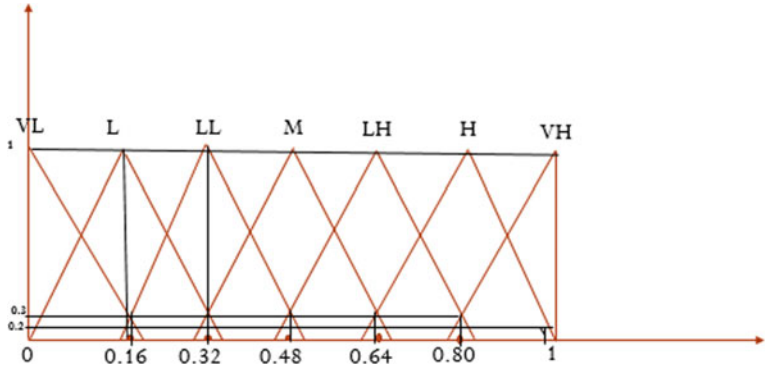


Fig. 8 Value of truth for route 4

As we may see, the given truth value for different routes indicates that route lifetime increases in different conditions according to Eq. 1.

One can see according to Table 7 the lifetime for route 2 has the greatest value among all possible alternative routes hence one can say that the next possible route is R4 and R3, R1.

Table 7 Route lifetime when linguistic variables are different

Routes	Function	Route lifetime (RL)
R1	(0.3,0.3,1,0) 0.4, (0.3,0.3,1,0) 0.4, (0.3,0.3,1,0) 0.4,(0.3,0.3,1,0) 0.4	0.400
R2	(1,1,1,0.3) 0.825, (0.3,0.3,1,1) 0.65, (0.3.0.3,1,1) 0.65, (1,1,1,0.3) 0.825	0.7375
R3	(0.3,0.3,1,0) 0.4, (0.3,0.3,1,0.2) 0.45, (0.3,0.3,1,0.2) 0.45,(0.3,0.3,1,0.2)0.45	0.4375
R4	(0.3,0.3,1,1) 0.65, (0.3,1,0.3,0.3) 0.475, (0.3,0.3,1,0.2) 0.45, (0,0.3,1,0.3) 0.4	0.49375

5 Conclusion

In the above work, it has been shown that network sustain life could be extended and this leads to enhanced throughput performance of the network. The simulation result shows significant increment of battery backup, which can be utilized to sustain long period of network lifetime. When the network nodes are mobile, this decision may more beneficial to overcome from the difficulty of route selection. However, the overall power of nodes depends on physical environment around the node and their network topology as well. With increase in power, the level of interference also increases as well. Hence by limiting internode interference, our proposed scheme can give better solution to the existing power consumption problems.

References

1. V. Kawadia, P.R. Kumar, Principles and protocols for power control in wireless Ad- hoc networks. IEEE, Sel. Areas Commun. **23**, 76–88 (2005)
2. C.E. Perkins, Ad-hoc Networking. Addison-Wesley (2002)
3. C. Long, B. Zang, H. Yang, Non cooperative power control for wireless Ad-hoc networks with repeated games. Sel. Areas Commun. **25**, 1101–1112 (2007)
4. V.W. Zheng, X. Zhang, D. Liu, D.K. Sung, A joint power control link scheduling and rate control algorithm for wireless Ad-Hoc network. in *Proceedings IEEE Conference on Wireless Communication and Networking* (2007), pp. 3636–3640
5. S. Guo, O. Yang, Minimum energy multicast in wireless Ad-hoc networks with adaptive antenna: MILP formulations and heuristics algorithms. IEEE Trans. Mobile Comput. **5**, 333–346 (2006)
6. Y. Chen, G.Y. Peiliang, Q.Z. Zhang, Power aware cooperative relay selection strategies in wireless Ad-hoc networks, in *Proceedings IEEE International Symposium on Personal, Indoor and Mobile Radio Communications* (2006), pp. 1–5
7. V.P Singh, K. Kumar, Literature survey on power control algorithms for mobile Ad-hoc network, Proc. ACM Digit. Libr. Int. J. Wirel. Pers. **60**(4), 679–685 (2011)
8. H. Pimentel, J. Martins, Ad-hoc network performance dynamics-A study, in *Proceedings of International Symposium on Wireless Pervasive Computing* (2007), pp. 46–58

9. J.W. Lee, R. Mazumdar, N. Shroff, Joint opportunistic power scheduling and end-to-end rate control for wireless Ad-hoc networks. IEEE Trans. Veh. Technol. **56**, 801–809 (2007)
10. T. Acharya, S. Chattopadhyay, R. Roy, Multiple disjoint power aware minimum connected dominating sets for efficient routing in wireless Ad-hoc network, in *Proceedings of IEEE Conference on Information and Communication Technology* (2007), pp. 336–340
11. Meenu, V. Jaglan, Optimal route selection by predicting nodes lifetime in a homogeneous wireless Ad-hoc network using genetic algorithm, J. Adv. Res. Dyn. Control Syst. **10**(02-Special Issue), 2216–2225 (2018)
12. G. Gupta,V. Jaglan, A.K. Raghav, Safety based virtual positioning system for Ad-hoc wireless network. Int. J. Eng. Technol. **7**(3.12), 1053–1055 (2018)

Spatial Correlation Based Outlier Detection in Clustered Wireless Sensor Network



Robin Kamboj and Vrinda Gupta

1 Introduction

Wireless sensor networks (WSNs) composed of small, compact, low energy, low cost, multifunctional sensors are placed in the wanted region. Each node can sense the physical condition of the environment surrounding it and also can transmit and receive the data. Currently, WSNs have been broadly used for tracking and surveillance, environmental monitoring, medical care, infrastructure management, and many other areas [1]. WSNs are usually used in military operations [2]. However, sensor nodes are generally spread out in the precipitous and hostile domain where it is tough to reach and power availability is restricted. Therefore, sensor nodes may be faulty and unreliable. Generally, hardware fault and software fault are the two types of fault mainly occurring in WSN. The fault in hardware is usually due to damage of sensor nodes or due to low energy, and then those nodes cannot collect and transmit the data. Software fault can happen just because of internal error in software. Sensor nodes, in this case, can sense, collect, and transmit further the sensed information, but their collected values often deviated from actual values. The observed values which are sufficiently deviated from the real values are called as anomalies or outliers [3, 4]. In the fields of surveillance like military operation, outliers cannot be ignored because in such kind of surveillance, very high precision is required. Therefore, anomalies or outlier detection is an important issue in the field of WSN.

There are generally two approaches used for detecting outliers, viz., centralized and distributed approach. In centralized approach [5], all the processing is done by a central node, for example, bootstrapping technique to detect intrusion [6] wherein

R. Kamboj (✉) · V. Gupta
Department of Electronics and Communication Engineering,
National Institute of Technology Kurukshetra, Kurukshetra, Haryana, India
e-mail: robin_31705123@nitkkr.ac.in; robinkamboj027@gmail.com

V. Gupta
e-mail: vrindag16@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_13

cluster head (CH) performs all the calculation of data after receiving the data from sensors and in distributed detection approach [7], all the processing is performed by the sensor nodes itself, for example, distribution Bayesian approach [8].

In a wireless sensor network, if the sensing range of any two nodes overlaps each other, the nodes are called as having spatial correlation. In this paper, spatial correlation based approach is proposed in which cluster head does the detection. The cluster head of each cluster compares the data of a node to the neighbors of that node. If data received by cluster head from a node is similar to its neighbor node, then cluster head treats it as normal event and forward the sensed data to the sink node, otherwise the sensed data is treated as an outlier.

Further, the manuscript is categorized as follows: The existing efforts in the field of outlier detection techniques are described in Sect. 2. The proposed algorithm and network model are covered in Sect. 3. Simulation results and discussion are presented in Sect. 4, followed by the conclusion in Sect. 5.

2 Related Work

A lot of work has been done in the area of outlier detection. The research works of some of the authors are described in this section. A survey on different outlier detection techniques for WSNs based on techniques such as nearest neighbor statistical artificial intelligence and classification-based approaches exists in literature [3, 4]. A work on centralized approach to detect outliers has been done in [5]. In this algorithm, outlier detection was performed by base station which has large storage and resource available. The base station collects the data from each node of the network and processes it to detect the real event and an outlier. But it has the drawback that large energy is consumed by using this algorithm.

Another work uses a density-based spatial clustering of application with noise (DBSCAN) technique for outlier detection. Another work uses a bootstrapping technique for intrusion detection [6]. In this technique, frequent item set mining was used to distinguish the normal event from outliers.

Based on distributed detection, Tzu-Liang et al. [7] proposed a PMC diagnosis model. In this paper, the outlier is detected through test results appearing in extending star. But this algorithm is applicable where the numbers of outliers are small. In this, it is found that the DA gets lowered when the number of outliers rises.

A distributed Bayesian algorithm (DBA) has been used by the authors in [8]. An assumption is made that prior fault probability of each node is the same and by using comparison between neighbor and node itself, the posterior probability of node fault is calculated. Hence, outliers are easily determined by using this posterior fault probability. But the problem with DBA algorithm is that it can determine only similar kind of outlier only, therefore it is not suitable where different types of outliers are present and also message complexity is very high. A distributed outlier detection based on credibility feedback (DODCF) algorithm has been used by the author in [9] who uses two steps to determine the outliers. First, initial credibility of each sensor node is calculated using distance between nodes and its neighbor. Then

final credibility of sensor node is calculated by using initial credibility and Bayesian theorem. But this algorithm has the application where it is required to improve the detection accuracy outliers. It has the disadvantage that the message complexity of this algorithm is still high. In most of the research works, the performance of outlier detection algorithms is determined with help of DA and FAR.

3 Proposed Work

3.1 Network Model

The network model composed of 400 sensors that are placed uniformly in the field of 100 × 100 units. The field is considered to be divided into four numbers of small clusters. Each cluster contains 50 sensor nodes and one cluster head (CH) as shown in Fig. 2. Each node is assumed to be having sensing range of 5 units and data transmission range is 50 units. Each node is placed in such a way that the sensing range of every node overlaps the sensing area of neighbor nodes as shown in Fig. 1. Here, the circles around the sensor nodes are representing the sensing range of nodes and number inside the sectors of circle represents numbers of nodes covering that particular area.

In a particular cluster, whenever an event is sensed by a node, the sensed data is transmitted to the cluster head. As event can be normal or outlier, therefore cluster head compares this data with all the data received by it at the same time to check whether the received data is a normal event or an outlier. As all the points in the surveillance area are wrapped by more than one node, therefore more than one sensor node measured data would be same in case of a normal event.

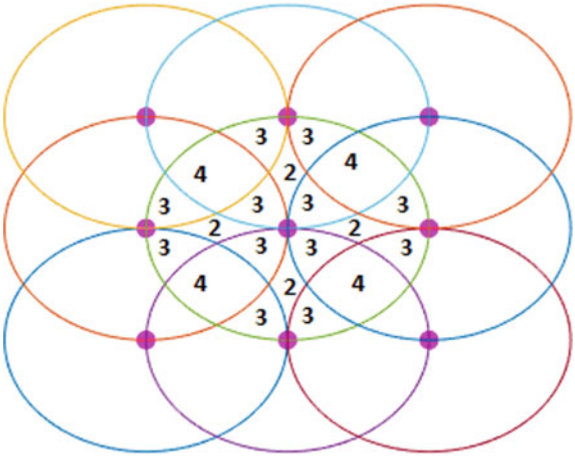


Fig. 1 Overlapping sensing area of sensor nodes

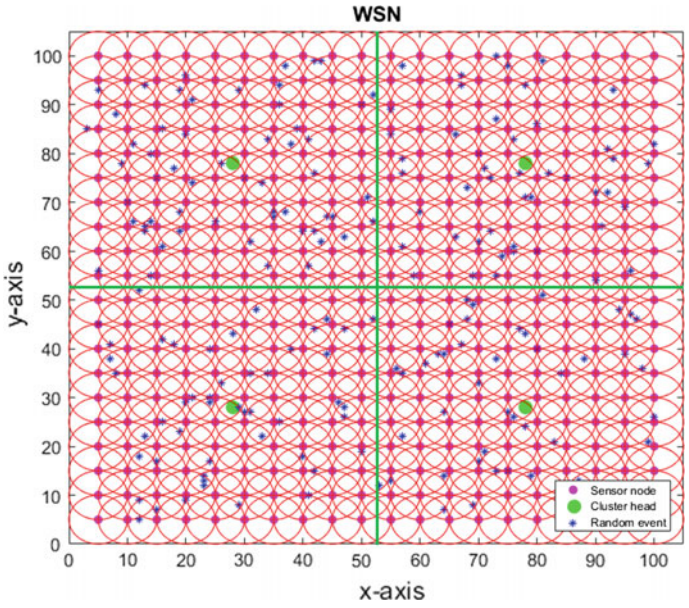


Fig. 2 Uniform distribution of wireless sensor node in four different clusters

Therefore, during comparison by cluster head, if a particular data is received more than one time, cluster head will treat it as a normal event, and if only single copy of data is present, then cluster head will treat it as an outlier and will not send that data to the sink for further purpose.

3.2 Proposed Algorithm

The proposed algorithm is explained in Sect. 3.2.1 with the help of a flow chart. The value observed by sensors in the surveillance field is represented by mathematical equation as given in [11]

$$x_i = A_i + w_i \quad i = 1, 2, \ldots k \tag{1}$$

where x_i represents the observed value of a node n_i, A_i represents the actual value, and w_i represents the additive white Gaussian noise which is having zero mean. After observing this, the node sends this data to CH. The CH compares this data of node n_i to the observed data of the neighbors of n_i, which it receives at the same time. By using the threshold t of the data, the CH determines whether the two compared data are similar or not. If $|x_i - x_j| \leq t$, data will be treated as similar and flag value f_{ij} is set to 1 otherwise to 0.
Therefore

$$f_{ij} = \begin{cases} 1 \text{ if } |xi - xj| \leq t \\ 0 \text{ if } |xi - xj| > t \end{cases} \tag{2}$$

where f_{ij} represents the flag value between node n_i and its neighbor node n_j.

Number of times a particular data is present in the memory of cluster head is represented by "count" which is expressed in Eq. (3).

$$\text{count} = 1 + \sum_{j=1}^{n(n_i)} f_{ij} \tag{3}$$

where $n(n_i)$ represents the number of neighbor nodes of node n_i. If the count ≥ 2, CH sends the data to the sink node by treating it as a normal event. And if count is less than 2, CH will delete the data by treating it as an outlier. In this way, by comparing the nodes data every time with respective neighbors node data, cluster head determines whether the received data is an outlier or a normal event.

3.2.1 Algorithm Flow Chart

See Fig. 3.

4 Results and Discussion

The performance of the proposed algorithm has been evaluated by using MATLAB software, version 14. Table 1 shows the confusion matrix of outlier detection as given in [6]. Terms used in this table are as follows: True positive (TP) means actual outlier is correctly predicted as outlier, false negative (FN) means actual outlier is wrongly predicted as normal, false positive (FP) means actual normal is wrongly predicted as outlier, and true negative (TN) means actual normal is correctly predicted as normal.

The performance metrics used for evaluation of the algorithm are detection accuracy and false alarm rate. Also, the message complexity of the proposed algorithm has been evaluated.

- Detection accuracy (DA): It is defined as the ratio of sum of the number of events which are correctly predicted as outlier and normal to the total number of events in the network.

$$\text{DA} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{FN} + \text{TN}} \times 100 \tag{4}$$

- False alarm rate (FAR): It is defined as the ratio of number of actual normal which are wrongly predicted as normal event to the total number of actual normal in the network.

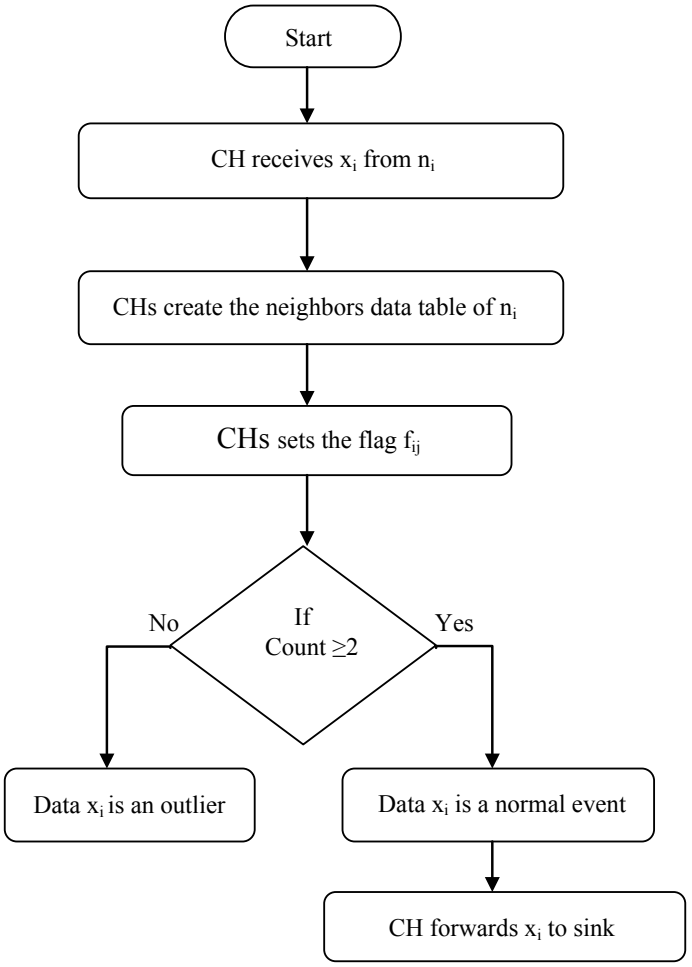


Fig. 3 Flow chart of proposed algorithm

Table 1 Confusion matrix of outlier detection

	Predicted as "outlier"	Predicted as 'normal'
Actual outlier	True positive (TP)	False negative (FN)
Actual normal	False positive (FP)	True negative (TN)

$$\text{FAR} = \frac{\text{FP}}{\text{FP} + \text{TN}} \times 100 \tag{5}$$

- Message complexity: It is defined as the number of times the data is exchanged between sensor nodes in each round of outlier detection. In proposed algorithm, sensor node directly sends the data to cluster head, therefore message complexity is come out to be $k\dot{n}$ where k is number of nodes in the network and $\dot{n}$ gives average number of nodes of each node. The message complexity of DODCF algorithm is $(2k + k')\dot{n}$ [11] where k' is number of nodes which determines as normal event. The message complexity of proposed algorithm is low as compared to DODCF algorithm.

The network model that has been described in Sect. 3.1 is used for simulation. Considering the number of outlier points to be in range 20–200, the threshold value t is fixed to 0.5.

Figure 4 shows that the detection accuracy of proposed algorithm is very high, nearly 100% as compared to DODCF algorithm [11]. It means all the outlier points are correctly determined in the proposed algorithm. As the number of outlier points increases, detection accuracy of proposed algorithm decreases very less as compared to DODCF algorithm [11]. This is because approximately 75% sensing area of a single node is covered by 3–4 nodes, which increases the accuracy of detection as shown in Fig. 1.

False alarm rate (FAR) of proposed algorithm is shown in Fig. 5. It shows that false alarm rate of proposed algorithm is also improved as compared to DODCF algorithm. As the number of outlier point increases above 100, FAR of proposed algorithm is almost same, but of DODCF algorithm, it increases to large extent.

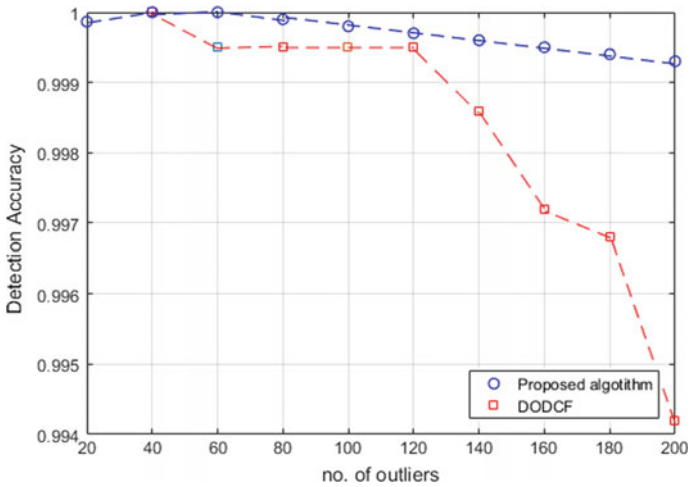


Fig. 4 Detection accuracy w.r.t. number of outliers

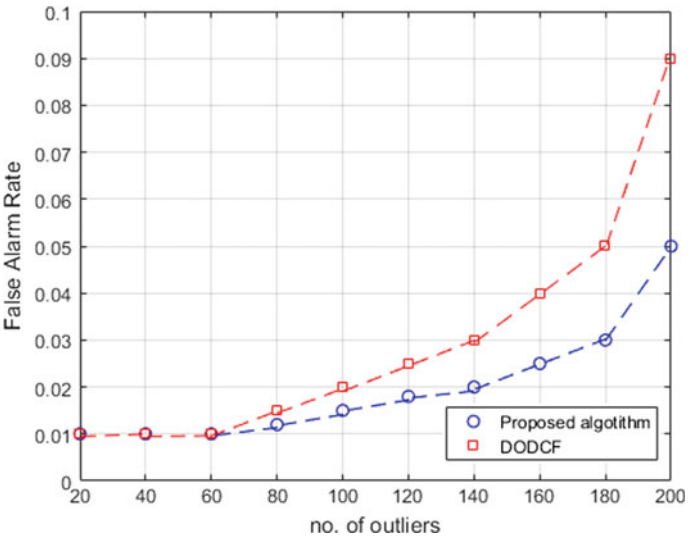


Fig. 5 False alarm rate w.r.t. number of outliers

5 Conclusions

This paper has proposed an outlier detection algorithm exploiting the fact that events measurement in a wireless sensor network is likely to be spatially correlated. The proposed algorithm, spatial correlation based outlier detection by cluster head for wireless sensor network has the advantage that it can detect outlier more accurately and low message complexity. This algorithm can be used to find any kind of outliers such as noise, fault, or error in measurements. In our proposed algorithm, we use only spatial correlation to compare the nodes data at CH, but spatio-temporal correlation can be carried out as future work.

References

1. Y.A. Bangash, Y.E. Al-Salhi, Security issues and challenges in wireless sensor networks: a survey. IAENG Int. J. Comput. Sci. **44**(2) (2017)
2. M.P. Durisic, Z. Tafa, G. Dimic, V. Milutinovic, A survey of military applications of wireless sensor networks, in *2012 Mediterranean Conference on Embedded Computing (MECO)* (2012), pp. 196–199
3. A. Ayadi, O. Ghorbel, A.M. Obeid, M. Abid, Outlier detection approaches for wireless sensor networks: a survey. Int. J. Comput. Sci. **129**, 319–333 (2017)
4. Z. Fei, B. Li, S. Yang, C. Xing, A survey of multi-objective optimization in wireless sensor networks: metrics, algorithms, and open problems. IEEE Commun. Surv. **19**(1), 550–586 (2015)
5. P.R. Chandore, D.P.N. Chatur, Hybrid approach for outlier detection over wireless sensor network real time data. Int. J. Comput. Sci. Appl. **6**(2), 76–81 (2013)

6. D. Barbara, Y. Li, J. Couto, J.-L. Lin, S. Jajodia, Bootstrapping a data mining intrusion detection system, in *Proceedings of the 2003 ACM Symposium on Applied Computing*, ACM Press (2003), pp. 421–425
7. K. Tzu-Liang, C. Hsing-Chung, J.J.M. Tan, On the faulty sensor identification algorithm of wireless sensor networks under the PMC diagnosis model, in *Proceedings of the International Conference on Networked Computing and Advanced Information Management*, Seoul, Korea, vol. 8 (2010), pp. 657–661
8. Y. Hao, Z. Xiaoxia, Y. Liyang, A distributed Bayesian algorithm for data fault detection in wireless sensor networks, in *Proceedings of the International Conference on Information Networking*, Cambodia, vol. 1 (2010), pp. 63–68
9. H. Feng, L. Liang, H. Lei, Distributed outlier detection algorithm based on credibility feedback in wireless sensor networks. IET Commun. **11**(8), 1291–1296 (2017)
10. X. Luo, M. Dong, Y. Huang, On distributed fault-tolerant detection in wireless sensor networks. IEEE Trans. Comput. **55**(1), 58–70 (2016)
11. A. Abid, A. Kachouri, A. Mahfoudhi, Outlier detection for wireless sensor networks using density-based clustering approach. IET Commun. **7**(4), 83–90 (2010)
12. T. Palpanas, D. Papadopoulos, V. Kalogeraki, D. Gunopulos, Distributed deviation detection in sensor networks, in *ACM Special Interest Group on Management of Data* (2003), pp. 77–82
13. Z. Feng, J. Fu, Y. Wang, Weighted distributed fault detection for wireless sensor networks based on the distance. Chin. Control Conf. Nanjing China 7, 322–326 (2014)
14. S.N. Das, S. Misra, Correlation-aware cross-layer design for network management of wireless sensor networks. IET Wirel. Sens. Syst. **5**(6), 263–270 (2015)

Guidelines for an Effective Network Forensic System



Rajni Ranjan Singh and Deepak Singh Tomar

1 Introduction and Background

Network forensics is a scientific process to classify malicious traffic from the incoming/outgoing network traffic log information. The objective is to identify origin of the attack [1].

1.1 Categorization of Network Forensics

Network forensic systems are categorized into two types based on their collection characteristics [2].

"Catch-it-as-you-can" in this system, network traffic is being captured and then analyzed by the specialized device installed at specific position in the network. This required enormous amount of storage.
"Stop-look-and listen" in this system, network streams are analyzed in memory and some relevant information is stored for future. This required faster processing.

R. R. Singh (✉) · D. S. Tomar
Maulana Azad National Institute of Technology, Bhopal, India
e-mail: ranjansingh06@gmail.com

D. S. Tomar
e-mail: deepaktomar@manit.ac.in

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_14

2 A Generalized Network Forensic System Architecture

Generally, a network forensic investigation utilizes following steps while investigating a cyber attack, Fig. 1 shows the generalized network forensic system framework.

Major phases/components of generalized forensic system framework have been discussed in the following section.

2.1 Capturing

Capturing is the most important phase of network forensics because there is only one chance to capture the evidence. It is needed to deploy an acquisition system at the most prominent position in the network in order to capture all the traffic data.

2.1.1 Acquisition System Deployment

Forensic preparation is needed to establish logging infrastructure to facilitate network investigations, therefore, firewall, intrusion detection system, or/and packet analyzer



Fig. 1 Generalized network forensic system framework

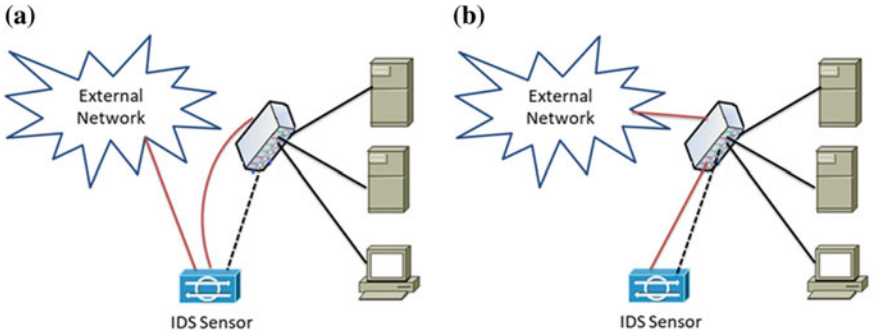


Fig. 2 IDS deployment positions, **a** IDS connected inline, **b** IDS connected to switch spanning port

can be deployed at different strategic positions in the network and configure them. Because we have only one chance to capture network traffic. It feasible to deploy acquisition system at network vantage point for watching all network traffic, there are two approaches for deploying acquisition system, one is to deploy "inline to the router" and another is to use SPAN (Switched port analyzer) capability of switch device. This capability allows capture/monitor traffic from one port on the switch to another (Fig. 2).

The following point should be kept in the mind while deploying acquisition system. Sensor deployment is discussed in the literature [3–7].

(1) It is feasible to deploy acquisition system at the network vantage point despite deploying inside subnet or after switches. The deployment at vantage point assures complete data collection because if we increase the distance from source, it affects signal strength, latency, causing timeouts, increases error rate, therefore, it is feasible to deploy capturing system at the minimum distance position between source and destination.
(2) To prevent a single point of failure and to assure complete data collection, despite deploying single acquisition system, it is recommended to deploy more than one instance of acquisition system. All the deployed instances work simultaneously in the same network at the same time. After the data collection, all the instances can compare the collected data, if they are identical, it means collection is accurate and verified.
(3) Management approval and authorization are required to monitor network traffic so that organizational privacy is not violated.

2.1.2 Packet Capturing System

The acquisition system should be able to collect an exact copy of the network packets in suitable format which can be supported by analysis system, like de facto standard

store the data in pcap or winpcap format. Many traffic analysis applications are capable to import tcpdump files (Fig. 3).

It is beneficial to use highly recommended applications and hardware's for the network traffic collection and these applications should work in read-only mode so that they cannot modify evidence. These applications and devices should not generate any kind of data in the network. If network interface card (NIC) of acquisition system is working in a read-only mode (NIC card without transfer capabilities), it prevents generating any kind of network traffic. Many devices like router modify the incoming/outgoing packet header fields such as hop count value or TTL and it may also fragment the packets based on MTU (maximum transmission unit), therefore, router may not be used as acquisition system.

The amount of data captured will be huge requiring enormous storage space, therefore, acquisition system should be capable to capture only relevant data by selective packet capturing, in order to reduce log size, for example, capturing only packet header and avoid capturing payload data is a simple method.

An ideal collation should follow the following characteristics. Evidence collection is discussed in the literature [1, 3, 5].

(1) Acquisition system should perform transparent, invisible/stealthy inline operation so that attacker doesn't know about the acquisition system. The acquisition system should passively capture all attackers' activities.
(2) The acquisition system must be secure, fault tolerance having limited access, and should allow legitimate encrypted remote connections.
(3) For maintaining the integrity of collecting data, it is required that acquisition system should have read-only network access, therefore, the collected evidence cannot be changed/modified during the collection/capturing process.

2.2 Preservation and Storage

The information collected in the form of network traffic logs and traces are stored on the read-only backup storage device and a secondary copy will be used for analysis purpose. A hash value of all collated data is calculated and preserved on read-only storage media. The most common method for hash calculation is MD5 and SHA1. The hash value can be calculated during capturing or after the completion of capturing (when the captured data have been saved on the disk). It guarantees that digital evidence has not been changed while it was acquired and investigator will be capable to prove same when the procedure is repeated on the original acquired data.

Following necessary information should be added while preserving evidence which is discussed in the literature [1, 7, 8].

(1) There should be a provision for signing and/or hashing of collected data (tools such as OpenSSL or PGP can be utilized for evidence signing).
(2) It is recommended to store backup on the write-once media.

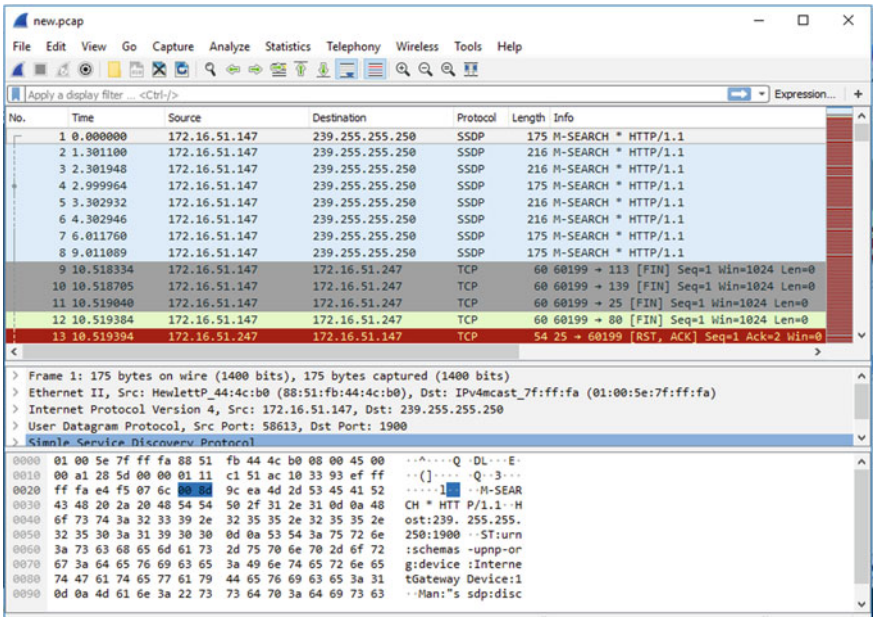


Fig. 3 A snapshot of pcap file, captured by open-source packet capturing tool Wireshark

(3) The operating system, additional application, configuration files, and forensic examiner activity logs are stored on a separate disk.
(4) For location identification, it is required that network location stamping should be added it may include geographic address where capturing is taking place (country, city, address, etc.), location of collector (floor, office, server room/data center, etc.), and point of attachment details (patch panel, hub/switch no., rack no, etc.).
(5) It is important to encrypt the data and deciphering key given to the analysis team so that only the analysis team can be able to decrypt the data.

2.3 Analysis

Data acquired from different sensor devices are integrated, classified, and clustered in order to select useful patterns and then collated data are searched methodologically for known attack patterns and behaviors using statistical, soft computing, and data mining approaches. In most cases, collected data is searched to find out any significant deviation from normal network behavior. Identified patterns are correlated with associated similar patterns and then to construct a timeline of various events to find out an outline of what happened and to recognize the objective and method of the attacker.

A generalized analysis system performs the following task. Traffic analysis is discussed in the literature [1, 4, 5, 8, 9].

2.3.1 Preprocessing

It is a preprocessing phase that prepares captured data for analysis. In order to improve the forensic analysis capability, it is required to transform the captured data into a standard format.

Following operating can be performed according to their applicability.

(1) Data may be acquired from multiple sensor devices which may be configured with different software platforms/operating systems and operating in diverse time zones, therefore, conversion from local time zone to standard global time zone is necessary example UTC (Coordinated Universal Time). In addition, it is required to normalize data, for example, data either in hostnames and/or in IP addresses.
(2) Data can be analyzed either by manually or by using automated software like Ethereal [10], Tcpflow [11], Net Intercept [12], etc., they are capable to import tcpdump files directly however most of the analysis software like soft computing software's requires data file in suitable format like CSV, XML, ARFF, etc. In most cases, where captured data is enormous, then it needed to store it into database. Therefore, it is required to prepare captured data according to the target analysis software and/or databases (Fig. 4).
(3) Most of the cyber attack signatures are splits between packets, therefore, it is required to sessionize the captured data. Once reassembly is complete, actual analyses take place.

2.3.2 Data Reduction

Data reduction is required to find out an exact item in the network traffic (online/offline). For extracting only web traffic, it is required to capture only packets targeted to port 80 and/or 8080. Software like Wireshark [10] provides various filters and GUI to help investigator to fetch items of interest (Fig. 5).

2.3.3 Protocol Parsing and Analysis

Captured packets are then dismembering at the various physical/DLL/Network/Transport level protocols, such as UDP or TCP, IP, and Ethernet. The resulting streams are then again dismembering with application-level protocol such as MNS chat, IRC, and HTTP.

Investigators commonly perform protocol analysis run strings to pull text from traffic stream and grep to find specific words or phrases in the recovered strings.

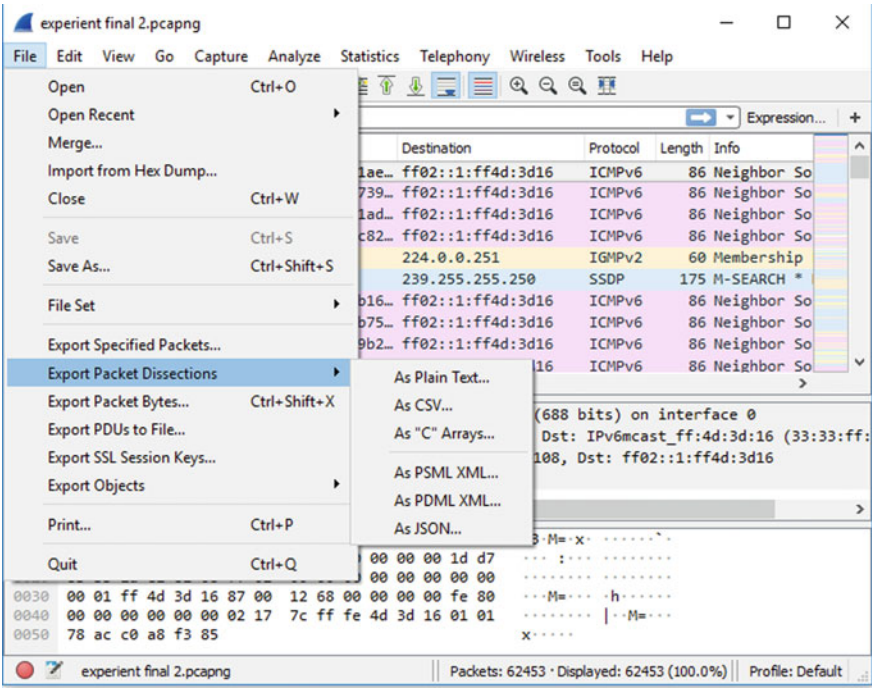


Fig. 4 Network traffic CSV file after conversion by Wireshark

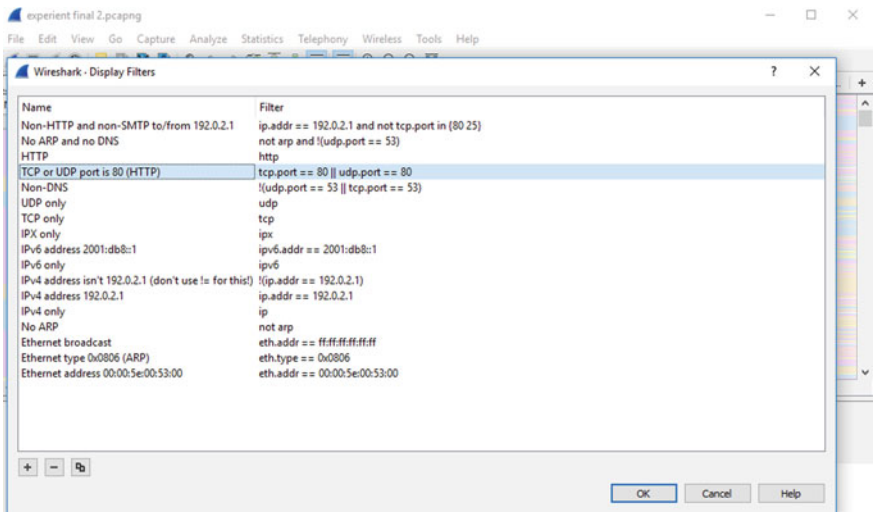


Fig. 5 Data reduction filters provided by Wireshark utility

String search can be performed at the logical or physical level, for example, finding a particular word "black mail" where black exists in one packet and mail contains by another packet the logical search of network traffic can be performed by reconstructing flows and then searching keyword that is split between packets. Techniques like machine learning, mining, clustering, hypothesis testing, statistical approach, anomaly detection, and histogram are widely used for traffic analysis purpose.

2.4 Investigation and Attribution

The information gathered from the identified attack patterns are used to discover attack source, target systems, time, and methods. Here, forensic examiner arranges the evidence in order to establish the links between intruder and the compromised systems. The network traffic captures, logs, and information gathered are utilized for the attribution of the cyber attack. The network log analysis should guide investigator to the origin of the attacks. Once the source is identified, following investigations are suggested which is discussed in the literature [1, 8, 13–16].

(1) Source IP address should be investigated for IP spoofing so that this should clear that the identified source is genuine or not. IP trackback techniques are best suited for this task.
(2) Investigate source IP address position mapping to find an intruder physically.

2.5 Forensic Reporting

The analysis report should be presented in a comprehensible language to the company management and legal officials. This includes various standard measures/guidelines utilize to arrive at conclusion. The methodical documentation should be included to fulfill the lawful requirements. The observations should be well written in order to utilize in future investigations and enhancement of product security. Discussed in the literature [1, 8, 13–15].

(1) Evidence should be relevant, authentic, and reliable.
(2) Documentation includes digital evidence, audit trails of digital evidence, associated digital evidence, examiner actions, packet losses, acquisition system performance history, etc.
(3) Include log evidence that shows the correct operation of the acquisition system.
(4) Acquisition system file access log may be useful in proving whether or not the captured data file was alerted.
(5) The message digest should be stored with other investigation activities on the read-only write-once memory. This information will be utilized in future for verification.

(6) Errors and lost/corrupted data and events during acquisition should be included while presenting evidence.

3 Conclusion

This paper offers a primer on the investigation of network attacks in the forensic arena. There are many conclusions drawn from this work.

- Recording everything perhaps violating organizational policies so it is not realistic.
- Digital evidence are fragile require proper handling.
- Acquisition requires enormous amount of processing and storage, therefore, selective packet capturing is a good alternative.
- In order to prevent a single point of failure and for validating log evidence, it is required to deploy more than one instance of acquisition system.
- Analysis of encrypted traffic is tedious and require further research.
- Selection of analysis methods, correlating events, and establishing links between source and destination require both technical and legal expertise and experience.

References

1. E.S. Pilli, R.C. Joshi, R. Niyogi, Network forensic frameworks: survey and research challenges. Digit. Investig. **7**(1–2), 14–27 (2010) (Elsevier)
2. S. Garfinkel, Network forensic tapping the internet. http://www.oreillynet.com/lpt/a/1733
3. B.J. Nikkel, A portable network forensic evidence collector. Digit. Investig. **3**(3), 127–135 (2006) (Elsevier)
4. M.I. Cohen, PyFlag–an advanced network forensic framework. Digit. Investig. **5**, S112–S120 (2008) (Elsevier)
5. E. Casey, Network traffic as a source of evidence: tool strengths, weaknesses, and future needs. Digit. Investig. **1**(1), 28–43 (2004) (Elsevier)
6. B.J. Nikkel, Generalizing sources of live network evidence. Digit. Investig. **2**(3), 193–200 (2005) (Elsevier)
7. B.J. Nikkel, Improving evidence acquisition from live network sources. Digit. Investig. **3**(2), 89–96 (2006) (Elsevier)
8. V. Corey, C. Peterman, S. Shearin, M.S. Greenberg, J. Van Bokkelen, Network forensics analysis. IEEE Internet Comput. **6**(6), 60–66 (2002)
9. C. Boyd, P. Forster, Time and date issues in forensic computing—a case study. Digit. Investig. **1**(1), 18–23 (2004) (Elsevier)
10. Ethereal/wireshark.https://www.wireshark.org/
11. Tcpflow. www.circlemud.org/jelson/software/tcpflow/
12. NetIntercept. www.securitywizardry.com/index.php/…/niksun-netintercept.html
13. M. Solon, P. Harper, Preparing evidence for court. Digit. Investig. **1**(4), 279–283 (2004) (Elsevier)
14. F. Buchholz, E. Spafford, On the role of file system metadata in digital forensics. Digit. Investig. **1**, 298–309, (2004) (Elsevier Ltd.)

15. F. Buchholz, E. Spafford, On the role of file system metadata in digital forensics. Digit. Investig. **1**(4), 298–309 (2004) (Elsevier)
16. M. Reith, C. Carr, G. Gunsch, An examination of digital forensic models. Int. J. Digit. Evid. **1**(3), 1–12 (2002)

Handling Incomplete and Delayed Information Using Optimal Scheduling of Big Data Stream



Ravi Kishan Surapaneni, Sailaja Nimmagadda and Roja Rani Govada

1 Introduction

The amount of data being produced and consumed has increased exponentially in this decade alone. With the penetration of Internet services to remote areas, many new users are able to access the Internet and consume data. This has led to vast amounts of data generation. Each second, millions of user requests have to be processed by the big data technologies without any inconvenience to the user. Smart devices and IoT (Internet of things) sensors also become the catalyst for data generation [1, 2]. Millions of devices simultaneously send data that they have captured to the servers where the data has to be processed in parallel for further use. Big data technologies have paved a way for handling these huge amounts of data. Whether the data may be structured or unstructured, it has to be processed within limited time [3]. Failing to do so can lead to congestion of data at the servers. The computational cost for processing data such as collected from IoT devices is very high [4]. This type of data is generally termed as stream data where the data comes as a continuous stream. Data stream mining incorporated with machine learning is used to process these kinds of data.

R. K. Surapaneni (✉) · S. Nimmagadda · R. R. Govada
VR Siddhartha Engineering College, Kanuru, Vijayawada, India
e-mail: suraki@vrsiddhartha.ac.in

S. Nimmagadda
e-mail: sailunims@gmail.com

R. R. Govada
e-mail: rojarani@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_15

Stream data can include data such as instant messaging, social network posts, emails, digital audio, or video. The general technique for stream data mining is to use a pipeline solution in which features of interest are extracted from the raw data source and low complexity classification algorithms are performed to generate applications [5]. Other techniques for stream data mining are by using online data mining method and incremental method. In both of those methods, the model is reconstructed and refined with the arrival of new data. Transient patterns are captured from the data for time-critical predictions [6]. The main disadvantage is that the assumptions made using these techniques are not always correct. It is due to the fact that in real-world scenarios, data is not always complete, it is often delayed and incomplete. But the algorithms assume that the data is complete which will cause the lack of value [7]. Incompleteness in the data may be due to storage issues and connectivity issues. Delayed information can be attributed as latency where higher latency will cause higher delay.

Scheduling is the best solution to overcome the problems arisen from incomplete and delayed information [8]. Scheduling applications on various data centers is an important task. Various applications require different types of scheduling algorithms. Depending upon the variation in the parameter, the scheduling algorithms are used. The resources are allocated to the various data centers so that some critical issues are rectified [9]. In this paper, we are proposing a novel optimal scheduling algorithm that can handle incomplete and delayed information. The proposed algorithm is used to effectively schedule the tasks in the data streams for data centers. We are using several measures such as volatility, Hurst exponent, and distance to select the task from the data stream. Enthalpy value is then computed based on the extracted feature for each data streams and the computed enthalpy value is taken as a feedback ID. Finally, we are using krill herd optimization algorithm for the optimal scheduling of tasks based on the generated feedback ID.

This paper is organized as follows. In Sect. 2, we are reviewing the related works in scheduling data streams. In Sect. 3, we are proposing an optimal scheduling algorithm and the proposed methodology. In Sect. 4, the results are presented and conclusions are provided in Sect. 5.

2 Related Work

Fong et al. [10] proposed a versatile lightweight feature assortment to mine streaming information on the fly by using accelerated particle swarm optimization (APSO) algorithm which was composed mainly the feature of assortment. Poli et al. [11] used fuzzy expert systems for extracting knowledge to make decisions in event streams. The design mainly depends on the collaboration of four adjustable models and a

chat-based portrayal of the administer. The outcomes are estimated based on the precision and execution against the gauge kinematic conditions. Valsamis et al. [12] proposed a settled machine learning calculations utilizing static datasets and multi-scan train approaches and distinguishes the best possibility to be utilized as a part of actualizing a solitary pass prescient approach, under ongoing requirements. A semantic-based approach for fragmenting sensor information management is proposed by Triboan et al. [13] by utilizing ontologies to perform phrasing box and affirmation box thinking, alongside logical tenets to derive whether the approaching sensor event is identified with a given succession of the movement. Liu et al. [14] also proposed an algorithm for dynamic assignment specifically for mobile Internet services in which every weight for stream query graph is calculated. Zhang et al. [15] proposed a MapReduce-based framework for stream data processing where the Map and Reduce daemons are running in loops. A dynamic data stream clustering algorithm (DDS) [16] is proposed by Tas which facilitates the density change with time by using a dynamic threshold.

3 Methodology

The input big data stream consists of a number of data streams and each data stream consists of a number of tasks. Initially, the input big data stream is analyzed and the task is selected by calculating the features such as volatility, Hurst exponent, and distance Fig. 1.

3.1 Features Based Task Analysis and Selection

The data in big streams and the number of tasks in data streams are analyzed and the tasks are selected for the scheduling by extracting the effective features such as volatility, Hurst exponent, and distance.

3.1.1 Volatility Measure

In the statistical investigation of the big data streams, volatility measures the time differences of data streams. Little time contrasts indicate high unpredictability while extensive contrasts mean low. Standard deviation is a statistical term that measures the amount of fluctuation around an average.

$$V_y = \sigma_m^2 = \frac{1}{m}\left[\Sigma t^2 - \frac{(\Sigma t)^2}{m}\right] \tag{1}$$

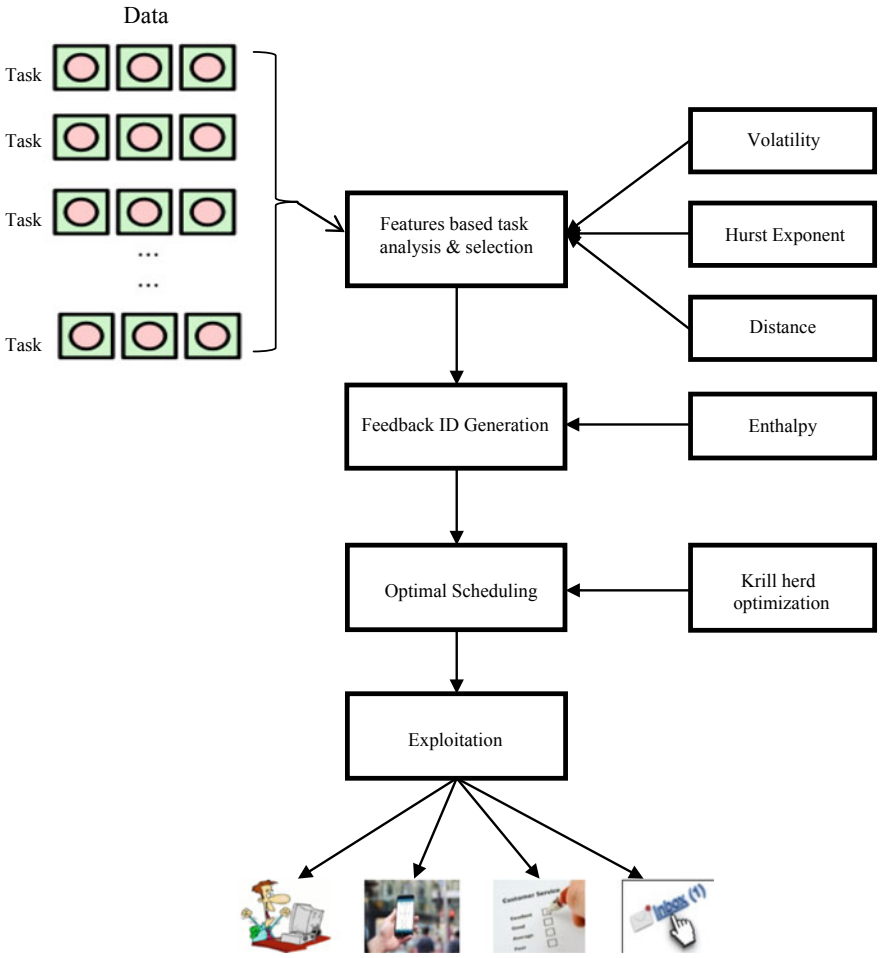


Fig. 1 Block diagram of proposed optimal scheduling in big data streams

where t is the tasks in the data streams, m is the total number of tasks, and V_y is the volatility measure.

3.1.2 Hurst Exponent

The long haul memory nature of time arrangement is evaluated in the Hurst exponential. The relative propensity of a stream is evaluated, revealing the implying that either a strong regression exists as shown by the last record toward a direction. For every data stream, $H = 0.5$ implies the time arrangement is uncorrelated, while $H > 0.5$ demonstrates the information with long-extend correlations, and $H < 0.5$ demonstrates the existence of long-range anticorrelation in data. The estimation of Hurst exponent can be computed by using Eq. (2),

$$E_d[D(n)/S_d(n)] = Cn^H \tag{2}$$

where $D(n)$ is the difference value, $S_d(n)$ is their standard deviation of tasks in data streams, $E_d[x]$ is the expected value, C is a constant, and H is called the Hurst exponent.

3.1.3 Distance

The previous descriptive measurements it might give us a general outline and portray a general state of the time arrangement information. Specifically, we are recognizing the time arrangement of huge information streams utilizing distance measure. The separation between the data streams (d_n) in the set of big data stream (B_d) is computed by the accompanying equation as

$$\widehat{D} = \sqrt{\sum_{k=1}^{n} (d(k) - d(k-1))^2} \tag{3}$$

where n is the number of data streams, $(d(k) - d(k-1))$ is the distance between two data streams, and D is the distance between data streams.

3.2 Feedback ID Generation

In this section, feedback ID is generated based on the enthalpy measure of extracted features such as volatility, Hurst exponent, and distance. This feedback ID generation is used to improve the optimization of data scheduling process.

3.3 Enthalpy Measure

Enthalpy is implied as a condition of data which depends just upon the current equilibrium state apparent through the inward energy of the big data stream (B_d). Here, the enthalpy is identified utilizing the computed measures for task analysis and selection. The enthalpy measure for the feedback ID generation is signified as

$$E_y = V_y + Z \tag{4}$$

where,

$$Z = (\widehat{D} * H) \tag{5}$$

The generated feedback ID is given as the input to the krill herd optimization algorithm. Using this feedback, ID krill herd optimization algorithm efficiently schedules the big data streams.

3.4 Optimal Scheduling of Big Data Streams

The optimal scheduling of big data streams in the proposed work is performed by utilizing the krill herd optimization algorithm. Here, the generated feedback ID is given as the optimizing parameter to the krill herd optimization algorithm. This feedback ID enhances the optimization of big data streams scheduling.

3.5 Krill Herd Optimization Algorithm

The Krill herd algorithm is an iterative heuristic strategy empowered by the inalienable phenomenon of the activities of krill herd. This is primarily utilized for settling the optimization tribulations. The Pseudocode of krill herd optimization algorithm is represented below.

Begin

(i) size of population (X) and iteration ($\widehat{I}_{\max}$) are defined.

(ii) Initialization

Set iteration counter $I = 1$;

Introduce the feedback ID (E_y), population data stream (B_d);

$\widehat{X} = 1,2,3,.....X$ Krill individuals haphazardly and every krill compares to a potential solution to the give issue;

Set the foraging speed F_s, the most extreme diffusion speed $Ds_{\max}$, and the most extreme induced speed $Is_{\max}$.

(iii) Fitness assessment

Assess every krill individual as indicated by the krill position

(iv)While $I < \widehat{I}_{\max}$ **do**

Sort the population/krill from best to most exceedingly awful.

For $i = 1 : X1$ do

Perform the accompanying motion calculations,

1) *Movement induced by other krill individuals*

2) *Foraging activity*

3) *Physical diffusion*

Refresh the krill singular position in the inquiry space.

Assess every krill individual as per its position.

End for i

Sort the populace/krill from best to most noticeably bad and locate the present best.

$\widehat{I}_{\max} = \widehat{I} + 1$.

End while

(v): Estimate the krill best result.

End

Algorithm description

Step 1:
In the beginning, initialization of feedback ID calculated by the equation and the big data streams $B_d = \{d_1, d_2, d_3, \ldots d_n\}$ randomly. Here, every data stream comprises of number of tasks $d = \{t_1, t_2, t_3, \ldots t_n\}$ randomly.

Step 2:
Fitness value is calculated from each krill individual based on the krill individual positions.

Step 3:
Next, the algorithm loop starts by arranging all the krill from best to worst individual.

Step 4:
After that, induced movement, foraging, and random diffusion are computed for all krill utilizing the accompanying equations.

(a) **Foraging motion update**

The foraging update is done by

$$F_y(t+1) = S_f \beta_x + \omega_i F_y(t) \tag{6}$$

$$\beta_y = \beta_y^{\text{food}} + \beta_y^{\text{best}} \tag{7}$$

where S_f is the foraging speed, ω_i is the inertia weight, and β_y^{best} is the finest solution of the yth krill individual.

(b) **Induced movement update**

It is the density maintenance of krill herd given every individual krill besides is given by

$$M_y(t+1) = M_{\max} \alpha_y + \omega_i + M_y(t) \tag{8}$$

$$\alpha_y = \alpha_y^{\text{total}} + \alpha_y^{\text{target}} \tag{9}$$

where $M_{\max}$ is the most extreme instigated speed, ω_i is the inertia weight, α_y^{total} is the local impact the yth krill individual has on its neighbors, and α_y^{target} is the best arrangement of the yth krill individual.

(c) **Physical diffusion update**

The third motion update is emulating the physical dissemination by random activity and is given as

$$D_y(t+1) = D_{\max}\left(\frac{1-i}{i_{\max}}\right)\delta \tag{10}$$

where $D_{\max}$ is the most extreme diffusion speed, is the random directional vector in $[-1, 1]$.

Step 5:
By utilizing distinctive parameters movements in between the time yth krill location is calculated in between the time to $t' + \Delta t'$ by the equation and used to compute every individual krill position.

$$K_y(t' + \Delta t') = K_x(t') + \Delta t' \frac{\mathrm{d}K_y}{\mathrm{d}t} \tag{11}$$

where $\Delta t'$ is a standout among the most extreme vital constants and would do well to be fine-tuned similar to the given real-world optimization. The reference equation

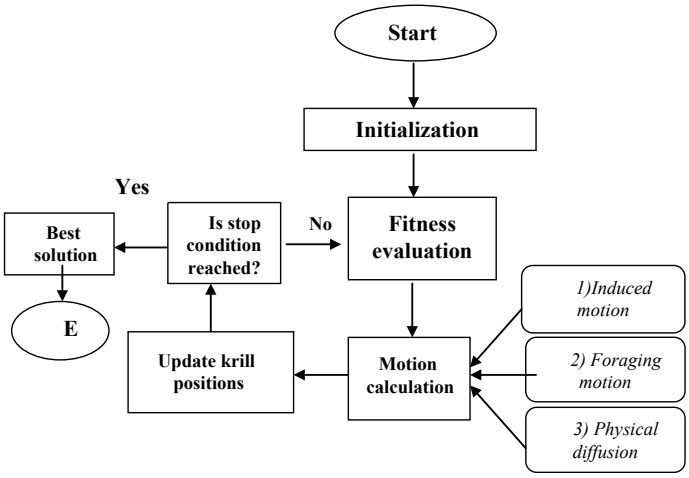


Fig. 2 Krill herd flow diagram

is used to update the krill individual's location to measure the objective function of krill individual toward the complete of the algorithm, the finest krill (solution) is occurred.

Step 6:
The halting criterion specifies whether the function evaluations are satisfied or not. While it is not met, the krill population has to be sorted from best to most unpleasant. After sorting, just compute the movement updates. The entire flow diagram is given in Fig. 2.

The proposed model effectively schedules the data streams and by that which effectively handles the incomplete information and delay. Finally, effectively scheduled data streams are exploited by big data stream applications.

4 Results

Our proposed algorithm was implemented on a system with 32 gb ram, 1 tb hard disk, and octa-core processor. All the codes were implemented in MATLAB, version type r2017a.

The results have shown that our proposed algorithm was able to perform better than krill herd and ABC algorithm in terms of computation time, schedule time, and throughput (Tables 1, 2 and 3).

Table 1 Computational time of our proposed system along with ABC algorithm and krill herd algorithm

Computational time	Proposed	ABC	KH
10,000	27.0622	27.9964	27.8362
20,000	62.5186	67.9983	63.2977
30,000	115.2352	121.4853	114.7863
40,000	179.4086	191.7466	188.4253
50,000	262.719	282.7938	266.3896

Table 2 Schedule time of our proposed system along with ABC algorithm and krill herd algorithm

Schedule time	Proposed	ABC	KH
10,000	2.4541	3.852	3.3521
20,000	4.392	10.2653	5.6012
30,000	8.5735	15.7192	9.3048
40,000	11.9376	21.6368	12.8389
50,000	15.1971	23.9512	15.2831

Table 3 Throughput of our proposed system along with ABC algorithm and krill herd algorithm

Throughput	Proposed	ABC	KH
10,000	4.07E+03	2.60E+03	2.98E+03
20,000	4.55E+03	1.95E+03	3.57E+03
30,000	3.50E+03	1.91E+03	3.22E+03
40,000	3.35E+03	1.85E+03	3.12E+03
50,000	3.29E+03	2.09E+03	3.27E+03

5 Conclusion

Optimal scheduling of tasks in data stream mining is one of the hot topics of research. We have proposed a novel optimal scheduling algorithm that can handle delayed and incomplete information. By incorporating several measures for feedback, ID generation can facilitate in finding the best task. The results have shown that our proposed model outperformed popular scheduling algorithms in terms of computational time, schedule time, and throughput. Our proposed model can effectively schedule tasks in data centers for processing.

References

1. M. Grzenda, K. Kwasiborska, T. Zaremba, Combining stream mining and neural networks for short term delay prediction, in *International Joint Conference SOCO'17-CISIS'17-ICEUTE'17*, León, Spain, 6–8 Sept 2017, *Proceeding* (Springer, Cham, 2017), pp. 188–197

2. Y. Qin, Q.Z. Sheng, N.J.G. Falkner, S. Dustdar, H. Wang, A.V. Vasilakos, When things matter: a survey on data-centric internet of things. J. Netw. Comput. Appl. **64**, 137–153 (2016)
3. D. Sun, G. Zhang, C. Wu, K. Li, W. Zheng, Building a fault tolerant framework with deadline guarantee in big data stream computing environments. J. Comput. Syst. Sci. **89**, 4–23 (2017)
4. S.K. Sharma, X. Wang, Live data analytics with collaborative edge and cloud processing in wireless IoT networks. IEEE Access **5**, 4621–4635 (2017)
5. F. Fu, D.S. Turaga, O. Verscheure, M. van der Schaar, L. Amini, Configuring competing classifier chains in distributed stream mining systems. IEEE J. Sel. Top. Signal Process. **1**(4), 548–563 (2007)
6. H. Wang, W. Fan, P.S. Yu, J. Han, Mining concept-drifting data streams using ensemble classifiers, in *Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ACM (2003), pp. 226–235
7. S. Parsons, Current approaches to handling imperfect information in data and knowledge bases. IEEE Trans. Knowl. Data Eng. **8**(3), 353–372 (1996)
8. Z. Tang, L. Jiang, J. Zhou, K. Li, K. Li, A self-adaptive scheduling algorithm for reduce start time. Future Gener. Comput. Syst. **43**, 51–60 (2015)
9. R. Sandhu, S.K. Sood, Scheduling of big data applications on distributed cloud based on QoS parameters. Cluster Comput. **18**(2), 817–828 (2015)
10. S. Fong, R. Wong, A.V. Vasilakos, Accelerated PSO swarm search feature selection for data stream mining big data. IEEE Trans. Serv. Comput. **9**(1), 33–45 (2016)
11. J.-P. Poli, L. Boudet, A fuzzy expert system architecture for data and event stream processing. Fuzzy Sets Syst. **343**, 20–28 (2017)
12. A. Valsamis, K. Tserpes, D. Zissis, D. Anagnostopoulos, T. Varvarigou, Employing traditional machine learning algorithms for big data streams analysis: the case of object trajectory prediction. J. Syst. Softw. **127**, 249–257 (2017)
13. D. Triboan, L. Chen, F. Chen, Z. Wang, Semantic segmentation of real-time sensor data stream for complex activity recognition. Pers. Ubiquitous Comput. **21**(3), 411–425 (2017)
14. W. Kun, Y. Yue, B. Liu, DAS: a dynamic assignment scheduling algorithm for stream computing in distributed applications, in *2016 IEEE Global Communications Conference (GLOBECOM)*, Atlanta, GA (2017), pp. 1632–1637
15. F. Zhang, J. Cao, S.U. Khan, K. Li, K. Hwang, A task-level adaptive MapReduce framework for real-time streaming data in healthcare applications. Fut. Gener. Comput. Syst. **43–44**, 149–160 (2015)
16. T.N. Hasan, A data mining approach for handling evolving data streams. OSR J. Comput. Sci. 28–33 (2016)

Twitter Sentimental Analytics Using Hive and Flume



Rupesh Kumar Mishra, Suman Lata and Soni Kumari

1 Introduction

Social media and Web 2.0 are two popular buzzwords, which have brought persistent changes in business-to-business communication, business-to-customer communication, and customer-to-customer communication. In the present era of social media, the Internet has built a participatory platform that makes an allowance for consumers to become the "media" themselves for collaborating and sharing information. Further, due to high usage of smartphones and the superfast Internet (4G and 3G), users are able to interact with social media platforms like Facebook, Twitter, Instagram, WhatsApp, etc. Social networking sites have become a well-established platform for users to express their feelings and give reviews on various topics, events, individuals, or products. The social media platform has become a famous, good, and interesting platform to put your views and ideas to people and to interact with people worldwide sitting at one place only. The amount of social media data being generated is growing rapidly in every 60 s. For every 60 s, there are 10,0000+ tweets, 695,000+ status updates, 1,10,00,000+ instant messages, 698,445 google searches, 217+ new mobile user, and 168,000+ emails. Bodnar (2010) states that more than 3 million photos are uploaded to Flickr, 5 million tweets, and a million new blog entries are posted on Twitter and other blog sites. These statistics distinctly demonstrate in our lives, with the pervasiveness of social media. Twitter is having more than 500 million users, out

R. K. Mishra (✉)
Computer Science and Engineering, Universitat of Jaume I, Castellón de la Plana, Spain
e-mail: al329127@uji.es; rupeshmishra.gn@jgannath.org

S. Lata
Department of Tourism and Travel Management, Central University of Jammu, Jammu, India

S. Kumari
Computer Science and Engineering, JEMTEC, Greater Noida, India

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_16

of which more than 332 million are active. Many tweets and search queries every day.

For managing large datasets, big data was introduced. Big data refers to a huge volume of data that cannot be stored and processed using the traditional approach in the given time frame. With such a large data, the volume and variety of data are being generated. While generating this dataset, a number of challenges are also generated. By 2020, data will grow from 4.4 zettabyte today. However, different kinds of data are being generated in form of MP3, Video, Pdf, Png, Email, etc. Structured data (Tabular), Semi-structured data (JSON, XML, CSV, EMAIL), and Unstructured data (log file, audio, video, image) are generated. Nowadays, almost every 80% of data generated is unstructured and rest is in structured form or in semi-structured form.

Paper Organization—In this paper, organization of the contents are as follows: In Sect. 2, the work related to this topic is done. In Sect. 3, we discuss about the methodology that has been used. In Sect. 4, we discuss about datasets analysis, the input data, the output data, and the result. In Sect. 5, we discuss about the conclusion of this study.

2 Related Work

"Sentimental Analysis" is also known as "Opinion mining". Nowadays, it has become popular trend as it helps in finding the reviews of a product. It uses ETL (Extract, transform, load) method. Due to high scope of big data, various researches have been done on data science. To implement the sentiment analysis, various methods and techniques are used in this study. In previous researches, some researchers have used classification methods such as Naive Bayes, Support Vector Machines, K- Nearest Neighbors, Pig, Spark, and Oozie. Continuous research is used to find the most efficient and effective method for streaming real-time data.

Various studies in previous literature that have used sentimental analysis are:

Mahalakshmi and Suseela [1] Social Sentiment Analysis and Data Visualization on Big: This method is composed of HDFS system based on Hadoop ecosystem and MapReduce functions for sentiment analysis. It uses R language for visualizing the results of sentiment and Mongo DB is used to save the Twitter status.

Mane et al. [2] Real-time Sentiment Analysis of Twitter Data using Hadoop: This method also uses the Hadoop. It uses emoticons for sentimental evaluation. Hashtags were not introduced at that time so they were not analyzed in their study.

Danthala [3] Twitter Data Processing using Apache Hadoop: This paper provides a way of analyzing data such as Twitter data using Apache Hadoop using MapReduce function by generating mapper, combiner, partitioner, and reducer which will process and analyze the tweets on a Hadoop. The analysis is done on tweets and tweets ids.

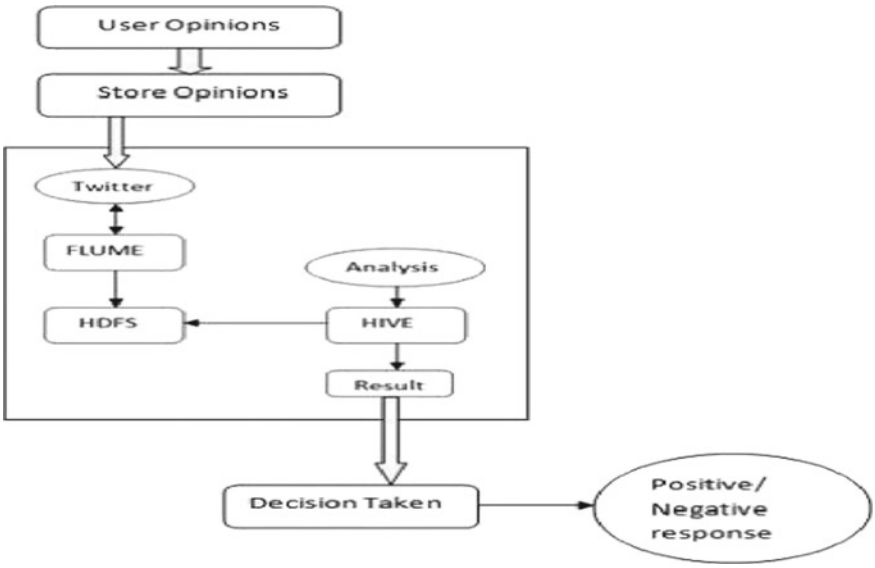
Judith Sherin Tilsha and Shobha [4] A Survey on Twitter Data Analysis Techniques to Extract Public Sentiment: In this paper, the insights of techniques that analyze the Twitter data were used. The machine learning algorithm is mainly used

in this study and as well as some other techniques were based on dictionary approach and ensemble classifiers.

Nadagoud [5] Market Sentiment Analysis for Popularity of Flipkart: In this paper, sentiment analysis on data sets is done using Big data Hadoop and some of its ecosystem components. The output data for tweets is categorized into three groups: positive, neutral, and negative reviews.

3 Methodology

To analyze the large and complex data, we require a tool that can easily handle this data in minimum time limit. Due to this, we are using Hive and Flume which is a open-source, scalable, fault tolerance framework for which no need to learn Java. Hive is similar to SQL (Structured query language) and used for data analysis transformation for large data.



As we have seen many abovementioned studies in previous literature which is related to this present study, so here are the various steps that show how to fetch data, process that data, and state in HDFS and further,

How to process this work by using different techniques. The following steps are:

1.

Create a Twitter Application.
2. Data fetching using Flume
3. Query using HQL.

(1) **Create a Twitter Application**:

The present study requires Twitter information by creating Twitter application for the analyses of data. Following are the various steps that are used to create Twitter application:

- First, open the connection of dev.twitter.com/app in Mozilla Firefox Browser and sign in the Twitter account and do some work with Twitter Application window where you will find create, delete, and manage Twitter Apps.
- In the next step, click on the Create New App button. At that point, you will get an application frame to fill your detail data.
- At this point, the new App will be created. The new app is utilized to make Consumer Key, Access Key, and Access Token Key. This will be used to edit in the Flume.conf file/record. While getting data from Twitter, these Consumer Key, Access Key, Access Token Key are used to fetch data which is lively tweeting in the account.
- These keys are used to Access Tokens tab and it can observe a button with the name of Create my access token. By Clicking this, we can produce the access token.
- Consumer keys (API Key), Consumer Secret (API Secret), and Access tokens are utilized to arrange the Flume operator.

(2) **Fetching the data using Flume**:

After creating an application in the Twitter developer site, we need to fetch the data from Twitter. For that,

- We will use the consumer key and secret key with the access token and secret values.
- Further, we can fetch data through Twitter that is required and it will be in JSON format and we will put this data in the HDFS in the location where we have saved all the data that comes from Twitter.
- The configuration file is used to get real-time data from Twitter. All the details or the points of interest need to filled in the flume-twitter.conf file, i.e., configuration file of Flume.

(3) **Query utilizing HQL**:

After setting the above design, the Flume is run, the Twitter data/information automatically will be saved into HDFS in different directories where we have set the storage path to save the Twitter data/information that was extracted by using Flume.

- To keep the data in the local file directory is not feasible because loading data in local file directory is a lengthy process.

- From the collected data, we will create a table "mytweets_raw" where the filtered data will be kept into a formatted structured so that we can clearly show that we have converted the unstructured data into structured data or in organized way.
- After loading the real-time data in Hive table, more tables are created like dictionary table which stores the polarity and the word and tweets_sentiment table which will contain all the tweets id and its sentiment. Many such more tables are created and different operations are done on data.

4 Datasets and Result Analysis

Before applying queries on the data, we need to make sure that the Hive table can appropriately translate the JSON formatted data by using JSON validator. Hive takes that input files which use a delimited row format, yet our fetched data is in a JSON format, which will not work. On the other hand, we can utilize the Hive SerDe interface to determine how to translate the data. SerDe is the interface which tells the Hive that how it should modify/change the data that Hive can process. For that, we have added a jar file "**hive-serdes-1.0-SNAPSHOT.jar**" into the directory/usr/local/hive/lib. This will be used by Hive shell to extract the clean data from the downloaded data into the Hive table (Fig. 1).

By using the Hive jar file and custom SerDe files, we can store unstructured data into Hive table name "mytweets_raw" in the structured format. The figure demonstrates the structured data stored in table named "mytweets_raw" and this is also our input data in which sentiment analysis is done.

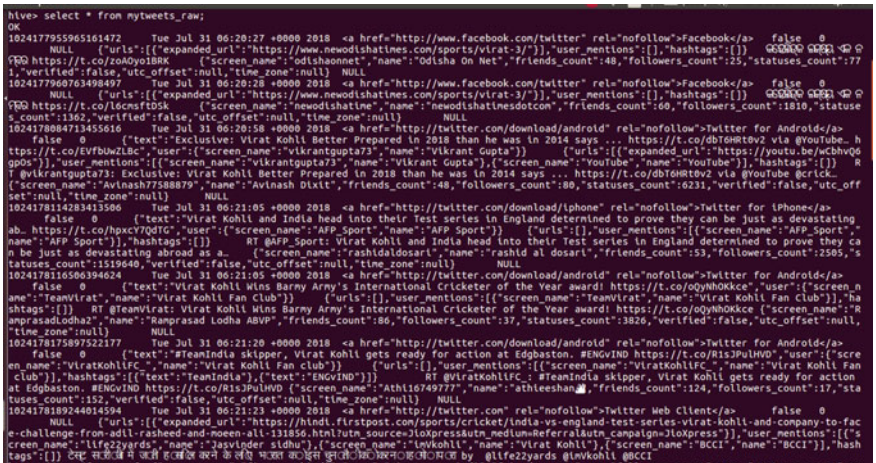


Fig. 1 Input data stored in Hive

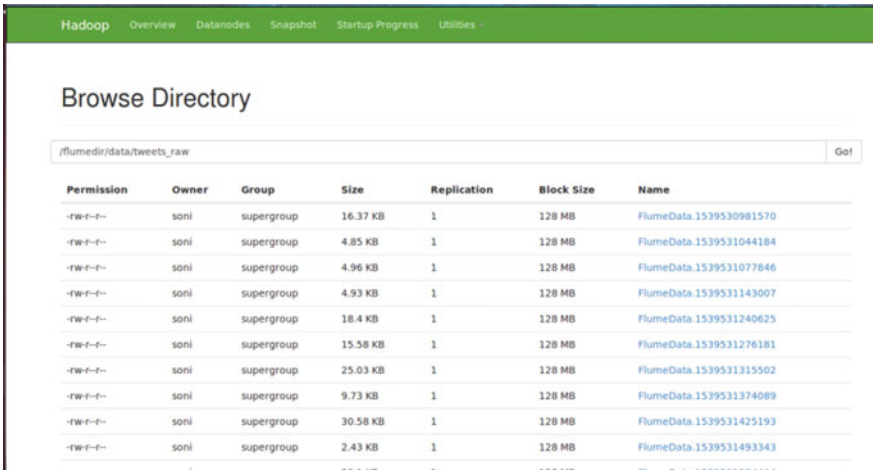


Fig. 2 Data in Flume directory

The set of data was taken from the social media platform Twitter using Twitter Streaming APIs (application program interface). Further, it was passed through Apache Flume. These fetched datasets (tweets) are stored in HDFS. Figure 2 shows the tweets data in Flume directory and it represents the list of Twitter data extracted which contains the keyword as specified in the configuration file. We can check the files by downloading them and seeing the tweets relating to the keyword. Here, our keyword or the data fetched from Twitter is of Virat.

The sentiments of the tweets were calculated by using polarity. Our output shows the sentiment of the tweets, for instance, positive, negative, or neutral in nature. The output table consists of the tweet id and the sentiment. As every tweet contains its unique id so it is easy to analyze the sentiment of every tweet (Fig. 3).

5 Conclusion

Hive and Flume are the tools of Big data Hadoop and they are efficient for extracting and loading the data. Hive is basically used for managing and querying structured data whereas Flume is used for collecting, aggregating, and moving large amounts of streaming event data. There are different methods of real-time streaming data by using codes or using MapReduce, etc. Using Apache Hive and Apache Flume, this work can be done easily and it utilizes less time too. The operations are performed on the stored data. Resulting in the present circumstances, this paper has used big dataand its tools by using Hadoop's environment. It is a Hadoop-based setup. Lastly, it provides facility for live streaming the tweets and storing it into the HDFS and analyzing using Apache Hive. In this way, here the handling time taken is additionally less contrasted with alternate strategies in light of the fact that Hadoop MapReduce

TWEET ID	SENTIMENT
1024177955965161472	NEUTRAL
1024177960763498497	NEUTRAL
1024178084713455616	POSITIVE
1024178114283413506	POSITIVE
1024178196848287745	NEGATIVE
1024179212901654528	NEGATIVE

Fig. 3 Sample of output

and Hive are the best techniques to process vast measure of information in less time. In comparison to Hive, somehow MapReduce is a time taking process. So, Hive and Flume are good and efficient tool.

References

1. R. Mahalakshmi, S. Suseela, *Big-SoSA: Social Sentiment Analysis and Data Visualization on Big Data*
2. S.B. Mane, Y. Sawant, S. Kazi, V. Shinde, *Real Time Sentiment Analysis of Twitter Data Using Hadoop*
3. M.K. Danthala, Tweet analysis: twitter data processing using Apache Hadoop. Int. J. Core Eng. Manag. (IJCEM)
4. S. Judith Sherin Tilsha, M.S. Shobha, A survey on twitter data analysis techniques to extract public opinion. Int. J. Adv. Res. Comput. Sci. Softw. Eng.
5. S. Nadagoud, K.D. Naik, Market sentiment analysis for popularity of Flipkart. Int. J. Adv. Res. Comput. Eng. Technol. (IJARCET)
6. Online Resource of Hive Available on: http://hive.apache.org/
7. Online Resource of Flume Available on: https://flume.apache.org/
8. A. Go, R. Bhayani, L. Huang. Twitter sentiment classification using distant supervision. CS224N Project Report, pp. 1–12 (Stanford, 2009)
9. J. Dean, S. Ghemawat, MapReduce: simplified data processing on large clusters, Commun. ACM **51**(1), 107–113 (2008)
10. K. Shvachko, H. Kuang, S. Radia, R. Chansler, The Hadoop
11. S.A. Bahrainian, A. Dengel, Sentiment analysis using sentiment features, in *Proceedings of WPRSM Workshop and the Proceedings of IEEE/WIC/ACM International Conference on Web Intelligence, Atlanta, USA* (2013)
12. Online resource related to big data Hadoop availabe on: https://data-flair.training/blogs/hadoop-tutorial/
13. K. Bodnar, *The Ultimate List: 300 + Social Media Statistics*. http://blog.hubspot.com/blog/tabid/6307/bid/5965/The-Ultimate-List-300-Social-Media-Statistics.aspx?source=Webbiquity (2010)

KKG-512: A New Approach for Kryptos Key Generation of Size 512 Bits Using Plaintext



Kamal Kumar Gola, Gulista Khan, Ashish Joshi and Rahul Rathore

1 Introduction

As we know that the field of cryptography provides strong protection for information technology, this technique also provides the confidentiality, integrity of message, and authentication of the sender. Cryptography depends on the three main components: first an algorithm for encryption, second an algorithm for decryption, and third is the key (public key or private key or secret key [1, 2]).

In secure management, encryption and decryption keys always play a very important role and the message's security is fully dependent on the security of the key that means if someone gets the key, then he/she can find the message. So it is necessary to secure the keys. So this work proposes an algorithm which provides security to the key. The proposed algorithm uses F1 to F8 function to increase the security of the key.

F1: ((a AND b) NOR a)
F2: ((a NOR b) AND b)
F3: (a NAND (a OR b))

K. K. Gola (✉) · G. Khan · R. Rathore
Department of Computer Science and Engineering, Faculty of Engineering, TMU, Moradabad, India
e-mail: kkgolaa1503@gmail.com

G. Khan
e-mail: gulista.khan@gmail.com

R. Rathore
e-mail: rahul.rathore15@gmail.com

A. Joshi
Department of Computer Science and Engineering, THDC-IHET, Tehri, India
e-mail: a.joshicse1986@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_17

F4: (a XNOR b) OR (a AND b)
F5: (NOT a) XOR (NOT b)
F6: ((a XOR b) OR a)
F7: ((NOT b) NAND a)
F8: (a XOR b) AND ((NOT a) OR (NOT b)).

2 Literature Review

In [3], the authors have proposed a new technique for the key generation with the help of orthogonal matrix in the hill cipher algorithm. Hill cipher is an application of a matrix-based polygraphic substitution. The proposed algorithm has reduced the disadvantage of non-invertible matrix in Hill cipher. This paper discovers the idea of generating key that is basically the reflection on a given plane.

In [4], the authors have proposed a novel key generation algorithm for Diffie–Hellman agreement that derives computational efficiency from constructing a parallel architecture. The authors have compared to the serial structure for traditional binary representation method, our algorithm is significantly more efficient on key generation and suitable for hardware implementation in an ephemeral-static mode for Diffie–Hellman agreement which is thought to be more secure than [5].

In [6], the authors have proposed a key generation algorithm using random generation. The advantage of this algorithm is the largest key size 512 bits. On the basis of 512 bits, this work also provides the internal key generation at receiver side which will give by the sender and store this key in the database at sender side. This algorithm also provides multiple time encryptions for the message or data.

In [7], the authors have been proposed a new symmetric key cryptography algorithm using extended MSA method: DJSA symmetric key algorithm. This algorithm uses the concept of random key generator to generate the initial key that is used for the encryption process. A substitution operation is also performed here that takes four characters from the input file then finds the corresponding character in the given random key matrix. Now, the encrypted message is stored in another file. A searching method was proposed by Nath in MSA algorithm to search the character in the random key matrix.

In [8], the authors have proposed a technique "Effect of Security Increment on Symmetric Data Encryption through AES Methodology". This technique is much more similar to Rijndael algorithm which is also a symmetric key algorithm. Rijndael algorithm is differing from this algorithm. This algorithm starts with block size of 200 bits while Rijndael algorithm works with block size of 128 bits.

In [9], the authors have proposed a technique for key security during transmission known as secure key transmission (SKT) approach. This approach uses the concept of MD5 algorithm, Gray code, XOR Operation, initial, and final substitution to secure

the key. The proposed approach also ensures the integrity of the key. The proposed approach shows a better performance in terms of key security and integrity.

3 Proposed Algorithm

Step-1 Let the plaintext is KryptosKeyGeneration512.
Step-2 Now convert the given plaintext into decimal number. Like a to z will be denoted by 0–25 and A to Z will be denoted by 26–51, 0–9 will be expressed as 52–61.
Step-3 Now convert the numbers into its binary number with eight bits. If the total number of bits in plaintext is less than 512 bits, then padding of 11111111 will be done to make a total of 512 bits. During decryption if the receivers get 11111111 in the data, then this has to be discarded at this stage. If the numbers of bits in plaintext is more than 512 bits, then consider the first 512 bits to generate the key (Table 1).
Step-4 Divide the bits into two parts known as LPT (A) and RPT (B). Each part having 256 bits.
Step-5 Now find the gray code of RPT (B) and XOR of LPT (A) and then apply the function-1 on LPT (A) and RPT (B).{For all function, the value of a is known as LPT (A) and value of b is RPT (B)}.
Step-6 Now taking the gray codes the function-1 which will become the value of new RPT (B) for the next round.
Step-7 Taking the XOR of the previous value of the RPT (B) which will become the new value of LPT (A) for the next round.
Step-8 Repeat step-7 and step-8 for the rest of the function according to Fig. 1.
Step-9 Now merge the LPT (A) and RPT (B) which will be the temporary key of size 512 bits.

Table 1 Symbol representation in binary number

K	00100100	**K**	00100100	**G**	00100000	**5**	00111001
r	00010001	**e**	00000100	**e**	00000100	**1**	00110101
y	00011000	**y**	00011000	**n**	00001101	**2**	00110110
p	00001111			**e**	00000100		
t	00010011			**r**	00010001		
o	00001110			**a**	00000000		
s	00010010			**t**	00010011		
				i	00001000		
				o	00001110		
				n	00001101		

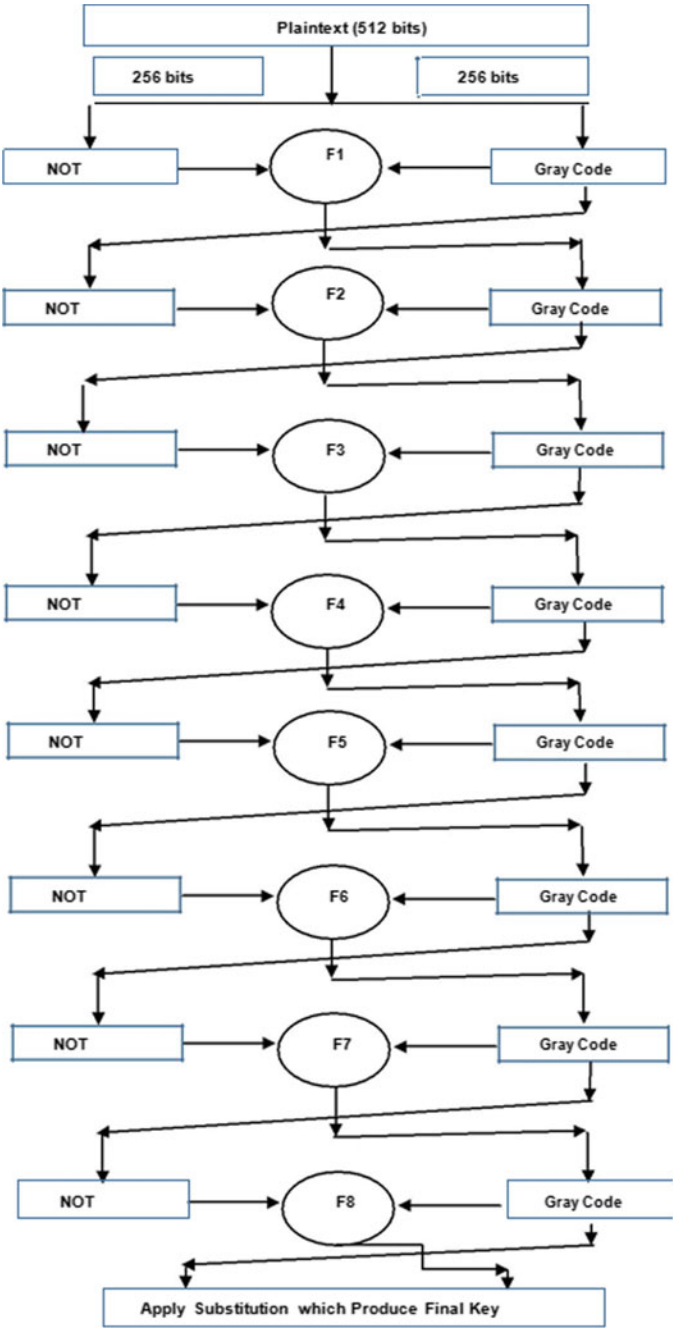


Fig. 1 Model for key generation

Step-10 Now apply the substitution operation on each group of eight bits and then convert the integer value into its equivalent symbol which will produce the final key of size 512 bits (Table 2).

4 Implementation

Step-1 Let the plaintext is KryptosKeyGeneration512.
Step-2 Now convert the given plaintext in decimal number as discussed in the proposed algorithm.
Step-3 Now performed this step as discussed in the proposed algorithm and then considers the first 512 bits to generate the key.
Step-4

00100100000100010001100000001111000100110000111000010010001001000000
10000011000001000000000010000001101000001000001000100000000001001100
00100000001110000011010011100100111010100110111011111111111111111111
11111111111111111111111111111111111111111111111111111111111111111111
11111111111111111111111111111111111111111111111111111111111111111111
11111111111111111111111111111111111111111111111111111111111111111111
11111111111111111111111111111111111111111111111111111111111111111111
11111111111111111111111111

Step-5 Divide the bits into two parts known as LPT (A) and RPT (B). Each part having 256 bits.
LPT (A)

00100100000100010001100000001111000100110000111000010010001001000000
10000011000001000000000010000001101000001000001000100000000001001100
00100000001110000011010011100100111010100110111011111111111111111111
11111111111111111111111111111111111111111111111111

RPT (B)

11111111111111111111111111111111111111111111111111111111111111111111
11111111111111111111111111111111111111111111111111111111111111111111
11111111111111111111111111111111111111111111111111111111111111111111
11111111111111111111111111111111111111111111111111

Step-6 Now find the gray code of RPT (B) and NOT of LPT (A), then apply the function-1 on LPT (A) and RPT (B).
Step-6.1

Table 2 Substitution table

Middle six bits																
Outer bits		**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
	0	1	3	5	7	9	11	13	15	17	19	21	23	25	27	29	31
	1	2	61	59	57	55	53	51	59	47	45	43	41	39	37	35	33
	2	5	4	6	8	10	12	14	16	18	20	22	24	26	28	30	32
	3	0	15	60	58	56	54	52	50	48	46	44	42	40	38	36	34
Middle six bits																	
Outer bits		**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**26**	**27**	**28**	**29**	**30**	**31**
	0	33	35	37	39	41	43	45	47	49	51	53	55	57	59	61	5
	1	31	29	27	25	23	21	19	17	15	13	11	9	7	5	3	1
	2	34	36	38	40	42	44	46	48	50	52	54	56	58	60	15	0
	3	32	30	28	26	24	22	20	18	16	14	12	10	8	6	4	2
Middle six bits																	
Outer bits		**32**	**33**	**34**	**35**	**36**	**37**	**38**	**39**	**40**	**41**	**42**	**43**	**44**	**45**	**46**	**47**
	0	0	2	4	6	8	10	12	14	16	18	20	22	24	26	28	30
	1	15	60	58	56	54	52	50	48	46	44	42	40	38	36	34	32
	2	1	3	5	7	9	11	13	15	17	19	21	23	25	27	29	31
	3	5	61	59	57	55	53	51	49	47	45	43	41	39	37	35	33
Middle six bits																	
Outer bits		**48**	**49**	**50**	**51**	**52**	**53**	**54**	**55**	**56**	**57**	**58**	**59**	**60**	**61**	**62**	**63**
	0	32	34	36	38	40	42	44	46	48	50	52	54	56	58	60	15
	1	30	28	26	24	22	20	18	16	14	12	10	8	6	4	2	0
	2	33	35	37	39	41	43	45	47	49	51	53	55	57	59	61	5
	3	31	29	27	25	23	21	19	17	15	13	11	9	7	5	3	1

a) Calculation of Function-1
a.1) F1: ((a AND b) NOR a)
{For all function, the value of a is known as LPT (A) and value of b is RPT (B)}
a.2) Gray Code of RPT (B)

1000000000000000000000000000000000000000000000000000000000000000000000000000000000
0000000000000000000000000000000000000000000000000000000000000000000000000000000000
0000000000000000000000000000000000000000000000000000000000000000000000000000000000
0000000000000000000000000000000000000000000000000000000000000000

a.3) Value of NOT Operation of LPT (A)

1101101111101110111001111111000011101100111100011110110111011011111111
0111110011111011111111111011111100101111101111101110111111111110110011
1101111111000111110010110001101100101011001001000000000000000000000000
0000000000000000000000000000000000000000000000000000000000000000

a.4) Value of (a AND b) Operation

1000000000000000000000000000000000000000000000000000000000000000000000000000000000
0000000000000000000000000000000000000000000000000000000000000000000000000000000000
0000000000000000000000000000000000000000000000000000000000000000000000000000000000
0000000000000000000000000000000000000000000000000000000000000000

a.5) Value of ((a AND b) NOR a) Operation which is Function-1

0101101111101110111001111111000011101100111100011110110111011011111111
0111110011111011111111111011111100101111101111101110111111111110110011
1101111111000111110010110001101100101011001001000000000000000000000000
0000000000000000000000000000000000000000000000000000000000000000

a.6) Value of NOT Operation of Gray RPT (B) (Step a.2)

0111111111111111111111111111111111111111111111111111111111111111111111111111111111
1111111111111111111111111111111111111111111111111111111111111111111111111111111111
1111111111111111111111111111111111111111111111111111111111111111111111111111111111
1111111111111111111111111111111111111111111111111111

a.7) Gray Code of Function-1 (Step a.5)

0111011000011001100101000000100010011010100010010001101100110110000000
1100001010000110000000001100000101110000110000110011000000000001101010
0011000000100100001011101001011010111110101101100000000000000000000000
0000000000000000000000000000000000000000000000000000

Step 6.2
b) Calculation of Function-2
b.1) F2: ((a NOR b) AND b) for this function the value of a will be the value of (Step a.6) and the value of b will be the value of (step a.7).
b.2) Value of (a NOR b) Operation

1000000000000000000000000000000000000000000000000000000000000000000000000
0000000000000000000000000000000000000000000000000000000000000000000000000
0000000000000000000000000000000000000000000000000000000000000000000000000
0000000000000000000000000000000000000000000000000

b.3) Value of ((a NOR b) AND b) which is Function-2

0000000000000000000000000000000000000000000000000000000000000000000000000
0000000000000000000000000000000000000000000000000000000000000000000000000
0000000000000000000000000000000000000000000000000000000000000000000000000
0000000000000000000000000000000000000000000000000

b.4) Gray Code of Function-2 (Step b.3)

1000000000000000000000000000000000000000000000000000000000000000000000000
0000000000000000000000000000000000000000000000000000000000000000000000000
0000000000000000000000000000000000000000000000000000000000000000000000000
0000000000000000000000000000000000000000000000000

b.5) Value of NOT Operation of Gray Code of Function-1 (Step a.7)

1000100111100110011010111111011101100101011101101110010011001001111111
0011110101111001111111111001111101000111100111100110011111111110010101
1100111111011011110100010110100101000001010010011111111111111111111111
1111111111111111111111111111111111111111111111111

Step 6.3
c) Calculation of Function-3
c.1) F3: (a NAND (a OR b)) for this function, the value of a will be the value of (Step b.5) and the value of b will be the value of (Step b.4)
c.2) Value of (a OR b) Operation

1000100111100110011010111111011101100101011101101110010011001001111111
0011110101111001111111111001111101000111100111100110011111111110010101
1100111111011011110100010110100101000001010010011111111111111111111111
1111111111111111111111111111111111111111111111111

c.3) Value of (a NAND (a OR b) which is Function-3

01110110000110011001010000001000100110101000100100011011001101100000011000010100001100000000011000001011100001100001100110000000000110101000110000001001000010111010010110101111101011011000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000

c.4) Value of NOT Operation of Gray Code of Function-2 (Step b.4)

01111111111

1100110100010101010111100000110011010111110011011001011010101101000000
1010001111000101000000001010000111001000101000101010100000000001011111
0010100000110110001110011101110111100001111011010000000000000000000000
0000000000000000000000000000000000000000000000000000

d.5) Gray Code of Function-4

1010101110011111111110001000010101011110000101011010111011111101110000
1111001000100111100000001111000100101100111100111111110000000011100000
1011110000101101001001010011001100010001000110111000000000000000000000
0000000000000000000000000000000000000000000000000000

d.6) Value of NOT Operation of Gray Code of Function-3 (Step c.5)

1011001011101010101000011111001100101000001100100110100101010010111111
0101110000111010111111111010111100011011101011101010101111111110100000
1101011111001001110001100010001000011110000100101111111111111111111111
1111111111111111111111111111111111111111111111111111

Step 6.5
e) Calculation of Function-5
e.1) F5: (NOT a) XOR (NOT b) for this function, the value of a will be the value of (Step d.6) and the value of b will be the value of (Step d.5)
e.2) Value of (NOT a)

0100110100010101010111100000110011010111110011011001011010101101000000
1010001111000101000000001010000111001000101000101010100000000001011111
0010100000110110001110011101110111100001111011010000000000000000000000
0000000000000000000000000000000000000000000000000000

e.3) Value of (NOT b)

0101010001100000000011101111010101000011110101001010001000000100011111
0000110111011000011111111000011101101001100001100000000111111110001111
0100001111010010110110101100110011101110111001000111111111111111111111
1111111111111111111111111111111111111111111111111111

e.4) Value of (NOT a) XOR (NOT b) which is Function-5

0001100101110101010100001111100110010100000110010011001010100101111110
1011100001110101111111101011110001101110101110101010111111111101000001
1010111110010011100011000100010000111100001001011111111111111111111111
1111111111111111111111111111111111111111111111111111

e.5) Gray Code of Function-5

0001010111001111111110001000010101011110000101011010101111111011100001
1110010001001111000000011110001001011001111001111111100000000111000 01
0111100001011010010010100110011000100010001101110000000000000000000000
0000000000000000000000000000000000000000000000000000

e.6) Value of NOT Operation of Gray Code of Function-4 (Step d.5)

0101010001100000000011101111010101000011110101001010001000000100011111
0000110111011000011111111000011101101001100001100000000111111110001111
0100001111010010110110101100110011101110111001000111111111111111111111
1111111111111111111111111111111111111111111111111111

Step 6.6
f) Calculation of Function-6
f.1) F6: ((a XOR b) OR a) for this function, the value of a will be the value of (Step e.6) and the value of b will be the value of (Step e.5)
f.2) Value of (a XOR b) Operation

0100000110101111111101100111000000011101110000010000100111110011011 10
1110100110010111011111101110110010101010111010111111101111111011011 10
0011101110001000100100001010100011001100110100110111111111111111111 11
1111111111111111111111111111111111111111111111111111

f.3) Value of ((a XOR b) OR a) Operation which is Function-6

0101010111101111111111101111010101011111111010101101010111111011101111
1110110111011111011111111110111011111011111101111111111011111111110111 1
0111101111011010110110101110110011101110111011101111111111111111111 1
1111111111111111111111111111111111111111111111111111

f.4) Gray code of Function-6

0111111100011000000000011000111111110000011111101111110000011001100 0
0001101100110000110000000001100110000110000110000000011000000000011000
1100011000110111101101111001101010011001100011001100000000000000000000
0000000000000000000000000000000000000000000000000000

f.5) Value of NOT Operation of Gray Code of Function-5 (Step e.5)

11101010001100000000011101111010101000011110101001010100000100011110
00011011101100001111111000011101101001100001100000000111111110011110
10000111101001011011010110011001110111011100100011111111111111111111
11111111111111111111111111111111111111111111111111

Step 6.7
g) Calculation of Function-7
g.1) F7: ((NOT b) NAND a) for this function, the value of a will the value of (Step f.5) and the value of b will be the value of (Step f.4)
g.2) Value of (NOT b) Operation

10000000111001111111111001110000000011111100000010000001111100110011
11100100110011110011111111100110011110011110011111111100111111111100111
00111001110010000100100001100101011001100111001100111111111111111111
11111111111111111111111111111111111111111111111111

g.3) Value of ((NOT b) NAND a) Operation which is Function-7

01111111110111111111100110001111111111100011111111111111111111111001
11111111011111111100000111111011110111111111111111111111100000011111001
11111110011111111111111111111101011101110111111110000000000000000000
00000000000000000000000000000000000000000000000000

g.4) Gray Code of Function-7

01000000001100000000010101001000000000010010000000000000000000000101
00000000110000000010000100000110001100000000000000000000100000100000101
00000001010000000000000000000001111001100110000000100000000000000000
00000000000000000000000000000000000000000000000000

g.5) Value of NOT Operation of Gray Code of Function-6 (Step f.4)

10000000111001111111111001110000000011111100000010000001111100110011
11100100110011110011111111100110011110011110011111111100111111111100111
00111001110010000100100001100101011001100111001100111111111111111111
11111111111111111111111111111111111111111111111111

Step 6.8
h) Calculation of Function-8
h.1) F8: (a XOR b) AND ((NOT a) OR (NOT b)) for this function, the value of a will the value of (Step g.5) and the value of b will be the value of (Step g.4)

h.2) Value of (a XOR b) Operation

1100000011010111111110110011100000001110111000001000000111110011000101111001000000111100011110111000000100100111100111111110001111101100010001110001000100001001000011001001000000000001001100011111111111

Step-8 Now make a group of eight bits and then apply the substitution operation on each group of eight bits and convert the bits into its equivalent symbol according to Step-9

0 49 59 8 5 33 1 1 41 7 5 9 15 2 1 9 31 0 16 1 1 6 25 9 1 1 1 1 1 1 11 1 41 5 57 15 33
0 13 41 33 57 9 3 48 1 57 25 16 35 19 25 33 5 57 11 1 1 1 1 1 1

Step-9 Convert the integer value into its equivalent symbol

aX7ifHbbPhfjpcbjFaqbbgzjbbbbbbbbbPf5pHanPH5jdWb5zqJtzHf5bbbbbbbb.

This is the Final key.

5 Results and Conclusion

In symmetric key cryptography, only one key is used to encrypt and decrypt the data and only those can party share the key who are involved in the communication. So key must be kept secret, and the parties who share a key trust upon each other not to disclose the key. In secure communication, keys play a very important role and information security is totally dependent on the security of the key that means if someone gets the key, then he/she can find the message. So it is necessary to secure the keys. To conclude, this work is an effort to introduce a new approach in the generation of symmetric keys to overcome the problem of weak keys. In proposed algorithm, plaintext is used to generate the key and this process is only done by the sender that means only the sender knows about the key. The proposed algorithm is much secure as compared to other existing algorithms because this work always generates a large key of size 512 bits. This work uses eight round function and substitution operation to generate the key which is unbreakable for the unauthorized users. In this work, only numeric value and alphabets are used to generate the key. So this work has a wide scope to generate the large key size up to 1024 bits having special symbols, operators, etc.

References

1. K.K. Gola, B. Gupta, Z. Iqbal, Modified RSA digital signature scheme for data confidentiality. Int. J. Comput. Appl. **106**(13), 13–16 (2014)
2. Z. Iqbal, K.K. Gola, B. Gupta, M. Kandpal, Dual level security for key exchange using modified RSA public key encryption in playfair technique. Int. J. Comput. Appl. **111**(13), 5–9 (2015)
3. F.H. Khan, R. Shams, F. Qazi, D.-E.-S. Agha, Hill cipher key generation algorithm by using orthogonal matrix. Int. J. Innov. Sci. Mod. Eng. (IJISME) **3**(3) (2015). ISSN: 2319-6386

4. Y. Chen, X. Chen, Y. Mu, A parallel key generation algorithm for efficient Diffie-hellman key agreement, in *International Conference on Computational Intelligence and Security*, ed. by Y. Cheung, Y. Wang, H. Liu (IEEE, Hong Kong, 2006), pp. 1393–1395
5. E. Rosorla, Diffie-Hellman key agreement method. RFC 2631 (1999)
6. K.K. Pandey, V. Rangari, S.K. Sinha, An enhanced symmetric key cryptography algorithm to improve data security. Int. J. Comput. Appl. (0975–8887) **74**(20) (2013)
7. D. Chatterjee, J. Nath, S. Dasgupta, A. Nath, A new symmetric key cryptography algorithm using extended MSA method: DJSA symmetric key algorithm, in *2011 International Conference on Communication Systems and Network Technologies*, 978-0-7695-4437-3/11 (IEEE 2011)
8. M.N. Islam, M.M.H. Mia, M.F.I. Chowdhury, M.A. Matin, Effect of security increment to symmetric data encryption through AES methodology, in *Ninth ACIS International Conference on Software Engineering, Artificial Networking, and Parallel/Distributed IEEE* (2008)
9. K.K. Gola, V. Sharma, R. Rathore, SKT: a new approach for secure key transmission using MGPISXFS, in *Information Systems Design and Intelligent Applications. Advances in Intelligent Systems and Computing*, vol. 433 (Springer, New Delhi, 2016)

HD-MAABE: Hierarchical Distributed Multi-Authority Attribute Based Encryption for Enabling Open Access to Shared Organizational Data



Reetu Gupta, Priyesh Kanungo and Nirmal Dagdee

1 Introduction

Cloud computing provides a convenient and flexible method for sharing and outsourcing the data. Any organization or data owner can share the data that may be beneficial for the use of the society and the individual. But, the data may contain valuable information or it may be sensitive due to privacy concerns, e.g., medical data of a patient. To keep the shared data confidential, the data owner imposes some access control mechanisms. In open-access systems [1, 2], data can be shared with internal and external users of the organization based on characteristics of users, i.e., based on the attributes possessed by them. In this paper, we propose an access control scheme, in which organization's internal users, as well as, authenticated external users can access the required data. For instance, in electronic healthcare management system, health records are kept on cloud in encrypted and protected mode, which can be accessed by the doctors for the treatment of the patient in emergency situations. The system should allow any eligible doctor to decrypt the record. The user of the data, i.e., the doctor, can be a registered doctor to the healthcare management system or he can be external to the system, but he should carry the authentication of being a doctor.

Attribute based encryption (ABE) [3] is a well-accepted access control technique for sharing the data in cloud environment. Using this technique, data is kept on

R. Gupta (✉) · N. Dagdee
Sushila Devi Bansal College of Technology, Indore, India
e-mail: reetu.gupta@sdbct.ac.in

N. Dagdee
e-mail: nirmal_dagdee@rediffmail.com

P. Kanungo
SVKM's NMIMS, Shirpur Campus, Shirpur, India
e-mail: priyeshkanungo@hotmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_18

cloud in encrypted form and the data owner can define the access policies based on various attributes. If a user satisfies the access policy, then the encrypted data can be decrypted by him. There are two major types of ABE: Ciphertext-Policy ABE (CP-ABE) [3] and Key-Policy ABE (KP-ABE) [4]. In CP-ABE scheme, ciphertexts contain access policy for encryption and the encrypted message. The user's secret keys are issued by attribute authority based on his attributes. A user is able to decrypt the ciphertext if his attribute secret keys satisfy the access policy. In KP-ABE scheme, a ciphertext consists of a set of attributes, and the user's secret key contains the access policy. A user can decrypt the ciphertext if the policy in his secret key is satisfied by the attributes mentioned in the ciphertext. The attributes used in ABE schemes are issued by some authority. There can be a single authority or multiple authorities for attribute management. In single authority ABE approaches, a centralized master authority issues secret keys to all the users [5–7]. This type of approach can be used when users for the shared data belong to single organization or domain. In case of open-access systems [1, 2], we will require multi-authority approach. Various multi-authority schemes [8–10] have been proposed in literature. These schemes allow users from multiple domains, but cannot handle issues of open access systems. In open-access systems, we need flexibility in defining access control policies and the system should be scalable as well. In this paper, we have proposed hierarchical distributed multi-authority attribute based access control scheme. In our scheme, data owner can mention general attributes in his access policy. These attributes can be issued by well trusted hierarchical structured attribute authorities. This hierarchy is established by trusted root authority. For example, doctor attribute issued by a hospital, which in turn is authorized by government medical council, is considered trustable by everyone. The attribute of professor issued by any university or college, which is also affiliated by government authority, is considered as trustable. In this way, we can make a large group of users to be eligible to access the system. Issuance of attributes from properly established and trusted hierarchical structured authorities will make the attribute management easy, scalable, and distributed.

The rest of the paper is arranged as follows: Section 2 summarizes the work already done in the field of cryptographic access control covering multi-authority attribute based encryption and hierarchical attribute based encryption. Section 3 discusses the data access control scenario and system model. Section 4 presents the preliminary knowledge about access structure. Section 5 describes the scheme's construction. In Sect. 6 the security of the scheme is analyzed. Finally, we have concluded the work in Sect. 7 by outlining its salient advantages.

2 Related Work

ABE technique [11] is an access control technique to encrypt data in one-to-many encryption formats. It is used for sharing the data in a controlled manner in realistic scenarios. The centralized approaches [3, 4, 12] allow the single key generation center to issue user's secret key and thus prone to single point of failure. Distributing the task

of key issuing to multiple authorities gave rise to multi-authority ABE [8] systems, in which each authority controls a set of attributes. It used the concept of a trusted central authority (CA) and global identifiers (GID), but security of encryption and privacy of users were two major concerns there. Security of encryption is breached as CA can decrypt every ciphertext. Also privacy is compromised if some colluding authorities pool their data by using user's GID. Chase and Chow [9] proposed improved multi-authority ABE by removing CA, but with one limitation. In their scheme, a user has to acquire at least one attribute from each authority. Li et al. [12] used above approach to design personal health records (PHR) sharing scheme, where users can be from public or private domains. But the limitation of having at least one attribute from each authority cannot be satisfied in this kind of real application of PHR sharing. Decentralized ABE [10] proposed by Lewko and Waters do not require a central authority. The setup also does not require any cooperation among multiple authorities. The approach used a hash function on user's Global Identifier (GID). Using GID helped to link a user's keys which are issued to him by different authorities.

Decentralized and multi-authority ABE schemes distribute key management tasks among multiple authorities of different domains. Wang et al. [13] proposed Hierarchical ABE (HABE) system for fine-grained access control in cloud storage services. To support hierarchical structure of departments inside an organization, this approach embeds a root authority, which manages lower-level authorities. These lower-level authorities further manage attributes and users at different levels. This approach served the data access needs in large-scale enterprises and allowed delegation of access rights with improved scalability. Another similar approach named HASBE (Hierarchical Attribute Set Based Encryption) was proposed by Wan et al. [14]. This work mixed ciphertext-policy attribute set based encryption (ASBE) [15] with hierarchy concept. Key management work in hierarchical setting was relaxed by key delegation. Huang et al. [16] used HABE approach to curtail heavy key management burden in collaborative data sharing among mobile devices. Luo et al. [17] proposed a scheme for friend discovery by using hierarchical multi-authority CP-ABE scheme. In [18], key delegation approach is investigated in single authority scheme to resolve the need of multiple authorities. In that scheme, central authority issued keys to top-level users, who can further delegate keys to lower-level users. This approach could work well in a corporate environment, but can't be used in open access systems, where higher-level authorities will authorize lower-level authorities for a subset or proper subset of their attributes. So it should not be permissible for a user to delegate a key in open access system. In this paper, we are merging key delegation approach with hierarchical multi-authority CP-ABE scheme, where the root authority will delegate his authorization rights to lower-level authorities for set of attributes. Our work for closed and open domain users is addressing the same issue, as the work [12] presented for public and personal domain.

3 Data Access Scenario and System Model

We consider a scenario of a research organization working in the field of basic sciences and atomic energy. The organization may be interested in sharing some research data with masses based on some access policy. The policy can be to make the shared data available to all the Ph.D. scholars enrolled with the organization and to all the professors who are working in the area of physics. This scenario is illustrated in Fig. 1. Here the motive of sharing of research data is to give benefit to the organization's scholars, as well as, to support research work of eminent professors of the field. In this scenario, standard ABE would not work, as the attributes may be issued by different authorities. For the attribute "Ph.D. scholar", the organization itself would act as an attribute authority, but the attribute "Professor in Physics" could be issued by any university or college on which the organization trusts.

But the research organization would not like to list out all the trusted universities/colleges. Instead, it would be appropriate to create a trusted hierarchy of universities and colleges. For example, Technical Education Council (TEC) (see Fig. 1), authorizes various universities and affiliated colleges in its hierarchy, to issue attributes of professors.

To model this kind of system, the data owner will outsource encrypted data on the cloud for sharing it with various users (see Fig. 2). It will also act as an attribute authority for closed domain users. The external users will get their credentials issued from attribute authorities that exist in a hierarchy. The access policy stated by the data owner will include attributes from open as well as closed domain.

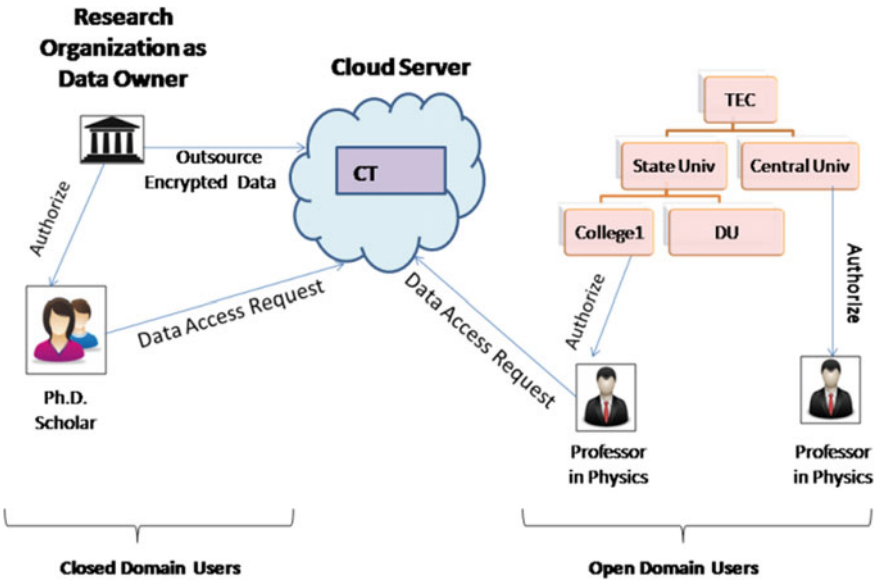


Fig. 1 Example scenario

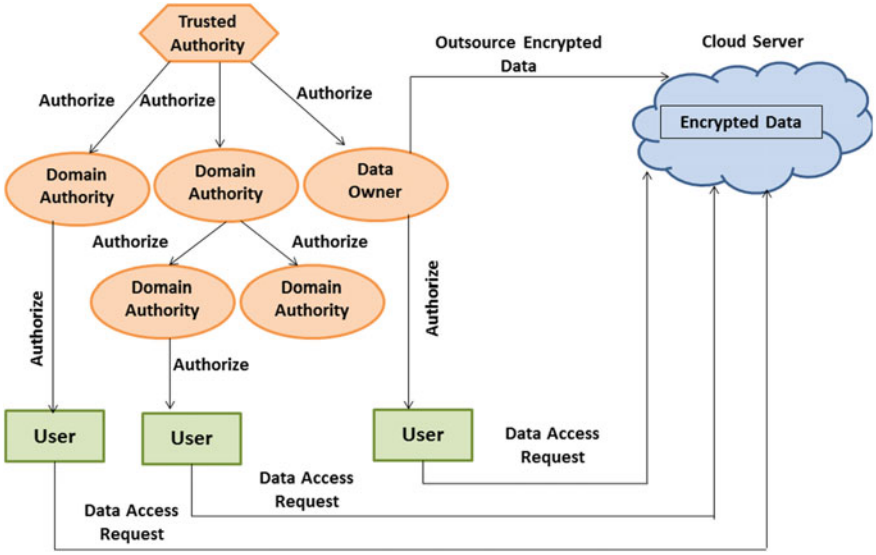


Fig. 2 System model

4 Access Structure

In the scenario, we stated in Fig. 1, the data owner/organization acts an attribute authority for organizational users and multiple attribute authorities (AAs) would be issuing attributes to external users. The data shared may be categorized in small or fine-grained records with tags based on some keywords. Each record can be accompanied by respective access structure (Π). The access structure can be defined by using $\vee$ (or) and $\wedge$ (and) operators over attributes. For example the access structure for sharing research data can be formed as:

$$(\text{Any Professor} \wedge \text{Specialization in Physics}) \vee \text{Ph.D. Scholar}$$

In our approach, we will use the fact that every Boolean formula can be reduced to Disjunctive Normal Form (DNF). DNF is ORing of various conjunctive which are basically ANDing of terms. So if n conjunctives are there, the access structure can be represented as:

$$\Pi = \Pi_1 \vee \Pi_2 \vee \ldots \vee \Pi_n$$

For simplicity, we assume that Π_1 is managed by hierarchical structure say TEC and Π_2 is managed by the data owner, i.e., research organization.

5 HD-MAABE Scheme Construction

This section describes the proposed HD-MAABE scheme which is built upon the CP-ABE scheme [5, 19] and HASBE scheme [14]. The scheme has five algorithms: (i) Global Setup algorithm (ii) Authority Setup algorithm, which comprises of RootAASetup and other Level Authority Setup, (iii) User key generation algorithm, (iv) Encryption algorithm, and (v) Decryption algorithm.

5.1 Global Setup Algorithm

The algorithm takes the input λ (security parameter) and generates GP (global parameters) for initializing the system. The GPs are shared among all AAs. The algorithm chooses two groups G and G_T (multiplicative groups) of prime order p and g (generator) of G. A bilinear map $e : G \times G \rightarrow G_T$ is also chosen. A secure hash function $\boldsymbol{H} : \{\mathbf{0}, \mathbf{1}\}^* \rightarrow \boldsymbol{Z}_p^*$ is used to map each user's identity string to a unique value $\boldsymbol{Z}_p^*$. So the system's global parameters would be GP = {G, G_T, g, p, e, H}.

Each attribute authority (AA) and the data owner (DO) are initialized locally. Each AA administers a particular attribute set in its domain.

5.2 RootAASetup Algorithm

This setup algorithm takes global parameters as input and generates public key PK and secret key SK of root attribute authorities. DO and root master authority can be considered as AA_R for a closed and open domain, respectively. We will use subscript R to denote root authority. Let root authority administers attribute set AS_R. It chooses two random exponents $\alpha_R, \beta_R \in Z_p^*$ and computes g^{β_R} and $e(g, g)^{\alpha_R}$. It then selects random values $t_{R,i} \in Z_p^*$ for each element in AS_R. The public key element and delegation key element of an attribute are denoted as:

$$PKA_{R,i} = g^{t_{R,i}} \tag{1}$$

$$DKA_{R,i} = g^{t_{R,i}^{-1}} \tag{2}$$

for i = 1, 2, … $|AS_R|$

The public and secret key components for root authorities will be

$$PK_R = \{h_1 = g^{\beta_R}, e(g, g)^{\alpha_R}, \{PKA_{R,i}\}, f_1 = g^{1/\beta_R}, \{DKA_{R,i}\}\} \tag{3}$$

$$SK_R = \{\alpha_R, \beta_R, t_{R,i}\} \tag{4}$$

for i = 1, 2, … $|AS_R|$

5.3 *Level1Setup Algorithm*

This setup algorithm takes public and secret key of root authority (PK_R, SK_R) and a set of attributes $AS_{L1} \subseteq AS_R$as input and outputs SK_{L1}. It picks up a random $r_1 \in Z_p$ and then generates

$$SK_0^{L1} = g^{\frac{\alpha_R - r_1}{\beta_R}} \tag{5}$$

and for each *j*th attribute in AS_{L1}

$$SK_j^{L1} = \left(DKA_j\right)^{r_1} = g^{\frac{r_1}{t_{R,j}}} \tag{6}$$

for j = 1, 2, … $|AS_{L1}|$

The secret key of level1 authority is denoted as

$$SK_{L1} = \{SK_0^{L1}, \{SK_j^{L1}\}_{j=1,2,\ldots|AS_{L1}|}\}$$

5.4 *Level2Setup Algorithm*

The inputs to the algorithm are PK_R, SK_{L1}, AS_{L2}. Here we assume attribute set $AS_{L2} \subseteq AS_{L1}$. Now it picks up a random $r_2 \in Z_p$ and generates

$$SK_0^{L2} = (SK_0^{L1}).(f1)^{-r_2} = g^{\frac{\alpha_R - r_1 - r_2}{\beta_R}} \tag{7}$$

and for each *j*th attribute in AS_{L2}

$$SK_j^{L2} = \left(SK_j^{L1}\right)\left(DKA_j\right)^{r_2} = g^{\frac{r_1 + r_2}{t_{R,j}}} \tag{8}$$

for j = 1, 2, … $|AS_{L2}|$

The output of algorithm SK_{L2}, i.e., the secret key of level2 authority is denoted as

$$SK_{L2} = \{SK_0^{L2}, \{SK_j^{L2}\}_{j=1,2,\ldots|AS_{L2}|}\}$$

5.5 UserKeyGen Algorithm

Let a user U wants to get AS_U to be issued from the authority at level2 and let $AS_U \subseteq AS_{L2}$. Now the algorithm will first choose $r_u = H(U)$. Then it computes parts of SK_U.

$$SK_0^U = (SK_0^{L2}) \cdot (f1)^{-r_u} = g^{\frac{\alpha_R - r_1 - r_2 - r_u}{\beta_R}} \tag{9}$$

$$SK_j^U = \left(SK_j^{L2}\right)\left(DKA_j\right)^{r_u} = g^{\frac{r_1 + r_2 + r_u}{t_{R,j}}} \tag{10}$$

for $j = 1, 2, \ldots |AS_U|$.

5.6 Encrypt Algorithm

Let the data owner wants to encrypt message M with access structure Π. Here we assumed Π as disjunctives of substructures Π_1 and Π_2, which are access structures for organizational users and outside users, respectively. Let us assume that each access structure consists of n attributes, which are issued by either the data owner or by the hierarchical structure of attribute authorities. Here we will represent ciphertext component encrypted with access substructures as $\varepsilon_{\Pi 1}$ and $\varepsilon_{\Pi 2}$. Each ciphertext component will have following contents:

$$C_0 = Me(g, g)^{\alpha_R s} \tag{11}$$

$$C_1 = (h_1)^s = g^{s\beta_R} \tag{12}$$

$$C_{R,j} = PKA_{R,j}^{s_i} \tag{13}$$

where a random $s_i \in Z_p^*$ is chosen for each attribute in n attributes, but the last term is assigned a value equal to $s - \sum_{i-1}^{n-1} s_i$ (according to unanimous consent control by modular addition scheme [19]).

5.7 Decrypt Algorithm

Let a user U, carrying secret key SK_U for attribute set AS_U, submits his keys for the decryption process. To decrypt ε_Π, the algorithm first calculates

$$\prod_{j=1}^{n} e(C_{R,j}, SK_j^U) = \prod_{j=1}^{n} e(g^{t_{R,j}s_i}, g^{\frac{r_1+r_2+r_u}{t_{R,j}}}) = e(g,g)^{(r_1+r_2+r_u)s} \tag{14}$$

Then it calculates

$$e\left(C_1, SK_0^U\right) = e\left(g^{s\beta_R}, g^{\frac{\alpha_R - r_1 - r_2 - r_u}{\beta_R}}\right) = e(g,g)^{\alpha_R s - (r_1+r_2+r_u)s} \tag{15}$$

By multiplying (14) and (15) and dividing C_0 by result, the algorithm outputs the message M as follows:

$$\frac{Me(g,g)^{\alpha_R s}}{e(g,g)^{(r_1+r_2+r_u)s}.e(g,g)^{\alpha_R s-(r_1+r_2+r_u)s}} = M$$

Thus, after running decryption algorithm, we get the message M which was encrypted with access structure Π.

6 Security Analysis

We analyze our HD-MAABE scheme in respect of collusion resistance, data confidentiality, and fine-grained access control.

6.1 Collusion Resistance

In data sharing systems, the prevention of collusion means to refrain the users from combining their attribute keys to access the data. By colluding the keys they can access the data, which they can't access when they operate individually. In our UserKeyGen() function, we used hash function on user's identity to map it to a unique value in Z_p^*. So if two users U_1 and U_2 collude their secret keys, they will have different r_u in secret keys

$$SK_0^U = g^{\frac{\alpha_R - r_1 - r_2 - r_u}{\beta_R}} \text{ and } SK_j^U = g^{\frac{r_1+r_2+r_u}{t_{R,j}}}.$$

In the decryption process Eq. (14) will not be solved, so decryption would fail in that case.

6.2 Data Confidentiality

The data confidentiality is being achieved in the scheme, as any user or application which does not satisfy the access structure Π, cannot decrypt the shared encrypted data. Equations (14) and (15) cannot result in the factor $e(g, g)^{\alpha_{RS}}$ without satisfying the respective access structure.

6.3 Fine-Grained Access Control

Our HD-MAABE scheme provides fine-grained access for users from closed and open domains. By using CP-ABE scheme and key delegation in a hierarchical setting, this kind of access control is being imposed. The data owner can create expressive and flexible access structure by using public or general attributes and organizational attributes.

7 Conclusion

This paper presents a Hierarchical Distributed Multi-Authority ABE (HD-MAABE) scheme. The scheme allows the data owner to share his data in cloud environment with users from open and closed domains. This kind of sharing in open access system benefits a large group of users. The attributes of external users are issued by the hierarchy of established trustworthy authorities and attributes of internal organizational users are issued by the data owner, i.e., the organization. The inclusion of these attributes in access structure makes the system reliable, flexible, and scalable. Due to hierarchical and distributed multi-authority system, the key issuance process becomes easy and scalable. New attribute authorities can be added to the system through a simple attribute delegation mechanism. This kind of open-access system can be used in any realistic application such as PHR data sharing, smart grid applications, IoT based applications, etc.

References

1. S. De Capitani di Vimercati, S. Foresti, S. Jajodia, P. Samarati, Access control policies and languages. Int. J. Comput. Sci. Eng. **3**(2), 94–102 (2007)
2. N. Dagdee, R. Vijaywargiya, Policy architecture for credential based access control in open access environment. J. Inf. Assur. Secur. **6**, 039–047 (2011)
3. J. Bethencourt, A. Sahai, B. Waters, Ciphertext-policy attribute-based encryption, in *IEEE Symposium on Security and Privacy* (2007), pp. 321–334

4. V. Goyal, O. Pandey, A. Sahai, B. Waters, Attribute-based encryption for fine-grained access control of encrypted data, in *13th ACM Conference on Computer and Communications Security* (2006), pp. 89–98
5. L. Ibraimi, M. Petkovic, S. Nikova, P. Hartel, W. Jonker, Mediated ciphertext-policy attribute-based encryption and its application, in *International Workshop on Information Security Applications* (2009), pp. 309–323
6. S. Yu, C. Wang, K. Ren, W. Lou, Attribute based data sharing with attribute revocation, in *5th ACM Symposium on Information, Computer and Communications Security* (2010), pp. 261–270
7. B. Waters, Ciphertext-policy attribute-based encryption: an expressive, efficient, and provably secure realization, in *International Workshop on Public Key Cryptography* (2011), pp. 53–70
8. M. Chase, Multi-authority attribute based encryption, in *Theory of Cryptography Conference* (Springer, Berlin, Heidelberg, 2007), pp. 515–534
9. M. Chase, S.S. Chow, Improving privacy and security in multi-authority attribute-based encryption, in *Proceedings of the 16th ACM Conference on Computer and Communications Security* (2009), pp. 121–130
10. A. Lewko, B. Waters, Decentralizing attribute-based encryption, in *Advances in Cryptology—EUROCRYPT 2011*, vol. 6632, Lecture Notes in Computer Science, ed. by K.G. Paterson (Springer, Berlin, Heidelberg, 2011), pp. 568–588
11. A. Sahai, B. Waters, Fuzzy identity-based encryption, in *Annual International Conference on the Theory and Applications of Cryptographic Techniques*, vol. 22 (Springer, Berlin, Heidelberg, 2005), pp. 457–473
12. M. Li, S. Yu, Y. Zheng, K. Ren, W. Lou, Scalable and secure sharing of personal health records in cloud computing using attribute-based encryption. IEEE Trans. Parallel Distrib. Syst. **24**(1), 131–143 (2013)
13. G. Wang, Q. Liu, J. Wu, Hierarchical attribute-based encryption for fine-grained access control in cloud storage services, in *Proceedings of the 17th ACM Conference on Computer and Communications Security* (2010), pp. 735–737
14. Z. Wan, J.E. Liu, R.H. Deng, HASBE: a hierarchical attribute-based solution for flexible and scalable access control in cloud computing. IEEE Trans. Inf. Forensics Secur. **7**(2), 743–754 (2012)
15. R. Bobba, H. Khurana, M. Prabhakaran, Attribute-sets: a practically motivated enhancement to attribute-based encryption, in *European Symposium on Research in Computer Security* (Springer, Berlin, Heidelberg, 2009), pp. 587–604
16. Q. Huang, Y. Yang, M. Shen, Secure and efficient data collaboration with hierarchical attribute-based encryption in cloud computing. Futur. Gener. Comput. Syst. **72**, 239–249 (2017)
17. E. Luo, Q. Liu, G. Wang, Hierarchical multi-authority and attribute-based encryption friend discovery scheme in mobile social networks. IEEE Commun. Lett. **20**(9), 1772–1775 (2016)
18. M. Horvath, Private key delegation in attribute-based encryption, in *Mesterproba Conference of the Budapest University of Technology and Economics for Graduating MSc and First Year PhD Students* (2015)
19. L. Ibraimi, Q. Tang, P. Hartel, W. Jonker, Efficient and provable secure ciphertext policy attribute-based encryption schemes, in *International Conference on Information Security Practice and Experience* (Springer, Berlin, Heidelberg, 2009), pp. 1–12

A Unified Platform for Crisis Mapping Using Web Enabled Crowdsourcing Powered by Knowledge Management



A. Vijaya Krishna, Somula Ramasubbareddy and K. Govinda

1 Introduction

Knowledge Management (KM) is a powerful strategic tool designed to capture, disseminate, and exploit knowledge within an organization or enterprise. Lacking business skills and organizational funds to grow, strengthen, and sustain their operations, many non-profit organizations struggle to maximize their impact [1]. In order to bridge these gaps, we intend to launch this Unified problem solving platform (UPSP) powered by crowdsourcing and knowledge management to collect real time crisis data. This platform helps to transform the non-profit sector and to build efficient sustainable organizations.

Crowdsourcing is a coinage for the act of taking a task traditionally performed by an employee or contractor and outsourcing it to an undefined, generally large group of people or community in the form of an open call. Crisis mapping initiatives and early warning systems hold promise for providing humanitarian actors with much-needed tools for conflict prevention, mitigation, and response. The field of dynamic crisis mapping "NeoGeography" which is essentially about "people using and creating their own maps, on their own terms and by combining elements of an existing toolset", to generate a meaningful view.

Crisis mapping initiatives, information communication technology, and early warning systems have the potential to prevent mass atrocities. Knowledge management is a powerful strategic tool designed to capture, propagate, and exploit knowledge within an organization or enterprise. Knowledge Management (KM) aids to develop timely policy to classify the crisis data collected from crowd. KM facilitates

A. V. Krishna · K. Govinda
SCOPE, VIT University, Vellore, Tamil Nadu, India

S. Ramasubbareddy (✉)
Information Technology, VNRVJIET, Hyderabad, Telangana, India
e-mail: svramasubbareddy1219@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_19

the organizations and volunteers to make effective use of the crisis data with real time organization. We are proposing a Unified Problem Solving (UPS) Platform to integrate all organizations and help them collaborate at a single place where reports are collected and refined.

2 Crisis Mapping

2.1 GIS to Neogeography

Using mainstream technologies, we can develop tools to encourage innovative practices to help people in various ways. The arena of dynamic crisis mapping is new and rapidly changing. The core drivers of this change are dynamic mapping tools, mobile data collection tools progress of new methodologies. Some experts at the cutting edge of this change call the results "NeoGeography," which is fundamentally about "people using and creating their own maps, on their own terms and by combining elements of an existing toolset." The revolution in applications for user-generated content and mobile technology delivers the basis for widely distributed information collection and crowdsourcing a term coined by Wired less than three years ago. The exceptional rise in citizen journalism is stark evidence of these revolutionary new methodologies for conflict trends analysis increasingly take spatial and/or inter-annual dynamics into account and thereby reveal conflict patterns that otherwise persist hidden when using traditional methodologies [2]. Until recently, traditional mapping tools were expensive and highly technical that required extensive training to produce static maps. Introduction of products like Google Earth and Virtual Earth powered and strengthened the notion of NeoGeography.

2.2 Crisis Mapping Approach

Real time data collected through crowdsourcing helps to mitigate the effects of natural disasters, reinforce international aid agency coordination, improve resource allocation, and develop timely policy. Crisis mapping exploration agenda can be categorized into the following three areas: Crisis Map Sourcing, Mobile Crisis Mapping, and Crisis Mapping Analytics.

Crisis Map Sourcing (CMS) seeks to further research on the challenge of envisioning disparate sets of data ranging from structural and dynamic data to automated and mobile crisis mapping data. The challenge of CMS is to develop proper methods and best practices for mashing data from Automated Crisis Mapping (ACM) tools.

Mobile Crisis Mapping platforms to add value to Crisis Mapping Analytics. The purpose of setting an applied research agenda for Mobile Crisis Mapping, or MCM, is to recognize that the future of distributed information collection and crowdsourcing

will be increasingly driven by mobile technologies and new info ecosystems. This presents the crisis mapping community with a host of pressing challenges ranging from data validation and manipulation to data security. If the persistent problem of data quality is not adequately resolved, then policymakers may question the reliability of crisis mapping for conflict prevention, rapid response, and [3] the documentation of human rights violations. Worse still, inaccurate data may put lives at risk. In other words, new and informative metrics are needed to be technologically advanced to identify conflicts in real time [2].

Crisis Mapping Analytics (CMA) is becoming increasingly important given the unprecedented volume of georeferenced data that is rapidly becoming available. Existing academic platforms like Warviews and operational MCM platforms like Ushahidi do not include features that allow practitioners, scholars, and the public to query the data and to visually analyse and identify the underlying spatial dynamics of the conflict and human rights data. This is largely true for Automated Crisis Mapping (ACM) tools as well. In addition, existing techniques from spatial econometrics need to be rendered more accessible to non-statisticians and built into existing dynamic crisis mapping platforms [4, 5].

3 Unified Problem Solving Platform (UPSP)

The major cause for failure of the system in survival is the existence of a functioning gap between the organizations despite their commitment to exceptional mission-driven work. Media is the major source for raising problems to the fore. Our proposal is to bridge these gaps by introducing a novel problem solving platform which unite all categories of people steered by a knowledge management (KM) system which dynamically collaborates NGOs for greater impacts, and to deal with the live problems in the society. This platform gives them an opportunity to illustrate their teething troubles to a wider audience who are ready to help (Fig. 1).



Fig. 1 Graph displaying current situation

3.1 Key Challenges

Bringing all kinds of NGOs and volunteers on a single platform. Easily deployable crowdsourcing and crisis mapping system which collects real time data using citizen journalism. Knowledge Management system recognizes the like minds and forms dynamic groups based on their interest to support the current crisis. Validating the limited report data collected through mobiles from diverged locations.

4 Architecture

Crowd source data collection is the main theme of the proposed platform. The platform uses the capabilities of the Google Maps API's to pinpoint the location of need of help and will be overlaid over the Ushahidi engine which was integrated with crowd source data collection and Knowledge management Engine. The platform uses a swift based web framework which makes substantial modules [6] flexible. It uses Model View Controller (MVC) architectural pattern. It aims to be secure, lightweight, and easy to use. The input data collected through various means is imperilled to categorization by various constraints. Category filters are shrouded on top of the spatial filters and this task is envisioned to be automated by knowledge classification module of KM system. Our KM system includes different processes like knowledge creation, knowledge classification, knowledge retrieval, knowledge sharing, and knowledge reuse. A successful KM system should support all these five processes (Fig. 2).

Models are used to represent a specific piece of data, or data chunks together, or an essential component which is an integral part of the system. Views are used as data to form rendering layers. Controllers are used as the "entry point". They also direct and control the process flow of the application, and handle how a URL is converted into an application function. Libraries are used as tools that operate on some form of pre-existing data, either in the form of an array (e.g., Session, Validation, Input) or some other data structure, such as ORM (database table) or Archive (file system). Helpers are used for simple, repetitive tasks, such as creating HTML tags, making a URI into a URL, or validating an email address. The file system is made up of a single directory structure that is mirrored in all directories along what we call the include path.

4.1 Knowledge Processes

In the knowledge management system, the knowledge is briefly represented by the SAO assembly. Where S stands for Subject, A stands for Action, and O stands for Object. Generally, the SAO structure can express a piece of knowledge. In this picture

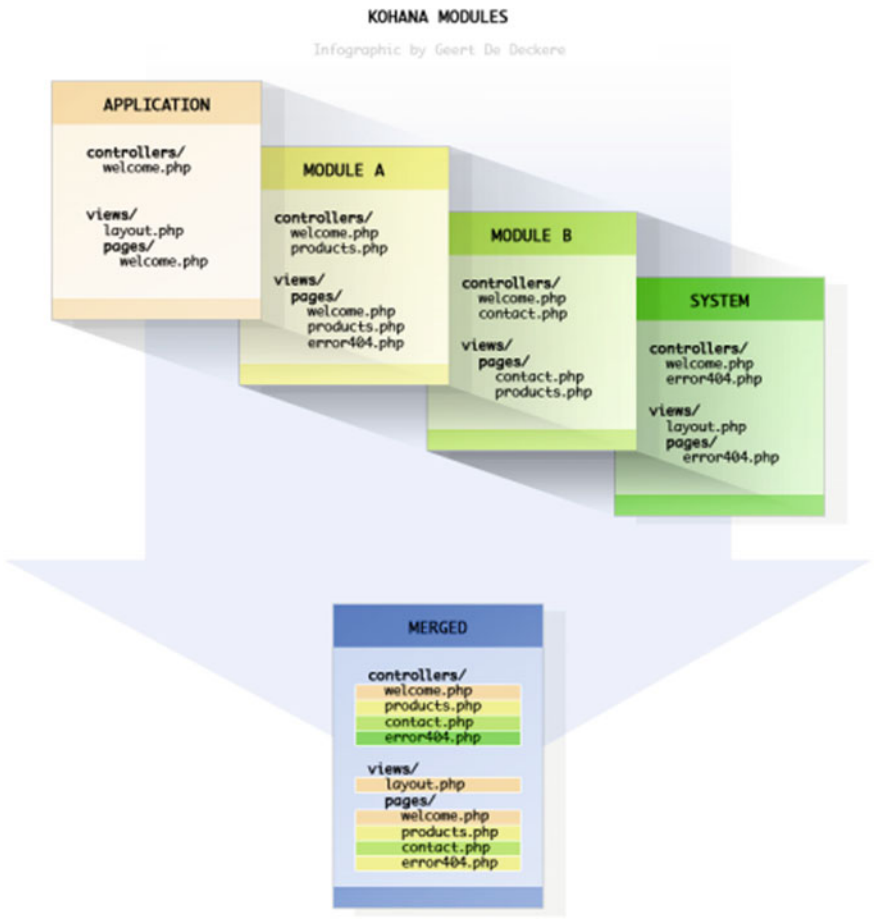


Fig. 2 Kohana architecture

these knowledge objects are filled by data from various mainstream data sources like mobile, email, forums, and base reports [7].

The knowledge extraction process involves capture of metadata and summary content of various crisis data items. We propose a form-based interface with various pre categorized filters to classify the crisis data thus making it easy for extraction. The standard approach to classification of knowledge improvement is to divide the domain area into classes of objects with shared properties. Then, identify the special classes which would have their own properties apart from inheriting the general properties. This is followed by populating concepts, relations, and attributes for each knowledge item in the knowledge base. Domain expert's input for classification and explanation of topics in a given domain is necessary at this phase.

4.2 Upsp Functioning

The problems which are highlighted come to notice of most of us through media or other organizations. In an ideal scenario, there might be others who need help, whose problem didn't rise to fore. The platform uses various API's to mine the problem data. Collects real time crisis data through citizen journalism via Forums, Chat Clients, Emails, Social Networks, Mobile Messaging, and other Crisis data collected through various means are mapped on to our platform using Google API's. The concept of overlays has been used to visualize the mapped reports and filter the domain-specific reports by KM system.

4.3 Upsp Layout

The main layout consists of a briefing of our mission integrated with all the basic features of the platform like report data overlays on the map API and timestamp overlays integrated with MIT's smile API so that a new user can know everything about our platform with a single click and advanced overlays. All the associated NGOs are streamed through a feed on the bottom left of the page. The statistics widget is a data visualization module to reveal the no of organizations involved, no of volunteers involved, no of problems reported, and no of the locations covered. A new user needs a privileged user account to post or report anything. So if it is the case of an organization, a new organization can fill in their basic details in the link us form with which we'll verify them and approach them later. Organizations list feed is streamed to all the new volunteers and in addition to this all recent reports from various social networks and other sources are also streamed in the mainline. Jquery is used wherever necessary to provide smooth transitions in the interfaces.

5 Filters and Classifications

A category filter has been implemented to help the KM system while retrieving the context-based report data. The main module where all the registered volunteers and NGOs browse the report data collected from various locations. We broadly divided the problem domain into multiple categories such as civilians, riots, deaths; property loss due to natural disasters, and government forces, etc., A new category will be added based on the user requests for a particular cause (Fig. 3).

All the categories are represented with a unique colour and the same overlays are replicated on the maps. The incidents reported are represented as a bubble that will grow its size based on the intensity of the problem. Volunteers can easily analyse the reports from the real time timeline below. The KM system implemented here is capable of updating the mainstream visualizations in run time. The custom analysis

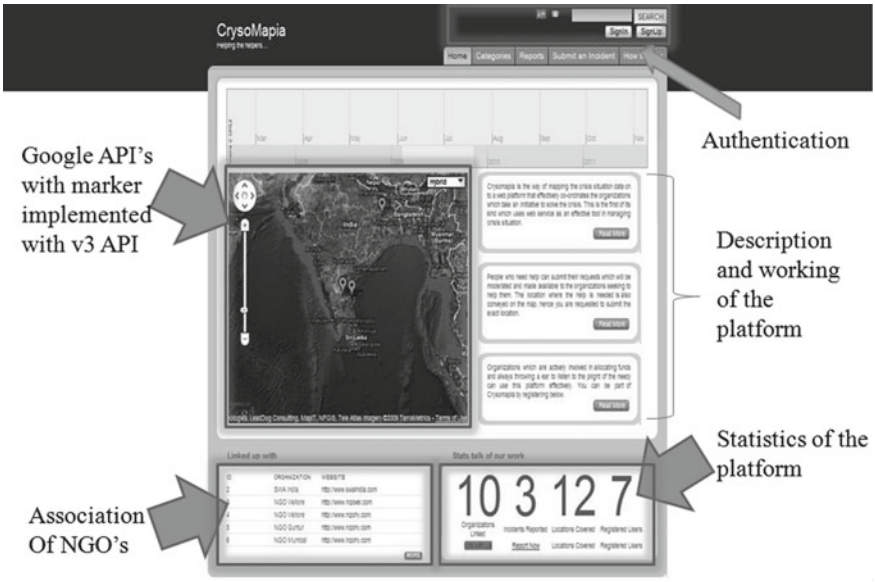


Fig. 3 Interface for the platform

is also accessible subjected to user willingness with a simple slider provided for ease of access. This plots a graph of the intensity of the reports over a timeline. NGOs can easily identify the problems or area which requires immediate attention (Fig. 4).

A layer overlay differentiates pictures, videos, and news articles associated with the specific incident [8]. This helps the volunteers to get a deeper understanding of the problem before getting into it. Maps instance is created using kohana API framework [6]. On hover, functions are enabled to zoom in and out of the locations easily. Hybrid overlays are provided to view the location in real time. Category click event fetches the reports associated with that category and the corresponding ones are displayed on the map. This category filter is very much handy for certain types

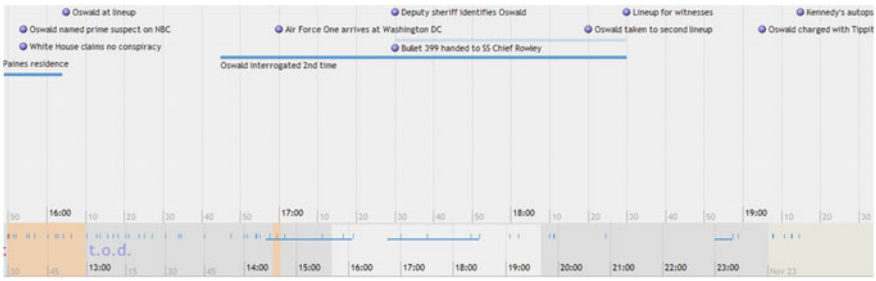


Fig. 4 Dynamic timeline to render runtime visualizations fetched from KM system

of NGOs and volunteers as they can access reports according to their interest in real time (Fig. 5).

An Incident Submit Module is one of the most important module as it's the key to fetch most important information of the system i.e., reports, so it should be as much user friendly as possible and at the same time should boast of all the required parameters (Fig. 6).

Report Data Feeding the KM System: This is one of the most important modules as it's the key to fetch most important information of the system, i.e., reports, and

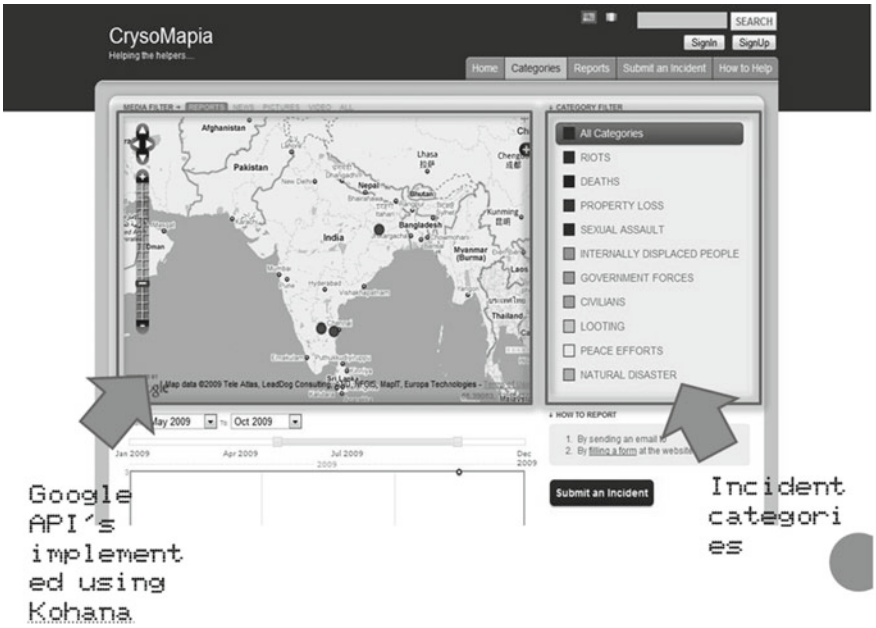


Fig. 5 Category reports

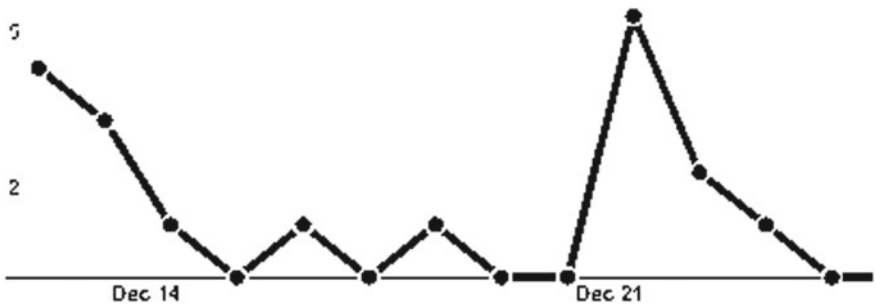


Fig. 6 Report data over a period of time depicting the spike in the problem data

should boast all the required parameters [9]. We categorized this report info into four types to mine it in the later stages.

5.1 Basic Information

We planned to collect some basic necessary information for every report without which SOA structure could not be formed, such as report title and a brief description in addition to timestamps. An aptana API is used to maintain the timestamps of each and every event regarding a report. Timestamps we collect will help the KM system place the reports on the timeline easily and accurately and helps volunteers to keep track of the problem from their mobile devices. ETS of the problem will attract the attention of the people easily if it's projected properly in the timeline. Category check will note the context of the report to project proper overlays and to divide the reports.

5.2 Location Information

Location info is a key to the report to assist local NGOs and volunteers. Built-in maps are provided by Google Maps API [10] with search tab which fetches the location nearest to the query origin available. This search feature helps the users to pinpoint the location with more accuracy. All lat long values are maintained separately for strategic advantage.

5.3 Media Information

All the media related info like videos and news articles associated with the report will help the volunteers and NGOs to get deeper insights about the problem. KM system facilitates users to upload the images related to the report supporting their cause. All the extra media related to the content will attract more NGOs and volunteers to support the report (Fig. 7).

5.4 Personal Info

This can keep track of various reports submitted and location they are interested. This info can be used later to assist volunteers and NGOs in collaborating with each other. Clustered users driven approach is used to authenticate mobile reports. More the support, more the trust.

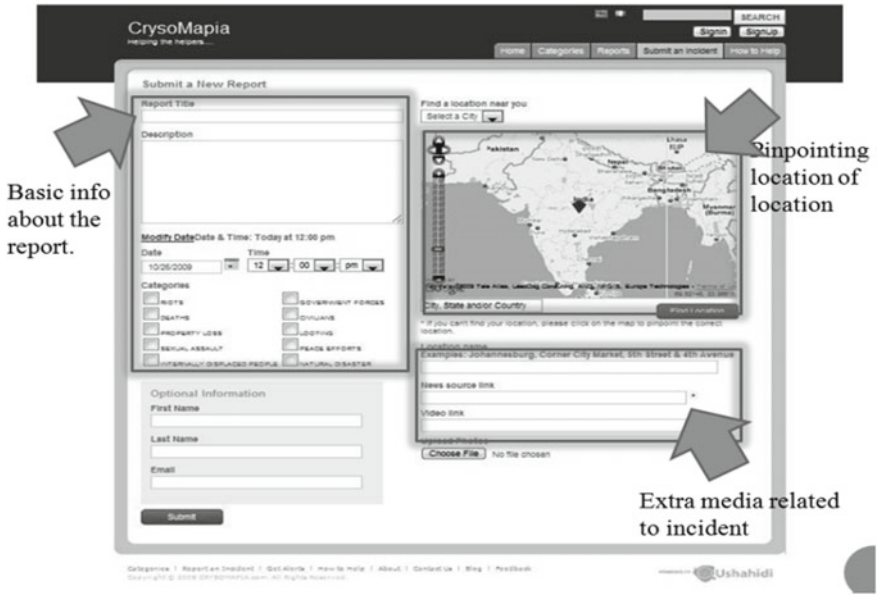


Fig. 7 Reports submission

6 Conclusion and Future Work

Crisis mapping using web enabled crowdsourcing powered by knowledge management system is proposed in this paper. The unified approach of crisis data management is discussed which intends to fill the void between dynamic crisis adaptability and operational gap as discussed in this paper. The knowledge management system periodically analyses the operational effectiveness of the groups and adapts the changes in the later groups to be formed. Real time collaboration ensures more participation and supervises problem eradication manifesto. The unified approach facilitates to extract the crisis data more easily like never before. Once the platform is launched training, the knowledge management system based on the enormous input data, Prediction of the crisis caused by chain reaction of events prior, and alerting accordingly remains as the topics for further research.

References

1. Crowd sourcing, attention and productivity, B.A. Huberman, Social Computing Lab, HP Laboratories, Palo Alto, D.M. Romero, Center for Applied Mathematics, Cornell University, Ithaca
2. HHI Harvard Humanitarian Initiative objectives. Jennifer Leaning, MD, SMH
3. A.K. Henrikson, P. Meier, On scale and complexity in conflict analysis, Jennifer Leaning Applied Professor

4. D. Fensel, *A Silver Bullet for knowledge Management and Electronic Commerce* (Springer, Berlin, Heidelberg, 2001)
5. http://irevolution.files.wordpress.com/2009/03/meier2006-scale-complexity-conflict.pdf
6. Helping the helpers by Wayne Mathews 6/00 JMG. NC State University
7. R. Richardson, A.F. Smeaton, Using WordNet in a knowledge-based approach to information retrieval
8. P. Meier, Jennifer, Applied Technology to Crisis Mapping and Early Warning in Humanitarian Settings (2009)
9. Ushahidi crowd sourcing platform architecture and APIs
10. http://code.google.com/apis/maps/documentation/reference.html

Web Image Authentication Using Embedding Invisible Watermarking



T. Aditya Sai Srinivas, Somula Ramasubbareddy, K. Govinda and S. S. Manivannan

1 Introduction

Hiding the handwritten signature with the web image will be merged to form a combined image with a secret key. In this work, we provide the security based algorithm which is defined as transformation of values using Fourier transformation method and changing the position of the pixels of the image by using permutations called as p-box technique, and at last by using cryptography-based algorithm like chaos which uses the sequence of pseudo-random numbers this three approaches on source and destination provide two secret keys by performing XOR operation to maintain image hiding. This algorithm performs better results than public key encryption and private key encryption algorithms like RSA, AES, 3DES, and IDEA and encryption for eclipse curve based (ECC). Handwritten signature can be embedded with the host image to form a combined/composite image. This embedded image can be encrypted and decrypted by using cubic based algorithm which is a highly secured algorithm. The embedded information should be in digitalized signal manner. This algorithm will provide the protection of data confidentiality over unauthorized networks. In this method, the image hiding and compression can be achieved. The system images can lose visualization when it is embedded to form combined image that is bytes can be reduced in the system image. In this hiding the image in another image can be done and the quality of the image can be recovered accurately. This algorithm can provide perfect ability to hide the image and it also resists from attacks like differential, statistical, and exhaustive. In the previously proposed methods, the image can be hidden and will provide authentication and protection of ownership.

T. A. S. Srinivas · K. Govinda · S. S. Manivannan
SCOPE, VIT University, Vellore, Tamil Nadu, India

S. Ramasubbareddy (✉)
Information Technology, VNRVJIET, Hyderabad, Telangana, India
e-mail: svramasubbareddy1219@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_20

2 Literature Survey

In the digital data to maintain security is an important issue from the third party systems. So for this purpose hiding the data in the image format various techniques are evolved. According to Poulamidas et al. [1] proposed that handwritten signature is embedded with the color images at both receiver and sender end for extracting the image from color images at destination. It will achieve the robustness of detection of copy with respect to the complexity, mark the integrity issue for the mark partial removal and verification of ownership is the weakness. Bandyopadhyay et al. [2] proposed that new method to embed the data in the form of binary image. This image includes figures, signature, and text scanned. It manipulates the pixels which are in flappable nature in order to embed the specific amount of digital data. Shuffling can be applied to the image before embedding the image due to maintaining uncertain capacity of embedding image. It also provides the digital signatures to protect from unauthorized parities and also provides authentication digital documents. Wu and Liu [3] proposed that in the automatic surveillance video (SV) systems, the concept is that authenticating the content of video is the primary task. In this digital videos and images cannot be directly as an authentication. So consider the sequence of images by using watermarking based algorithm to provide the authentication for the digital videos and images as this algorithm can test the data in authenticated manner. Bartolini et al. [4] proposed that many algorithms for the watermarking of digital data. But in this paper the watermarking is used further for chart image and map classes. These images are in the format of homogenous structure of pixels and mapped color and also in binary format. The above characteristics are used to design the watermarking for continuous tone photographic images. This algorithm divides the image into homogenous regions and adds the watermarking signals to the pixel locations on boundary of various regions. The presence of signals in watermarking image determines on the basis of correlated detector. It will also find the synchronization errors in shuffling pixels of the image and cropping the image in the presence of noisy data. Masry [5] proposed that it is a basic approach for embedding the image. In this the algorithm follows three main steps firstly, it uses the mask sobel filters, secondly, least significant bit (LSB) for each pixel in the image, and finallythe information can be hidden using fuzzy logic approach and ASCII code which is called the connectivity of gray levels. The three images can be hidden as a single image as authenticating and identifying the text, embedding the data hiding the information in digitalized images. It also provided the compression of the data and reserving the space for memory. Alwan et al. [6] digitalized watermarking is a newly emerged technique for hiding the image. The image can be modified and altered for the watermarking image which can be detected in the fragile watermarking system. Lim et al. [7] proposed that it is channel and signal capacity based mechanism for the images. It will add the additional data which is essential for data distortion or to remove noise in the signals and also it will provide the highly flexible data rates for audio signals up to 4–6 bps.

Bhattacharyya et al. [8] proposed that data hiding is the method which consists of a sequence of bits that can be embedded with the host image by the technique called deterioration with a small visualization it means that extracting the image afterward. Tsai et al. [9] proposed that it will hold larger volumes of data with low degradation perceptual. The data hiding can overcome security-based attacks over a reliable data for the compression and resizing of image. Johnson and Jajodia [10] proposed that the data hiding can be done using the "RSA algorithm" which provide key generated algorithm for encryption and decryption mechanism in image hiding at sender and receiver.

3 Proposed Method

In this project, we propose a technique for embedding watermark bicolor image into a color image. At the source, the handwritten signature image (bicolor image) is encoded at the end of the color image. Double folded security of handwritten signature can be achieved (over the entrusted network) by accumulating the starting point of encoding the image data which is depended on the size of the images; secondly, the starting point (byte location) of bicolor image encoding in the color image is stored within a four-byte block in the encoded form. This four-byte location encoding is done by public key. At the target, firstly, the starting point of encoding bicolor image data (location) is decoded by private key and then, secondly, starts extracting the encoded bicolor image data from the color image. This technique requires knowledge of the color image for the recovery of the handwritten signature image. At the receiver, the algorithm reconstructs the original handwritten signature image. Embedding high volume of information into images without causing perceptual distortion that is quite challenging. Consider the problem of image in image hiding, in which an image, called the handwritten signature image is to be embedded into another image called the host image to get a composite image. It is a useful Hybrid data hiding technique resulting in invisible watermarking.

In this project, two algorithms for encoding and decoding purposes are used. They are

- Hybrid digital embedding Algorithm at source/client-side
- Handwritten Signature Extraction Algorithm from Hybrid Composite Image file at the target server.

3.1 The Discrete Cosine Transform (DCT)

The discrete cosine transform (DCT) used to split the image into parts of differing spectral sub-bands. "The DCT is similar to the discrete Fourier transform which

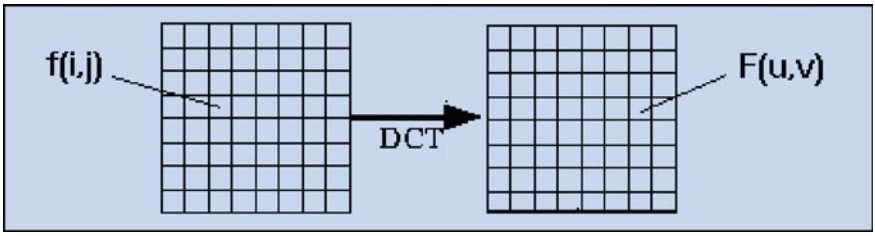


Fig. 1 DCT function

transforms a signal or image from the spatial domain to the frequency domain as shown in Fig. 1".

3.2 DCT Encoding

The standard equation for one-dimensional "DCT" is defined as Fig. 2.

And further, the relative inverse one-dimensional "DCT" transform function is simple "$F^{-1}(u, v)$" as per Fig. 3.

The standard equation for two-dimensional "DCT" is defined as Fig. 4.

And further, the relative inverse two-dimensional "DCT" transform function is simple "$F^{-1}(u, v)$" as per Fig. 5.

Hence, the eight point DCT be like Fig. 6.

In this work "DCT methodology" is used to drop unimportant pixel data that do not have any contribution in visibility of image.

The basic operation of the DCT is shown in Fig. 7.

$$F(u) = \left(\frac{2}{N}\right)^{\frac{1}{2}} \sum_{i=0}^{N-1} \Lambda(i).cos\left[\frac{\pi.u}{2.N}(2i+1)\right] f(i)$$

Fig. 2 Standard equation for DCT

$$\Lambda(i) = \begin{cases} \frac{1}{\sqrt{2}} & \text{for} \xi = 0 \\ 1 & \text{otherwise} \end{cases}$$

Fig. 3 "DCT" transform function

$$F(u,v) = \left(\frac{2}{N}\right)^{\frac{1}{2}} \left(\frac{2}{M}\right)^{\frac{1}{2}} \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} \Lambda(i).\Lambda(j).cos\left[\frac{\pi.u}{2.N}(2i+1)\right] cos\left[\frac{\pi.v}{2.M}(2j+1)\right].f(i,j)$$

Fig. 4 2-D standard "DCT" function

$$\Lambda(\xi) = \begin{cases} \frac{1}{\sqrt{2}} & \text{for} \xi = 0 \\ 1 & \text{otherwise} \end{cases}$$

Fig. 5 Inverse transform "DCT" function

$$\Lambda(\xi) = \begin{cases} \frac{1}{\sqrt{2}} & \text{for} \xi = 0 \\ 1 & \text{otherwise} \end{cases}$$

Fig. 6 8 point "DCT" function

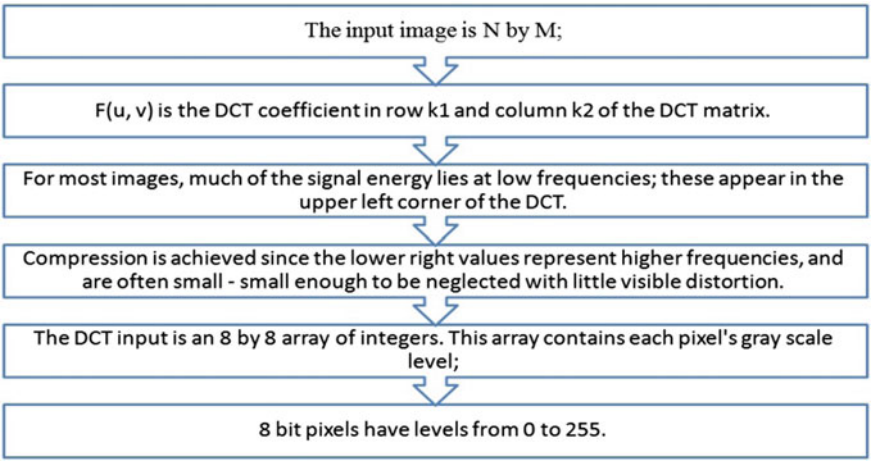


Fig. 7 DCT operation

4 System Design

Figure 8 describes the architectural design of whole system in two different domains like client environment or user-end and another one is authenticating domain or server end. At the user-end user delivers the host image and handwritten signature is provided. This data flows through the network to the server end in order to carry out embedding process. At the server end server program receives the data that came over the network. The application server program which is ready with "Hybrid Digital

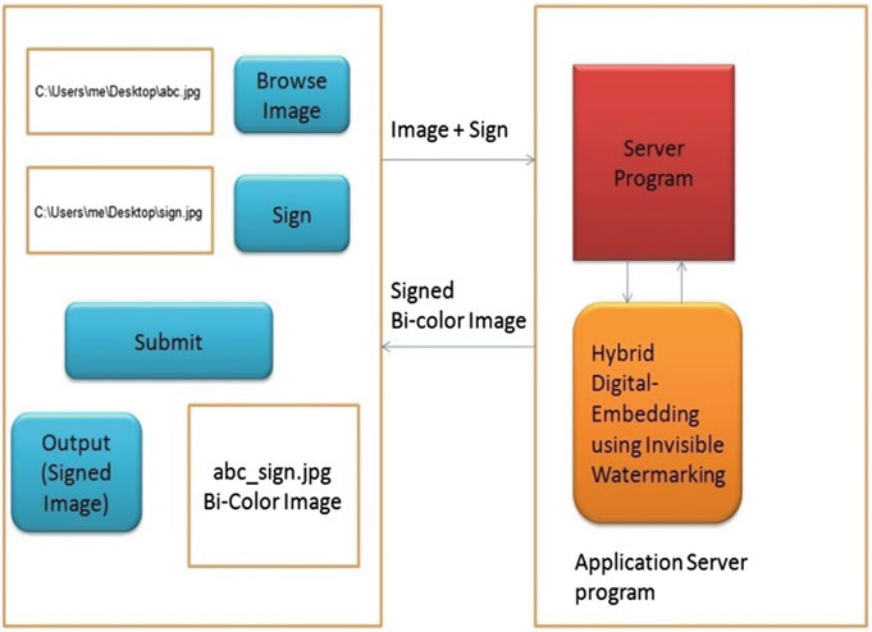


Fig. 8 System architecture of web image authentication

Embedding using Invisible Watermarking" algorithm performs the embedding of host image and handwritten signature and develops the hybrid image. The resultant hybrid image is then sent back to the user, and here onwards users can publish it over the web as copyright content.

5 Implementation

5.1 Embedding of Image

In this paper, we have focused mainly on handwritten image and system image. System image is maximum number of pixel based image and multicolored whereas handwritten image is bicolor. A system image can be given to the discrete cosine based transformation (DCT), This can discard the bits which will cause visual degradation to the system image and also provide the larger pixel based area. The system image can be embedded with the handwritten for security and authentication purpose.

The handwritten signature-based image can be embedded by using the handwritten signature inserted algorithm (HSIA). This HSIA can be used for inserting the images and it is also used for the invisible and visible water marking concepts in the invisible digitalized water marking image technique, used for the security of handwritten

signature-based image. HSIA verifies whether the system image is greater than the handwritten signature-based image and if it satisfies the above criteria the system image can be embedded with the handwritten signature image and can be obtained at one particular point with transformation (DCT).

The handwritten signature-based image and system image can be embedded with the HSIA by using the invisible digitalized water marking technique to obtain the composited image.

5.2 Encoding

The combined image can be encoded by getting the initial point of the embedded handwritten signature-based image which is in bicolor format in the system image. The transformation (DCT) mechanism can obtain the initial point in the combined image. To obtain the original system image inverse transformation can be done to the combined image with pixel.

The address values of the initial point can be obtained by using cryptography based algorithm like RSA algorithm. It is an integer based format to encode the image so it is difficult for the user to address values in the form of keys by using character-based format and it provides high security to encode and to get the initial point of the combined image.

The encoding mechanism converts the combined image into encrypted format as an image that cannot be extracted by any party and it can be done by only the authorized user and a key is generated for the encryption of image and is stored. The key which is generated can be verified at the time of decoding mechanism.

Finally, the image can be obtained as an ordinary image but as the handwritten signature-based image which has encryption initial point as shown in Fig. 9.

5.3 Decoding

The decoding mechanism can be done by decryption of the initial point of the handwritten signature-based image in the combined image. Firstly the character-based format key is entered to start the decryption mechanism. A new key is generated by the RSA cryptography algorithm for encryption image. Then the encryption image is subjected to the decryption mechanism. The key which is generated is verified prior. If the key value is matched then the initial point is decoded. If it is not matched with key-value then immediately it displays the error message to enter the correct key value.

It is very difficult for the unauthorized user to obtain the key values which are encryption of very larger integers and character-based key format. Verification can be done twice for decryption mechanism at the starting stage.

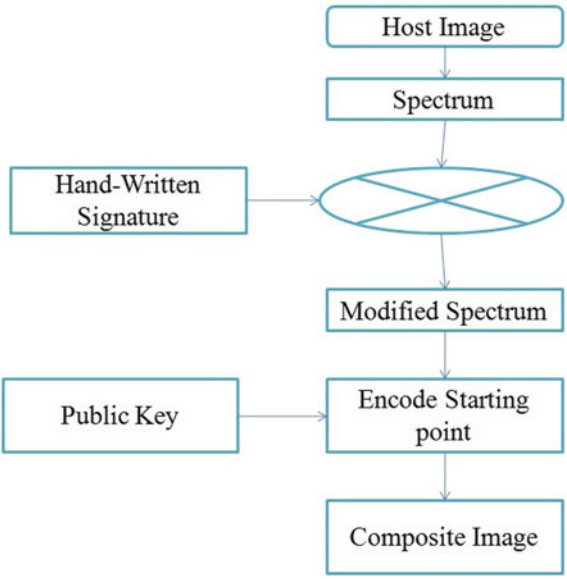


Fig. 9 Flowchart for encoding of image

5.4 Extraction of Image

The handwritten signature-based image can be extracted by using handwritten signature extraction algorithm (HSEA). It mainly used for extracting the images accurately and efficiently.

The handwritten signature can be retrieved from the combined image without degradation in the visibility of the image. The handwritten signature-based image which is obtained is quality based image and it is original. The extracted image can be used by authorized users.

Unauthorized users can obtain the extracted image which is highly confidential and authenticated. The system image and the handwritten signature-based image can be obtained with quality based pixels. Security is given as a higher priority to the system (Host) image in Fig. 10.

6 Result Analysis

In this section, different image formats are analyzed under the parameter of the resultant compressed size of image after encryption. This analysis has come with the interesting findings. We considered image formats like BMP, JPG, PNG, GIF, etc. And these image formats tested against different image sizes like 200 KB, 500 KB, 1000 KB, 1500 KB, and 2000 KB while testing the results we found that JPG format faced little difficulties in compression as shown in Fig. 9. For JPG image of 200 KB

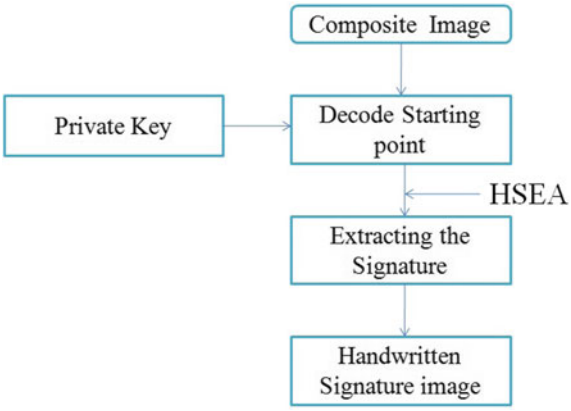


Fig. 10 Flowchart for decoding of image

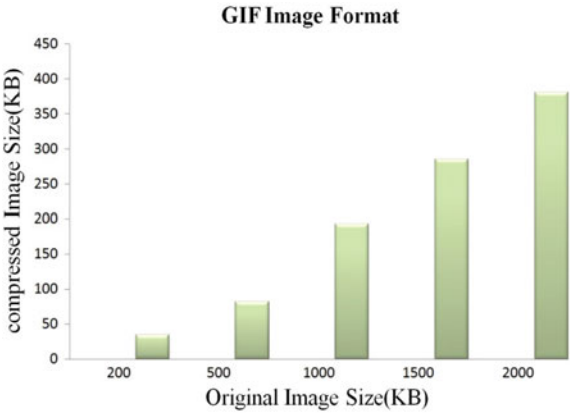


Fig. 11 Compression ratio of BMP image format

size, it started very slow with 9% as depicted in Fig. 11 and further it raised up to 70% but not increased than that. BMP image format results are well accepted for every size of image format as shown in Figs. 12 and 13 with constant average compression rate above 90%. GIF format enables the stable average 80% compression ratio on every possible size as shown in Fig. 14. The compression rate for PNG image format lies between 80 and 90% as depicted in Figs. 15 and 16 (Table 1).

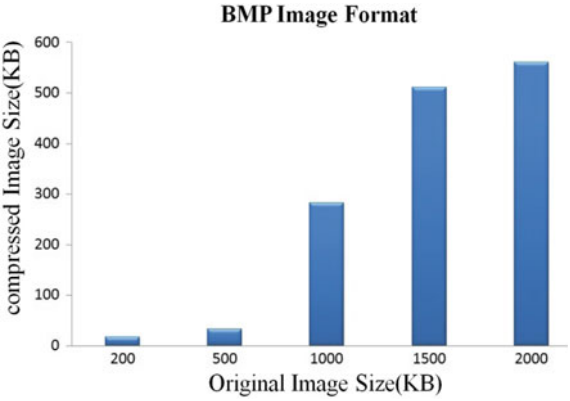


Fig. 12 Compression ratio of GIF image format

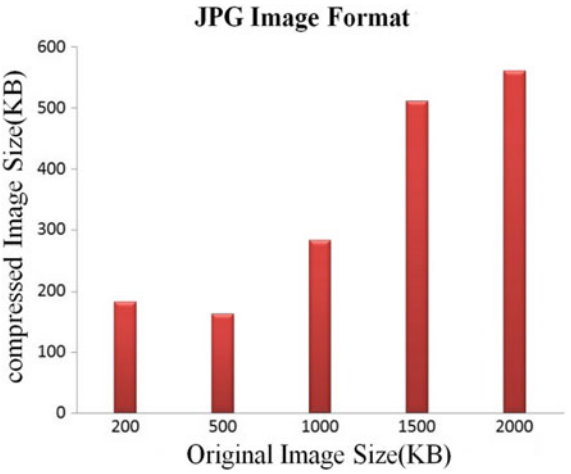


Fig. 13 Compression ratio of JPG image format

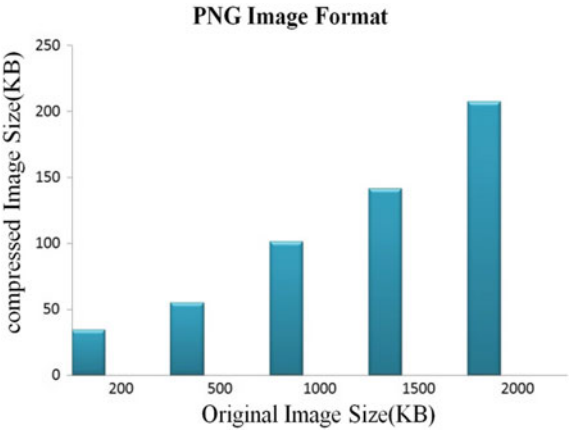


Fig. 14 Compression ratio of PNG image format

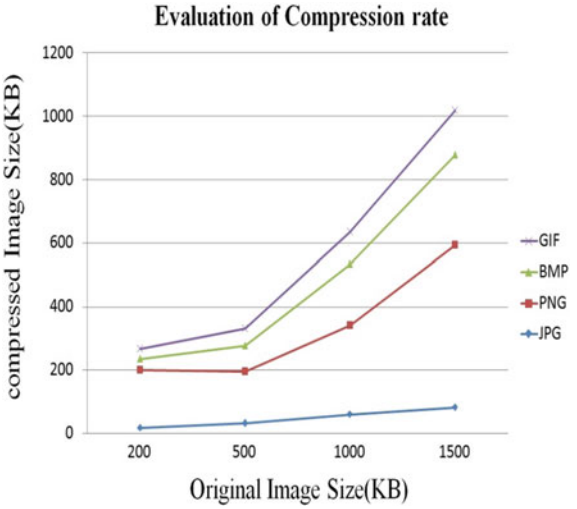


Fig. 15 Evaluation compression rate of image formats

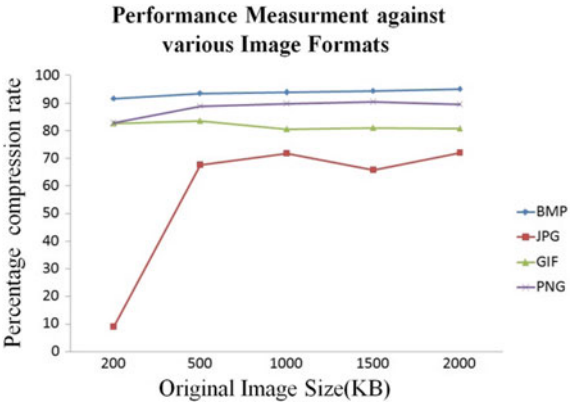


Fig. 16 Performance analysis of BMP, JPG, GIF, PNG formats in terms of percentage compression rate

Table 1 Compression of image formats (Size in KB)

Type/size	200	500	1000	1500	2000
BMP	16.7	32.1	58.6	81.7	96.5
JPG	182	162	282	511	560
GIF	34.4	82	193	285	381
PNG	34	54.6	101	141	207

7 Conclusion

This work provides web image authentication, integration, and increased bandwidth efficiency through encoding and image compression techniques. The visual quality of the handwritten signature image is maintained after the extraction has been done. The key generation is used, which increases the security of the bicolor image and provides robustness. This project developed undistracted encryption to gain the same quality measures before and after the coding along with resultant lower space of the image, this leads an absolute methodology to protect various image formats as well as various size image objects. The resultant hybrid image provides the user the copyrighted or protected content which is safe to publish on the web. This technique enhances the two major parameters, the first one is the sustained image quality after encoding/embedding and the second one is compression ratio which is the main focus of all previous efforts taken before in this field. In this work, compression ratio has increased in tremendous way. Further, invisible digital watermarking has proved it's significan that the extracted image at the target shows excellent visual quality. The idea explored in this literature can be used to implement video compression as well as video decoding to enhance video authentication. This concept helps to reduce video size which will eventually increase video streaming on the real time network.

References

1. Hybrid digital embedding using invisible watermarking. IEEE (2008)
2. S.K. Bandyopadhyay, D. Bhattacharyya, A.J. Pal, Secure delivery of handwritten signature. ACM Ubiquity **7**(40) (2006)
3. M. Wu, B. Liu, Data hiding in binary image for authentication and annotation. IEEE Trans. Image Process. **12**, 696–705 (2003)
4. F. Bartolini, A. Tefas, M. Barni, I. Pitas, Image authentication techniques for surveillance applications. IEEE Proc. **89**(10) (2001)
5. M.A. Masry, A watermarking algorithm for map and chart images, in *Proceedings of the SPIE Conference on Security, Steganography and Watermarking of Multimedia Contents VII* (2005)
6. R.H. Alwan, F.J. Kadhim, A.T. Al-Taani, Data embedding based on better use of bits in image pixels. Int. J. Signal Process. **2**(2), (2005). ISSN 1304-4494
7. Y. Lim, C. Xu, D.D. Feng, Web based image authentication using invisible fragile watermark, in *Pan-Sydney Area Workshop on Visual Information Processing (VIP 2001)* (Sydney, Australia, 2001)
8. D. Bhattacharyya, D. Choudhury, S.K. Bandyopadhyay, Bi-color nonlinear data embedding and extraction of handwritten signature, in *IEEE Electro Information Technology Conference, EIT-2007* (Illinois Institute of Technology, Marriott O'Hare Chicago, Illinois, U.S.A., May 17–20, 2007)
9. C.-L. Tsai, K.-C. Fan, C.-D. Chung, T.C. Chuang, Reversible and lossless data hiding with application in digital library, ICME '04, in *2004 IEEE International Conference* (11–14 Oct. 2004), pp. 226–232
10. N.F. Johnson, S. Jajodia, Exploring steganography: seeing the unseen. IEEE Comput. **31**(2), 26–34 (1998)

Frequent Item Set, Sequential Pattern Mining and Sequence Prediction: Structures and Algorithms



Soumonos Mukherjee and R. Rajkumar

1 Introduction

Evolutionary research on the field of data mining by using the statistical quantitative approach in the historical massive databases is the most viable practice in the field of data science research on business analytics, healthcare analytics, market trend prediction and so on. The market basket analysis is a promising analytical approach to determine affinity of customers towards the product line, customer behaviour segmentation and making essential business decisions for production and future strategies. This is undergone along with analysing the frequent patterns of data and falls under the frequent item set mining subdomain. Similarly the pattern mining is actually referred to a large class of data mining algorithms for determining frequent item set, subgraph, frequent episodes and sequential pattern mining. Pattern mining and prediction is thus the largest area of data mining and warehousing domain research.

Our paper provides a survey on various pattern mining and item set mining algorithms and draw comparative classification among the pioneer and recent developmental approaches. For structured documentation, we have segmented our paper into three subparts.

S. Mukherjee (✉)
Data Science and Analytics, EPITA (École Pour l'Informatique et les Techniques Avancées), Paris, France
e-mail: soumonosvit@gmail.com

R. Rajkumar
School of Computer Science and Engineering, VIT, Vellore, India
e-mail: vitrajkumar@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_21

2 Related Techniques and Structures

2.1 Support

Support is the absolute measure of frequency in a dataset of sequence or item sets. Support implies the associativeness among two or more items or within a pattern. In the item set and pattern mining algorithms, system takes an input-minimum support as an user defined value. It evaluates the item sets and patterns with this value as a threshold. The min sup is, therefore, a deciding criterion for FIM and SPM algorithms analysis. As the min sup is decreased, the accuracy of mining is subjected to increase.

2.2 Confidence

In the item set or sequence set databases, confidence is the measurement of the frequency of items when correlation occurs. For example, to say for in a consumer database, if it is observed that 40% of purchases consisting milk also consisted bread, the correlated frequency of milk and bread is called confidence which is 40%. While working on transaction database, the algorithms take input as evaluation threshold, an user-defined value of min-confidence. The pattern has to have a confidence value not less than min-confidence to be frequent or interesting.

2.3 ID-List

This is an index data structure introduced by vertical algorithms like E-Clat for FIM and Spade for SPM. ID-list is actually the list against each individual item of a dataset which comprises of the index of the item positions of a pattern in the item set or the sequences. FIM algorithms denote this structure as TID-list and in case of SPM these are denoted as SID-list. The ID-list structure allows the algorithms to count the support of each pattern or item set after every iteration without scanning the database every time. The time complexity of ID-list is $O(N*sup(S)_{I\text{-}N})$ where N is the number of items and S is the sequences.

2.4 Bit-Vector

This is an encoding data structure for large ID-list optimization. The bit-vector is used as the ID-list of pattern S_N, say, ID-list(S_N) is transformed into B-list(S_N), so that the i-th bit represents the k-th item set of N-th sequence and i-th bit is set to 1 if the same item set of the sequence belongs to S_N and otherwise is set to 0.

2.5 Projected Database [1]

It is a summary data structure used as in memory virtual copy of databases. The structure stores all the item sets and their value in memory by scanning the database once and can append the new item sets on the previous structure. This is introduced to fight two major drawbacks: the cost of repeated scanning of database to count the support and the direct access of the database to the computation to prohibit non-existent patterns from being generated.

2.6 SAX

The symbolic aggregate approximation algorithm was proposed by Lin et al. [2]. This is a method proposed for converting the time series into sequence databases. The algorithm takes one or more time series as input. It can be floating point decimal numbers or a string, marked with separator, the number of segments and the number of symbols. The algorithm gives output as a piecewise approximate aggregation (PAA) where each segment is replaced as the average of its data point. The overall representation of output is a sequence database which can then be work for sequential pattern mining to be undergone.

2.7 Diff-Sets [3]

This is an enhancement data structure applied to improve the time efficiency in vertical algorithms as E-Clat in the replacement of ID-list structures. These node data structures keep track of only the differences of TIDs from the generated patterns. The memory cost of diff-sets is just a fraction of that of TID-list which drastically cuts down the size of memory for storing the whole database with TID-lists. The optimization gave birth to dE-Clat algorithm as an advancement to the traditional E-Clat.

2.8 Item Set Tree

It is a tree structure designed [4] to perform incremental data mining in an online and dynamic database setting for processing updates in database, handling queries, perform join and insertion operations and to process the input transactions. The insertion complexity is O(1) for each item and the complexity of support counting is O(N) where N is the cardinality of the item domains.

2.9 Compressible Prefix-Tree

This is typically a synopsis data structure for designing algorithms for FIM that can handle the stream of massive scale of data in online and updatable database structures. The CP-Tree represents an approximation where the nodes of the modified prefix-tree hold concise synopsis [5] of information by which the support of several item sets can be traced together. The structure can be upgraded and optimized to maximize the accuracy and the trace of the output can be linearly transformed to produce the actual support counts of the frequent item sets in the database. The size of the tree reduces with increasing number of supports counted by the same.

2.10 Compact Prediction Tree

The most advanced and accurate data structure for building a predictive model that can predict the future items of the sequence is a compressed prediction tree (CPT) [6] which is an integrated approach of a modified prefix-tree and a bit-vector list which is represented as inverted index list. One more element of the algorithm working with a CPT is a lookup table which is an array-based structure for performing the prediction of any future item set at a constant time constraint. The CPT has been proved to outnumber all the previous noble approach structures like dependency graph and other Markovian models in terms of accuracy of prediction.

3 Module-1: Frequent Item Set Mining

Keeping equivalence with the broader aim of data mining to vividly understand and analyse the historical data and predict the future, and with a view of finding frequently co-occurring item sets from massive scale of historical transaction databases the first approach was developed by Agarwal and Srikant who named it as a new category of data mining domain: Large item set mining. Although it has been considered as a frequent item set mining in the later development. The primary goal was to revolutionize business analytics with running algorithms on customer purchase and transactional databases to find the trend and analyse customer behaviour. This helped big business houses to reshape their production, assembly lines and future business strategy. The unknown and so far undiscovered association between various products helped different segments of market to co-promote products in a structured and successful way. On today's scenario frequent item set mining are the key part of data analytics in network traffic analysis, malware and fraud detection, bioinformatics and several more areas to find interaction, occurrence, relationships and inclinations between particular attributes of the research datasets. The paper introduces the key

concepts, problem types, tools and mechanisms and gradually developed algorithms to find frequently occurred item sets from databases.

3.1 Traditional Algorithms for FIM

A priori [7] was the first designed algorithm for FIM. It came up with a standard horizontal representation as input. The further developed algorithms considered the same input and output for working approach but, however, the required database representation, the data structures, statistical methods and paradigms changed with the up coming more efficient algorithms in later stages. The FIM is trivially an enumeration problem and can be fairly classified by using broadly two types of approach: Algorithms working on a breadth first search (BFS) and those on depth first search (DFS) approach.

A priori: The BFS horizontal algorithm, developed by Agarwal and Srikant takes an input of a horizontal transaction database with tuples and attributes, then performs a breadth first search in the dataset. During that process it takes the frequent item sets of (n-1) size, say In-1 and to generate substantial frequent item sets of length n, say In. It discovers all 1-item sets in the database and then it appends the items in all possible permutations to form two-item sets. Then it performs an analysis whether each element of the two-item sets are frequent. If any item comes out to be non-frequent, the algorithm discards the item set for violation of the downward closure property. The algorithm takes a user-defined min support as threshold input and checks whether the generated item sets have the support count more than that, those who have the support value not less than the min sup are considered as frequent. And considers all frequent two-item sets to generate three-item sets. The process iterates until there are no more item sets to be generated in the database.

Drawbacks: However, this approach has proved to stand with certain limitations. It iterates the process of appending and generation without considering the database elements and that gives birth to item sets that do not even occur in the dataset. Moreover, it has to count the support of the item sets after each iteration. Both of these processes are costly in terms of time and memory usage. A vivid analysis of time complexity by Chang and Lee [8] has proved it to be $O(m^2n)$ where m is the number of distinct items and n is the number of transactions.

E-Clat: The enhancement was needed on A priori for improving the huge toll of memory expense and it came as E-Clat [9], an algorithm using DFS approach and taking a vertical representation of database as input. The vertical algorithms such as E-Clat avoid the multiple scanning of the database by scanning the database only once and generating a list of occurrence of each item which is called a TID-list. By considering the list and the min sup value as the decision criteria, it identifies all n-item sets that are frequent, then searches all the possible combinations to find the (n+1)-item sets that are extended from the n-item sets and shares all the elements of the particular former item set except the last element. This operation works as a loop and repeats until all the frequent item sets are discovered. The candidate generation

and search operation are done without scanning the database multiple times. This is an advantage where it outperforms A priori in terms of memory and time efficiency.

Drawbacks: The E-Clat algorithm eliminates the major problems faced by BFS algorithms by avoiding multiple scans of database but especially when run on dense datasets, the time complexity for generation of the TID-list from the principle database can be costly and that exhausts huge amount of time and memory. To decrease the size of the TID-list, some mechanisms have been proposed with introducing a data structure called diff-sets. Another drawback of BFS algorithm which is not addressed by this one is the generation of item sets that do not occur in the database. This is caused because of not accessing the database while iterating the candidate generation process.

3.2 Pattern-Growth Algorithm for FIM

The development aiming towards solving the hitch of a generation of item sets that do not appear in the database, pattern-growth algorithms came in the light with an alternative view of DFS. It came with a new data structure called projected database and minimized the time complexity of scanning the database and make projections by using an optimization called pseudo-projection. This method uses pointers to represent the database and avoids making copy of the whole database. Many algorithms have been developed using pattern-growth principles. FP-Growth algorithm [10] uses prefix-tree structure to reduce space complexity. H-mine algorithm uses hyper-structure and LCM uses an integrated merge-datasets mechanism to merge the elements of identical databases and further compression of projected database to optimize the memory utility. It also introduces occurrence-delivery, an array-based efficient support counting process.

3.3 Dynamic Algorithms for FIM

All the above discussed traditional algorithms share a common limitation. They are strictly batch algorithms. So the algorithm has to be run every time the database gets updated in an online setting. This incurs is a huge gap in the time efficiency and thus they need to recompute the whole item set mining operation from scratch even when a small change occurs in a database. Addressing this problem there are certain algorithms designed to be used in dynamic databases. They are discussed in a nutshell below.

Incremental algorithms: CanTree [11], Pre-FUFP [12] are some of the algorithms which update the set of the frequent item sets keeping pace with the changes in the database. They use canonical order trees and modified FP-tree structures, respectively, as buffer and appends the decision set without being recomputed.

Stream mining algorithms: The era of big data analytics has introduced the need for scalable algorithms to be developed for frequent item set mining where they can handle high-frequency massive streams of data in a fast way. Stream processing algorithms are mainly approximation algorithms. They maintain fabricated prefix tree structures and apply several optimization techniques to force a tighter upper bound on the error of the output with each iteration. Algorithms like estDec [13] and estDec+ use compressible prefix tree to trace the support counts of the item sets.

Recent advancements: The recent developments of scalable and fast algorithms for finding frequent item sets from datasets are quite remarkable. Usage of set enumeration tree as the most recent approach has set the advanced algorithms apart and to succeed over the former ones. Diffnodesets and Negnodesets are the designed advanced data structures to work over the nodes of modified prefix-tree and are, respectively, used in DFin [14] and NegFin [15] algorithm.

4 Sequential Pattern Mining

Pattern mining gained its point of interest in the researchers from its ability to show hidden and essential trends and patterns in data with having an application of large number of fields for automated system development for database analysis. Most appreciated and practiced pattern mining algorithms, e.g.,: frequent item set mining, episode mining, subgraph mining work on sequential and transactional databases to perform statistical analysis on most frequent and rare trends classification, but it deliberately avoids the sequential ordering of events and transactions from its range of consideration. Where the ordering of events is an important factor for mining, the above mentioned are failed to produce necessary results. From this point the sequential pattern mining algorithms pose their importance. These are the algorithms for extracting 'interesting' subsequence from the sequential data where further to clarify the rhetoric nature of the quoted word, the interestingness of the subsequence is relative and can be measured through various signals such as frequency of occurrence, recentness, length and size, relevance, similarity index and many more. The paper follows a description of a typical problem statement for frequent sequential mining and continues with various techniques, methods and algorithms used to solve real world scenarios.

4.1 Sequential Data: Problem Statement

While researching we can find out that in data mining applications there are largely two types of data that are tagged as sequential. It can either be a time series which is an ordered list of numerical data projected over line of timestamps or it can be one or many sequences of symbols and letters. The sequential pattern mining algorithms originated from an aim to find the important subsequences from sequential data. Later

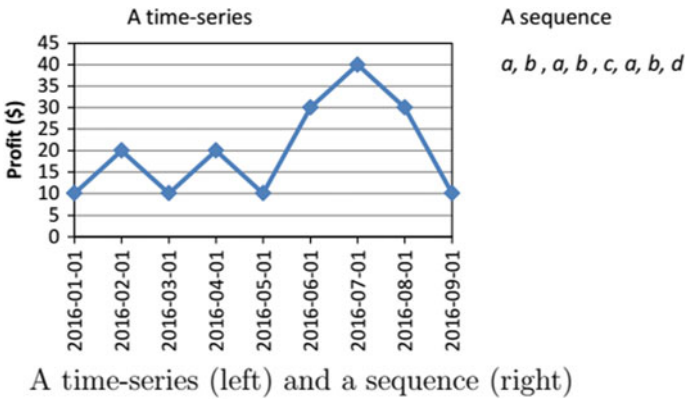


Fig. 1 Sequential pattern mining data types [16]

in applied science domains it was discovered that once discretized as a sequence by a neat preprocessing, time series data which are mainly historical projected data series on weather report, census, cumulative investment or import–export, share market stock-price and revenue data, etc. can be represented and analysed using these algorithms. The other form which is often transactional data of supermarkets, multi-agent decision poles, webpage click stream data, etc. are also analysed using the type of approach. Below, we have exemplified the two types of data one showing a time series representation of annual profit by some company and other one is a sequence of four distinct characters (Fig. 1).

4.2 Pioneer Algorithms

With GSP [17] being the first algorithm developed to solve sequential pattern mining problems, all the algorithms proposed in this topic are seen to follow any of the two distinct paradigms: Breadth first search (BFS) and Depth first search (DFS). We can classify them on that aspect and draw the prospects and drawbacks of both of them and can observe how they gave a passage towards the further development and minimizing the discrepancies.

Breadth First Search algorithms for SPM: The primary algorithms such as GSP, seen to use BFS proceeds from first identifying all the 1-sequences (the sequences having only 1 candidate item) and then perform S-extension and i-extension to move on to identify and generate two-sequences then three-sequences and so on. The process iterates till there is no subsequence left in the sequence. This algorithm was inspired by A priori. It takes an n-count for the initial subsequence length and recursively mines the larger subsequences by generating potentially frequent pattern with a length of n+1. After generating a new sequence say Sn, the algorithm checks for

all its subsequences with up to n-1 items. And if any of them are observed as non-frequent, it deletes the Sn from the list of most frequent patterns by the downward closure property check. Else than that it outputs the final Sn. They essentially use horizontal representation of databases. This approach taken by such kinds of algorithms as GSP is proved to have several limitations.

Drawbacks: BFS algorithms:

Time inefficiency: The BFS algorithms recurrently search and scan the database to calculate the support count of the generated incremental sequences. This leads to a multiple database scan problem and incurs huge costs even after using optimization techniques.

False pattern generation: GSP has been seen to have generated some candidate sets that are not existent in the databases. This causes wastage of time and memory triggered by its generation of sequences by appending and associating smaller subsequences without keeping a track of the database transformation steps. GSP later has to discard those patterns after a scan from the final sequence list and it causes unnecessary toll on time and memory.

Memory inefficiency: The level-wise approach followed by the BFS algorithms has a typical disadvantage. It's that they have to keep all n-sized candidate sets in the memory to generate the n+1 long subsequence. This causes a huge consumption of memory in each step and reduces the efficiency of the parallel operations carried on.

Depth First Search SPM algorithms: The insufficiency faced by the BFS algorithms aimed to develop the betterment and came in the form of a new approach and set of algorithms. Depth first search is the paradigm, roughly obeyed by almost all the popular and recent algorithms for sequential pattern mining. The principle difference made by these algorithms is that they start to use a vertical representation of the data which can be obtained by preprocessing of the horizontal database with only a single scan. Vertical representation holds where each item is occurring in an item list of data. This is denoted by the ID-list of the database which promises mainly two advantageous properties. The calculation of the support count for each item set can be obtained by just marking the distinct identifier in the ID-list and the sequence generation with extension operations can be done by a join operation without scanning the whole database repeatedly. These properties on the first-hand save a lot of cost in both memory and time from that of the BFS algorithms. Spade [18] is the first algorithm that used DFS as its paradigm and appeared to be the immediate solution to the problems caused by GSP. The later developed ones such as Spam [19] and Spade followed the same method to generate the whole transformed search space with scanning the database only once at O(n) cost and can generate all the frequent sequential patterns just by applying Joint operation on the singular ID-list item sets and do not repeatedly scan the database. DFS was proposed as the best possible alternative of the BFS algorithms and was seen to outperform all the primitive algorithms which followed BFS.

Drawbacks: Here are some limitations faced by the state of the art primitive vertical algorithms using DFS.

Large sized ID-list: The major fall outfaced by the Spade algorithm was discovered when analysed dense and long sequence databases. ID-list became too large and the joining operation became very costly both in terms of time and space.

Large-numbered Joins: It was observed that the primitive DFS based algorithms generate a huge number of candidate sets and performing joint operation on each of them causes unnecessarily huge cost.

Bit-vector enabled algorithms: The above-mentioned limitations were discovered when running the algorithms in massive datasets. The solution to the large ID-list problem came up with the introduction of an optimized structures for the ID-list which was Bit-vector. The transformed representation was brought up with the development of Spam algorithm where B-list corresponding a bit-vector containing all the item set of the previous list and having value set to 1 or 0, respectively, based on the presence of the presence or absence of the generated sequence in the ID-list. This efficiently saves considerable cost of running the ID-list form of dense datasets by setting most of the bit values as 1. The modified bit-vector enabled faster version Spade algorithm came up as Prism [20] and BitSpade [21]. Spam further developed to give birth to a new algorithm called Fast [22] which used a new support count optimization called index sparse ID-list. It decreased the memory usage more.

CMAP based algorithms: The second drawback caused by the innumerable join operation was resolved by the design of a new data structure called the Co-occurrence map (CMAP). An approach called co-occurrence pruning resulted in storing all the two-sequences in a single scan over the database. The later generated patterns are analysed using this by considering the last two items of the sequence Sn, if not a potential frequent two-sequence, the sequence was to be directly discarded from the list. This development was offered by CM-spade [23] and CM-span [23] which evidently outperformed the Spade and Spam algorithms in terms of both time and memory efficiency.

Drawbacks: Even these improved DFS based algorithms could not address one major incompetency. The generation of false redundant candidate patterns which did not appear in the database. This limitation was caused by the strategy of generating candidate sets by combining smaller item sets and without keeping the track of the database by accessing it.

4.3 Pattern-Growth SPM Algorithms

The problem of false candidate sets and the cost in deleting them after generation exhausted the settings of sequential pattern mining. Addressing this problem, Prefixspan [24] was developed. It efficiently avoids the problem by recursively scanning the database. To deal with the cost for scanning datasets repeatedly, a modified structure is introduced. This is called projected database. It reduces the input database size by DFS, considering the larger sized patterns. Prefixspan algorithm follows a pattern-growth approach. The algorithm has two major subsequent paradigms to be followed. First it scans the whole original database for calculating the support score

for each candidate set and marks the item sets as frequent if their support comes out to be greater than or equal to the min sup threshold. The second step involves depth-first search using the frequent sequential patterns discovered in the first step. If the sequence is depicted as Sn, the algorithms generally create projected database for Sn and find the candidates among them, those can be appended by i-extensions and s-extensions to generate new subsequence with n+1-sequence. This process recursively repeats the two steps to find out all the frequent sequential patterns. The generation of duplicate item sets is avoided by appending each item to a total order such as lexicographical order.

Drawbacks: The repeated scan of databases and creating database projections after each iteration can turn out to be costly in terms of runtime. The memory also can be exhausted if brute force scan and generation methods are applied to create a projected database, the worst case being scanning and imitating the whole database.

4.4 Special Cases of Sequential Pattern Mining

Uncertain pattern mining: This is an extension of sequential pattern mining to run across a database with patterns having uncertain value and support mainly distorted or inaccurate caused by a noisy data collection pipeline or medium. So far two major models are proposed to provide solution towards mining of uncertain sequential patterns. They are briefed below.

Expected support count model: The model uses a typical database with all the items with a particular expected probability of existing in a particular pattern. Which can be defined by $P(n_i, S)$, S being the particular candidate set and n is the sequence. The expectation of a pattern is calculated as the product of expected values of support count of the sequence wherever they occur in the whole database. The noisy sequential pattern mining algorithms developed based on GSP or Spam works on this principle.

Probabilistic pattern mining model [25]: This is one which works considering two user-defined threshold, the first being the min-confidence threshold and the other to be the min support. A pattern is regarded as frequent if it occurs in more transactions than the min support and considered space is more than min-confidence value. Seq-U-Prefixspan [26] is an algorithm developed inspired by Prefixspan uses probabilistic model for sequential pattern mining.

Fuzzy pattern mining: This is another extension of sequential pattern mining where the database has a parallel projection of the sequences having numerical values in an [0, 1] interval and fuzzy membership functions are used to convert those values into nominal values such as ‘extreme’, ‘moderate’, ‘less’. There are a variety of algorithms differentiated by the ways of counting the occurrence of patterns in the converted fuzzy sequences that have been proposed. Speedyfuzzy [27] is one such variant that identifies a sequential pattern as frequent if the membership values of each individual sequence in a pattern are greater than 0. Minifuzzy [27] offers modification

with an user-defined threshold membership value and analyse the data to give output to those patterns which have all the items having greater value than the threshold.

Weighted sequential pattern mining [28]: A variant of sequential pattern mining method considers a preprocessed database structure where each sequence has got a weight associated with them. The aim is to find sequential patterns that have minimum normalized weights. The importance of the patterns is judged on the basis of associated weights and that poses an insufficiency in the output which gives way to an enhancement measuring utility of patterns. We mention that one below.

High utility sequential pattern mining [29]: This is another case of weighted pattern mining approach. Here the data is processed to have two constraints into every sequence in the database. An weight and quantities of items both are considered for every sequence. The calculation of utility for the pattern is derived as the sum of all maximum utility values of the pattern in all the sequences they occur. The calculation of utility is not monotonic like the support or confidence counts as in the previous algorithms. The algorithms that mine the sequential pattern based on high utility try to propose a tight upperbound to deal with the fuss of calculating utility values.

5 Module-3: Sequence Prediction

Predicting the next item in the sequence, given a considerable set of real world historical sequence data for training the model is one important task to be developed in this current era of predictive analytics and its inevitable application in the field of stock market prediction, text and keyword prediction, customer behaviour prediction, product recommendation systems and many more. The goal was to design a memory-efficient data structure that is incrementally updatable to store the training sequences and to offer a scalable and time-efficient algorithm with high accuracy to predict the possible next items in the sequence set. The main challenge in designing the predictive model was that sequential pattern mining is time-consuming when it comes to discovering and extracting patterns and using them for prediction. One more disadvantage is that patterns generally ignore rare cases. Here, we will talk about some of the methods and structures to model the prediction systems.

5.1 Traditional Models

In the process of gradual development, several algorithms have been proposed which use various data structures such as Dependency graphs(DG) [25], Prediction by partial matching (PPM) [26], All-K-Order Markov(AKOM) [27], Transition directed acyclic graph and many other. Optimizations are applied to increase the accuracy of the models. Several compression algorithms have been applied and machine learning models like neural networks have been used to predict the sequences.

Limitations: The main set back faced by these above-discussed models are:

(1) PPM models of 1 or 2 order and dependency graph share the Markovian assumption as a principle. Accordingly they consider that the predicted sequence should be completely dependent on the previous sequences.
(2) In considering the training data, the PPM models show the limitation of considering a limited number of training attributes and this reduces the accuracy.
(3) For All-K-Order Markov model has an exhaustive time and space complexity to consider up to k previous sequences for prediction.
(4) Models like sequential pattern mining are costly to update as there is no efficient stream processing algorithm and most of them end up as batch algorithms.
(5) Not any of the models are lossless.

5.2 Advanced Model

Aiming towards solving the drawbacks of the existing models, Gueniche et al. proposed an algorithm with a new data structure: compact prediction tree which promises a compression of training sequence which is absolutely lossless by using the similarity between the subsequences. CPT [28] faces a limitation of higher time and space complexity. However, developments have been proposed to improve the size of the CPT by Frequent subsequence compression and simple branches compression. The enhanced technique is regarded as CPT+ algorithm.

Analysis: The space complexity is improved from the previous approaches drastically. The cost for insertion becomes O(s) where s is the length of the sequence and the worst case space complexity becomes $O(N*avg(s)_{1\text{-}N})$, where N is the number of sequences and 'avg(s)'is the average of the length of all sequences. In reality, the space complexity is subjected to decrease as sequences overlap.

5.3 Comparative Performance Analysis

The CPT+ algorithm is compared with four state of the art algorithms: Dependency graph (lookup window = 4), PPM (order-1), AKOM(order-5) and traditional CPT.

The data: The analysis is done against three individual variants of dataset. The description is listed below (Table 1).

Table 1 Dataset description

Dataset	Description	Seq-count	Unique items	Avg item count/seq
FIFA	Webpage click	28,978	3301	1.04
SIGN	signs	730	276	1.79
BIBLE	characters	32,529	76	4.78

Comparison: prediction accuracy: The result of the accuracy analysis is listed in Table 2.

Comparison: Space occupancy (size): The analysis result with respect to the number of nodes (size of the structure) (Table 3).

Comparison: Time efficiency (training time): The comparative report is listed below (Table 4).

Comparison: Time efficiency (prediction time): The results are denoted below in tabular form (Table 5).

Table 2 Accuracy analysis

Dataset	CPT+	CPT	DG	PPM	AKOM
FIFA	37.98	35.65	25.45	24.33	26.65
SIGN	33.32	31.15	9.67	8.32	5.65
BIBLE	73.55	68.73	7.06	3.14	7.15

From the above-mentioned data, an inference can be drawn that CPT+ has shown considerable amount of advancement from the previously existing models and its thrice the magnitude more accurately from DG, PPM and AKOM on an average scale

Table 3 Space requirement analysis

Dataset	CPT+	DG	PPM	AKOM
FIFA	18,218	11,200	10,198	76,782
SIGN	1015	999	1151	8279
BIBLE	322	145	154	5432

The CPT+ occupies similar amount of space as DG and PPM. Its size is more than two magnitudes less than AKOM. This is a trade-off for generating more accurate results

Table 4 Training time analysis

Dataset	CPT+	DG	PPM	AKOM
FIFA	0.148	2.626	0.176	10.373
SIGN	0.005	1.428	0.011	3.212

In terms of training time, CPT+ is faster than all of the previous state of the art algorithms

Table 5 Prediction time analysis

Dataset	CPT+	DG	PPM	AKOM
FIFA	0.004	1.121	0.012	0.721
SIGN	0.002	0.521	0.004	0.211

The CPT+ shows promising improvement in prediction time than previous models

6 Conclusion

Throughout the paper, we have presented a review work on the study of different data structures, algorithms and integrated mechanisms of finding the frequent item sets, discovering hidden sequential patterns from the transaction databases. We have discussed the basic and primitive algorithms, the gradual development and improvement of empirical algorithms, advanced approaches and relevant methods. The extensions and special cases of frequent item set and sequential pattern mining problems are discussed. We have also looked upon prediction of sequence items based on training dataset and sequence. We have studied the classical approaches and the most accurate and most recent algorithms. We have drawn a comparison among the discussed algorithms. The paper should provide a suitable overview on the research so far in this field. The study should act as a comprehensive guide in brief.

References

1. J. Han, J. Pei, Y. Ying, R. Mao, Mining frequent patterns without candidate generation: a frequent-pattern tree approach. Data Min. Knowl. Disc. **8**(1), 53–87 (2004)
2. J. Lin, E. Keogh, Wei, R. Srikant, Fast algorithms for mining association rules, in *Proceedings of 20th International Conference on Very Large Data Bases, VLDB 1994* (Santiago de Chile, Chile, 12–15 September, 1994), pp. 487–499
3. M. Hegland, The apriori algorithm—a tutorial. Math. Comput. Imaging Sci. Inf. Process. **11**, 209–262 (2005)
4. M.J. Zaki, Scalable algorithms for association mining. IEEE Trans. Knowl. Data Eng. **12**(3), 372–390 (2000)
5. J. Han, J. Pei, Y. Ying, R. Mao, Mining frequent patterns without candidate generation: a frequent-pattern tree approach. Data Min. Knowl. Discov. **8**(1), 53–87 (2004)
6. C.K. Leung, Q.I. Khan, Z. Li, T. Hoque, CanTree: a canonical-order tree for incremental frequent-pattern mining. Knowl. Inf. Syst. **1;11**(3), 287–311 (2007)
7. C.W. Lin, T.P. Hong, W.H. Lu, The pre-FUFP algorithm for incremental mining. Expert. Syst. Appl. **31;36**(5), 9498–9505 (2009)
8. J.H. Chang, W.S. Lee, Finding recent frequent item sets adaptively over online data streams, in *Proceedings of 9th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (Washington DC, USA, 24–27 August, 2003), pp. 487–492
9. Z.-H. Deng, DiffNodesets: an efficient structure for fast mining frequent item sets. Key Laboratory of Machine Perception (Ministry of Education), School of Electronics Engineering and Computer Science, Peking University, Beijing 100871, China
10. N. Aryabarzan, B. Minaei-Bidgoli, M. Teshnehlab, negFIN: an efficient algorithm for fast mining frequent item sets. Expert. Syst. Appl.
11. An introduction to sequential pattern mining by P. Fournier-Viger at http://data-mining.philippe-fournier-viger.com/introduction-sequential-pattern-mining
12. R. Srikant, R. Agrawal, Mining sequential patterns: generalizations and performance improvements, in *The International Conference on Extending Database Technology* (1996), pp. 1–17
13. M.J. Zaki, SPADE: an efficient algorithm for mining frequent sequences. Mach. Learn. **42**(1–2), 31–60 (2001)
14. J. Ayres, J. Flannick, J. Gehrke, T. Yiu, Sequential pattern mining using a bitmap representation, in *ACM SIGKDD, International Conference on Knowledge Discovery and Data Mining* (2002), pp. 429–435

15. K. Gouda, M. Hassaan, M.J. Zaki, Prism: an effective approach for frequent sequence mining via prime-block encoding. J. Comput. Syst. Sci. **76**(1), 88–102 (2010)
16. S. Aseervatham, A. Osmani, E. Viennet, bitSPADE: a lattice-based sequential pattern mining algorithm using bitmap representation, in *The International Conference on Data Mining* (2006), pp. 792–797
17. E. Salvemini, F. Fumarola, D. Malerba, J. Han, Fast sequence mining based on sparse id-lists, in *The International Symposium on Methodologies for Intelligent Systems* (2011), pp. 316–325
18. P. Fournier-Viger, A. Gomariz, M. Campos, R. Thomas, Fast vertical mining of sequential patterns using co-occurrence information, in *The Pacific-Asia Conference on Knowledge Discovery and Data Mining* (2014), pp. 40–52
19. J. Pei, J. Han, B. Mortazavi-Asl, J. Wang, H. Pinto, Q. Chen, U. Dayal, M.C. Hsu, Mining sequential patterns by pattern-growth: the prefixspan approach. IEEE Trans. Knowl. Data Eng. **16**(11), 1424–1440 (2004)
20. M. Muzammal, R. Raman, Mining sequential patterns from probabilistic databases. Knowl. Inf. Syst. **44**(2), 325–358 (2015)
21. Z. Zhao, D. Yan, W. Ng, Mining probabilistically frequent sequential patterns in large uncertain databases. IEEE Trans. Knowl. Data Eng. **26**(5), 1171–1184 (2014)
22. C. Fiot, A. Laurent, M. Teisseire, From crispness to fuzziness: three algorithms for soft sequential pattern mining. IEEE Trans. Fuzzy Syst. **15**(6), 1263–1277 (2007)
23. J.H. Chang, Mining weighted sequential patterns in a sequence database with a time-interval weight. Knowl.-Based Syst. **24**(1), 1–9 (2011)
24. C.F. Ahmed, S.K. Tanbeer, B.S. Jeong, A novel approach for mining high-utility sequential patterns in sequence databases. Electron. Telecommun. Res. Inst. J. **32**(5), 676–686 (2010)
25. V.N. Padmanabhan, J.C. Mogul, Using prefetching to improve world wide web latency. Comput. Commun. **16**, 358–368 (1998)
26. J. Cleary, I. Witten, Data compression using adaptive coding and partial string matching. IEEE Trans. Inform. Theory **24**(4), 413–421 (1984)
27. J. Pitkow, P. Pirolli, Mining longest repeating sub-sequence to predict world wide web surng, in *Proceedings of 2nd USENIX Symposium on Internet Technologies and Systems* (Boulder, CO, 1999), pp. 13–25
28. T. Gueniche, P. Fournier-Viger, V.S. Tseng, Compact prediction tree: a lossless model for accurate sequence prediction, in *Advanced Data Mining and Applications. ADMA 2013*, ed. by H. Motoda, Z. Wu, L. Cao, O. Zaiane, M. Yao, W. Wang. Lecture Notes in Computer Science, vol. 8347 (Springer, Berlin, Heidelberg)
29. T.C. Truong, P. Fournier Viger, A survey of high utility sequential pattern mining, in *High-Utility Pattern Mining: Theory, Algorithms And Applications* (2019)

Study of AdaBoost and Gradient Boosting Algorithms for Predictive Analytics



Pritika Bahad and Preeti Saxena

1 Introduction

Predictive analytics utilizes statistical methods and machine learning methods to formulate a prediction about possible future outcomes [1]. One of the most promising areas where it is applied nowadays to make a change is healthcare. Emergence and use of IoT enabled wearable health devices has set a new research dimension in Big data analytics. In healthcare industry, data analytics is used to provide patient-centric prescription to improvise patient satisfaction. Big data analytics in health care facilitates analysis of the voluminous datasets of patients to discover hidden relationships in datasets and to develop predictive models using statistical methods and machine learning techniques to deliver improved healthcare services. Predictive analytics applications in healthcare can be used to determine which patients are at high risk of developing lifetime illnesses such as diabetes, Alzheimer's disease, or asthma on the basis of patient history. These applications can assist in reducing time of diagnosis, well-timed detection, and finally, increasing the life span of patient.

Diabetes is a chronic lifelong disease. The changing diet habits and daily physical activity has increased the number of diabetic patients all over the world. As per the World Health Organization (WHO) [2] report, there will be approximately 350 million diabetes affected people worldwide by year 2030. Diabetes is grouped into two types as Type 1and Type 2 diabetes. Type 1 diabetes is related with an abnormal increase in the level of glucose (hyperglycemia) in blood due to the inadequate production of insulin by pancreas. The cells fail to respond effectively to insulin

P. Bahad (✉) · P. Saxena
School of Computer Science and Information Technology, Devi Ahilya University, Indore 452001, India
e-mail: bahad.pritika@gmail.com

P. Saxena
e-mail: preeti_ms@rediffmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_22

produced by pancreas in Type 2 diabetes. Diabetes can be detected with several automatic diagnosis systems.

Many researchers exploited a range of machine learning classification algorithms as Naive Bayes, Support Vector Machine, Nearest Neighbor, and Decision Tree for disease prediction on the basis of risk factors and disease matrix. Researchers have demonstrated that machine learning algorithms [3–5] improved the ability to diagnose different diseases. SVM algorithm has been successfully implemented for liver disease diagnosis [6]. A combined Naïve Bayes classifier and K-Nearest Neighbor approach [7] is used for prediction of heart disease to get better accuracy. Using machine learning methods, a predictive model is designed to identify the risk of development of Alzheimer's disease from mild cognitive impairment [8]. Ensemble machine learning techniques combine the predictions of several machine learning classification techniques to improvise effectiveness over a single machine learning classification. For early diagnosis of diabetes, ensemble machine learning classification algorithms can be utilized. Literature [9–12] reveals that boosting is a very compelling and broadly used ensemble machine learning method for predictive analytics.

The current study follows to build a model and evaluate AdaBoost and Gradient boost ensemble machine learning algorithms to predict diabetes disease on the basis of health-related human matrices. The proposed predictive analytics model is evaluated on PIDD—Pima Indians Diabetes Dataset.

The remainder of the paper is structured as follows: Sect. 2 presents the background of AdaBoost and Gradient Boosting techniques. The proposed experimental predictive analytics model and evaluation results are shown in Sects. 3 and 4, respectively. The Sect. 5 concludes the study and suggests future directions.

2 Ensemble Machine Learning Techniques

Ensemble machine learning techniques use multiple base learners to obtain better predictive accuracy. The basic insight is to improve the generality of individual classifiers by training multiple base classifiers on different datasets (sampled from the original dataset) and combining the results. Averaging the outputs of several classifiers helps to reduce the variance component and/or the bias in the classification. The key sources of difference in actual and predicted values are noise, variance, and bias. Ensemble machine learning helps to minimize variance and bias. Bagging and boosting are widely used ensemble learning methods. Bagging reduces the variance whereas Boosting minimizes variance as well as bias of the classification.

Bagging that means Bootstrap Aggregating is an ensemble method introduced by Zhou [9]. Bootstrap sampling is used to obtain the data subsets for training the base learners. A bootstrap sample is attained by subsampling the training data set with replacement where the size of a sample and the training dataset is same. Bagging combines the base learners by majority voting and highly voted class is predicted.

Let A and C indicate the instance space and the set of class labels, respectively. Given a training data set $D = \{(a_1, c_1), (a_2, c_2), \ldots, (a_N, c_N)\}$, where $a_i \in A$ and $c_i \in C$ ($i = 1, \ldots, N$). L represents the number of base learners and the h_i's are the base learners. A function h(a) classify new test sample by returning the class *c* that obtains the highest number of votes from the base models $h_1, h_2, \ldots, h_T$ and indicator function I(w) returns 1 if w is true and 0 otherwise. Bagging combines the base learners by majority voting and the highly voted class is predicted. Random Forests [11] is one of the popular bagging methods. The algorithm 1 summarizes bagging.

Algorithm 1 : Bagging

1 **Input:** Dataset $D = \{(a_1, c_1), (a_2, c_2), \ldots, (a_N, c_N)\}$,
Base learner L and
Number of learning iteration T

2 For t=1 to T :
(a) Create a bootstrap sample from D
D_t= Bootstrap(D)
(b) Train a base learner $h_t$from the bootstrap sample
$h_t = L(D_t)$

3 **Output :** $H(a) = \max_{c \in C} \sum_{t=1}^{T} I(h_t(a) = c)$

In contrast to bagging, boosting assigns weights to data. The weights of misclassified samples are increased to focus the learning algorithm on specific samples. Typically, Boosting generate more accurate results than Bagging.

2.1 AdaBoost

AdaBoost (Adaptive Boosting) is one of the popularly used boosting algorithms. In AdaBoost, the weak learners are decision trees with a single split termed as decision stumps. AdaBoost works by weighting the samples. It increases the weight of samples that are difficult to classify and lowers the weight for those samples that are easy to classify. The process is repeated until the algorithm identifies a model that correctly classify these samples. The final classifier is linear combination of base classifiers from each stage. AdaBoost.M1 [9] is the most successful form of the AdaBoost algorithm that was used for binary classification problems.

Algorithm 2 summarizes AdaBoost.M1. Initially, equal weights w_i are assigned to all the training samples. The weights distribution at the tth learning iteration is denoted as D_t. The algorithm generates a base learner h_t from the training data set D_t by applying the base learning algorithm. The training samples are used to test h_t. Weights of the incorrectly classified samples are increased and updated weight distribution is calculated. AdaBoost generates another base learner from the training samples by again calling the base learning algorithm. This process is continued T times and final learner is generated by weighted majority voting of T base learners. AdaBoost minimizes an exponential loss function.

Algorithm 2 : AdaBoost.M1

1 **Input :** Dataset $D = \{(a_1, c_1), (a_2, c_2), \ldots, (a_N, c_N)\}$,
Base learner L and
Number of learning iteration T

2 Initialize equal weight to all training samples $w_i = 1/N$, $i = 1,2,\ldots, N$

3 For t=1 to T :

(a) Train a base learner h_t from D using D_t to training sample using w_i.
$h_t = L(D, D_t)$

(b) Compute error of h_t as
$$err_t = \frac{\sum_{i=1}^{N} w_i I(h_t(a_i) \neq c_i)}{\sum_{i=1}^{N} w_i}$$

(c) Compute the weight of h_t as
$$\alpha_t = \log\left(\frac{1 - err_t}{err_t}\right)$$

(d) Set $w_i \leftarrow w_i . \exp[\alpha_t I(h_t(a_i) \neq c_i)]$

4 **Output :** $H(a) = \text{sign} \sum_{t=1}^{T} \alpha_t h_t(a)$

2.2 Gradient Boost

Gradient boosting or gradient tree boosting is an ensemble machine learning technique employed for regression and classification problems. The objective of gradient boosting is to minimize a loss function by stirring in the reverse direction of the gradient. A loss function is a measure that tells us "how good" a model is at making predictions for a given set of parameters. The loss function to be used depends on the type of problem. For example, for binary classification of disease diagnosis, the loss function is a measure based on failing to diagnose the disease of a sick person. AdaBoost identifies the "shortcomings" (samples that are difficult to classify) by high-weight data samples whereas gradient boosting identifies "shortcomings" by gradients.

Algorithm 3 : Gradient Boost

1 **Input :** Dataset D ={$(a_1, c_1), (a_2, c_2), \ldots, (a_N, c_N)$},
Base learner L,
Number of learning iterationT,
Step size Υ, and
Loss function to minimize Loss(c_i,Υ)

2 Initialize the model with a constant value
$f_0(a)=\text{argmin}_{\Upsilon} \sum_{i=1}^{N} \text{Loss}(c\ ,\Upsilon)$

3 For t=1 to T :

(a) For i =1,2,…, N compute the pseudo residuals
$$r_{im}=-\left[\frac{\partial \text{Loss}(c_i,f(a_i))}{\partial f(a_i)}\right]_{f=f_{t-1}}$$

(b) Fit base learner to pseudo residuals r_{it} giving terminal regions
R_{jt}, j=1,2,…,J_t

(c) For j=1,2,…,J_t computestep magnitude multiplier
$$\Upsilon_{jt}=\text{argmin}_{\Upsilon} \sum_{a_i \epsilon R_{jt}} \text{Loss}(c_i, f_{t-1}(a_i) + \Upsilon)$$

(d) Update $f_t(a) = f_{t-1}(a) + \sum_{j=1}^{J_t} \Upsilon_{jt} I(a \epsilon R_{jt})$

4 Output $\hat{f}(a) = f_T(a)$

Gradient Boost [12] is summarized in Algorithm 3. Gradient descent minimizes complicated loss functions that cannot be minimized directly. The loss function to minimize is represented as *Loss*. Initialize the model with a single prediction value $f_0(a)$ with the mean of training target values. For first iteration t = 1, compute the gradient of *Loss* with respect to prediction value $f_1(a)$ and fit a base learner to the gradient components. Compute step magnitude magnifier and update the prediction value $f_1(a)$. The procedure is continued recursively until the computation of final predictive function $f_T(a)$.

Boosting techniques use cross-validation and Out-of-bag (OOB) methods to determine the optimal number of boosting iterations. OOB allows on-the-fly computation without the need of repeated model fitting.

3 Methodology

The diagrammatic representation of the proposed two-class predictive analytics model is shown in Fig. 1.

The selection of training data set greatly affects the reliability of a predictive analytic model. To determine optimal number of iterations and to prepare training and test data, Train–Test split and K-fold cross-validation method are used. The K-fold cross-validation method is used to reduce the bias associated with random sampling. Feature extraction techniques are used to find out a meaningful cause–effect relationship between features. AdaBoost with Naïve Bayes base learner, AdaBoost with Decision Tree as base learner, and Gradient Boosting are candidate prediction classifiers. Different loss functions can be used for gradient boosting classifier as per

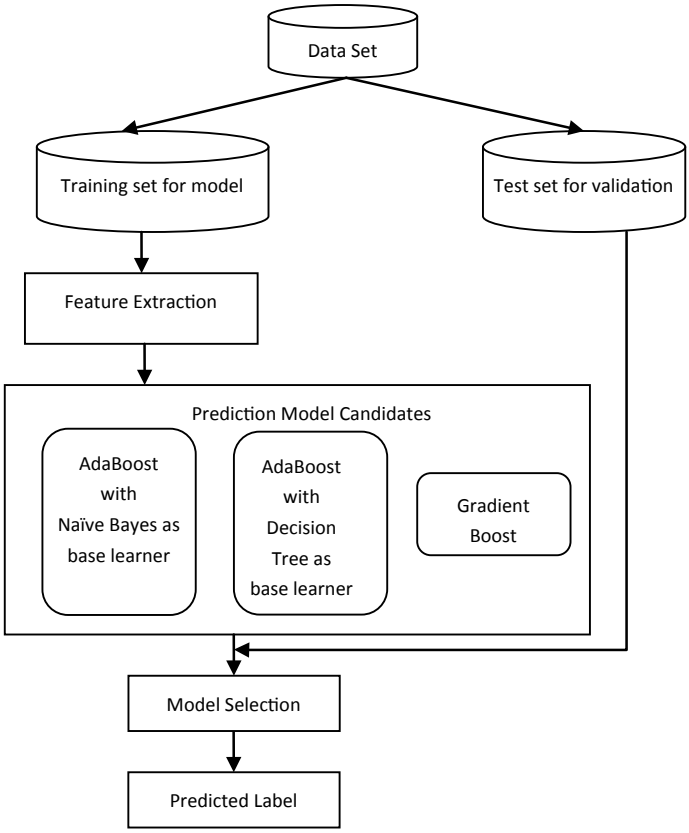


Fig. 1 Two-class predictive analytics model

classification problem requirement. Accuracy matrices are used to assess classifier,s performance.

4 Experimental Results

All experiments have been performed on a Core™ processor Intel® CPU i3-5005U 2.00 GHz with 4 GB RAM. AdaBoost and Gradient Boost ensemble machine learning algorithms available in Python are used for analysis, classification, and prediction. The two-class predictive analytics model is tested to predict whether an individual would develop diabetes on the basis of health matrices. The dataset utilized in this study PIDD (Pima Indian Diabetes Database) is obtained from open Machine Learning Repository accessible at Kaggle [13]. This diabetes dataset consists of 768 records of female patients. Each record consists of nine attributes as eight risk factors

Table 1 Specification of PIDD dataset

Attribute	Type	Range
Pregnancies (number of times pregnant)	Integer	0–17
Plasma glucose (mg/dL)	Integer	0–199
Diastolic blood pressure (mm Hg)	Integer	0–122
Triceps skinfold thickness (mm)	Integer	0–99
Serum insulin (mu U/ml)	Integer	0–846
Body mass index (kg/m^2)	Float	0–67.1
Diabetes pedigree function	Float	0.078–2.42
Age (years)	Integer	21–81
Outcome	Integer	1 = Tested positive for diabetes
		0 = Tested negative for diabetes

and one target class as "outcome". Out of 768 records, 500 are labeled as 0 (non-diabetic) and 268 as 1 (diabetic). It is a binary classification problem where all the input variables are numeric and have different scales. PIDD Dataset specification is shown in Table 1.

The performance accuracy of candidate boosting classifiers used in the proposed two-class predictive analytics model can be done using various available evaluation metrics. To check the prediction accuracy, the test data is provided and widely used evaluation matrices as accuracy, precision, sensitivity, specificity, F1-score, ROC, and AUC are used. The confusion matrix is a table that shows correctly and incorrectly classified samples of each class by candidate classifier. For a binary classification problem, a 2×2 confusion matrix is formed. For multiclass classification problem with m different classes, a $m \times m$ confusion matrix is generated. The confusion matrix for binary classification of diabetes diagnosis is shown in Table 2.

Table 2 Confusion matrix for binary classification of diabetes diagnosis

		Predicted	
		Not a diabetes disease patient	Diabetes disease patient
Actual	Not a diabetes disease patient	True negative (TN)	False positive (FP)
	Diabetes disease patient	False negative (FN)	True positive (TP)

Accuracy is calculated as the ratio of the number of correctly predicted samples to the total number of samples for a given test data set. It is not a good measure for balance data set.

$$Accuracy = (TP + TN)/(TP + FP + TN + FN) \tag{1}$$

Precision is calculated as the ratio of all correctly predicted samples to all predicted samples.

$$Precision = TP/(TP + FP) \tag{2}$$

Sensitivity, Recall, or True Positive Rate is the ratio of all correctly predicted samples to all samples that should be predicted.

$$Sensitivity = TP/(TP + FN) \tag{3}$$

False Negative Rate (FNR) is the fraction of the false negative samples to all samples that should be predicted.

$$FNR = FN/(TP + FN) \tag{4}$$

False Positive Rate (FPR) is fraction of the number of misclassified negative samples to the total number of negative samples.

$$FPR = FP/(FP + TN) \tag{5}$$

Specificity or True Negative Rate is the fraction of the number of correctly classified negative samples to the total number of negative samples.

$$Specificity = TN/(TN + FN) \tag{6}$$

F1-score is defined as a weighted average of precision and recall.

$$\begin{aligned} F1\text{-}score &= 2 * (precision * recall)/(precision + recall) \\ &= 2 * TP/(2 * TP + FN + FP) \end{aligned} \tag{7}$$

The obtained individual classification results for PIDD dataset with single Train–Test split dataset is shown in Table 3. Observation shows that AdaBoost with Decision tree and Gradient Boosting had better accuracy result than AdaBoost with Naïve Bayes.

The five-fold cross-validation method is applied on PIDD dataset to measure the performance of candidate models. Four folds are used to train the classifiers and the left-out fold is used to estimate the classification error. The accuracy scores in Table 4 reflect that the performance of all candidate models is improved over Train–Test split.

Table 3 Results of AdaBoost and gradient boost classifier with single train–test split

Classifier	Accuracy	Precision	Recall	F1-Score
AdaBoost with Naïve Bayes as base learner	0.57	0.44	0.83	0.58
AdaBoost with decision tree as base learner	0.73	0.71	0.45	0.55
Gradient boost	0.74	0.75	0.45	0.55

Table 4 Accuracy of AdaBoost and gradient boost classifier with five-fold cross-validation

Classifier	Accuracy
AdaBoost with Naïve Bayes as base learner	0.63
AdaBoost with decision tree as base learner	0.77
Gradient boost	0.78

Receiver Operating Characteristic Curve is another widely used performance measure of classifier. ROC curve allows visualizing classifier's behavior and represents the trade-off of the classifier between the number of correct positive predictions and incorrect positive prediction. Area under the ROC Curve (AUC) also conveys models prediction performance. AUC values vary from 0.5 to 1.0 and larger value represents better performance. The steepness of ROC should be high as it represents high true positive rate while less false positive rate.

The ROC curve in Fig. 2 represents the result of the diabetes classifier. ROC curves of different ensemble machine learning models considered in the study are

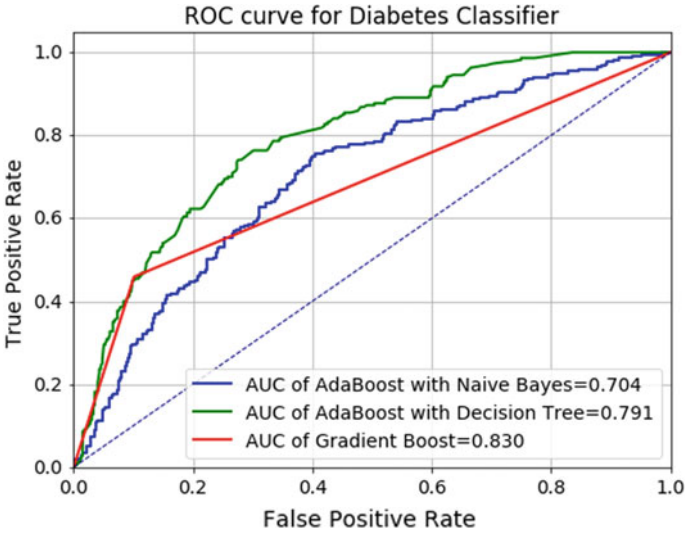


Fig. 2 ROC curve for diabetes classifier

shown with different colored lines. Results show that Gradient Boosting has better classification accuracy over AdaBoost.

5 Conclusion

Boosting is based on building successive classifiers on the fitted version of the training set which is generated on the basis of the error rate of the previous classifier. The study suggests Gradient Boosting focuses on minimization of loss function and achieves better prediction accuracy over AdaBoost. Gradient Boosting effectively handles heterogeneous feature dataset. The performance of a prediction model depends on the characteristics of the dataset; the base classifier and the loss function. The proposed predictive model works well for balanced and imbalanced low-dimensional data set. The study can be further extended to perform online predictive analytics on high-dimensional multiclass dataset.

References

1. E.B. Donald, A. Ahmed, Y.K.L. Raymond, Predictive analytics: introduction. IEEE Intell. Syst. **30**, 6–8 (2015)
2. WHO-Diabetes programme Home Page. http://www.who.int/diabetes. Last accessed 17 Mar 2019
3. N. Jayanthi, B. VijayaBabu, N. Sambasiva Rao, Survey on clinical prediction models for diabetes prediction. J. Big Data **4**(26), 1–15 (2017)
4. M. Fatima, M. Pasha, Survey of machine learning algorithms for disease diagnostic. J. Intell. Learn. Syst. Appl. **9**, 1–16 (2017)
5. K. Kourou, T.P. Exarchos, K.P. Exarchos, M.V. Karamouzis, D.I. Fotiadis, Machine learning applications in cancer prognosis and prediction. Comput. Struct. Biotechnol. J. **13**, 8–17(2015)
6. E.M. Hashem, M.S. Mabrouk, A study of support vector machine algorithm for liver disease. Am. J. Intell. Syst. **4**(1), 9–14 (2014)
7. E.Z. Ferdousy, M.M. Islam, M.A. Matin, Combination of Naïve Bayes classifier and K-nearest neighbor (cNK) in the classification based predictive models. Comput. Inf. Sci. **6**(3), 48–56 (2013)
8. X.Y. Qu, B. Yuan, W.H. Liu, A predictive model for identifying possible MCI to AD conversions in the ADNI database, in *2nd International Symposium Proceedings of the on Knowledge Acquisition and Modeling*, vol. 105(3) (IEEE, Wuhan, P.R. China, 2009), pp. 102–105
9. Z.-H. Zhou, *Ensemble Methods: Foundations and Algorithms*, 1st edn. (Chapman & Hall/CRC, 2012)
10. L. Breiman, Bagging predictors. Mach. Learn. **24**(2), 123–140 (1996)
11. S. Pouriyeh, S. Vahid, G. Sannino, G.D. Pietro, H. Arabnia, J. Gutierrez, A comprehensive investigation and comparison of machine learning techniques in the domain of heart disease, in *22nd IEEE Symposium on Computers and Communication Workshops—ICTS4eHealth* (2017)
12. J.H. Friedman, Greedy function approximation: a gradient boosting machine. Ann. Stat. **29**(5), 1189–1232 (2001)
13. Pima Indians Diabetes Database. https://www.kaggle.com/uciml/pima-indians-diabetes-database. Last accessed 17 Mar 2019

Enhancing Privacy and Security in Medical Information with AES and DES



Nikhil Khandare, Omkar Dalvi, Valmik Nikam and Anala Pandit

1 Introduction

Security of data has become a major requirement to create an efficient system. Since information regarding a person tells much about his nature and personal details. This information is called as data about that person. Such data may reveal much of the personal details and can result in serious consequences, if misused. If such data gets accessed by an attacker it may result in great loss to the person to which the data belongs. To avoid such situations, not only the system should be secure but the data that is processed in the system should be secure. This can be made possible by applying encryption on the data. The data should be made secure in such a way that it can be accessed by authenticated entities. It could be implemented by applying encryption to the data. Encryption is the process of converting readable information into non-readable format. Encryption technique should be chosen in a way to make the attacker difficult to attack on the encrypted data or to identify the data.

N. Khandare (✉) · O. Dalvi · A. Pandit
Department of Master of Computer Application, Veermata Jijabai Technological Institute, Mumbai, India
e-mail: nbkhandare@mc.vjti.ac.in

O. Dalvi
e-mail: omkarpd95@gmail.com

A. Pandit
e-mail: aapandit@mc.vjti.ac.in

V. Nikam
Department of Computer Engineering and Information Technology, Veermata Jijabai Technological Institute, Mumbai, India
e-mail: vbnikam@it.vjti.ac.in

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_23

AES and DES are block ciphers, these algorithms encrypt the data in blocks. In health center all the reports are scanned and stored in database, to the programmer these reports are images (rectangular arrangement of pixels in m-rows and n-columns), where each pixel is having greyscale value (intensity) between 0 to 2^n for n bit image. Now this rectangle of pixels (matrix of values) can be divided into block and each block can be encrypted separately by using block cipher, this was the motivation behind choosing AES and DES (block cipher). In this paper an attempt is made to secure communication among the entities in health center and encrypting medical records along with maintaining a remote database of records.

2 Related Work

Medical information card, credit card size plastic card was used for conveniently carrying the individual's medical information in pocket. Physician would type the information and this information would be printed by laser on the card. Information printed on card was readable by naked eye and was not requiring special reading equipment, there was no security feature added to these cards and on theft or loss, the confidentiality of data was compromised [1].

Health Insurance Portability and Accountability Act (HIPAA) of 1996 has defined "protected health information (PHI)" as individually identifiable health information [2]. Protected health information also includes the following,

- Past, present or future medical or physical health or condition of person.
- Past, present or future payments for provision of health care to the individual.
- Name, Address, Date of birth, Social security number.

HIPAA also defines de-identified health information as the information that neither discloses the identity of individual nor it provides a basis to identify the individual. There is no restriction on disclosure of de-identified health information [2].

The Picture archiving communication system (PACS) at the University of California San Francisco Medical Center consolidate images and associated data from various scanners into a centralized data archive and transmit them securely to remote display stations for review and consultation purposes, Symmetric key cryptography was one of the technique used [3].

Data Encryption standard (DES) stands for Data Encryption Standard. It comes under Block cipher [4]. The size of the block in DES is 64 bits. The length of the key that is actually used is 56 bits and remaining 8 bits are not used in the encryption process. DES performs 16 rounds in both encryption and decryption process. Substitution and permutation process play a major role in working of DES. The diagrammatic representation of working of DES is given in Fig. 1.

Permutation is processed by using p-box as shown in Fig. 2, For decryption the same steps are executed in reverse order.

Data encryption standard was used for encrypting the medical images along with discrete wavelet transform for compression and using this technique security of

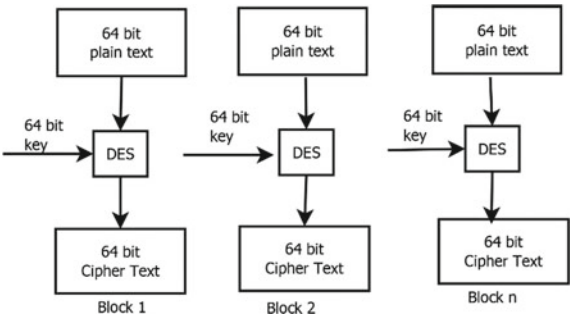


Fig. 1 Working of data encryption standard

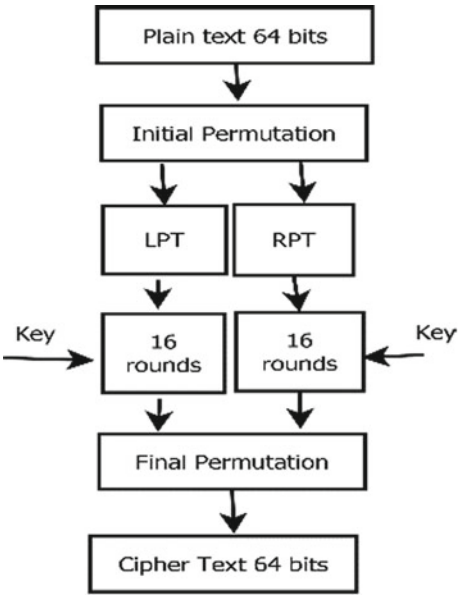


Fig. 2 Working of data encryption standard 2

medical images were enhanced and also transmission rate was improved [5]. Crypt-analysis of DES was done in [6, 7]. A new protection principle: the transformed masking method was proposed for enhancing the security of block ciphers [8].

AES was given by national institute of standards and technology in 2001 [9, 10]. AES stands for Advanced Encryption Standard. It works on symmetric key encryption and block cipher. It starts with initialization process. The key size and plain text is taken as byte in AES. It requires 16 byte of key size. The plain text of 16 bytes is taken as 4 * 4 matrix and in further process is taken as a matrix. The number of rounds that are performed depends upon the size of the key taken in AES algorithm. The diagrammatic representation of AES in shown in Fig. 3, and steps shown in Fig. 3 are executed in reverse order for decryption.

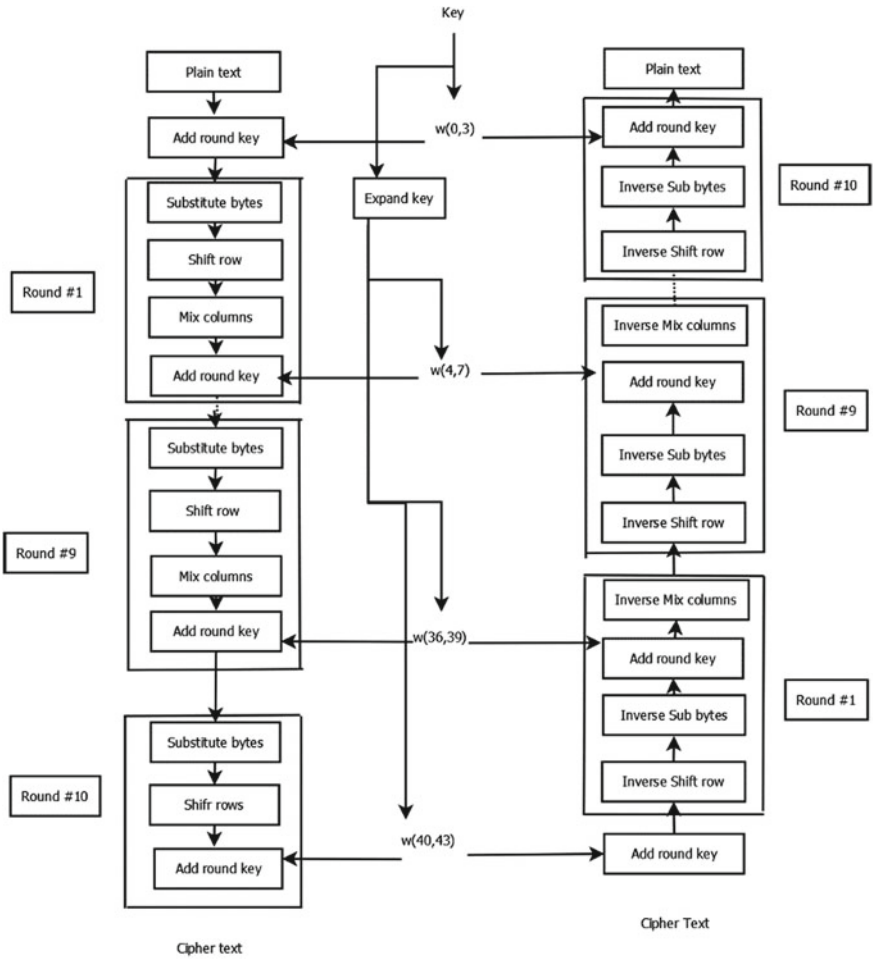


Fig. 3 Working of advanced encryption standard algorithm

The strengths and weaknesses of the three most popular symmetric-key cryptosystems: AES, DES, and 3DES was discussed in [11]. Combination of classical techniques and machine learning technique was used to generate an attack in AES [12]. Improving the level of security by combining One-Time Pad (OTP) encryption technique with AES to encrypt the data before loading it into data warehouse was done in [13]. Medical Image encryption was done using hierarchical diffusion and non-sequential encryption, diffusion property was achieved by using image shuffling and pixel modification [14].

AES technique was used for secure transmission of medical images in radiology.

All the images of 3D volume of patient is taken and these images were used to reconstruct radiograph digitally, The Digitally Reconstructed Radiograph (DRR)

is divided into four parts of equal sizes. Zigzag pattern was applied to all the 16 quadrants and each quadrant was given as input for encrypting to block of AES [15].

This research focuses on protecting the individually identifiable health information. In past AES and DES were used for securing medical data, however little attention is paid to securing the communication between entities in health center (to the best of authors knowledge) along with encrypting and storing the medical reports. This paper proposes the way to securely exchange the information in health center. Rest of the paper is organized as follows, Sect. 3 discusses proposed work, Implementation and results were discussed in Sect. 4, and conclusion is discussed in Sect. 5.

3 Proposed Work

3.1 Block Diagram of Health Center and Secure Communication Between Entities

For a secure sharing of medical information, AES and DES algorithms are used for communication between entities. Health center and communication between entities is given as follows (Fig. 4),

1. Registered user can log in into the system by using (username, password and IP address) combination.
2. New user can get himself registered and then login to the system.
3. Entities taking part in health center are broadly classified as,

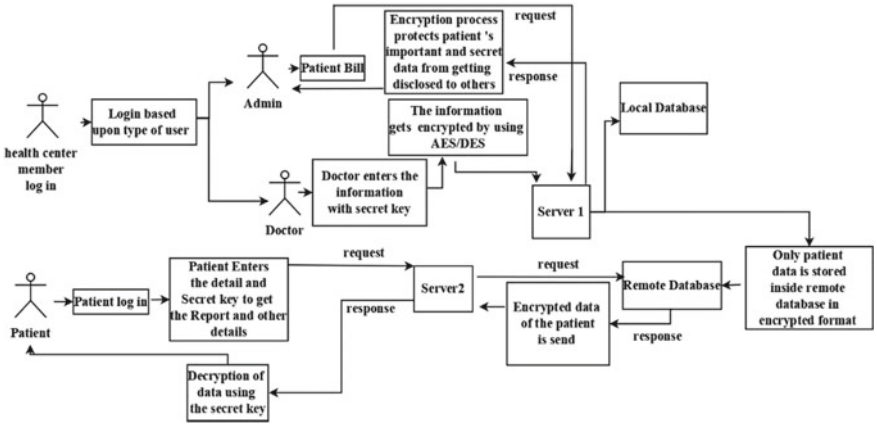


Fig. 4 Health center and secure communication between entities

a. Staff: After successful login, staff member can get information about workload assigned to them, the member can also get access to duty and shift allocation for the next day. No individually identifiable information is accessible by staff.
b. Admin: Roles and responsibilities of admin are, Patient bill, Staff Payroll, Doctors payment, Stock balance and entry. In Patient bill module the personal information, information related to reports, health condition and doctor's advice are encrypted. Only cost, payment and other charges included are visible and can be processed by the admin. This can be made possible by AES/DES technique and by providing access restriction.
c. Doctor: Doctor will get the module of Add Reports, The Add Reports module furthers has different module such as diagnosis, vital sign, Discharge Summary and patient history. All this information comes under protected health information, thus confidentiality is maintained.
d. Director: Director is provided with the modules, Add Department, Add wards, Add new recruits, Checkout transaction and status of Action on complains.

4. The remote database is to store patient data so that it would not be possible for the patient to visit the health center to checkout his report. Patient is able to find his/her report at any time and from anywhere.
5. If the patient wants to access to the reports stored in database, he must be an authenticated user and he should have a key.

3.2 Securing Communication in Health Center with AES and DES

(1) Doctor: login into the System with his (Username, Password and IP address) combination (Fig. 5).
(2) Doctor enters the data about the Patient into its System module, the Data will contain the Patients health reports, Medicine, Vital Sign etc. With this Data a Secret key is added.
(3) Encryption process of AES/DES is performed on the data using the secret key.
(4) After encryption the data is sent further to server were the data is stored on the local database and another copy of that data is stored on remote database.
(5) Patient logs in through his device.
(6) Patient enters his details with the secret key and a request is sent to the server.
(7) Server then retrieves the encrypted data in response.
(8) The encrypted data is sent to the patients' device were the data decryption process is performed using AES/DES.
(9) Patient gets the requested data in the form of response.

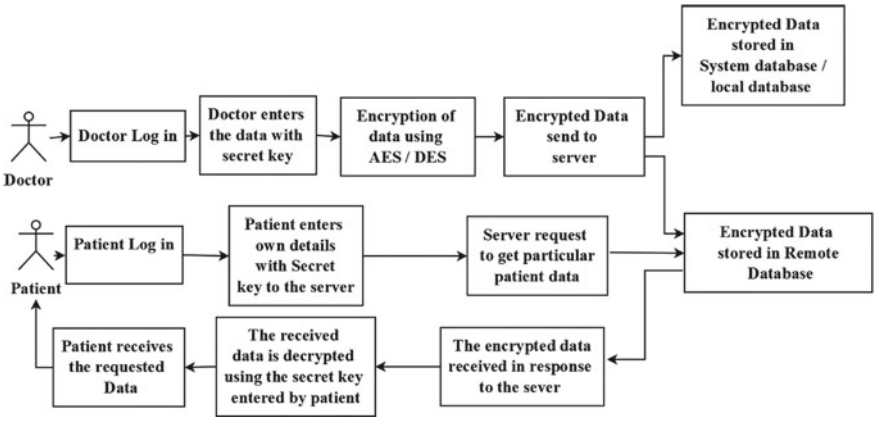


Fig. 5 Securing communication in health center with AES and DES

4 Implementation and Results

This section discusses the implementation of secure sharing of data in health center using AES and DES. It mainly discusses securing the communication between parties in the system by using AES and DES. Figure 6 shows the login screen of the system in which the entities can login using username, password and IP address. Person can login as doctor, staff, director and admin. In case the user forgets the password, he or she can reset the password using forgot password option. New user can register by providing the details asked by the system.

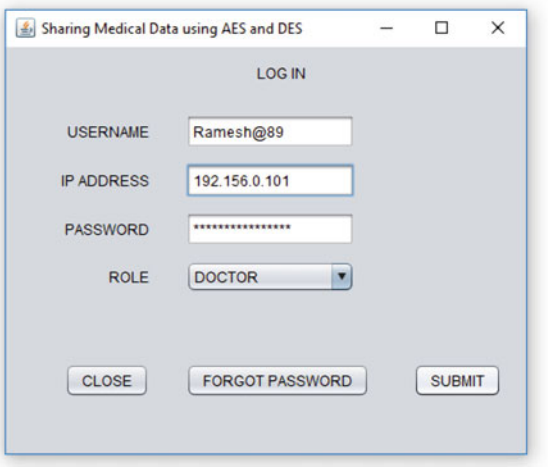


Fig. 6 Login screen of sharing medical data using AES and DES

Consider an example where the party A wants to communicate medical record with party B. It will encrypt and upload using advanced encryption standard algorithm as shown in Fig. 7.

Similarly the data to be sent can be encrypted using the Data Encryption standard algorithm as shown in Fig. 8. User can browse the medical record to be sent and choose the encryption algorithm as DES and then upload and send.

Further comparison is made between AES and DES algorithm, As the size of data increases the time to encrypt also increases, the pattern of change in time required

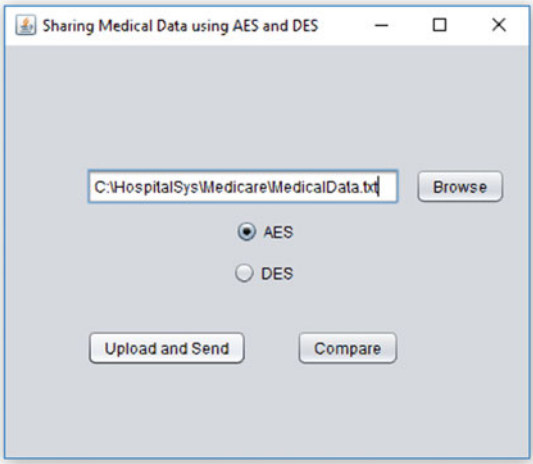


Fig. 7 Encryption of medical reports using AES algorithm

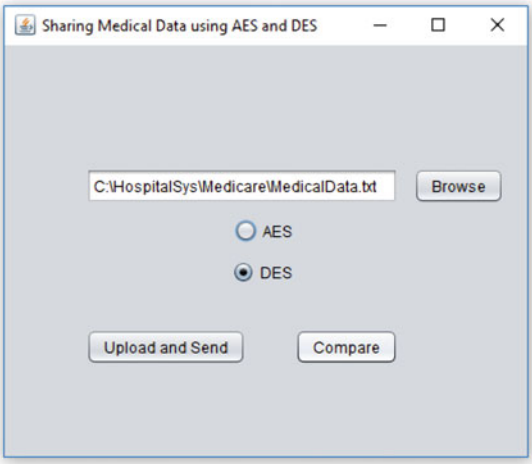


Fig. 8 Encryption of medical reports using DES

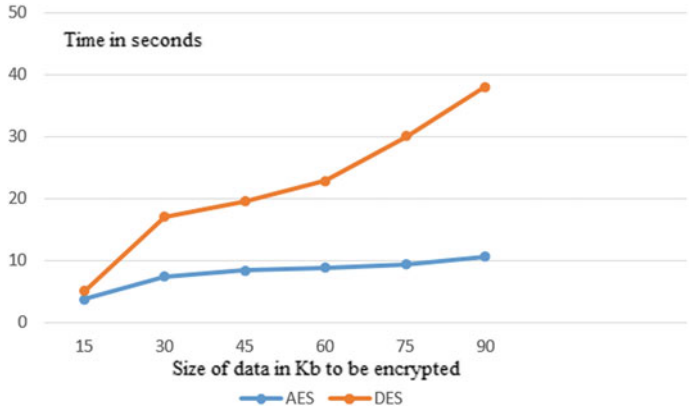


Fig. 9 Comparison between AES and DES (size of data in Kb versus time in seconds)

for encryption was observed and the time required for encryption (in seconds) is represented by y-axis and size of data to be encrypted (in kb) is represented by x-axis. For 90 kb data, time to encryption was 38 s for DES, whereas encryption time was 11.5 s for AES (Fig. 9).

5 Conclusion and Future Scope

Secure way of communication between the entities in health center was proposed in this paper. Communication in the health center is briefly explained and the way to secure the transmitted messages and data was proposed. Encryption was not only limited to text messages but the medical prescription, pathology reports, blood test report to name the few can be encrypted. For encryption two algorithms namely AES and DES were used, further the comparison was made between the two algorithms for time required to encrypt the same size data. For 90 kb data the time required for encryption by DES algorithm was 38 s however for the same size data time to encrypt by AES algorithm was 11.5 s.

The proposed work can be extended to encrypting the communication between entities and individually identifiable information by using public key cryptographic methods like RSA algorithm and Elliptic curve cryptography, which require lower key sizes. Comparison of encryption using symmetric and asymmetric algorithms can be made as an extension to this work. Also level of security achieved using the two types of algorithm (symmetric and asymmetric) can be compared. Network performance and ability to withstand various attacks can be discussed.

References

1. M.E. Dusek, Medical information card. U.S. Patent No. 5,171,039. 15 Dec. 1992
2. Assistance, HIPAA Compliance. Summary of the HIPAA Privacy Rule (2003)
3. S.T.G. Wong, A cryptologic based trust center for medical images. J. Am. Med. Inform. Assoc. **3**(6), 410–421 (1996)
4. D. Coppersmith, The data encryption standard (DES) and its strength against attacks. IBM J. Res. Dev. **38**(3), 243–250 (1994)
5. P.P. Dang, P.M. Chau, Image encryption for secure internet multimedia applications. IEEE Trans. Consum. Electron. **46**(3), 395–403 (2000)
6. E. Biham, A. Shamir, *Differential Cryptanalysis of the Data Encryption Standard* (Springer Science & Business Media, 2012)
7. M. Matsui, The first experimental cryptanalysis of the data encryption standard, in *Annual International Cryptology Conference* (Springer, Berlin, Heidelberg, 1994)
8. M.-L. Akkar, C. Giraud, An implementation of DES and AES, secure against some attacks, in *International Workshop on Cryptographic Hardware and Embedded Systems* (Springer, Berlin, Heidelberg, 2001)
9. NIST, AES. Advanced encryption standard. FIPS Publication 197 (2001)
10. P. Chown, Advanced encryption standard (AES) ciphersuites for transport layer security (TLS). No. RFC 3268 (2002)
11. H. Santhi et al., Study of symmetric-key cryptosystems and implementing a secure cryptosystem with DES, in *Information Systems Design and Intelligent Applications* (Springer, Singapore, 2019), pp. 299–313
12. A. Gohr, S. Jacob, W. Schindler, CHES 2018 side channel contest CTF–solution of the AES challenges, January 30, 2019 (2019)
13. S. Gupta, S. Jain, M. Agarwal, DWSA: a secure data warehouse architecture for encrypting data using AES and OTP encryption technique, in *Soft Computing: Theories and Applications* (Springer, Singapore, 2019), pp. 505–514
14. J. Chen et al., Medical image cipher using hierarchical diffusion and non-sequential encryption. Nonlinear Dyn. 1–22 (2019)
15. P. Prabhu, K.N. Manjunath, Secured transmission of medical images in radiology using AES technique, in *Computer Aided Intervention and Diagnostics in Clinical and Medical Images* (Springer, Cham, 2019), pp. 103–112

A Comprehensive Review on Unsupervised Feature Selection Algorithms



Anala A. Pandit, Bhakti Pimpale and Shiksha Dubey

1 Introduction

Multivariate attributes and high dimensionality are the challenging issues of the data sets available for analysis as everyone is trying to gather as much data in as many forms as possible. Since the features in high dimensional data are highly correlated, it deteriorates the execution of the algorithm. Feature selection [1–4] and feature extraction [5] are two of the ways to address this issue. Feature selection is capable of choosing a small subset of relevant features from the original ones by removing noisy, irrelevant, and redundant features [6]. There are various feature elimination/feature extraction techniques available to eliminate the unnecessary attributes and to select relevant attributes that would contribute to the output. Feature selection methods are generally categorized as three main groups including filters, wrappers, and embedded methods [7]. Filter methods select the most relevant features, i.e., the feature is selected based on their intrinsic properties without using any clustering algorithm [8] whereas wrapper methods involve selecting set of features as search problem; their various combinations are prepared and assessed. Last method, the embedded method, identifies the best feature that would contribute to the accuracy of the model while the model is being developed. The optimality of a feature subset is measured by an evaluation criterion [9]. In this paper, we are trying to evaluate if there is suitability between the type of data sets and algorithms based on internal and external

A. A. Pandit (✉) · B. Pimpale · S. Dubey
Veermata Jijabai Technological Institute, Matunga, Mumbai, India
e-mail: aapandit@mc.vjti.ac.in

B. Pimpale
e-mail: bspimpale_p18@mc.vjti.ac.in

S. Dubey
e-mail: sadubey_p18@mc.vjti.ac.in

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_24

evaluation measures. The purpose is to apply feature selection and reduction algorithms to unsupervised data sets and check their efficiency and performance on the basis of different parameters. For data sets without any class labels, mainly clustering algorithms like K-means [10] can be applied. A set of variety of data sets are chosen for these experiments, consisting of continuous/discrete data values, image/text, and binary data. Ideal feature selection/extraction algorithm is identified and selected for each type of data set. The organization of rest of the paper is as follows: Sect. 2 elaborates on the related work, the experimental results and discussion are in Sect. 3, and conclusion and future work are given in Sect. 4.

2 Related Work

To improve the quality of clustering, the reduction in data is useful. To reduce the data set, various data reduction techniques are used. Three main types of data reduction methods are *Dimensionality Reduction* makes sure that the number of attributes in the data sets is reduced. In *Numerosity Reduction* techniques, original data is replaced by an alternative smaller representation of data, and *Cardinality Reduction* where reduced representation of original data is obtained by applying transformations [11]. One of the widely used methods to eliminate irrelevant features is Dimensionality reduction that can be achieved using either feature extraction or feature selection techniques. Some algorithms used to achieve these are described below.

2.1 Methods for Feature Selection

In this method, the numbers of features are reduced; however, the features are used in their original form. Following are some of the feature selection techniques.

Low Variance. The underlying idea in this method is that if a feature is constant (i.e., it has 0 variance), then it cannot be used for finding any interesting patterns and can be removed from the data set. Consequently, a heuristic approach to feature elimination is to first remove e features whose variance is below some (low) threshold [12].

Laplacian score (LS). He et al. [2] proposed an algorithm called Laplacian score (LS). A Laplacian score is calculated for all features to reflect its locality preserving power, which indicates how individual feature preserves similarity between data and adjacent instances in the graph.

Spectral Feature Selection (SPEC). Zhao and Liu [13] proposed a novel algorithm of spectral feature selection for both supervised and unsupervised learning, which assist the combined study of supervised and unsupervised feature selection and the function is based on general similarity matrix.

Multi-Cluster Feature Selection (MCFS).

Cai et al. [1] proposed an algorithm named Multi-Cluster Feature Selection (MCFS), which identifies a smaller set of features, using a method similar to spectral clustering, such that the multi-cluster structure of the data is conserved. Highly scored features are selected based on their calculated feature coefficient.

Unsupervised Discriminative Feature Selection (UDFS). Yang et al. [3] suggested a new unsupervised feature selection algorithm that works in batch mode to select different features. It performs mutual feature analysis and utilizes distinct information and local structure of data distribution simultaneously.

Unsupervised Feature Selection Using Non-negative Spectral Analysis (NDFS). Li et al. [4] proposed a new unsupervised feature selection algorithm that mutually derives nonnegative spectral analysis and feature selection, in which features are selected by a combined framework of nonnegative spectral analysis and l_2, 1-norm regularized regression.

There are many more algorithms like Unsupervised forward orthogonal search (FOS) [14] and Variance Minimization Criterion to Feature Selection Using Laplacian Regularization (LapAOFS and LapDOFS) [15]. We will focus only on the algorithms mentioned above.

2.2 Methods for Feature Extraction

Principal Component Analysis. Karl Pearson invented PCA in 1901 [16]. PCA is an unsupervised machine learning algorithm that aims to reduce the number of features within the data set while maintaining the variability of the data. It discovers a new set of features called components. PCA picks the projection with the most elevated difference [17]. These are a combination of original features, which are uncorrelated with one another.

Various other algorithms like Singular Value Decomposition [18], Linear Discriminant Analysis [19, 20], and Canonical Correlation Analysis [21] are also available; however, we will evaluate only PCA as that is the most popular algorithm.

2.3 Evaluation Metrics

As mentioned earlier, there are two types of evaluation metrics related to clustering: internal and external. External evaluation depends on prior knowledge (for example, classes of the data) of the data. Internal evaluation depends on the intrinsic detail of the data alone.

Internal Clustering Evaluation Metrics. The basic objective of clustering is to group closely related objects together. Objects not similar to each other should be grouped in different clusters. To evaluate the quality of these clusters two main Internal Evaluation metrics are used, i.e., *Compactness* and *Separation* [1, 22].

Compactness. This parameter evaluates how similar the objects are to each other in a given cluster. It is measured using various methods like distance between center and average or farthest distance of the point; or pairwise maximum/average distance between two points [22].

Separation. Dissimilar objects should be in different clusters for good quality clustering. This means that clusters that contain dissimilar objects must be well separated. The distance between clusters is found using various methods like minimum distance between centroids or minimum distance between objects in different clusters [22].

Calinski Harabasz index (CH). CH evaluates the cluster validity based on the ratio between the within-cluster dispersion and the between-cluster dispersion. Clustering quality is better if the value of CH index is high [23].

Davies–Bouldin index (DB). The goal of DB index [24] is to identify sets of clusters that are compact and well separated. Based on the quantities and features inherent to the data set the evaluation metric identifies how well the clustering is done. The lower value of the index indicates better clustering result.

Silhouette index (S). The silhouette index value determines how cohesive a specific cluster's cluster elements are as compared to cluster elements from other cluster's cluster elements. The value is between -1 and $+1$. A higher positive value indicates better clustering outcome [25].

Sum of Squared Errors (SSE). SSE is the sum of the squared differences between each cluster element and the mean of all the elements in the cluster. It is a measure of how distinct that cluster is from all other clusters. If every object in the cluster is same then SSE should be zero. Smaller the SSE, the clusters are compact [26].

Dunn index. The Dunn Index (DI) [27–29] is a measure for evaluating a feature selection/extraction algorithm. A higher value of DI indicates good clustering output. The computation cost of DI is directly proportional to the number of dimensions.

R-squared index. R-squared is also known as the coefficient of determination which is nothing but how well a model explains and predicts future outcomes and is also used to compare the goodness of two or more models. The value of R-squared is between 0 and 1, where 1 indicates an exact fit, and 0 indicate that the model fails to accurately model the data [30].

2.4 External Clustering Evaluation Metrics

Accuracy. This metric determines how accurately the cluster label is as compared to the original label of the data set [31, 32]. The range of accuracy is from 0 to 1. Accuracy should be close to 1 if well-separated and compact clusters are formed.

Normalized Mutual Information (NMI). NMI is a measure for determining the quality of clustering. It is an external measure because we need the class labels of the instances to determine the NMI. Normalized mutual information is a normalization of the mutual information (MI) score to scale the results between 0 (no mutual information) and 1 (perfect correlation) [31, 33].

Table 1 Description of the data set

Domain/Type	Data set	# of features	# of classes
Image	COIL20	1024	20
	ORL	1024	40
Text (binary)	Basehock	4862	2
	Relathe	4322	2
Biomedical/Multivariate	Lung	3312	5
	Lymphoma	4026	9

Rand index. Rand index [31, 33] is used to measure how similar two data clusters are. Rand Index ranges from 0 to 1 where 0 is used to indicate that two data clusters are different and 1 implies that the clusters are the same.

3 Experiments

For our analysis, we have used six publicly available data sets. The experiments are performed on an Intel(R) Core(TM) i7-7500 CPU @ 3.40 GHz with 16.00 GB memory and 64-bit, x64-based processor, Windows 10 operating system.

3.1 Data Sets

The data sets used to perform these sets of experiments are six publicly available data sets. To ensure a variety of data types, we have used two data sets containing images (COIL20, ORL), two data sets containing text (Basehock, Relathe), and two multivariate data sets from biomedical domain (Lung, Lymphoma), Table 1 outline the description of these data sets.

3.2 Experiment Setting

To determine the optimal value of the number of features to select, we conducted the experiments once with different number of features to retain, viz., 50, 100, 150, 200 and 250,300,350. For these selected features, the k-means algorithm was repeated 20 times with random initialization for validating the results. Average values were calculated for each of these runs and are presented in the next section.

3.3 Results and Discussion

Following Figs. 1, 2, 3, 4, and 5 show the experimental results for different number of features retained for 'COIL20' data set. Figure 1 provides the output result of Rand Index. The results of Dunn Index are shown in Fig. 2. Figure 3 presents the output results for Silhouette Index. The results of DB index and sum of squared error are shown in Figs. 4 and 5, respectively.

The observations are consistent for the remaining data sets for different set of features but due to space constraints, we have provided results for only 'COIL20' data set on all algorithms. Based on the results of internal and external evaluation metrics performed, the number of features retained was found to be 100. As observed from Figs. 1, 2, 3, 4, and 5, the optimal number of features are 100 for most of the external and internal evaluation metrics, and hence the remaining experiments are performed by retaining 100 features. We compared the execution time of PCA, UDFS, NDFS, SPEC, Lap score, Low variance, and MCFS on all six data sets as shown in Fig. 15.

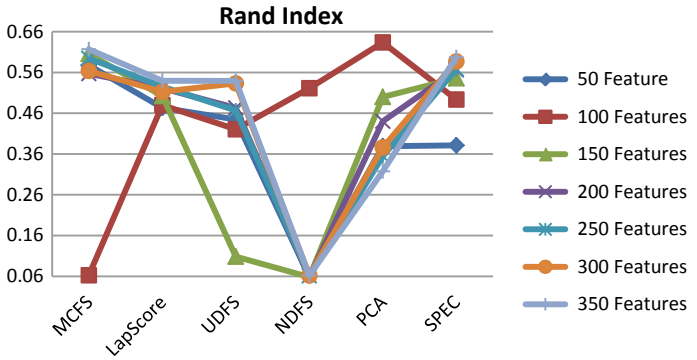


Fig. 1 Rand index (higher the better)

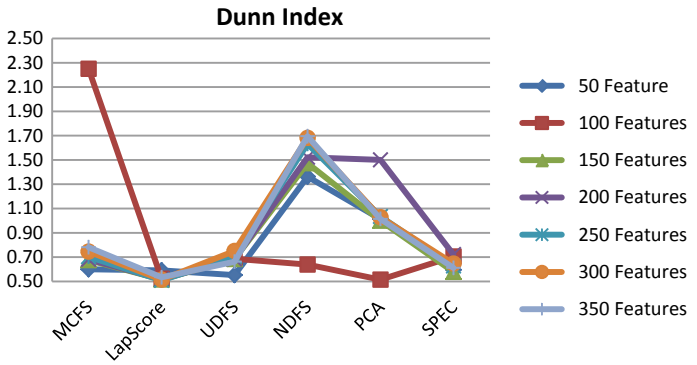


Fig. 2 Dunn index (higher the better)

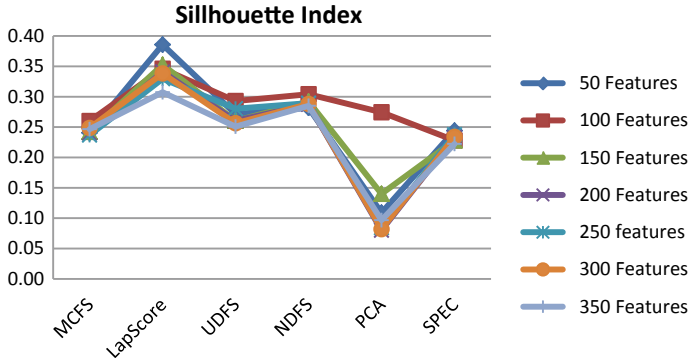


Fig. 3 Sillhouette index (higher the better)

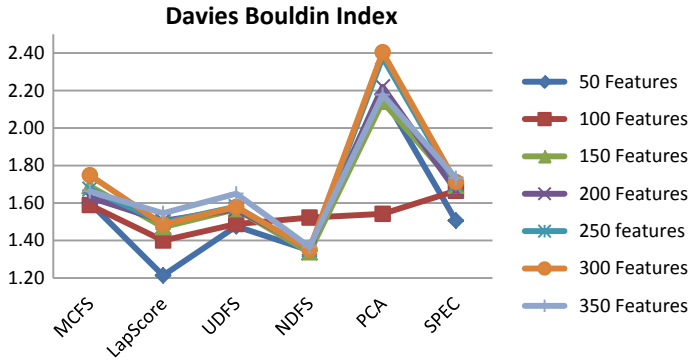


Fig. 4 Davies Bouldin index (lower the better)

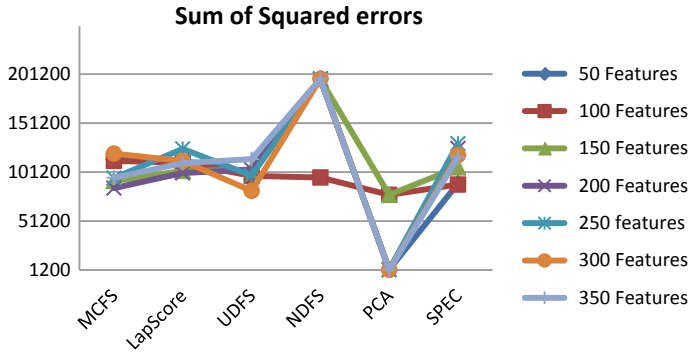


Fig. 5 Sum of squared errors (lower the better)

NMI						
DATASET	COIL20	ORL	LUNG	LYMPHOMA	BASEHOCK	RELATHE
UDFS	0.6454	0.701	0.37	0.5732	0.00633	0.0021
MCFS	0.7337	0.75	0.6	0.66	0.006339	0.0002
SPEC	0.7098	0.74	0.3	0.451	**0.024**	0.011
NDFS	0.7542	0.7581	0.52	**0.712**	0.0056	**0.056**
LOW VARIENC	0.72999	0.778	0.61	0.692	0.00633	0.0087
LAP SCORE	0.71215	0.696	0.46	0.562	0.02183	0.008763
PCA	**0.7864**	**0.7812**	**0.6429**	0.7042	0.00633	0.0087

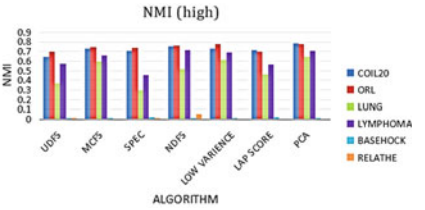


Fig. 6 NMI of all algorithms

Given below are the following set of figures based on the evaluation parameters and a comparative graph has plotted along with for data.

In the tables of Figs. 6, 7, 8, 9, 10, 11, and 12, the higher value of the output implies a better performance and hence is highlighted. Figure 6 shows the output results of NMI parameter where it measures the entropy. Figure 7 shows the output

ACCURACY						
DATASET	COIL20	ORL	LUNG	LYMPHOMA	BASEHOCK	RELATHE
UDFS	0.55	0.483	0.5133	0.4796	0.5	0.5444
MCFS	0.65	0.554	0.61	0.58	0.5	0.5388
SPEC	0.58	0.55	0.46	0.451	***0.569***	0.5439
NDFS	0.64	0.5641	0.57	***0.626***	0.52	0.525
LOW VARIENC	0.62	0.59	0.68	0.607	0.5007	***0.5466***
LAP SCORE	0.5501	0.481	0.613	0.503	0.5052	0.546
PCA	***0.6811***	***0.5975***	***0.7233***	0.6192	0.5007	***0.5466***

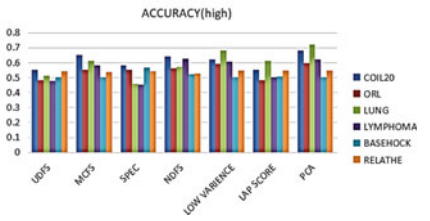


Fig. 7 Accuracy of all algorithms

RAND						
DATASET	COIL20	ORL	LUNG	LYMPHOMA	BASEHOCK	RELATHE
UDFS	0.4207	0.313	0.1693	0.2277	-5.022E-06	-4.2E-05
MCFS	0.0625	0.0189	0.32866	0.3851	1.6241E-05	-0.00206
SPEC	0.4937	0.3769	0.1538	0.1168	***0.0163***	-0.00091
NDFS	0.5214	0.4411	0.3555	0.3144	0.00918	***0.00058***
LOW VARIENC	0.5182	0.4575	0.4275	***0.4038***	-5.022E-06	0.00028
LAP SCORE	0.4797	0.2979	0.3093	0.2208	7.7652E-05	0.000283
PCA	***0.6337***	***0.4923***	***0.5614***	***0.4038***	-5.022E-06	0.000283

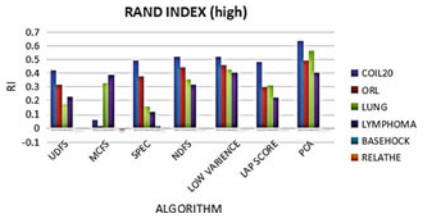


Fig. 8 Rand index of all algorithms

SILLHOUETTE						
DATASET	COIL20	ORL	LUNG	LYMPHOMA	BASEHOCK	RELATHE
UDFS	0.293	0.1193	0.094	0.084	0.953	0.9256
MCFS	0.26	0.162	***0.129***	0.072	***0.966***	0.681
SPEC	0.228	0.133	0.08	0.0582	0.2754	0.885
NDFS	0.304	0.148	0.101	***0.125***	0.245	0.45
LOW VARIENCE	0.285	0.137	0.098	0.074	0.944	0.9415
LAP	***0.346***	0.132	0.118	0.05	0.8211	***0.973***
PCA	0.2743	***0.1751***	0.1136	0.074	0.9533	0.9501

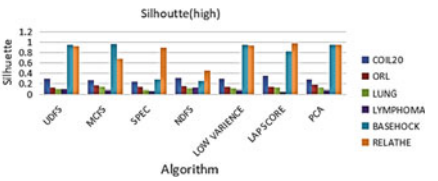


Fig. 9 Silhouette index of all algorithms

DUNN INDEX						
DATASET	COIL20	ORL	LUNG	LYMPHOMA	BASEHOCK	RELATHE
UDFS	0.6864	0.6008	1.013	0.9962	1.03	1.0599
MCFS	***2.2501***	***1.7149***	1.0209	***1.2375***	1.6049	***4.9159***
SPEC	0.7047	0.67	***1.2136***	1.0325	***6.4526***	2.2909
NDFS	0.6387	0.7586	0.4981	0.8815	3.125	2.98
LOW VARIENCE	0.6202	0.6484	0.8056	1.0293	1.037	1.0361
LAP	0.5217	0.4424	0.77	0.9354	3.0822	1.03019
PCA	0.5147	0.6221	0.8027	1.0293	1.0335	1.0399

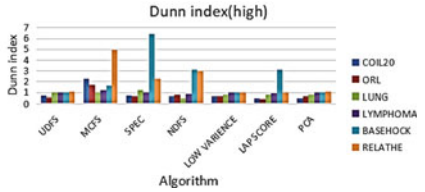


Fig. 10 Dunn index of all algorithms

CALINSKI HARBASZ						
DATASET	COIL20	ORL	LUNG	LYMPHOMA	BASEHOCK	RELATHE
UDFS	240.71	15.202	17.068	6.2194	673	401.472
MCFS	177.263	***21.272***	19.39	5.088	***691***	232.33
SPEC	155.629	18.362	14.8111	6.237	261.4692	275.95
NDFS	189.591	17.264	20	***6.66***	568.123	879.34
LOW VARIENCE	250.127	15.0653	21.95	5.35	448.177	372.42
LAP	***312.638***	19.637	***27.7***	5.0853	128.881	***966.562***
PCA	165.55	17.818	23.7196	5.3584	545.56	418.638

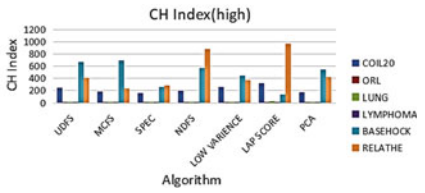


Fig. 11 CH index of all algorithms

r Squared						
DATASET	COIL20	ORL	LUNG	LYMPHOMA	BASEHOCK	RELATHE
UDFS	0.101	0.0931	0.1185	**0.172**	**0.5004**	0.0924
MCFS	**0.2049**	**0.2023**	0.0868	0.0485	0.4983	0.4572
SPEC	0.0923	0.11	0.1475	0.1082	0.4303	0.0921
NDFS	0.1256	0.1157	0.1666	0.1496	0.312	**0.493**
LOW VARIENCE	0.1283	0.1016	0.2112	0.0918	**0.5004**	0.476
LAP SCORE	0.1155	0.088	0.0975	0.0762	0.5	0.47605
PCA	0.0814	0.114	**0.251**	0.0918	**0.5004**	0.476

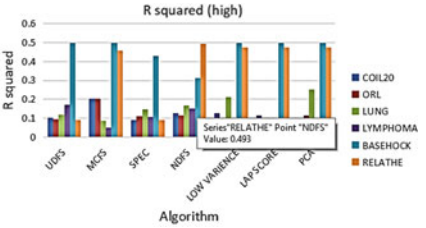


Fig. 12 R-squared index for all algorithms

results of Accuracy parameter. The output results of Rand Index are presented in Fig. 8. The results of Silhouette Index that indicates how well separated the clusters are shown in Fig. 9. The output results of Dunn Index which indicate quality clusters for each data set is shown in Fig. 10. Figure 11 shows the output results of CH Index for each data set which can help us decide on optimal number of clusters. Figure 12 shows the output results of R-Squared index. Since the lower value indicates a better performance in Figs. 13, 14, and 15 the lowest values have been highlighted. Figure 13 shows the output results of DB Index that indicates the features inherent to the data set for good quality clustering. The output results of SSE that indicate the compactness of the cluster are shown in Fig. 14.

Based on the results presented in Figs. 6, 7, 8, 9, 10, 11, 12, 13, 14, and 15, we can make several observations and interpretations like choosing the right features and the right number of features that will aid to improve the quality of clusters obtained. From the results obtained, it can be said that PCA algorithm performs well for low-dimensional image data sets (‘Coil20’ and ‘Orl’, which on plotting the data distribution, were found to be continuous and linear in nature), whereas

DB index						
DATASET	COIL20	ORL	LUNG	LYMPHOMA	BASEHOCK	RELATHE
UDFS	1.488	1.8006	2.472	2.3478	0.031	0.9837
MCFS	1.589	***1.5966***	2.381	2.432	***0.0212***	1.661
SPEC	1.665	1.731	2.658	2.495	1.9592	0.8271
NDFS	1.522	1.684	2.52	2.129	1.568	0.0758
LOW VARIENCE	1.441	1.7392	2.5	2.491	0.0367	0.03818
LAP SCORE	***1.4***	1.7611	***2.08***	2.467	0.2293	***0.016***
PCA	1.5419	1.6247	2.264	***0.4911***	0.0299	0.0319

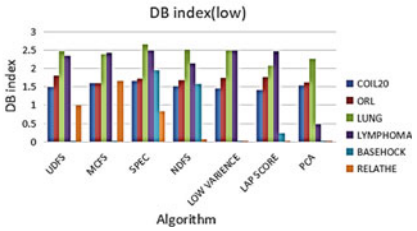


Fig. 13 DB index of all algorithms

SSE						
DATASET	COIL20	ORL	LUNG	LYMPHOMA	BASEHOCK	RELATHE
UDFS	97246	99312	655	1603	4987	654
MCFS	112628	98094	644	960	4987	3235
SPEC	88415	117287	815	1003	***4288***	***652***
NDFS	95668	103845	609	1252	4897	3677
LOW VARIENCE	122884	108312	1167	851	4987	3368
LAP SCORE	110694	***93885***	***539***	***707***	4983	3368
PCA	***77976***	121525	1387	851	4987	3386

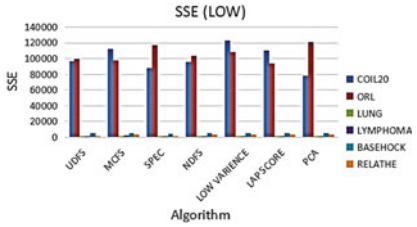


Fig. 14 SSE of all algorithms

Execution time						
DATASET	COIL20	ORL	LUNG	LYMPHOMA	BASEHOCK	RELATHE
UDFS	46.86	7.312	351.61	2283.89	213.768	205.55
MCFS	14.06	5.188	2.414	1.543	9.2032	5.749
SPEC	24.66	4.58	1.323	0.82	151.62	69.257
NDFS	38.17	15.696	221.94	324.48	149.561	74.561
LOW VARIENCE	8.9	16.315	3.73	5.2055	8.6722	7.6541
LAP SCORE	5.858	3.513	***0.92***	0.6601	9.3296	4.735
PCA	***5.3726***	***3.423***	1.054	***0.6404***	***4.941***	***3.2051***

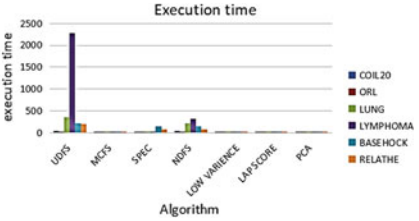


Fig. 15 Execution time of all algorithms

for high dimensional binary and text data sets, its performance degrades. Also, the unsupervised feature selection algorithms, namely, Laplacian Score, and NDFS work well on high dimensional, binary data sets like 'Lymphoma', 'Lung', and 'Relathe'. The SPEC algorithm performs best on text data set ('Basehock').

For image data sets, (continuous distribution) evaluation metrics like Normalized Mutual Information (NMI), Accuracy, Rand Index, Sum of Squared Error and Execution Time can be suitably applied. For text data sets (discrete distribution) NMI, Accuracy, Rand Index, Dunn Index, Sum of Squared Error, and R-Square evaluation metrics are appropriate. Last, for multivariate data sets (discrete distribution) NMI, Accuracy, CH Index, DB Index, and Silhouette prove to be beneficial. Depending on the type of data set different evaluation metrics should be selected appropriately.

4 Conclusion and Future Work

We have compared different unsupervised algorithms on six different data sets and concluded that for most of the algorithms, internal as well as external evaluation metrics give optimum values when one retains 100 features. The Image data sets like 'COIL20' and 'ORL', which are continuous in nature; PCA algorithm performs well. MCFS provides good results only for 'ORL' data set. Further, for multivariate data set like 'Lung' Laplacian score works well on different evaluation metrics and on the data set 'Lymphoma', NDFS works better. For text data sets like 'BASEHOCK', SPEC works the best and for another text data set 'Relathe' NDFS and Laplacian score works better. However, SSE, Silhouette Index, NMI, and Accuracy are good performance evaluation metrics for any type of data. We also conclude that knowing your data and appropriately using the tools for the task at hand is important.

For further work, alternative algorithms for feature extraction can be applied to extend the conclusions on different varieties of data sets using additional evaluation metrics.

References

1. D. Cai, C. Zhang, X. He, Unsupervised feature selection for multi-cluster data, in *16th ACM SIGKDD International Conference on Knowledge Discovery And Data Mining on Proceeding* (ACM, Washington, DC, USA, 2010), pp. 333–342
2. X. He, D. Cai, P. Niyogi, Laplacian score for feature selection, in *18th International Conference on Neural Information Processing Systems* (ACM, Canada, 2005)
3. Y. Yang, T. Shen, Z. Ma, Z. Huang, X. Zhou, l2, 1-norm regularized discriminative feature selection for unsupervised learning, in *22nd International Joint Conference on Artificial Intelligence on Proceeding* (AAAI, Barcelona, Spain, 2011)
4. Z. Li, Y. Yang, J. Liu, X. Zhou, H. Lu, Unsupervised feature selection using nonnegative spectral analysis, in *26th AAAI Conference on Artificial Intelligence on Proceeding* (AAAI, Canada, 2012)
5. C.O.S. Sorzano, J. Vargas, A. Pascual Montano, A survey of dimensionality reduction techniques. National Centre for Biotechnology (CSIC), ArXiv (2014)
6. J. Miaoa, L. Niub, A survey on feature selection. Elsevier **91**, 919–926 (2016)
7. S. Kashef, H. Nezamabadi-pour, B. Nikpour, Multilabel feature selection: a comprehensive review and guiding experiments. Wiley Period. **8**(2) (2018)
8. S. Solorio-Fernandez, J.A. Carrasco-Ochoa, J.F. Martinez-Trinidad, A review of unsupervised feature selection methods. Artif. Intell. Rev. 1–42 (2019)
9. L. Yu, H. Liu, Toward integrating feature selection algorithms for classification and clustering. IEEE Trans. Knowl. Data Eng. **17**(4), 491–502 (2005)
10. J. Wu, *Advances in K-means Clustering: A Data Mining Thinking* (Springer, 2012)
11. Soft computing and intelligent information system homepage. https://sci2s.ugr.es/sites/default/files/files/…/Cap6%20-%20Data%20Reduction.ppt. Last accessed 1 Apr 2019
12. Chris Albon homepage. https://chrisalbon.com/machine_learning/feature_selection/variance_thresholding_for_feature_selection/. Last accessed 1 Apr 2019
13. Z. Zhao, H. Liu, Spectral feature selection for supervised and unsupervised learning (SPEC), in *24th International Proceedings of the Conference on Machine Learning 2007* (ACM, USA, 2007), pp. 1151–1157

14. H.L. Wei, S.A. Billings, Feature subset selection and ranking for data dimensionality reduction. IEEE Tran. Pattern Anal. Mach. Intell. **29**(1), 162–166 (2007)
15. X. He, M. Ji, C. Zhang, H. Bao, A variance minimization criterion to feature selection using Laplacian regularization. IEEE Trans. Pattern Anal. Mach. Intell. **33**(10), 2013–2025 (2011)
16. K. Pearson, On lines and planes of closest fit to systems of points in space. Philosophical magazine. Lond., Edinb. Philos. Mag. J. Sci. **2**(11), 559–572 (1901)
17. R. Ravi Kumar, M. Babu Reddy, P. Praveen, A review of feature subset selection on unsupervised learning, in *3rd IEEE International Conferences on Advances in Electrical, Electronics, Information, Communication and Bio-Informatics (AEEICB17)* (IEEE, Chennai, India, 2017)
18. K. Modarresi, Unsupervised feature extraction using singular value decomposition, in *ICCS 2015 International Conference on Computational Science*, vol. 51 (Elsevier, USA, 2015), pp. 2417–2425
19. R.O. Duda, P.E. Hart, D. Stork, *Pattern Classification*, 2nd edn. (Wiley, USA, 2000)
20. K. Fukunaga, *Introduction to Statistical Pattern Classification*, 2nd edn. (Academic Press, San Diego, California, USA, 1990)
21. H. Hotelling, Relations between two sets of variates. Biometrika **28**(3–4), 321–377 (1936)
22. P.N. Tan, M. Steinbach, V. Kumar, *Introduction to Data Mining*, 1st edn. (Addison-Wesley Longman Inc, USA, 2005)
23. E. Rendon, I. Abundez, C. Gutierrez, S. Zagal, A. Arizmendi, E. Quiroz, H. Elsa Arzate, A comparison of internal and external cluster evaluation indexes, in *5th WSEAS International Conference Proceeding on Computer Engineering and Applications and 11th Proceeding of the American Conference on Applied Mathematics* (ACM, USA, 2011), pp. 158–163
24. T. Calinski, J. Harabasz, A dendrite method for cluster analysis. Commun. Stat. **3**(1), 1–27 (1974)
25. P.J. Rousseeuw, Silhouettes: a graphical aid to the interpretation and evaluation of cluster analysis. J. Comput. Appl. Math. **20**, 53–65 (2011)
26. W. Kwedlo, A clustering method combining differential evolution with the k-means algorithm. Pattern Recognit. Lett. ACM **32**, 1613–1621 (2011)
27. J.C. Dunn, A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters. J. Cybern. **3**(3), 32–57 (1973)
28. J.C. Dunn, Well-separated clusters and optimal fuzzy partitions. J. Cybern. **4**(1), 95–104 (1974)
29. The comprehensive R archive network homepage. https://cran.r-project.org/web/packages/clv/clv.pdf. Last accessed 1 Apr 2019
30. C. Cameron, A.G. Frank, An R-squared measure of goodness of fit for some common nonlinear regression models. J. Econ. **77**(2), 329–342 (1997)
31. M. Halkidi, Y. Batistakis, M. Vazirgiannis, On clustering validation techniques. Intell. Inf. Syst. **17**(2–3), 107–145 (2001)
32. Binghamton University homepage. http://www.cs.binghamton.edu/~lyu/SDM07/DR-SDM07.pdf. Last accessed 1 Apr 2019
33. S. Aghabozorgi, A. Shirkhorshidi, T. Wah, Time-series clustering—a decade review. Inf. Syst. Elsevier **53**, 16–38 (2015)

On the Prerequisite of Coprimes in Double Hashing



Vivek Kumar

1 Introduction

Dictionary operations are pervasively used in large number of computer applications. Given its expected constant run time of $O(1)$ in searching, much of the research is devoted to the advancement of hashing [1] techniques. Collision is a major problem in hashing and over the years many collision resolution techniques have been developed. Open addressing is one of them which makes sure that all the elements are stored in the table itself. It does that by constantly probing the table for empty positions. Double hashing is an open addressing technique which uses two hash functions for hashing. An important consideration is that the second hash function $h_2(k)$ must be relatively prime to the hash table size m so that the hash table can be searched thoroughly. The subsequent sections will discuss the open addressing scheme and double hashing technique, and the final section will discuss the mathematical proof for the condition of the relative primality.

2 Open Addressing

In the open addressing [2] hashing scheme, all the elements are stored in the table itself. It implies that there are no chains that are used and each table slot contains either an element or NIL. To search for a particular element a series of probes are done. Thus, the number of elements that can be stored in the hash table is no more than the table size, which concludes that the load factor can never exceed 1. The operation of insertion is performed by a probing sequence. The probe looks for an empty position into which the new element is inserted. The order of probe is not

V. Kumar (✉)
THDC Institute of Hydropower Engineering and Technology, Tehri, Uttarakhand, India
e-mail: vivek9837@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_25

necessarily <0, 1, 2, 3, …, $m-1$> , but it depends on the element to be inserted. The extended hash function to include the probe sequence is given as

$$h : U \times \{0, 1, 2, \ldots, m-1\} = \{0, 1, 2, \ldots, m-1\} \tag{1}$$

The probe sequence is

$$<< h(k, 0),\ h(k,\ 1),\ h(k, 2), \ldots, h(k,\ m-1) >> \tag{2}$$

3 Double Hashing

One important consideration in open addressing is the possible number of probe sequences that can be generated by the open addressing scheme. The linear probing and quadratic probing produces m possible permutation of probe sequences. Even though an idealistic open addressing scheme must produce any of the $m!$ permutation, in practice such kind of scheme does not exist. Double hashing has the capacity of producing $\Theta(m)$ possible probe sequence, which even though is far less then $m!$ but still is better than linear probing and quadratic probing. As the name suggests the double hashing scheme uses two hash functions instead of one. The hash function is used in the following form:

$$h(k,\ i) = (h_1(k) + i h_2(k)) \bmod m \tag{3}$$

where both $h_1(k)$ and $h_2(k)$ are auxiliary hash functions and m is the table size. The initial probe starts from the table slot $T[h_1(k)]$ and the subsequent probes are offset from the initial probe position by an amount of $h_2(k)$ mod m. The value i ranges between 0, 1, 2, …, $m-1$. The probe sequence must possess the desired property that the entire table must be searched for a successful or unsuccessful search. Next section will discuss and prove mathematically that why the second hash function $h_2(k)$ and the hash table size m must be relatively prime for the condition that the table is searched in its entirety.

4 Prerequisite of Coprimes

As the previous section proposed that the second hash function $h_2(k)$ and the hash table size m must be relatively prime, we will discuss the inverse, i.e., what would have been the range of search, had $h_2(k)$ and m were non-coprime. Let us say that a gcd $d > 1$ exists for $h_2(k)$ and m. It is evident that if the entire table is not searched the search will loop to its initial position before completion of the table, which means that only a particular subsection of the table will be searched. The initial probe position

for $i = 0$ is $T[h_1(k)]$, if this position is already filled, the next probe position will be for $i = 1$, i.e.,

$$h(k, 1) = (h_1(k) + h_2(k)) \bmod m \tag{4}$$

which is $T[h(k, 1)]$ subsequently if this position is filled the next probe will be at position $T[h(k, 2)]$ and so on until there is a successful hit or an unsuccessful search is confirmed. Let us say if it is possible for the probe sequence to halt midway and loop back to initial probe, then for a probe sequence i

$$h_1(k) = h_2(k, i) \tag{5}$$

Mathematically,

$$h_1(k) = h_2(k, i) = (h_1(k) + ih_2(k)) \bmod m \tag{6}$$

which is only possible when

$$ih_2(k) = qm \tag{7}$$

for some $q > 1$. It was assumed that $h_2(k)$ and m were not coprime and had a gcd $d > 1$. Using the relation of gcd and lcm, $l > 1$ we have

$$h_2(k) \times m = l \times d \tag{8}$$

$$l = \frac{(h_2(k) \times m)}{d} \tag{9}$$

$l > 1$ must be equal to Eq. (7) for $ih_2(k)$ to be a multiple of m

$$ih_2(k) = \frac{(h_2(k) \times m)}{d} \tag{10}$$

$$i = \frac{m}{d} \tag{11}$$

which implies that from starting index $i = 0$, i will reach to index m/d before rolling back to the initial probe position. It means that only $1/d$th of the table will be searched for any unsuccessful search, which is certainly not desired. It may also be noted that had $h_2(k)$ and m were coprime, i.e., if $\gcd(h_2(k), m) = d = 1$ then from Eq. (7)

$$i = \frac{m}{d} = m \tag{12}$$

which implies that the entire table would have been searched.

5 Conclusion

Section 4 laid emphasis on the relative primality of the hash function $h_2(k)$ and m. It was also proved that if gcd of $h_2(k)$ and m was any $d > 1$ then $1/d$th of the table would have been probed leading to the failure of the scheme. Subsequently, it was proved that on the grounds of relative primality of $h_2(k)$ and m, all of the hash tables could have been probed. Hence, the requirement of relative of $h_2(k)$ and m has been necessitated.

References

1. H.P. Luhn, Keyword in context index for technical literature. Am. Doc. **11**(4), 288–295. (1960) ISSN: 0002
2. A.M. Tenenbaum, Y. Langsam, M.J. Augenstein, Data Structures Using C. Prentice Hall, pp. 456–461, p. 472 (1990) ISBN 0-13-199746-7

Multilingual Machine Translation Generic Framework with Sanskrit Language as Interlingua



Promila Bahadur

1 Introduction

1.1 Background

In this paper, we have presented a generic framework for machine translation targeted for multiple languages. The framework is responsible for generating translation of multiple target languages from single source language. To supplement this process, Sanskrit language is introduced as Interlingua apart from this we also need general world knowledge, domain-specific knowledge, i.e., specialized knowledge, context and cultural knowledge [1–4]. The Two-Way model was proposed for translation from source to target language in [5]. The EtranS software developed on the Two-Way model reported impressive success rate of about ninety percent for translation from English to Sanskrit language for simple and compound sentences [6–8]. Further, the observations from the result analysis of EtranS [9] signify that the Sanskrit provides a very robust and nonambiguous translation.

Two-Way ExT model for translation is proposed in the paper. This model illustrates three major benefits. First, the amount of work required to bridge the gap between languages (source and the target) is reduced. Second, the neutral representation of the source language is done which can be further mapped with multiple target languages. Last, to obtain accurate translation and nonambiguous translation Interlingua will be a pivot element.

The paper is divided into five sections. In Sect. 2, we outline essence and properties of Sanskrit language with illustration on structure of Sanskrit language and word formation process. Further, comparison is done between grammar of Sanskrit language and the Context-Free Grammar. In Sect. 3, the translation model 'Two-Way

P. Bahadur (✉)
SRM University, Lucknow 226028, UP, India
e-mail: promilabahadur@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_26

ExT' model is presented. Section 4 discusses the algorithmic steps and flowchart to carry out translation process. Section 5 discusses the conclusion and future scope.

2 Sanskrit Language as Interlingua

As the Sanskrit language is an ancient Indian language and mathematical in approach formulated by sages. The characteristics of Sanskrit language make it a strong contender for Interlingua. Sanskrit grammar is mechanical in approach. The mathematical approach satisfies the basic criteria for being presented as Interlingua. The detailed discussion is presented on Sanskrit grammar are discussed in Sect. 2.1. The general mechanical features are as follows:

A. It is a synthetic language, i.e., the language is not evolved by civilization. It is developed by sages like Panini in the year 2500 B.C.
B. It is calculative, mathematical, and mechanical in approach, as the foundation of Sanskrit grammar lies on four thousand sutras or formulas. New words can be formed with the help of these formulas. As shown in Table 1 there is a Sanskrit word for each of the modern-day inventions. The new words are derived from root words. The selection of root word is based on the characteristics of the product to be renamed.
C. Each new word is derived from a root[1] word and contains the embedded grammatical information. We can formulate this as follows word (new) formed = root

Table 1 Sanskrit translation for modern-day English words

English word	Sanskrit word
Adopter	संयोजिका
Caption	शीर्षकवाक्यम्
Motherboard	आधारफलकम्
Chip	सूक्ष्मस्मृतिका
Computer	सङ्गणकम्
Laptop	अङ्कसङ्गणकम्
Facebook	मुखपुस्तकम्
Twitter	अभिकुजकम्

[1]Root word is a name given to different nouns, verbs, etc. The nomenclature is based on basic characteristics of the object to be named. While deriving a new word, the properties of the object are taken into consideration.

Table 2 Showing different translation combinations

1	हरिः पुस्तकम् पठति
2	हरिः पठतिपुस्तकम्
3	पठतिहरिः पुस्तकम्
4	पठतिपुस्तकम् हरिः
5	पुस्तकम् हरिः पठति
6	पुस्तकम् पठतिहरिः

word + vibhakti. The property of 'vibhakti' is that it is added to the root word as per the grammatical/characteristics nature of the product for example book is considered as neutral gender, therefore, it is named as पुस्तकम्. The properties to be added decided on the basis of state, number, and action of the word.

D. Sanskrit is free from word ordering. This language does not bind the user to follow Subject–Verb–Object order strictly through SVO is desirable. Each word in Sanskrit has embedded grammatical information like preposition, number, action, etc. S-V-O order is compulsory for almost all of the evolved languages [1, 5, 10]. For example, "Hari reads a book" the translation in Sanskrit is "हरिः पुस्तकम् पठति". The translation can have following five different combinations apart from the given one and all of them have same meaning. We can also analyze that for three words in Sanskrit the combination of translated sentences could be 3!; i.e., six different sentences as shown in Table 2. Section 2.1 discusses the advance properties of Sanskrit grammar, which makes it a strongly typed language.

2.1 Structure of the Sanskrit Language and Word Formation in Sanskrit Language

Structure of the Sanskrit language and word formation in the language is given in the following steps.

A. New words are derived in Sanskrit on the basis of their properties. The words in Sanskrit are derived from root word, i.e., 'dhatushabd' (धातुशब्द)or root words are responsible for derivation of new words then based on formulas vibhakti is applied and a new word is formed. Sanskrit language is mature, comprehensive, and elegant. The primary structure of Sanskrit language is shown in Fig. 1.

B. While performing morphological analysis of the Sanskrit language, we can see that each word that is derived from a root word has characteristics of verb, form, tense, and number. Sanskrit words are self-sufficient and strongly

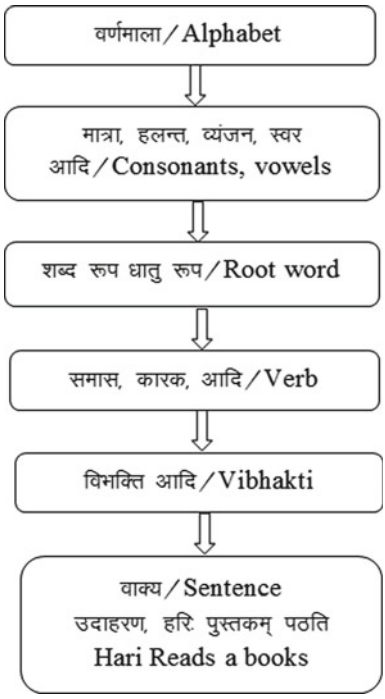


Fig. 1 Structure of Sanskrit language

typed, therefore, it can express themselves independently, e.g., sitA (सीता)depicts singular number, third person, feminine gender, and pratham vibhakti. gacCati (गच्छति)states present tense, first person, and singular. Therefore, gacCati (गच्छति)alone depicts 'is going' and sitAgacCati (सीतागच्छति)states 'Sita is going' or 'Sita goes'.

C. To specify verb ordering, in the Sanskrit language there are 27 words; that means for each of the three dhatu (tenses), there are nine forms. For example, 'read' has 27 forms of verb and each form is sufficient to portray the exact meaning.
D. Twenty-seven forms of noun are addressed in Sanskrit. The differences are primarily based on number and vibhakti (विभक्ति).[2] For example, noun '*Geeta'* can be written in 21 different forms, that take into account the different types of vibhakti (विभक्ति)(prathm, dritya (प्रथम्, द्वितीया), etc.) and the number (singular, dual, plural). The derived noun word carries information about its relation with ecosystem of sentence.
E. Singular, Dual, and Plural numbers can be addressed in Sanskrit.
F. Each word is strongly typed in Sanskrit language that makes it free from word order.

[2]Seven 'vibhakti' are provided by Sanskrit language that informs word with its action.

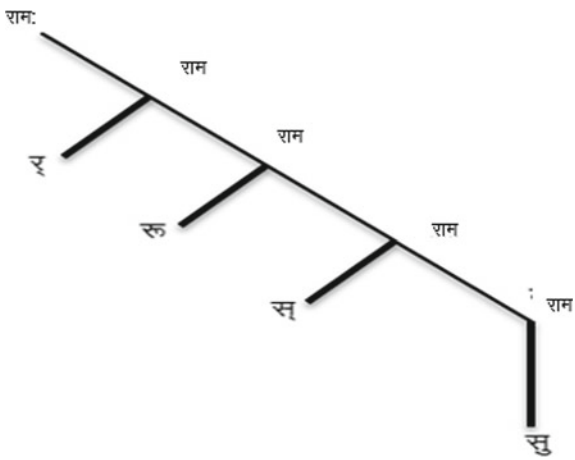


Fig. 2 Derivation tree of ramH

2.2 Advance Discussion Over Sanskrit Grammar

This section highlights the mathematical features of the Sanskrit grammar, which makes it free from word ordering. Panini Ahsthadhyayi gave mathematical model of grammatical description as shown in Fig. 2.

Panini method of verb forms is purely mechanical and unrelated of semantics. Panini proposed eight different relationships between two or more meaningful units; as discussed in the Table 3.

Sanskrit grammar has an approach for 'generative grammar', i.e., it is applicable for 'bottom-to-top' approach or 'beginning-to-end' procedure. The grammar is concerned more about correct usage of the relationship. The procedure of generative grammar is called integration. For example, word 'gamana' is built from root word 'gam' and suffix 'ana'.

Panini observed that meaning of A can be obtained by means of synonym of B and of B by means of C and so on. Panini observed that grammar should be focused on relationships among words and meanings can be relatively obtained. For example, words 'rajnahpurusah' can be written as 'rajapurusah' means 'king's man'. Karaka is a general name to deal with relational meaning. Relational meaning bridges the noun–verb relationship, i.e., it deals with Relational meaning their names and meanings related with noun–verb relations. The idea behind relational meaning is that it strengthens the noun in context to verbs so that each word in noun becomes strongly typed. This property of Sanskrit grammar makes each words carry its meaning therefore changing the order of the words does not affect the meaning of the sentence. Please be informed that Sanskrit grammar does not have concept of prepositions as we have in English grammar. The following properties of the karaka allow Sanskrit grammar to be free of word ordering. In fact, the karaka means doer. It is used to refer

Table 3 Showing different relationships and their definition

Relationship name	Definition	Example
Karaka (broadly divided into six categories)	It between finished word, i.e., verb and other non-verb	sHodanampacati. (he cooks rice)
Genitive	Is between two case-inflected words or between a word and a genitive case	For example,rajnahpurusah. (king's man)
Samanadhikarana or syntactic relationship	This relationship is marked by agreement in numbers, case, gender, person, etc.	nilahghatah 'a blue jar'
Mahavakya	This relationship is between two or more sentences forming compound or complex sentences	
Stem–suffix	The relationship between stem and root forming suffixes	Kumbhkarah (Potmaker)
Partitive	This relationship is between two inflected words or verbs	sHpathaticapacati (he studies and cooks)
Proverb–verb	This relationship is between proverb and verb	sHanugacchati (he goes after)
Upapada	The relationship between noun or verb with other word	pitrasaha (with father)

1. Any type of relationship meaning
2. Any item or person participating in an action
3. A word standing for such an item (Table 4).

Karaka is delimited to six varieties, which were enumerated. This is a very interesting factor to notice that concept of enumeration floated and structured several years back in the Sanskrit grammar. These names are used to refer

1. A particular type of relational meaning as further defined by rules in this section.

Table 4 Showing different translation combinations

S. no.	Karaka categories	Relationship	Example
1	Apadana	The fixed of relation to moving away	Odanampacati'he cooks rice'
2	Sampradana	The item one has in view through the karman	
3	Karana	The effective means	
4	Adhikarana	The location	
5	Karman	The item directly reached by the agent	
6	Kartr	The independent one	

Table 5 CFG and Sanskrit: a comparison

Sanskrit	Context-free grammar
Sanskrit grammar is based on formulas 'sutra' or production rule	Context-free grammar can be defined as $G =$ (Vn, Vt, P, S)
The non-terminals of CFG can be seen in 'pratyahar'	Vn: Capital letters A, B, C, D,… help in representing a finite set of non-terminals
The 'swar' and 'vyanjan' in Sanskrit are similar to terminals	Vt: Small letters a, b, c, and d represent finite set of terminals
Formation of word have start symbol, e.g., hariH	S: S belongs to Vn.Start symbol of the grammar is starting non-terminal
Production rule can be seen in sutra or formula work	P: CFG have set of rules or productions

2. An item participating in an action in the way indicated by definition.
3. A word standing for such an item.

Karak/verb link syntactic meanings with case endings which also help to define the nonlinguistic features, i.e., we can link them with heterogeneous names, as explained in Table 4. To deal with exceptions name is extended with use of *apadana*, etc. by phrasing additional rule or that particular feature is linked with heterogeneous name karman/doer to allow passivation, when verb concerned is intransitive.

2.3 Comparison of Sanskrit Grammar with Context-Free Grammar

Translational Equivalence: This requires a set of rules that are represented by a model to translate sentences from one language to another (source to target) within the framework of Machine Translation. These rules are formulated on the basis of the grammar of language. A comparative study between CFG and Sanskrit grammar is given below [2, 11] in Table 5.

We can consider the derivation process of the root word 'ram' as shown in Fig. 2. The word 'ram' goes through a series of derivation to reflect the information as third person, singular in number and gender as male.

राम–अदन्तहस्वअकारान्तपुल्लिंग शब्द

=>राम +सु

=>राम +स् (उकार की उत् संज्ञा एव लोपं)

=>राम +रू

Context-Free Grammar has a subset Synchronous Context-Free Grammar (SCFG), which provides foundation to the Two-Way ExT translation model having Sanskrit, as Interlingua. Linguistics representation of syntax is done in SCFG. The productions P, terminals T, and non-terminals N(N, T, and P) combine to form CFG [12–15].

$$P = \{N \rightarrow \{N \cup T\}^*\}$$

Terminals are derived from writing non-terminals recursively [13, 15]. The left side sequence of terminal and non-terminals replace the productions on right. Our model has syntactic categories formed by non-terminal symbol and terminal symbols represent words. Therefore, the sentence has a start symbol 'S' then it searches for different routes to rewrite symbols. This process is continued until either all the possibilities are explored or the sentence is generated as input. In Sect. 4 detailed discussion is done on this.

3 Two-Way ExT (Extended) Translation Model

The Two-Way ExT. model is an extension of Two-Way model. This model targets to achieve translation in multiple languages reporting minimum error. The Two-Way ExT model is generic and can be used for translation from one (source) language to Sanskrit language (Intermediate language) and from Sanskrit Language to Target language as shown in Fig. 3.

In this model the source sentence is divided into several tokens; then it is verified that if the sentence is syntactically correct or not with the help of set of rules, i.e., rule base. If the sentence is correct, further morphological analysis and mapping is done with the Interlingua, i.e., Sanskrit Language. Then these tokens are mapped with syntactic structure of the target language. Then the morphological analysis is done. The output in target language is obtained. The translation process follows Top-to-Bottom, then Bottom-to-Top approach iteratively from source to Sanskrit language and then from Sanskrit to target language. The advantage of this model is that by introducing Sanskrit as Interlingua we can obtain near accurate translations in multiple target languages at a time [1, 5, 7, 16, 17].

4 Algorithmic Steps for Translation Model

The flow chart and algorithmic steps for the translation are outlined in Fig. 4 and Table 6, respectively. As shown in Table 6, translation is carried out using a five step process. We can consider mapping multiple languages with the Interlingua to obtain output in multiple languages. The flowchart, as shown in Fig. 4, presents a translation process. We can see that the words are extracted from the sentence. Then grammatical information is gathered by the dictionary for the words extracted. The extracted

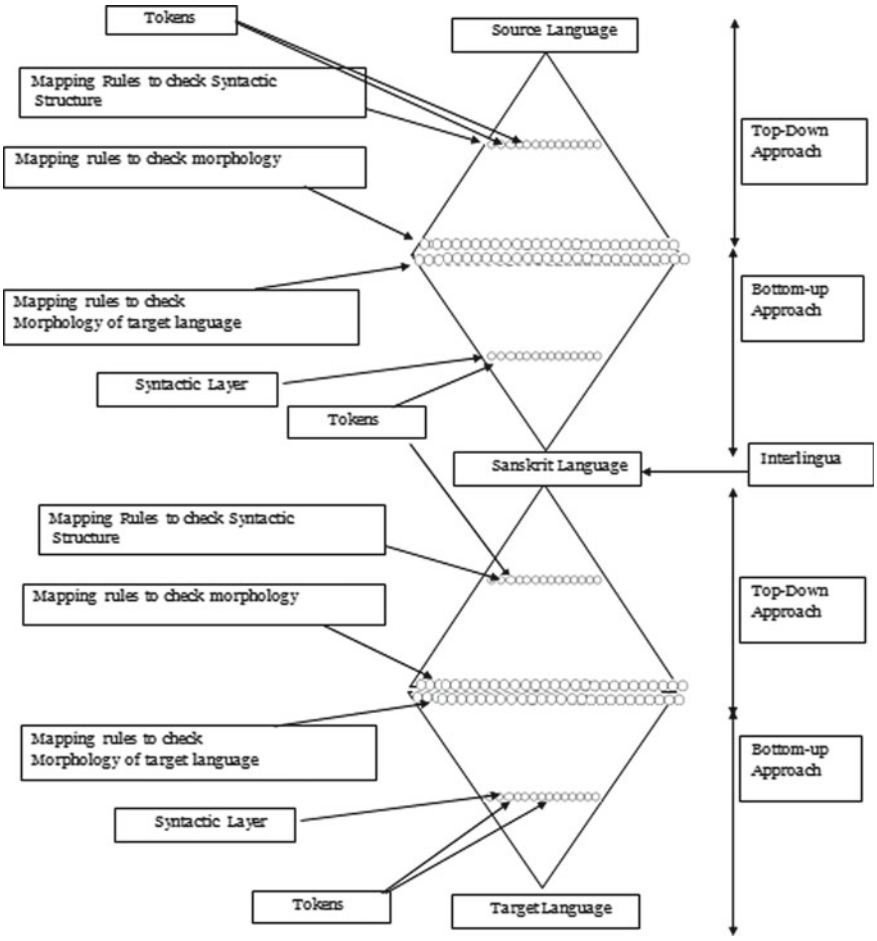


Fig. 3 Two-Way ExT model

information is checked with the rule-base file. If the sentence is grammatically correct then further the mapping is done with the intermediate language Sanskrit. Further, information is gathered for the Target language; mapping is done in order to achieve the desired language translation.

Figure 5 shows a simple parsing process for the sentence 'Bird flies'.

Bird: N
Flies: V

The Architectural model and Functional model are shown in Fig. 6 and Fig. 7, respectively. In Fig. 6, the architectural model shows the process of mapping from

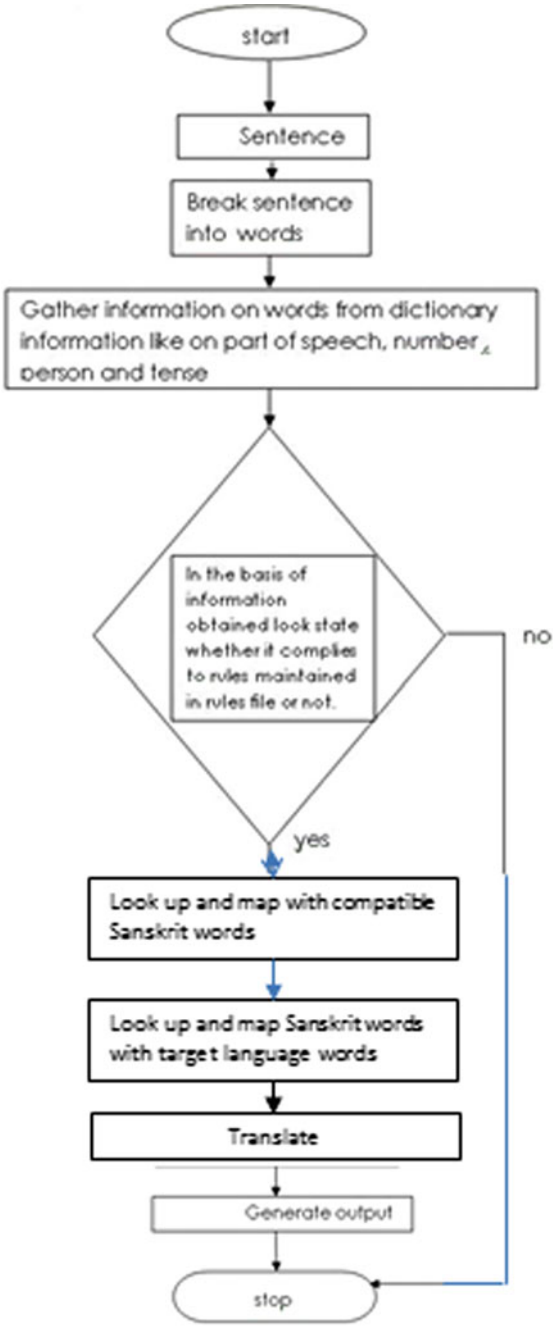


Fig. 4 Flowchart for translation

Table 6 Showing different backup states of parsing process

Step	Parsing process current state	Parsing process backup state	Remarks
A	Sentence(S)1		
B	(Noun Phrase)(Verb Phrase)NP VP		(Sentence)S rewritten to NP VP
C	(Noun) (Verb Phrase)(N VP)1	(Noun) (Verb Phrase) ((N VP)1)	(Noun Phrase)NP rewritten
D	(Noun) (Verb Phrase)((N VP)2)	(Noun Verb)((N V)1)	The backup state remains
E	(())		Success

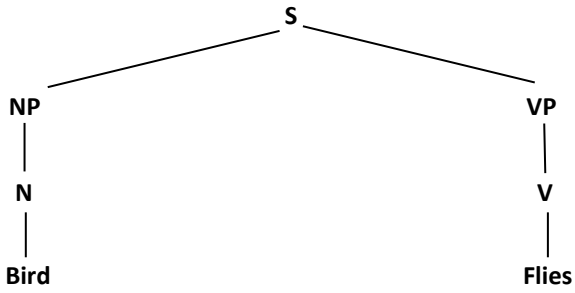


Fig. 5 A tree representation of 'Bird'

source language to Sanskrit language followed by mapping of Sanskrit Language with the target language. In Fig. 7, the functional model is outlined. This model outlines a general framework of translation process. A tree is formed by breaking sentence in major subparts. Each basic unit of tree is node which represents grammatical phrases. The major components of nodes are Noun Phrase or Verb Phrase further the leaves represent noun, adjectives, parts of speech like noun, verb, adjective, etc. Top-down parsing is started from the sentence of source language and ending up to symbol list. The categories of words are stored in multilingual lexicon.

The approach used for generating possibilities list is simple top-down parsing algorithm. Symbol list and a word position in the sentence is represented by the first element is a current state. The backup state represents remaining element. Table 7 shows Backup states of parsing process [11].

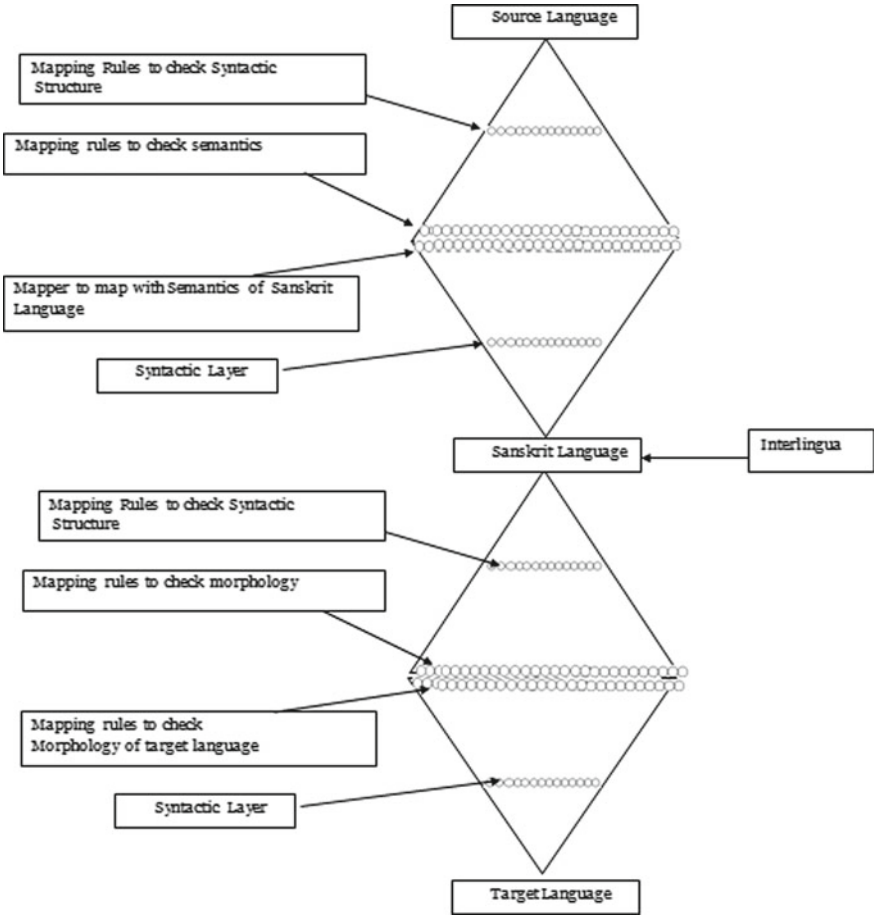


Fig. 6 Architectural model

5 Conclusion and Future Work

In this paper, we have introduced 'Two-Way (ExT) Extended Model' for translation process having Sanskrit as Interlingua. The EtranS software developed earlier [10] does translation from English to Sanskrit language with ninety percent accuracy. The functionality of the EtranS software is being extended to generate the output in multiple languages.

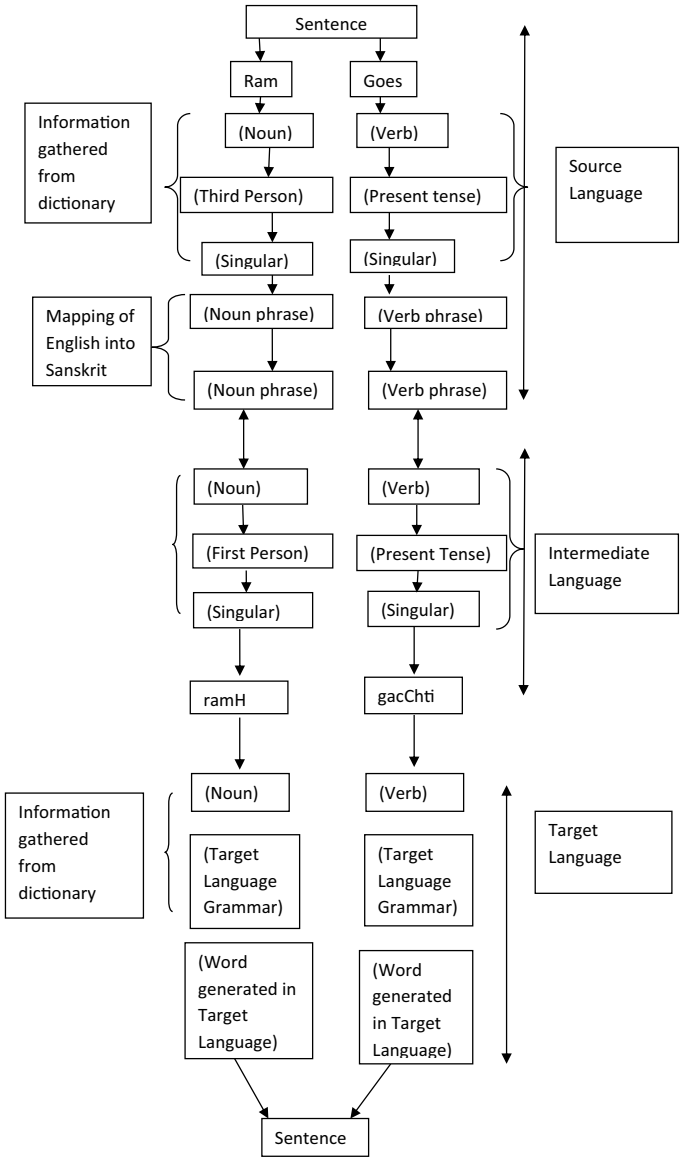


Fig. 7 Functional model

Table 7 Showing different backup states of parsing process

Step-1	Take the sentence in L1 language as input
Step-2	Break the sentence into words/tokens
Step-3	Gather relevant information like part of speech, tense, number, etc. is gathered from morphemes
Step-4	Syntactic analysis is performed just to check whether the information obtained via morphemes can be put together to form correct sentence and to determine structural roles played by each word in the sentence including the role of phrases
Step-5	If Step-4 is true then the semantic analysis of the sentence would be done else the system will generate an error message Semantic analysis is bit typical process here as the morphemes obtained from L1 are used by lookup dictionary to provide morphemes of Sanskrit Language keeping in view that Sanskrit Language follows a totally different semantic roles and structure in comparison to L1
Step-6	The source sentence is now in intermediate language, i.e., Sanskrit language
Step-7	The tokens are generated from the Sanskrit language
Step-8	Semantic analysis is again challenging process here as the morphemes obtained from Sanskrit Language are used by lookup dictionary to provide morphemes of L2 keeping in view that L2 follows a totally different semantic roles and structure in comparison to Sanskrit Language
Step 9	The target language is generated from intermediate language

References

1. A. Bharati, V. Chaitanya, R. Sangal, Natural language processing—a Paninian perspective
2. T. Dhanabalanetal, Tamil to UNL En-Converter, in *Proceedings of ICUKL 2002* (Goa, India, 2002)
3. N. Mishraetal, An unsupervised approach to Hindi word sense disambiguation (HCI, 2009), p. 327
4. S.D. Joshi, J.A.F. Roodbergen, PatanjaliVyakarna-Mahabhashya vol. 3,10,11 (Publications of the Center of Advanced Study in Sanskrit)
5. P. Bahadur, A. Jain, D. Chauhan, EtranS—English to Sanskrit machine translation, in *International Conference and Workshop on Recent Trends in Technology, (TCET) 2012*, Proceedings published in International Journal of Computer Applications® (IJCA) (2012)
6. P. Bahadur, English to Sanskrit machine translation-EtranS system. Int. J. Comput. Appl. Inf. Technol. **3**(II) (ISSN: 2278-7720) (2013)
7. T. Dhanabalan, T.V. Geetha, UNL Deconverter for Tamil, in *International Conference on the Convergence of Knowledge, Culture, Language and Information Technologies, Convergences 2003* (Alexandria, December 2–6, 2003)
8. P. Bahadur, A. Jain, D. Chauhan, EtranS—A complete framework for English to sanskrit machine translation, in *IJACSA Special Issue on Selected Papers from International Conference and Workshop on Emerging Trends in Technology* (2012)
9. J. Allen, *Natural Language Processing* (Pearson Educations, Pearson, 1995)
10. A. Seethaetal, Improving performance of English–Hindi CLIR system using linguistic tools and techniques (HCI, 2009), p. 261
11. P. Bahadur, A. Jain, D. Chauhan, Architecture of English to Sanskrit machine translation, in *AI Intelligent Systems Conference (IntelliSys), 2015 IEEE Technically Co-Sponsored SAI Intelligent Systems Conference 2015* (held at November 10 &11, 2015 in London, UK, 2015), pp. 616– 624

12. A. Lopezacm, Statistical machine translation. Comput. Surv. **40**(3) (Article 8, Publication date: August 2008)
13. A.V. Aho, J.D. Ullman, *The Theory of Parsing, Translation and Compiling* (Pearson Education)
14. C. Armentano-Oller et al., Open source Portugese-Spanish machine translation, PROPER 2006, in *Lecture Notes in Artificial Intelligence*, vol. 3960 (Springer, 2006), pp. 50–59
15. E. Komatsu et al., English generation from interlingua by example-based method, in *Proceedings of the 15th International Conference on Computational linguistics—Volume 1* (Kyoto, 1994), pp. 363–368
16. A. Pathak, P. Acharya, R. Balabantaray, A case study of Hindi–English example-based machine translation, in *Proceedings of ICEMIT 2017*, vol. 3 (2017)
17. P. Bahadur, A.K. Jain, D.S. Chauhan, English to Sanskrit machine translation, in *ICWET 2011* (ACM, Bombay, 2011)

Machine Learning Approach for Crop Yield Prediction Emphasis on K-Medoid Clustering and Preprocessing



Huma Khan and S. M. Ghosh

1 Introduction

The pressures arising from natural resource constrictions increase in fragmentation of capitals, climatic variations rate is very high, rise in input expenditure and postharvest damages pose massive challenges to sustainable agricultural development. Henceforth, crop yield prediction is the vital research area that may be helpful for India's GDP, for the policymaker it may be less troublesome in the diverse condition.

Work has been carried out in the field of crop yield prediction, still need improvement, to enhance the prediction accuracy, earlier methods may have taken less features which can affect the crop yield.

In this research, we have taken Chhattisgarh State Crop production data of different districts in past years. Earlier machine learning used crop yield prediction that involved four phases as shown in Fig. 1.

The first phase is preprocessing in which data reduction is done, second phase is feature selection in which it is needed to identify relation between dependent and independent variable, at last we need to choose the classification algorithm that will work as prediction model. Some prediction models also use clustering before applying any classification as prediction model. Clustering can reduce the error of various data prediction tasks. Clustering algorithm is cast off to mine how the data are structured. Depending on the previously demarcated metric numeric data points in one cluster are by delineation greatly analogous to one other than to numeric data points from different groups. One valuable way of observing at this is thinking of clustering as [1]: Ponder a dataset that is acquired by doing selection on a group of

H. Khan (✉)
CVRU, Bilaspur, Chhattisgarh, India
e-mail: humakhancse2018@gmail.com

S. M. Ghosh
Department of Computer Science and Engineering, CVRU, Bilaspur, Chhattisgarh, India

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_27

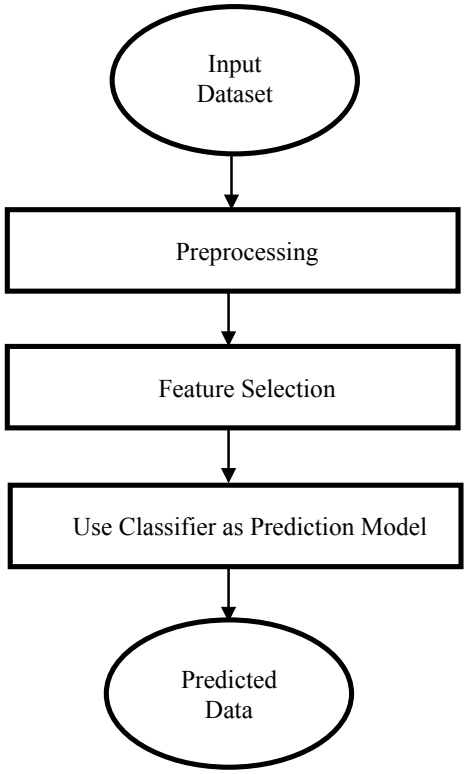


Fig. 1 Phases of prediction using machine learning

distributions $\{D_1, D_2, \ldots, D_k\}$ with their weights $\{w_1, w_2, \ldots, w_k\}$ such that $\mathfrak{E}w_i = 1$, i.e., from every distribution D_i, a datapoint is picked with likelihood w_i. The conception behind clustering is to categorize these disparate distributions that might have produced it and assign points in the dataset into dissimilar groups consequently.

Let us deliberate an illustration of regression task (Fig. 2): Presume that primarily we cluster the dataset into k groups by means of an algorithm, e.g., k-means. A distinct regression model is then trained on every one of these clusters (other than linear regression different classification algorithms can be applied). Each such model lets say "Cluster Model". Altogether the k cluster models can be assumed as developing a much compound model that we named as "Prediction Model". Prediction model symbolized as PM_k, with the subscript representing the number of cluster models in the prearranged prediction model (which in turn will be perceptibly identical to the number of clusters). To recapitulate, to train a "prediction model", the following stages are charted:

1. Group the training dataset into k fragments.
2. For each fragment, train a discrete predictor by means of the points exclusive to that cluster as its training set.

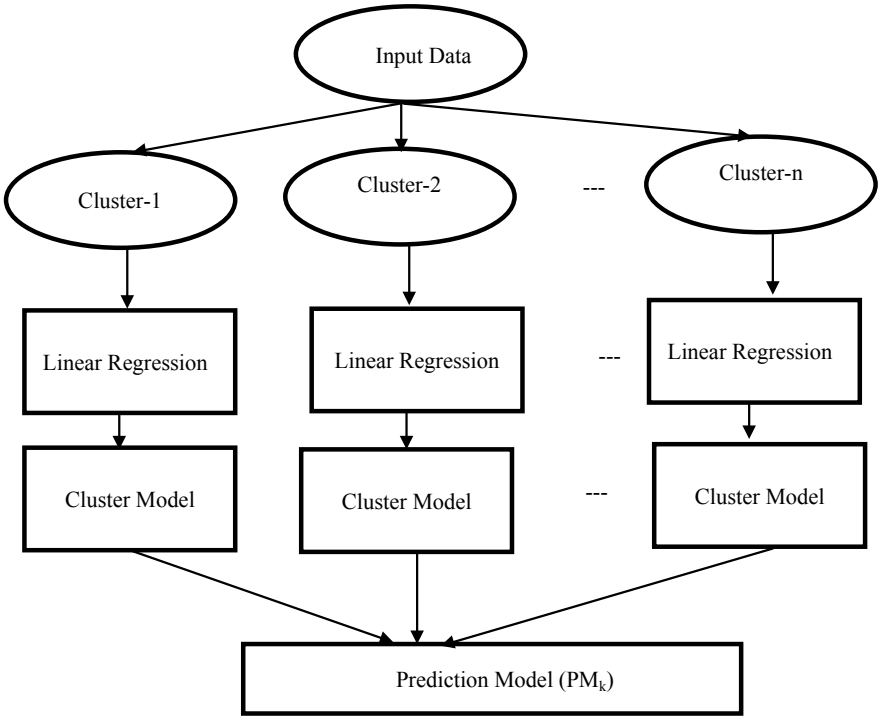


Fig. 2 "Prediction Model" incorporated with clustering

3. Each one of such predictor characterizes a model of the cluster, and henceforth is named as the cluster model.

As soon as a prediction model is attained forecast of a point from the input dataset (test) would encompass as shown in Fig. 3.

Forecasting result from the input dataset (test) would, as a result, involve the following two steps:

1. Categorize the cluster to which the test point be appropriate.
2. Apply the cluster model of the recognized cluster to make the forecast for that data point.

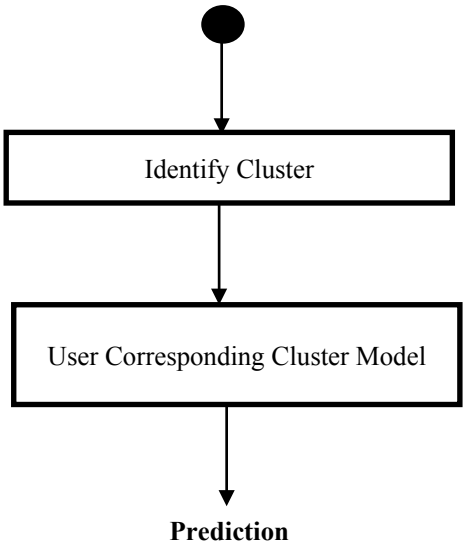


Fig. 3 Charting a test point to a group (cluster) forecasting

It must be noticed that PM1 would merely be predictor fit on the complete dataset (for the overhead illustration it would be fitting a linear regression model on the dataset, we can think of the whole dataset as single cluster).

2 Background Study

Niketa Gandhi et al. had visualized the data using Microsoft excel and applied different data mining algorithms using WEKA tool for crop yield prediction. Investigational results concluded that LADTree & J48 attained the utmost sensitivity, evaluating parameters of algorithm specificity and accuracy [2]. Wu Fan et al. proposed an architecture that encompasses three sections. Major section is a MapReduce-based climate data processing structure, which runs and computes large datasets on assemblage of computers, further in next part is to discover akin years, henceforth "nearest neighbor", which customs weather distances. The preceding chunk is to construct ARMA model based on the "nearest year" and acquire the forecast number [3]. Monali Paul et al. applied K-Nearest Neighbor (KNN) and Naive Bayes (NB) machine learning classifier over the soil dataset which took from the soil testing laboratory Jabalpur, M.P. Accuracy is attained by assessing the datasets. Both algorithms have been applied over the training dataset and their enactment in terms of accuracy is calculated alongside the forecast done in the Input (test) dataset [4].

S. No.	Author/Title/Publication	Algorithm Used	Description	Accuracy (%)
1.	Niketa Gandhi et al./Rice Crop Yield Prediction Using Artificial Neural Networks/IEEE 2016 [5]	Artificial Neural Network	The dataset was sorted out using WEKA machine leaning tool. A MLP Neural Network model was developed. Cross-validation method was used to authenticate the data.	97.5
2.	Rakesh Kumar et al./Crop Selection Method to Maximize Crop Yield Rate using Machine Learning Technique/IEEE 2015 [6]	Crop Selection Method (CSM), Machine Learning	Author projected a technique termed CSM (Crop Selection Method) to decide crop selection delinquent, and yield complete benefit of net crop yield rate over time of year and afterward accomplishes maximum economic progression of the country.	95.2
3.	Niketa Gandhi et al./PredictingRice Crop Yield Using Bayesian Networks/IEEE 2016 [2]	BayesNet and NaiveBayes	The constraints designated for the research were precipitation, min. temperature, average temperature, max. Temperature, reference crop evapotranspiration, area, production, and yield for the Kharif season (June to November) for the years 1998–2002. The experimental evaluation was done using the WEKA machine learning tool. The classifiers cast off in the research were BayesNet and NaiveBayes.	BayesNet −97.53 NaiveBayes −84.69 For Rice Crop {resection}

(continued)

(continued)

S. No.	Author/Title/Publication	Algorithm Used	Description	Accuracy (%)
4.	Leisa J. Armstrong et al./Rice Crop Yield Prediction in India using Support Vector Machines/IEEE 2016 [7]	SVM (Support Vector Machine)	This research deliberates the investigational results attained by smearing SMO classifier using the WEKA machine learning tool over the dataset of 27 districts of Maharashtra state, India. The dataset measured for the rice crop yield forecast was traced from openly available Indian Government records.	Accuracy 78.76

3 Dataset Used

For our research we have taken Chhattisgarh Government Meteorological data, Department of Agrarian Meteorology Indira Gandhi Agricultural University, Raipur Station: Labhandi Monthly Meteorological Data. We have taken data for the years 2010–2017.

Table 1 is the fragment of Meteorological Data, yellow-colored column is the dependent variable and blue-colored columns are dependent variable. Dataset contains crop yield production of different districts of Chhattisgarh and rainfall data for the same.

4 Methodology and Experimental Result

For our present study, we have proposed to use Principal Component Analysis (PCA) for preprocessing of input dataset. Preprocessing is meant for dimensionality reduction for the dataset. Dimensionality lessening can be explained as follows:

Suppose any dataset $\boldsymbol{E}$ represented in the form of matrix with dimension $\boldsymbol{n} \times \boldsymbol{m}$ containing n data vectors $y_{i,}$ where ($i \in \{1, 2, \ldots, n\}$) with dimensionality DI. Presume supplementary that this dataset has inherent dimensionality d where $d <$ DI, and frequently $d <<<$ DI, at this point, in scientific stipulations, inherent dimensionality of dataset means that the points in dataset $\boldsymbol{E}$ are mendacious on or nearby various dimensionality d, i.e., entrenched in the DI dimension. Dataset dimensionality lessening methods transmute dataset $\boldsymbol{E}$ with dimensionality DI into a fresh dataset $\boldsymbol{F}$ with dimension d, although absorbent the geometry of the data as greatly as possible.

Table 1 Meteorological Data

ID	Area	Rainfall	Production
1	906	103.93	508
2	5	103.93	2
3	13	103.93	20
4	7797	103.93	2854
5	554	103.93	916
6	20242	103.93	40570
7	83	103.93	211
8	224	103.93	89
9	12189	103.93	2591
10	487	103.93	59
11	4269	103.93	1192
12	239376	103.93	401486
13	464	103.93	207
14	13965	103.93	2630
15	17	103.93	25

Principal Component Analysis (PCA): In this method, variables are transmuted into a new set of variables, which are linear arrangement of inventive variables. These transmuted new sets of variables are identified as principal components. They are acquired in such a way that first principle component accounts for most of the possible dissimilarity of original data afterward which every consequent factor has the maximum likely variance.

Another principal component needs to be orthogonal to the earlier principal component. In other words, it does its best to apprehension the variance in the data that is not apprehended by the first principal component. For 2D dataset, there can be only two principal components. Underneath is a snapshot of the data and its first and second principal components. We can notify that second principal component is orthogonal to first principal component (Fig. 4).

PCA coefficients, also identified as loadings, for the $n \times m$ data matrix Y. Rows of Y resemble annotations and columns resemble variables. The constants matrix is $q \times q$. Every column of coeff comprises constants for one principal component, and the columns are in downward order of factor variance. PCA centers the data and uses the Singular Value Decomposition (SVD) algorithm (By Default). For implementation we have used Matlab 2016 (Fig. 5).

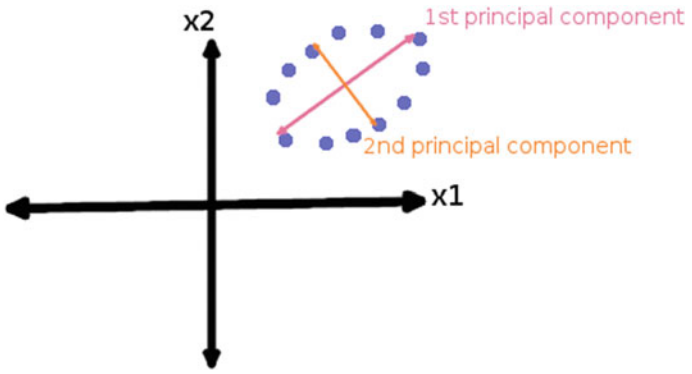


Fig. 4 First and second principal component

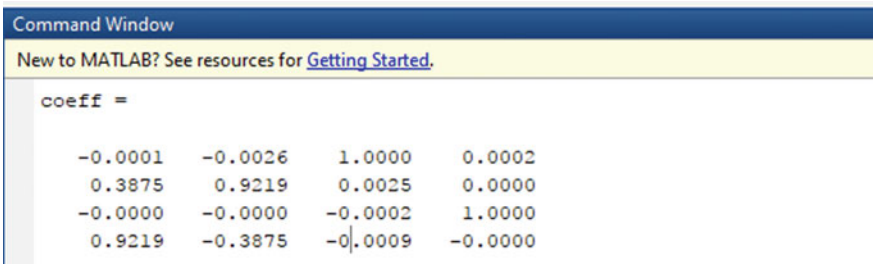


Fig. 5 Coeff value after PCA applied over dataset

From Fig. 6, we have proposed to use K-medoid as clustering algorithm, we have performed experimental evaluation as, applied k-means and k-medoid clustering over input dataset (Fig. 7).

There are some bottlenecks of *k*-nearest neighbor prediction that are as follows:

1. The upfront algorithm has a time complexity $\boldsymbol{O(n \log(k))}$, not worthy if the dataset is hefty.
2. It is computationally affluent to discover the *k*-nearest neighbors when the dataset is very bulky.
3. Performance is subject to the number of dimensions that we have (obscenity of dimensionality) $\Longrightarrow$ feature Selection.
4. If dimensions are more, we have much more examples we requisite to imprecise a hypothesis.
5. This is particularly wicked for *k*-nearest neighbors, i.e., if the number of dimensions is very high the nearby neighbors can be very far away.

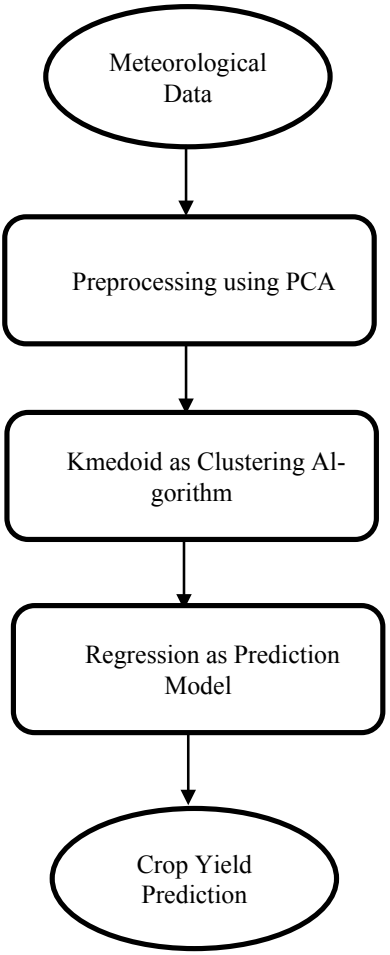


Fig. 6 Proposed system flow

6. The number of examples that we have in a volume of space diminishes exponentially with the number of dimensions.
7. K-means has problems when clusters are differing
    a. Sizes
    b. Densities
    c. Non-globular shapes

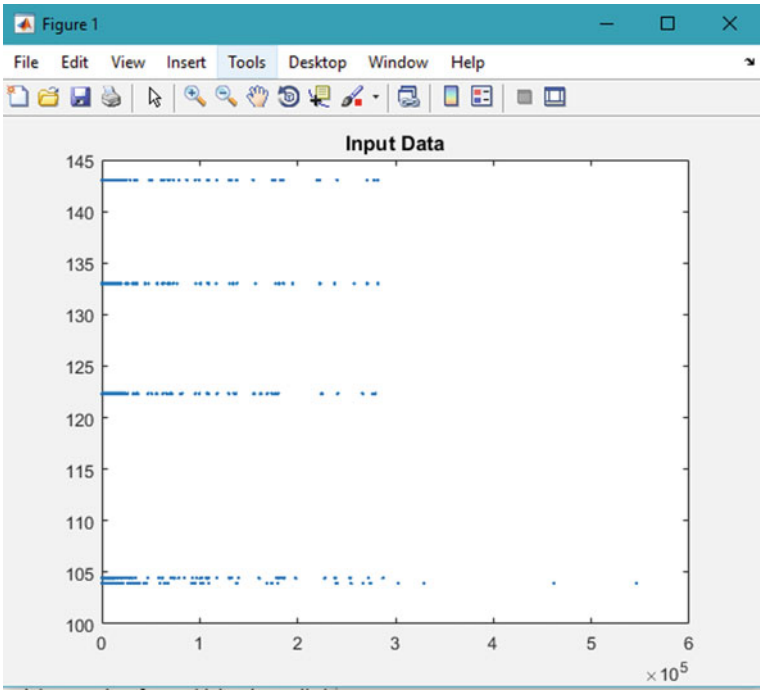


Fig. 7 Input dataset

8. Problems with outliers
9. K-means algorithm is slow and scales ailing in case of time complexity it takes for huge number of data points.

K-nearest neighbours - Algorithm
Step-1. Training: Store all the historical data
Step-2. Forecast: h(x_{new}) Test
Let be $x_1, \ldots, x_k$ the k more analogous samples to xnew
h(x_{new})= conglomerate forecasts($x_1, \ldots, x_k$)
Step-3. The arguments of the algorithm are the number k of neighbours and the procedure for uniting the forecasts of the k samples
Step-4. The value of k has to be attuned (cross validation)

Algorithm

Algorithm: K-Medoid

1. Initialize: arbitrarily select(without auxiliary) k of the n data points as the medoids
2. Associate each data point to the neighboring medoid.
3. While the cost of the configuration diminutions:
    1. For every medoid m, for every non-medoid data point o:
        1. Exchange m and o, re-calculate the cost (sum of distances of points to their medoid)
        2. If the total cost of the configuration greater than before in the previous step, untie the exchange

Figure 8 shows the plot of K-means clustering of input dataset, and Fig. 9 shows the K-medoid clustering of input dataset. From Figs. 8 and 9 we saw that in case of k-means clustering number of outliers is greater than k-medoid clustering. For both the clustering we have used distance function as follows:

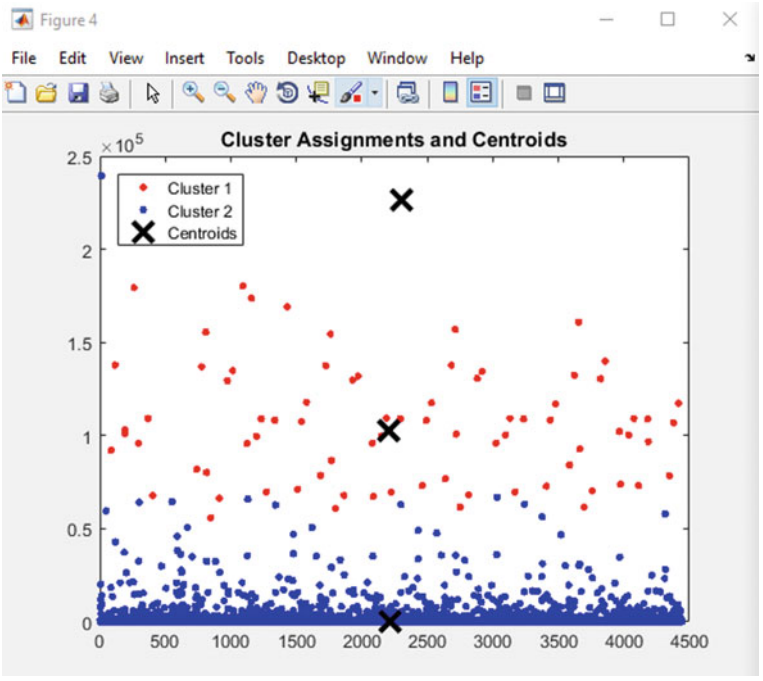


Fig. 8 K-means clustering

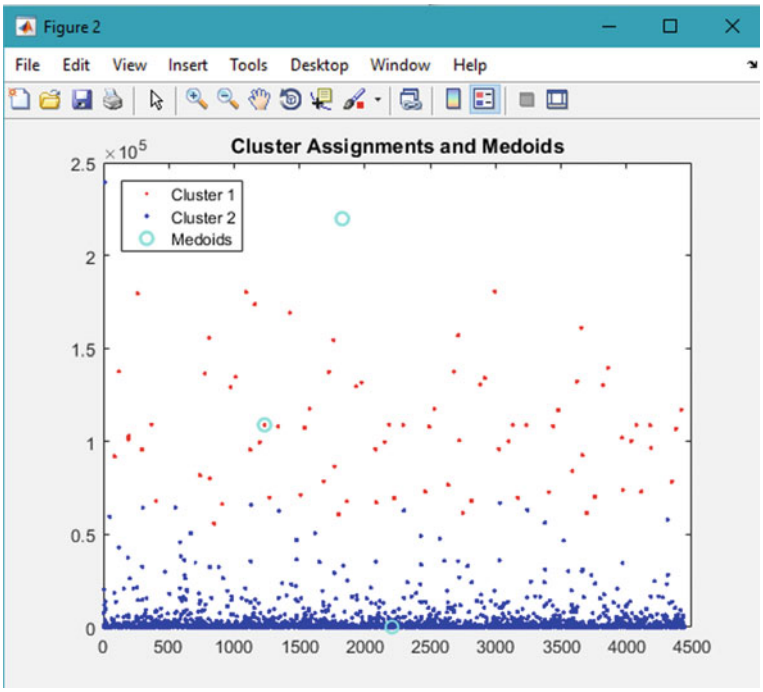


Fig. 9 K-medoid clustering

$$d(x, c) = \sum_{j=1}^{p} |x_j - c_j| \tag{1}$$

Every centroid is the median component-wise of the data points in that cluster.

5 Conclusion

The weights emerging from regular asset limitations, expanding fragmentation of possessions, frequent climatic change, rising input expenses, and postharvest misfortunes represent a tremendous challenge to continuing agricultural development. In this paper, we have gone through different literature tabular comparisons and we can conclude that there is need for improvement over accuracy of crop yield production, by opting different features and better preprocessing and classification algorithms we can achieve more accuracy. In this paper, we have applied PCA for dimension reduction and from experimental result it is concluded that K-medoid clustering performs better than K-means; henceforth, for proposed prediction model we will use K-medoid as clustering model. In future we will use regression as prediction model after clustering.

References

1. Dasgupta, S. Learning mixtures of Gaussians, in *40th Annual Symposium on Foundations of Computer Science* (IEEE, 1999), pp. 634–644
2. N. Gandhi, Predicting rice crop yield using Bayesian networks (IEEE, 2016)
3. Wu Fan1, C. Chong, G. Xiaoling, Yu Hua, Prediction of crop yield using big data, in *2015 8th International Symposium on Computational Intelligence and Design* 978-1-4673-9587-8/15 $31.00 © 2015 IEEE https://doi.org/10.1109/iscid.2015.191, p. 255
4. M. Paul, S.K. Vishwakarma, A. Verma, Analysis of soil behaviour and prediction of crop yield using data mining approach, in *2015 International Conference on Computational Intelligence and Communication Networks* 978-1-5090-0076-0/15 $31.00 © 2015 IEEE https://doi.org/10.1109/cicn.2015.156, p. 766
5. N. Gandhi, L.J. Armstrong, Rice crop yield forecasting of tropical wet and dry climatic zone of India using data mining techniques, in *2016 IEEE International Conference on Advances in Computer Applications (ICACA)*
6. R. Kumar, Crop selection method to maximize crop yield rate using machine learning technique (IEEE, 2015)
7. L.J. Armstrong, e Rice crop yield prediction in India using support vector machines (IEEE, 2016)

Combined Approach to Classify Human Emotions Based on the Hand Gesture



Samta Jain Goyal, Arvind Kumar Upadhyay and Rakesh Singh Jadon

1 Introduction

Face and hand detection and its identification with its prominent features is a crucial step in any facial emotion recognition system based on the hand gesture or hand position surrounding the face. Various approaches are presented in the literature in terms of static or temporal identification of the face, hand, and features. When the face and hand are located, it must be modeled so that it can be represented in an appropriate manner. The facial representation could be based on the facial geometry that encompasses some unique features of homogeneity and diversion across humans [1]. It could also be based on characteristics that appear after some transformation with mathematical emotions modeling texture, position, and gray-level information. After that the feature vector is built by extracting features. It can be represented either holistically or locally [2, 3].

2 Tools and Techniques

Our own maintained data set is used for the evaluation of the proposed method. It contains different features with different people. These images are taken from the professional Nikon Camera. There is total of 469 images. Neutral position inherits all characteristics of a basic emotion and it is included in the annotation of the database as a seventh basic emotion. My own data set is used for facial emotion recognition

S. J. Goyal (✉) · A. K. Upadhyay
Department of Computer Science and Engineering, ASET, Amity University, Gwalior 474005, Madhya Pradesh, India
e-mail: samtajaingoyal@gmail.com

R. S. Jadon
Department of Computer Applications, MITS, Gwalior 474001, MP, India

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_29

based on hand gesture surrounding the face area which has six basic facial emotion classes. The framework and the variation of the multi-class are proposed in [4–6] and are generally used to discriminate the component vectors.

2.1 Libraries and Tools Used

2.1.1 OpenCV

The library used for transforming images function is open CV. It is mainly used to convert the image to gray scale. This library is open source and it can be used for many functions of images and has a wider variety of implementations in algorithms. The languages supported by open CV are Python and C++. This is a complete framework that finds use along with other libraries for forming a pipeline to extract any image or as a method for detection [7, 8]. It supports a very large function range and it also includes themes for extraction of feature descriptors.

2.1.2 Dlib

Another powerful library for processing images is Dlib. It finds use in concurrence with c++, python, and several other tools. The key function that this library has is providing face detection and hand detection along with feature extraction and feature matching, etc. [9, 10]. It also supports various different domains like threading, machine learning, networking, and GUI.

2.1.3 Python

Python is an extremely useful language for scripting, mostly finding use in solving problems in statistics including algorithms for machine learning. It has various functions for utility, which help it in preprocessing [11]. This preprocessing is faster, and all the platforms support it. C++ integration and libraries for other images are very simple and there are built-in functions and inbuilt libraries manipulating and storing data of every type [12, 13]. It gives a numpy framework and pandas which help to manipulate the data in accordance with the user's need. Other important features set is possible for being created with the use of numpy arrays including data that is n dimensional.

2.1.4 Scikit-Learn

It is a library for machine learning in Python, which includes matplotlib, numpy, and a large group of algorithms for machine learning [14]. Its API is much easier for using

and understanding. What's multiple functions for analyzing and plotting the data. Another important feature it has is that a set of useful features can be created with the use of its function for reducing features, selecting features and feature importance functions [15, 16]. The algorithm that it provides can thus be used to classify and regress problems and subproblems.

2.1.5 Jupyter Notebook

For combining python with all our libraries we use jupyter notebook in our implementation. Even when the computations that are complex need more time for completion, this is still interactive [17–22]. Images and plots are instantly displayed. For all our requirements we can use it as a one-stop shop wherein most of the libraries like open CV, scikit-learn, and Dlib can be easily integrated.

3 Implementation

This work uses an approach that is static to use features that are extracted emotions through machine learning. We focus on feature extraction using image processing libraries and python via algorithms for machine learning used for prediction. We divide this implementation into three parts mainly. The first part of the image is preprocessed, and hands and face are detected. Detect hands and face the Dlib library's inbuilt methods are used. After detecting face and the hand, we extract the useful features of the face and the interesting regions from it. Various other features can be used for detecting emotion. In this research we focus mainly on the points on the face that lie around the mouth, eyes, and eyebrows, etc. This generates the problem of classification in multi-classes and not multi-label. The difference is very subtle as the feature set melay belongs to multiple labels but can have only a single class that is unique. Figure 1 visualizes the pipeline of processing.

Does the data set we developed includes only the face front and image of hands and it excluded every file which did not include an emotion or had a neutral emotion? For some of the images the conditions for illumination and lighting where different. Pipeline for processing for each image remained the same even when the conditions of illumination were different. We directly extract the emotion name from the filename of the image in which the format .jpg is used as a standard database, the number of classes was balanced for every emotion. In Table 1 we can see various emotion classes distribution.

469 files database is stored in a directory and every file is processed for creation of the feature set. Immediately after the file is picked up, the file name is parsed for extraction of the label for emotion. This label for emotion is then appended to a list

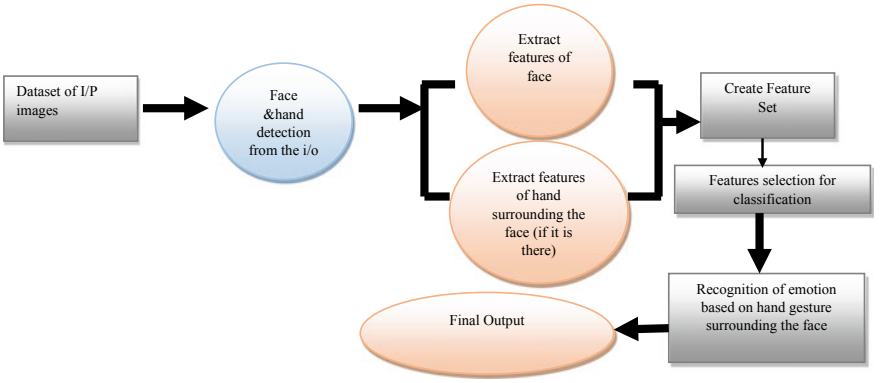


Fig. 1 Implementation pipeline

Table 1 Number of images per class for our own data set

Emotion	Number of images depicting the emotion
1. Anger	79
2. Contempt	52
3. Disgust	69
4. Fear	49
5. Happy	79
6. Sadness	48
7. Surprise	93

that forms our multiple classes for target. For hand and face detection the image is then processed and the feature is then predicted. The derived feature from every file is added to a list which is then transformed into a numpy array with the dimension $327 \times 68 \times 2$. In addition, we have the classes for target in the shape of a numpy array.

In the input image, it is given that landmark normalization method is applied to get each landmark on the face and hands for the proper coordination and in continuous manner. This coordination plays a vital role in display of the surface, in accordance with calculation. Let T_m^n is the result tracked for the mth landmark to the nth landmark for emotion calculation from the input image:

$$T_m^n = \left\{(x_0, y_0)_m^n, (x_1, y_1)_m^n, (x_2, y_2)_m^n, \ldots, (x_{N-1}, y_{N-1})_m^n, (x_N, y_N)_m^n\right\} \quad (1)$$

where $(x_l, y_l)_m^n$ is the mth to nth landmark coordinate position for the lth image with the *N number of emotions.* The average landmark position with respect to each landmark can be calculated from all the input images. Suppose $\left(\mu_{x0}, \mu_{y0}\right)_m^n$ denotes the average landmark position of the mth landmark of the emotion sequences.

Suppose $(\delta_{x0}, \delta_{y0})_m^n$ denotes the displacement of the mth landmark to nth landmark of the emotion sequence, with respect to the average landmark position.

$$(\delta_{x0}, \delta_{y0})_m^n = \left(\mu_{x0} - x_0, \mu_{y0} - y_0\right)_m^n \tag{2}$$

The sequence of human's facial emotion can be obtained with the help of displacement corresponding to each landmark. Hence to get the desired result we require to sum up those landmark—positions in every sequence of the input image. After calculating T_m^n, we can elaborate the transformed result of the tracked landmark obtained by Eqs. (1) and (2).

$$T_m^n = \left\{\left(x_0 + \delta_{x0}, y_0 + \delta_{y0}\right)_m^n, \left(x_1 + \delta_{x1}, y_1 + \delta_{y1}\right)_m^n, \ldots, \left(x_N + \delta_{xN}, y_N + \delta_{yN}\right)_m^n\right\} \tag{3}$$

From Eq. (3), normalized result can be obtained in such a way that all the emotion sequences evolves as per the displacement in the prospering frames.

4 Experimental Results

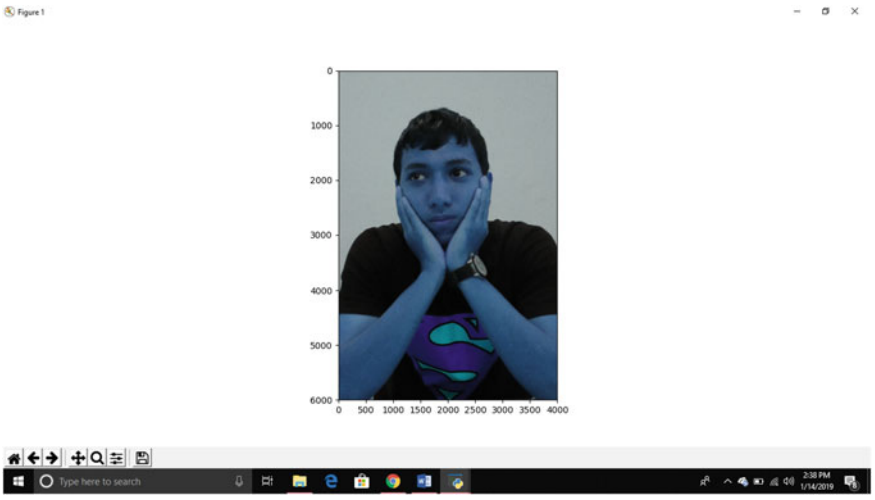
Human emotion classification is used to perform on structured and unstructured data. It's basically used to categorize output into number of specified classes. The objective of this phase is to classify and recognize human emotions based on his or her hand gesture surrounding the detected human face. Here, classifier is basically an algorithm that is used to map input data into the specific category of an output emotion class such as Happy, Surprise, Anger, and so on.

The Hybrid SVM is trained by the *SVMlight* toolbox. For this experiment the SVM and kernel specifications are in the proposed approaches that achieves much better recognition outputs than the existing approaches. In this experiment, it shows how SVM-based binary classifiers successfully combined to capture the multi-class classification problem. Many research has proposed collaborative methods using binary classification algorithms to solve K-class classification problems. We apply Decision Directed Acyclic Graphs (DDAG) algorithm in a K-class problem. The DDAG tree consists of $K(K-1)/2$ nodes which embed binary classifiers between the p-th and q-th pair of classes. The K-leaves of the tree show that predicted class labels of those samples which are unknown. The total depth of the DDAG is $K-1$. Given a test sample x, starting at the root node, the binary decision function is calculated. Then it transfers either left or right depending on the resulted value. Hence, we go to a way before access a leaf node which shows the predicted class. At the time of training, the DAG approach is necessarily the same as the One Versus One (OvO) approach producing $K(K-1)/2$ binary classifiers. However, during evaluation, only $K-1$ decision nodes are traversed. This systematic approach of

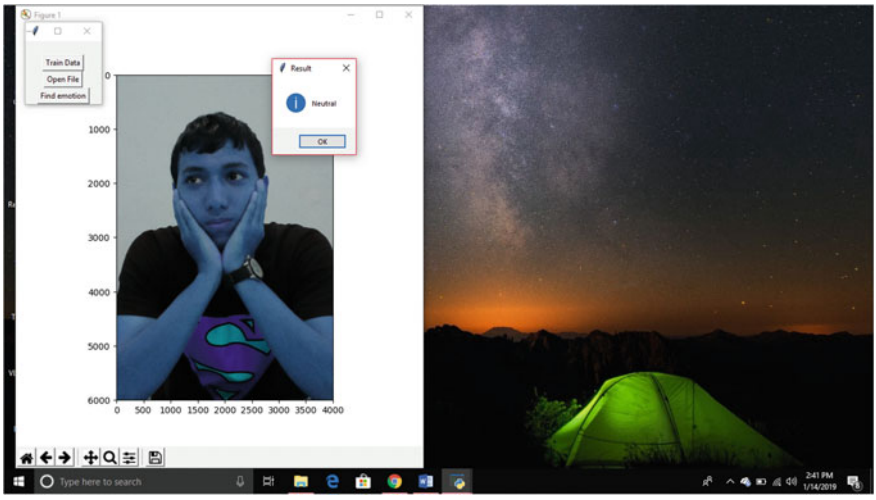
evaluation in DDAG outperforms the evaluation in OvO scheme by a factor of $K/2$ times in speed.

The results by applying SVM to our present data set. we then interpret the results with the use of confusion Matrix and the metric for accuracy. We also cross-validated the data set for removing any biases. The split value was selected and then splits, thus, resulting are bound to have similar number of images as the test set. The areas of result are as follows:

(a) INPUT IMAGE:



(b) OUTPUT EMOTION BASED ON FACIAL EXPRESSION AND HAND GESTURE



5 Comparison Matrix

Through our experiments, we found that the performance of SVM including linear kernel was better than different kernels. The emotion recognition performance is shown through the RPF, while poly had resulted as good as linear kernel. It was our effort to keep the set of test percentage similar for data that is cross-validated and split in order to have uniform results. The mean score for cross-validation was also almost similar to the score of accuracy obtained by the split. In Table 2 we can show the heat map confusion matrix from our classification of multi-class results. By analyzing further, the confusion matrix, we observe that the weight of diagonals is higher and hence each class except for class for fear may have a certain degree of misclassifications.

This comparison matrix is based on two measurements.One is accuracy of the proposed system an another is based on F1-Score which is calculated through the following parameters.Accuracy is basically a ratio of correctly predicted observation to the total observations. It is the most intuitive performance measure, whereas F1-Score is the weighted average of Precision and Recall. It is more useful than the accuracy because it work more efficiently when we have an uneven class distribution. Standard result of comparison matrix is shown in Table 3.

Table 2 Report of predicted values versus actual values

Actual	Predicted							
	Anger	Contempt	Disgust	Fear	Happy	Sadness	Surprise	All
Anger	9	0	2	0	0	0	0	11
Contempt	1	3	0	0	0	0	0	4
Disgust	3	0	12	0	0	1	1	17
Fear	0	0	0	5	0	0	0	5
Happy	1	0	0	0	15	0	0	16
Sadness	2	0	0	0	0	8	1	11
Surprise	0	1	1	2	0	2	12	18
All	16	4	15	7	15	11	14	82

Table 3 Standard result of comparison matrix is as follows

Classification Algorithms	Accuracy (%)	F1-Score
Naïve Bayes	80.11	0.6005
Stochastic Gradient Descent	82.20	0.5780
K-Nearest Neighbors	83.56	0.5924
Decision Tree	84.23	0.6308
Random Forest	84.33	0.6275
Support Vector Machine	84.09	0.6145

Through this, we are able to deduce the actual number of emotions predicted for every class. A lot of samples are wrongly classified into classes like sad, contempt, fear, and disgust. For every sample the detection of fear was accurate. In some of the cases the samples for sadness and disgust were classified wrongly as anger. This was the result of the feature overlap between these emotions. Going further there was a consideration for all features of facial point, 78 in number for determining the features which help actually in determination of emotions. In every micro emotion there was new edition or difference because of the jaw. As a result, decision was taken ignoring the position of the jaw in face. We only continued with mouth, eyes, hands, fingers, nose, and eyebrows for our experiments. The set of features still had similar dimensions of $327 \times 68 \times 2$. The ratio of split for test and train was kept the same for the same data set. We focus on the score for cross-validation because it considers the complete database for training and hence removes all biases. Then we compared the actual results with the predicted results and analyzed them for the present scenario.

6 Conclusion

Developing a human emotion recognition system is a vital issue in HCI. These recognition systems are used in many real-time applications which are related to machine. This method also uses trajectory-based approach to handle the shape of the object. This work first converts the given set of frames into the fixed-length vectors and then tang an SVM from the vectored manifolds. DTW is not only used to get shape information but also used to classify the gestures based on predefined classes of hand gesture and facial emotion. Feature extraction is an important segment of this implementation. These all weak classifiers are put into multi-class AdaBoost in order to take the distinct feature vectors.

In this, our experiments show that the facial and the hand gestures landmarks are dependent on each other. In our experiment, the prototypical facial emotions are processed further by using unimodal distributions. Our assumptions turn out to be true as the facial emotions are recognized successfully. The facial emotion evolves over time, and also independent of the ethnic group, age, and gender for each class of emotion. The result of the newly developed system gives 93% accuracy for the new database. There is no other database for this experiment.

Acknowledgements Authors are thankful to Amity University Gwalior and MITS for providing lab facility for the present research work.

References

1. X. Zhao, S. Zhang, Facial expression recognition based on local binary patterns and kernel discriminant isomap. Sensors **11**, 9573–9588 (2011)
2. S. Moore, R. Bowden, Local binary patterns for multi-view facial expression recognition. Comput. Vision Image Underst. **115**, 541–558 (2011)
3. G. Zhao, X. Huang, M. Taini, S.Z. Li, M. Pietikäinen, Facial expression recognition from near-infrared videos. Image Vision Comput. **11**, 607–619 (2011)
4. D. Ghimire, J. Lee, Histogram of orientation gradient feature-based facial expression classification using bagging with extreme learning machine. Adv. Sci. Lett. **17**, 156–161 (2012)
5. A. Dhall, A. Asthana, R. Goecke, T. Gedeon, Emotion recognition using PHOG and LPQ features, in *Proceedings of the IEEE International Conference on Face and Gesture Recognition and Workshop* (Santa Barbara, CA, USA, 21–25 March 2011), pp. 878–883
6. T. Jabid, MdH Kabir, O. Chae, Robust facial expression recognition based on local directional pattern. ETRI J. **32**, 784–794 (2010)
7. R. Zhi, M. Flierl, Q. Ruan, W.B. Kleijn, Graph-preserving sparse nonnegative matrix factorization with application to facial expression recognition. IEEE Trans. Syst. Man Cybernet.-Part B Cybernet. **41**, 38–52 (2011)
8. I. Kotsia, I. Buciu, I. Pitas, An analysis of facial expression recognition under partial facial image occlusion. Image Vision Comput. **26**, 1052–1067 (2008)
9. D.-T. Lin, Facial expression classification using PCA and hierarchical radial basic function network. J. Inf. Sci. Eng. **22**, 1033–1046 (2006)
10. Z. Wang, Q. Ruan, Facial expression recognition based orthogonal local fisher discriminant analysis, in *Proceedings of the International Conference on Signal Processing (ICSP)* (Beijing, China, 24–28 October 2010), pp. 1358–1361
11. G. Littlewort, M.S. Bartlett, I. Fasel, Dynamics of facial expression extracted automatically from video. Image Vision Comput. **24**, 615–625 (2006)
12. H.-C. Choi, S.-Y. Oh, Realtime facial expression recognition using active appearance model and multilayer perceptron, in *Proceedings of the International Joint Conference SICE-ICASE* (Busan, Korea, 18–21 October 2006), pp. 5924–5927
13. T.F. Cootes, G. Edwards, C.J. Taylor, Active appearance models, in *Proceedings of the European Conference on Computer Vision,* vol. 2 (1998), pp. 484–498
14. T.F. Cootes, C.J. Taylor, D.H. Cooper, J. Graham, Active shape models—their training and application. Comput. Vis. Image Underst. **61**(1), 38–59 (1995)
15. J. Daugman, Two-dimensional spectral analysis of cortical receptive field profiles. Vision. Res. **20**, 846–856 (1980)
16. J. Daugman, Uncertainty relation for resolution in space, spatial frequency and orientation optimized by two-dimensional visual cortical fields. J Opt. Soc. Am. A **2**(7), 1160–1171 (1985)
17. G. Donato, M.S. Bartlett, J.C. Hager, P. Ekman, T.J. Senjowski, Classifying facial actions, IEEE Trans. Pattern Anal. Mach. Intell. **21**(10), 974–989 (1999)
18. P. Ekman, *Telling Lies: Clues to deceit in the Marketplace, Politics, and Marriage* (W.W. Norton, New York, USA, 1991)
19. P. Ekman, W.V. Friesen, Constants across cultures in the face and emotion. J. Pers. Soc. Psychol. **17**(2), 124–129 (1971)
20. P. Ekman, W.V. Friesen, *The Facial Action Coding System: a Technique for the Measurement of Facial Movement* (Consulting Psychologist Press, San Francisco, 1978)
21. S. Pranjali, V.S, Ubale, Hand gesture recognition system: a survey. Int. J. Inven. Eng. Sci. (IJIES), **3**(3) (2015) (ISSN:2319-9598)
22. Dr. K.S. Kinage, S. Mutha, Study on hand gesture recognition. Int. J. Comput. Sci. Mob. Comput. (2015)

Detecting Gene Modules Using a Subspace Extraction Technique



Pooja Sharma, D. K. Bhattacharyya and Jugal K Kalita

1 Introduction

A gene is at the core of any activity occurring in the living body. Genes code for proteins that is associated with metabolic and functional activities. This production of proteins from genes involves a series of steps. One of the outcomes is the expression of genes leading to the formation of mRNA. Scientists use mRNA expression levels of genes for understanding cellular events. This type of study usually involves gene expression analysis. To become functionally active, a gene needs to act in coordination with other genes and hence the study of groups of functionally similar genes is more interesting than that of individual genes. When a gene coordinates its activities in a group, the group of genes is said to be coexpressed. A number of studies have been proposed in the literature to perform anlaysis of gene coexpression analysis and modular structure of gene networks. For example, Tulika et al. [15] used the concept of border genes to identify interesting genes on Alzheimer's Disease. Medina and Pilav [6] used Topological Overlap Measure (TOM) [3] to identify modules. This work was effectively applied to Type 1 Diabetes dataset. Yang et al. [17] proposed a module extraction technique and extended it to characterize properties of biomarkers associated with cancer. Most of these methods rely either on Pearson, Spearman coexpression similarity measures. However, it has been established that statistical biases introduced by these measures result in different networks for the same data [9].

P. Sharma (✉)
Department of Information Technology, BIT Sindri, Dhanbad 828123, Jharkhand, India
e-mail: poojasharma0303@gmail.com

D. K. Bhattacharyya
Department of Computer Science and Engineering, Tezpur University Napaam, Tezpur 784028, Assam, India

J. K. Kalita
Department of Computer Science, University of Colorado at Colorado Springs, CO 80933-7150, USA

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_30

In order to overcome this biases introduced by the coexpression similarity measures, researchers started studying the behavior of genes under different conditions. Ahmed et al. [5] made a study on the coexpression pattern of genes suggesting that two genes may be connected over multiple sets of samples among the given dataset. This leads to the formation of a multi-edge network, where the strength of each edge is determined by the variations among the sample pairs it is made up of. They proposed a measure called Topological Subspace Overlap Matrix (TSOM), an extension of TOM, over multi-edge networks. They extended their TSOM network to find groups of coexpressed modules and validated them. This provided a framework to analyze the effectiveness of TSOM over some disease datasets and studied their progression stages. In this paper, we propose a module extraction technique based on the subspace overlap matrix. Our technique has been applied to Parkinson's Disease dataset and the modules obtained are found to be biologically significant in terms of p-values.

2 Materials and Method

Parkinson's Disease (PD) is a chronic disorder, which only worsens with time. Around one million people are affected by this disease in the United States. At present, there is no treatment for this disease and the average life expectancy of a person diagnosed with the disease is 7–14 years. Effective and target-specific drugs can be used for better locomotory movement in patients thereby assisting them to lead a better life. In this paper, we proposed a method for extracting modules from both control and disease stage of Parkinson's Disease. The coexpressed module members can then be used by biologists to design target-specific drugs that may have the potential to ease the symptoms of the disease.

We use the GSE 8397 dataset entitled Expression Profiling of the Parkinsonian Brain. This dataset was prepared using chip A and B of the Affymetrix HG_U133 array and the samples were analyzed for 47 instances. Out of these 47 instances, 24 are of disease states and 23 are of control state. The raw dataset comprises the expression values of 22,273 genes. A log2 transformation of these expression values was used to normalize the range of values. In order to determine a subset of active genes, we used the variance parameter as suggested by [8]. An adaptive discretization strategy is used to convert the $p \times q$ matrix to a $p \times (q \times (q-1)/2)$ matrix, where p represents the number of genes and q represents the number of conditions. Since, we are interested in finding multiple edges between genes, we have used the same line of work as discussed in [5]. For each pair of conditions (i, j) for gene g_g in both the stages, we calculate

$$\omega = \arctan(\mathrm{GM}(g_g, i) - \mathrm{GM}(g_g, j)). \tag{1}$$

The arctan values for each sample pair is then discretized using a threshold value η, which is calculated based on the standard deviation of the expression matrix. The

Table 1 Gene expression value for dummy genes g_1 and g_2 for seven samples

Gene name	S1	S2	S3	S4	S5	S6	S7
g_1	0.33	0.29	0.13	0.72	0.11	0.03	0.86
g_2	0.45	0.14	0.34	0.56	0.56	0.5	0.59

arctan values closer to η are given the same discretized value. In this way, we get a new discretized matrix, D_m of order $p \times (q \times (q-1)/2)$.

The concept of multi-edge networks has been borrowed from [5]. Using this concept, the entry for each pair of genes in D_m is analyzed to obtain sample subsets. These sample subsets represent those conditions under which the D_m matrix shows the same value for two genes. Let us understand this with an example.

Example 1 Suppose we consider two genes g_1 and g_2 with expression values for seven samples as given in Table 1. Using Pearson correlation coefficient, we see that correlation value between g_1 and g_2 is 0.422. If we use a correlation threshold value 0.5 to consider the pair (g_1, g_2) as connected, we see that these two genes do not satisfy this requirement, and hence they are not connected (low correlation is shown with dotted line in Fig. 1a).

To determine whether g_1 and g_2 are connected under subset of conditions, we find the absolute difference between the expression values of the genes at each condition. If the difference between the values is less than some threshold, say, $\Gamma = 0.9$ (as used in [5]), we can use the same label for the two genes. This is shown in Table 2.

In Table 2, we see that g_1 and g_2 are connected via three sets of samples, viz., $\{S1, S2, S4\}$, $\{S3, S7\}$, and $\{S5, S6\}$ (Fig. 1b). Using Pearson correlation, we see that the subsets of samples comprising the three edges have a value greater than 0.5, hence all the three edges connect genes g_1 and g_2. The same process is repeated for each pair of genes for all pairs of conditions giving rise to a multi-edge network or graph (Fig. 1b).

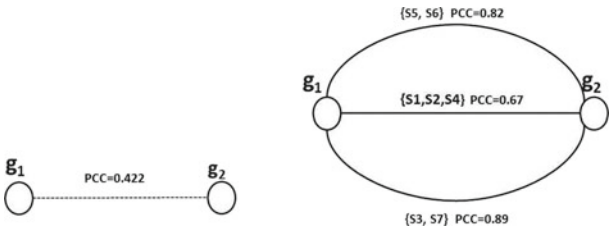


Fig. 1 **a** Genes g_1 and g_2 are shown connected with dotted line since Pearson correlation value between them is less than 0.5. **b** Genes g_1 and g_2 are shown to be connected via three subset of edges

Algorithm 1: Algorithm for module extraction from gene gene network

Input : $Adj = \{V, E\}$ (Gene gene network); $T_{SbOM} = \{V, E'\}$ (TSOM network); γ (Constrained neighbor score); δ (Simson Index threshold);
Output: $Modules = \{C_1, C_2, \cdots, C_N\}$

Initialize clusterExpNode $= V$, NodeList $= V$, Modules $= NULL$, count $=1$;
.....Calculate constrained node score for each gene.....
foreach $g_v \in V$ **do**
  $CnS(g_v) = \frac{|N_{CN}(g_v)|}{degree(g_v)}$
end
.....Module expansion.....
while |NodeList| > 2 **do**
  choose g_a from NodeList such that $CnS(g_a) > CnS(g_k) \forall g_k \in NodeList$ and $CnS(g_a) > \gamma$
  $partialCluster = partialCluster \bigcup g_a$;
  choose g_y such that $T_{SbOM}(partialCluster, g_y) > T_{SbOM}(partialCluster, g_x) \forall g_x \in \{V - g_a\}$
  $SI(g_y, partialCluster) = \frac{N(g_y \cap N(g_k)}{min(|N(g_y)|, |N(partialCluster)|)}$;
  while $SI(g_y, partialCluster) > \delta$ **do**
    $SI(g_y, partialCluster) = \frac{N(g_y \cap N(g_k)}{min(|N(g_y)|, |N(partialCluster)|)}$
    $clusterExpNode = V - partialCluster$;
    $partialCluster = partialCluster \bigcup g_y$;
    choose another g_y such that $T_{SbOM}(partialCluster, g_y) > T_{SbOM}(partialCluster, g_x) \forall g_x \in clusterExpNode$;
  end
  Mark $partialCluster$ as C_{count} only when $|partialCluster| \geq 3$;
  $Modules = Modules \bigcup C_{count}$;
  $count++$;
  NodeList $= V - g_a$;
end
Return $Modules$;

Fig. 2 Module extraction algorithm

Table 2 Labels for dummy genes g_1 and g_2 for seven samples using our edge finding approach

Gene name	S1	S2	S3	S4	S5	S6	S7
g_1	c1	c1	c2	c1	c3	c3	c2
g_2	c1	c1	c2	c1	c3	c3	c2

It is well-established that some information may be lost in a biological network due to the presence of spurious interactions or due to incompleteness of data. To alleviate this issue, Ravasz [3] proposed the concept of a topological overlap matrix incorporating *m*th- order neighbors (in addition to direct neighbors) between any pair of nodes. Studies have also shown that two proteins with a higher topological overlap are more likely to belong to the same functional group in a PPI network [1]. However, the use of Topological Overlap Matrix (TOM) has been restricted to single edge networks, whether weighted or unweighted. In this paper, we analyze the gene network from multi-edge perspective, and hence we use a measure called TSOM as proposed in [5]. We calculate the Topological Subspace Overlap Metric, T_{SbOM} for a pair of nodes a and b in a network with n nodes as follows.

$$T_{\text{SbOM}_{ab}} = 0.5 \times \alpha + 0.5 \times \beta. \tag{2}$$

where α and β represent the similarity between neighbors with direct connection and are given by Eqs. [3] and [4], respectively.

$$\alpha = \frac{\sum_{i=1}^{V} |E^{p}_{g_a g_i} \cap E^{q}_{g_i g_b}|}{\sum_{i=1}^{V} \min(|E^{p}_{g_a g_i}|, |E^{q}_{g_i g_b}|)} \tag{3}$$

where $E^{p}_{g_a g_i}$ represents the edge between genes g_a and g_i; p and q represent the sets of samples for which the Pearson correlation coefficients between (g_a, g_i) and (g_i, g_b) are maximum; and V is the total number of genes in GM.

$$\beta = \frac{|E^{r}_{g_a g_b}|}{sm} \tag{4}$$

where $1 \leq r \leq E_{g_a g_b}$ and $E^{r}_{g_a g_b}$ is the edge between g_a and g_b with as maximum number of sample subsets and sm is the total number of samples in the original matrix, GM. A factor of 0.5 is assigned to both the terms so that equal emphasis is given to edges directly connecting the two nodes g_a and g_b, and also the common edges between both g_a and g_b. This calculation gives rise to a subspace overlap matrix, T_{SbOM} having the same order as GM, i.e., $V \times E$, but with more interesting values indicating connectivity between gene pairs.

2.1 Network Module Extraction

After the subspace overlap matrix is obtained in the previous step, we convert it into a 0–1 matrix, Adj using a threshold value.

$$\text{Adj}_{g_a, g_b} = \begin{cases} 1 & \text{if } T_{\text{SbOM}_{g_a, g_b} > n_{th}} \\ 0 & \text{otherwise} \end{cases}$$

The choice of nth is made wisely using the requirement that two nodes which are connected by more than 50% connectivity have to be taken to be present in the network. This is also supported by many other sources [8]. A network module, in general, is a set of nodes with high topological and neighborhood similarity. To describe the module formation process, the following definitions are required (Fig. 2).

Definition 1 (*Nodedistance*) The distance between two nodes, Nodedistance(g_v, g_k) comprises shortest path from each node, to every other node, g_j, $\forall g_j = \{V - \{g_v, g_k\}\}$

$$\text{Nodedistance}(g_v, g_k) = \sum_{g_j} |(\text{Adj}_{(}g_v, g_j) == 1) \& (\text{Adj}_{(}g_k, g_j) == 1)| \tag{5}$$

Definition 2 (*Constrained neighbor*) A node g_k is a constrained neighbor of node g_v only if (i) Nodedistance(g_v, g_k) >= 1 and (ii) degree$(g_k) \geq 2$.

Definition 3 (*Constrained node score*) The constrained node score of a node g_v is the ratio of the number of constrained neighbors of g_v in the network to the degree of g_v. Mathematically, it is represented as

$$\mathrm{CnS}(g_v) = \frac{|N_{\mathrm{CN}}(g_v)|}{\mathrm{degree}(g_v)} \tag{6}$$

where $N_{\mathrm{CN}}(g_v)$ is the constrained neighbor set of g_v.

Definition 4 (*Simpson Index*) The Simpson Index for a pair of nodes g_v and g_k is given by the ratio of the common neigbors of g_v and g_k to the minimum number of neigbors of g_v and g_k. Mathematically, it is given as

$$\mathrm{SI}(g_v, g_k) = \frac{N(g_v) \cap N(g_k)}{\min(|N(g_v)|, |N(g_k)|)} \tag{7}$$

where $N(g_v)$ and $N(g_k)$ represent the sets of common neighbors of g_v and g_k.

Definition 5 (*Seed node*) A node g_v is chosen as a seed node for module extraction if $\mathrm{CnS}(g_v) \geq \gamma$, where γ is a user-defined threshold.

The module extraction process begins by taking the T_{SbOM} and the Adj matrix as input along with two user-defined thresholds, γ and δ. γ is used during the seed selection process. A node g_i with $\mathrm{CnS}(g_i) \geq \gamma$ is chosen as the seed node. The seed is then expanded to form a cluster. During the cluster expansion process, a node g_j with $T_{\mathrm{SbOM}_{g_i,g_j}} \geq T_{\mathrm{SbOM}_{g_i,g_k}} \forall g_k$ is chosen as a possible candidate for seed pair expansion. The membership of g_j in the *partialCluster* is further strengthened by the Simpson Index measure. If $\mathrm{SI}(g_i, g_j) > \delta$, then g_j can be added into the *partialCluster* and the expansion process continues with the next gene with highest T_{SbOM} value. The process stops when there is no node g_l with $SI(partialCluster, g_l) > \delta$. The elements in the *partialCluster* are declared modules if $|partialCluster| \geq 3$. A minimum limit of three elements has been set in order to declare a cluster a module, because lower size clusters cannot be used effectively for p-value analysis and for inferring the behavior of unknown genes. This constraint has been suggested in [14], where it was assumed that a cluster of proteins has to be of minimum size three. The new cluster formation process begins with the node having the next highest CnS score. This process also ensures the nonexclusive nature of genes in real life.

3 Experimental Results

We implemented our method in MATLAB running on an HP Z 800 workstation with two 2.4 GHz Intel(R) Xeon (R) processors and 12 GB RAM. The main objective of this work is to find functionally enriched modules from gene networks in terms of their p-values. We perform our module extraction technique at different γ and δ values and the p-value of these sets is reported in Table 3.

In order to establish our module extraction method, we have compared it with the method discussed in [8] on the same dataset. This work uses two parameters, CCT and SST, one being the topological threshold and the other is the functional threshold, respectively. We have also used another gene module extraction technique called Module Miner [7] which uses the spanning tree concept to identify modules. Table 4 reports the top p-value obtained from our proposed method and two other existing methods both for control stage and disease stage of Parkinson's Disease.

As can be seen from Table 4, the best p-value obtained in both control and disease stage is given by our proposed method. Thus, we can say that our module extraction process proposed here can take us a step further toward the use of multi-edge information when constructing the gene–gene network and thereafter in finding functionally similar modules.

4 Conclusion

In this paper, we have explored the properties of genes expressed at different time points under different conditions and extended the discussions to define multi-edge

Table 3 p-values of modules obtained using different thresholds in both control and disease stages

ModuleType	p-value			
	$\gamma = 0.03$			
δ	0.2	0.4	0.6	0.8
Control	6.281E−17	5.449E−16	6.251E−16	6.898E−16
Disease	6.983E−8	9.205E−8	1.378E−7	8.262E−8
ModuleType	p-value			
	$\gamma = 0.05$			
δ	0.2	0.4	0.6	0.8
Control	1.825E−17	4.638E−17	4.662E−15	3.154E−13
Disease	4.202E−8	7.397E−8	8.725E−8	4.880E−8
ModuleType	p-value			
	$\gamma = 0.07$			
δ	0.2	0.4	0.6	0.8
Control	4.958E−17	6.648E−17	1.259E−15	6.776E−9
Disease	4.202E−8	4.241E−8	6.988E−8	9.740E−8

Table 4 Best p-value reported using my proposed method and other existing works

Stage	Proposed work	Existing work [8]	Module miner [7]
Control	4.638E−17	4.13E−16	3.67E−10
Diseased	7.397E−8	1.79E−10	5.45E−6

gene–gene networks. We extract modules from the network at different stages of the disease. The use of multiple edges highlights the presence of even those edges which gets eliminated because of higher thresholding in Pearson correlation coefficient. The obtained gene network is more reliable and informative and hence leads to better module extraction and, therefore, would enhance the findings of biologists.

References

1. A.M. Yip, S. Horvath, Gene network interconnectedness and the generalized topological overlap measure. BMC Bioinform. **8**(1), 22 (2007)
2. D. He, Z.-P. Liu, M. Honda, S. Kaneko, L. Chen, Coexpression network analysis in chronic Hepatitis B and C hepatic lesions reveals distinct patterns of disease progression to hepatocellular carcinoma. J. Mol. Cell Biol. **4**(3), 140–152 (2012)
3. E. Ravasz, A.L. Somera, D.A. Mongru, Z.N. Oltvai, A.-L. Barabási, Hierarchical organization of modularity in metabolic networks. Science **297**(5586), 1551–1555 (2002)
4. G.F. Berriz, O.D. King, B. Bryant, C. Sander, F.P. Roth, Characterizing gene sets with funcassociate. Bioinformatics **19**(18), 2502–2504 (2003)
5. H.A. Ahmed, P. Mahanta, D.K. Bhattacharyya, J.K. Kalita, Module extraction from subspace co-expression networks. Netw. Model. Anal. Health Inform. Bioinform. **1**(4), 183–195 (2012)
6. I.R. Medina, Z. Lubovac-Pilav, Gene co-expression network analysis for identifying modules and functionally enriched pathways in type 1 diabetes. PLoS ONE **11**(6), e0156006 (2016)
7. P. Mahanta, H.A. Ahmed, D.K. Bhattacharyya, J.K. Kalita, An effective method for network module extraction from microarray data. BMC Bioinform. **13**(13), S4 (2012)
8. P. Sharma, D.K. Bhattacharyya, J. Kalita, Disease biomarker identification from gene network modules for metastasized breast cancer. Sci. Rep. **7**(1), 1072 (2017)
9. S. Kumari, J. Nie, H.-S. Chen, H. Ma, R. Stewart, X. Li, M.-Z. Lu, W.M. Taylor, H. Wei, Evaluation of gene association methods for coexpression network construction and biological knowledge discovery. PLoS ONE **7**(11), e50411 (2012)
10. S. van Dam, U. Võsa, A. van der Graaf, L. Franke, J.P. de Magalhães, Gene co-expression analysis for functional classification and gene–disease predictions. Brief. Bioinform. bbw139 (2017)
11. S.M.M. Hossain, S. Ray, A. Mukhopadhyay, Preservation affinity in consensus modules among stages of HIV-1 progression. BMC Bioinform. **18**(1), 181 (2017)
12. S. Ray, S. Biswas, A. Mukhopadhyay, S. Bandyopadhyay, Detecting perturbation in co-expression modules associated with different stages of HIV-1 progression: a multi-objective evolutionary approach, in *2014 Fourth International Conference of Emerging Applications of Information Technology (EAIT)* (IEEE, 2014), pp. 15–20
13. S. Ray, U. Maulik, Identifying differentially coexpressed module during hiv disease progression: a multiobjective approach. Sci. Rep. **7** (2017)
14. T. Nepusz, Yu. Haiyuan, A. Paccanaro, Detecting overlapping protein complexes in protein-protein interaction networks. Nat. Methods **9**(5), 471–472 (2012)
15. T. Kakati, H. Kashyap, D.K. Bhattacharyya, THD-module extractor: an application for CEN module extraction and interesting gene identification for Alzheimers disease. Sci. Rep. **6**(1), 38046 (2016)
16. V. Deshpande, A. Sharma, R. Mukhopadhyay, L.N. Thota, M. Ghatge, R.K. Vangala, V.V. Kakkar, L. Mundkur, Understanding the progression of atherosclerosis through gene profiling and co-expression network analysis in Apob tm2Sgy Ldlr tm1Her double knockout mice. Genomics **107**(6), 239–247 (2016)
17. Y. Yang, L. Han, Y. Yuan, J. Li, N. Hei, H. Liang, Gene co-expression network analysis reveals common system-level properties of prognostic genes across cancer types. Nat. Commun. **5**, 3231 (2014)

A Web Portal to Calculate Codon Adaptation Index (CAI) with Organism Specific Reference Set of High Expression Genes for Diverse Bacteria Species



Piyali Sen, Abdul Waris, Suvendra Kumar Ray and Siddhartha Sankar Satapathy

1 Introduction

The nonarbitrary usage of synonymous codons which belong to akin group is common across genomes, a phenomenon called Codon Usage Bias(CUB). Several mutational factors such as genome G+C% and strand asymmetric nucleotide composition [1, 2, 9–11, 15] are known to influence CUB. In addition to these mutational factors, selective forces also influence CUB though in variable strength [21]. During the process of translation, certain codons are used more frequently in comparison to other synonyms for faster and/or accurate translation in high expression genes than rest of the genes in a genome. This is considered as the primary selection factor influencing CUB [7] which is a common phenomenon in genomes of bacteria. Based on this concept, several mathematical formulas have been proposed for measuring codon usage bias [19]. Among those formula, Codon Adaptation Index (CAI) proposed by Sharp and Li (1987) is a popular measure. The CAI can be described as follows:

To what extent codon usage of a gene is adapted toward the codon usage of highly expressed genes in a genome. The CAI index is defined as the geometric mean of these values.

$$\mathrm{CAI} = \exp\left(\frac{1}{L}\sum_{k=1}^{L}\ln(w_k)\right) \tag{1}$$

P. Sen (✉) · A. Waris · S. S. Satapathy
Department of Computer Science and Engineering, Tezpur University, Napaam, Tezpur 784028, Assam, India
e-mail: piyalisen18@gmail.com

S. S. Satapathy
e-mail: ssankar@tezu.ernet.in

S. K. Ray
Department of Molecular Biology and Biotechnology, Tezpur University, Napaam, Tezpur 784028, Assam, India

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_31

Here, w_k is the relative adaptedness of kth codon, i.e., the ratio of the usage of the codon to that of the most frequent codon among the synonymous codons. L is the length of the gene in terms of number of codons. Amino acids methionine and tryptophan codons are not considered while calculating L. Mathematically, w_k is expressed as

$$w_k = \frac{\text{RSCU}_k}{\text{RSCU}_{\text{max}}} \tag{2}$$

where RSCU_k is the relative synonymous codon usage of a codon k and RSCU_{max} is the RSCU values for the most frequently used codons for an amino acid. RSCU is defined as follows:

$$\text{RSCU}_i = \frac{X_i}{\frac{1}{n}\sum_{i=1}^{n} X_i} \tag{3}$$

Here, X_i is the count of the ith codon among n number of synonymous codons for an amino acid.

Theoretically, possible maximum and minimum values for CAI are 1.0 and 0.0, respectively. CAI is often considered for predicting gene expression from CUB. Higher the CAI value, more is the expression level of a gene. Sharp and Li [24] showed that CAI values of the genes correlate with their expression levels in *Escherichia coli*. Therefore, CAI is a popular method for predicting expression level of the genes theoretically.

2 Limitations in Existing Implementation of CAI

Implementation of CAI is available in several software such as CodonW [13, 30], INCA [26], CAIcal [16], EMBOSS [18], CAI Calculator 2 [27], and DAMBE7 [29]. However, there are several difficulties in using existing software for calculating CAI. One of the crucial step in calculating CAI is to select the reference set of high repression genes and then to calculate the relative adaptedness of the codons, i.e, the w_k values as given in Eq. 2. The above methods have employed various approaches to do so, and have their own limitations. In CAIcal, there is dependency of reference database tables. If the organism for which CAI value is to be calculated is not present in the database, then the user have to generate the reference database table. This approach is not user-friendly for providing information about codon usage in reference set of high expression genes. It has been reported by Xia [28] that the EMBOSS [18] and CAI Calculator 2 [27] software provide erroneous result, might be because of some implementation issues.

These organism specific w_k values are available in existing software only for few organisms. Therefore, users can directly calculate CAI values for the genes of only those organisms. For calculating CAI for the genes of an organism, CodonW [30] provides alternative indirect approach using correspondence analysis [6].

2.1 Using Correspondence Analysis for Estimating Codon Usage in High Expression Gene Set May Not Be Appropriate

Correspondence analysis, a kind of multivariate statistical technique may be used to distinguish codon usage in highly and lowly expressed genes [4, 5]. Therefore, user may use correspondence analysis option available in CodonW software to estimate w_k values statistically in terms of a file called cai.coa and then use this file to calculate CAI. However, when there is only mutational bias with very low selection pressure, there can be huge difference in codon bias among genes and therefore, reference set estimated by correspondence analysis may not necessarily be high expression genes. Therefore, there is a need of a database of organism specific reference set of high expression genes for calculation of CAI values.

2.2 Considering Default E. coli Reference Set May Generate Erroneous Result

High expression gene set of *E. coli* is also suggested to be used as a default set for calculating CAI. *E. coli* is an organism with strong selection on codon usage [4] and (G+C)% around 50.0. However, (G+C)% of the organism varies widely among bacteria from as less as around 17.0% to more than 75.0% [17]. Furthermore, though the selected codon usage bias is universal among organisms [25], it varies from organism to organism [21, 23] and also differs among bacteria phylogeny [20]. While selection on codon usage is very strong in *E. coli*, *Bacillus subtilis*, and *Saccharomyces cerevisiae* (yeast), it is very low in several other organisms [21]. Therefore, calculating CAI considering default E. coli reference set can generate erroneous result.

Alternatively, if the user is familiar with any high-level computer programming languages, they may calculate w_k values from high expression gene sequences separately and input the values to CodonW to calculate CAI. These approaches may be complicated for naive users and not suitable for researchers unfamiliar with high-level programming languages. Keeping these constraints in calculating CAI in the command-driven CodonW and other software in view, we developed a web portal that provides online tool for calculating CAI.

3 Codon Adaptation Index (CAI) Web Portal

3.1 Reference Set of High Expression Genes Available in the Web Portal

We downloaded the bacterial genomes from the NCBI site [31]. Then,we extracted a set of genes known to be highly expressed and widely conserved across organism [24] from these genomes using Python scripts and made available in our portal as the reference gene sets. Ribosomal protein genes; outer membrane protein genes such as rplA, rpmB, and rpsA; elongation factor genes such as tufA, tufB, and fusA; and regulatory/repressor genes such as dnaG and araC are some of the example genes considered in the high expression gene sets. At present, we have provided high expression gene sets for 684 unique species of bacteria in our database. These bacteria belong to 29 different phylogenetic groups and with coding region (G+C)% between 23.65 and 74.68 as shown in Table 1. High expression gene sets for *E. coli* [8], *Saccharomyces cerevisiae*(yeast) [3], and *Homo sapiens* (human) [14, 22] available in our portal are based on the experimental expression data.

3.2 Server Configuration and Language Used for the Web Portal

Our web portal is launched in an IBM System x3630 M4 server with CentOS 6.10 operating system. The web portal is developed using Python programming language.

3.3 Description of How to Use our Web Portal

Keeping limitations of the available softwares and lack of reference set of high expression genes for large number of organisms in view, we envisaged this web portal. It is designed to simplify the computation. It is very simple to use and accessible in internet from any computer. It is designed not to have any limitation on the input genome sequences length. The user interface of the portal provides a two-step process to calculate CAI.

Step I: Upload the file(in fasta format) containing all the genes. The first step is to input the nucleotide sequence of the genes whose CAI values are to be calculated in a single file in FASTA format.

Step II: Select one of the following for High Expression Gene Set. Second step of the calculation is to provide additional input file of reference set of high expression genes. The web portal provides three simple options for input of this reference set.

Table 1 Details of bacteria whose reference set of high expression genes available in the web portal

Sl no.	Bacterial Group	No. of Organisms	Maximum (G+C)%	Minimum (G+C)%
1	Acidobacteria	1	61.1	61.1
2	Actinobacteria	80	74.5	46.31
3	Alphaproteobacteria	86	72.09	30.37
4	Aquificae	7	52.24	32.03
5	Bacteroidetes	35	66.73	27.25
6	Betaproteobacteria	59	70.36	37.78
7	Chlamydiae	10	44.31	36.1
8	Chlorobi	8	57.66	45.06
9	Chloroflexi	5	60.94	47.85
10	Cyanobacteria	10	62.86	40.4
11	Deferribacteres	3	43.2	31.08
12	Deinococcus-Thermus	11	70.23	63.01
13	Deltaproteobacteria	28	74.68	37.48
14	Dictyoglomi	2	33.99	33.81
15	Elusimicrobia	1	40.69	40.69
16	Epsilonproteobacteria	16	44.89	27.19
17	Fibrobacteres	1	48.89	48.89
18	Firmicutes	124	69.29	28.34
19	Fusobacteria	3	34.69	26.2
20	Gammaproteobacteria	137	70.46	23.65
21	Gemmatimonadetes	1	64.49	64.49
22	Nitrospirae	1	34.16	34.16
23	Planctomycetes	2	57.91	55.46
24	Spirochaetes	17	52.08	27.7
25	Synergistetes	2	64.45	45.75
26	Tenericutes	18	40.66	23.96
27	Thermodesulfobacteria	2	42.61	30.67
28	Thermotogae	11	47.08	30.73
29	Verrucomicrobia	3	65.47	45.85

Option a: Provide nucleotide sequences of high expression gene set as an input file. The user can input the reference set of high expression genes in the form of a FASTA file.

Option b: Select the name of the organism from a drop down list box. At present, the web interface provides a list of 684 bacteria species whose high expression gene set is available in our database. User can select the name of the organism from this list corresponding to the organism whose gene sequences were uploaded in the first step.

Option c: Select the list of genes from the uploaded file in Step I. In the third alternative, the web interface shows the gene informations from the uploaded file

in the first step. User can select multiple genes know to be highly expressed from the displayed list. Those selected genes will be considered as the reference set while calculating CAI.

Integrity checks: Before processing, the web portal examines the accuracy of the input sequences. These include presence of internal stop codons, presence of accepted start (i.e., NTG, ATN) [12] and stop codons, presence of non-IUPAC characters. If any of these problems are found, CAI is calculated for those sequences with potential errors with appropriate warning messages. Therefore, sequences that generate warnings should be carefully checked.

Based on these, output file is generated in terms of excel file. Along with the CAI values, the output file also contains additional information about the genes such as length in terms of number of amino acids, G+C%. Once the result is produced, no input sequence files are retained in the server to avoid any possible misuse of the users data.

4 Conclusion

The web portal is available for free for academic and research purpose in our web server at http://14.139.219.242:8003/cai. At present, the web portal can be used for calculating CAI as per universal genetic code table. Future scope lies for consideration of other available versions of genetic code. We believe that our web portal will be helpful for biologists working on molecular evolution.

References

1. S.L. Chen, W. Lee, A.K. Hottes, L. Shapiro, H.H. McAdams, Codon usage between genomes is constrained by genome-wide mutational processes, in *Proceedings of the National Academy of Sciences,* vol. 101(10) (2004), pp. 3480–3485
2. A. Frank, J. Lobry, Asymmetric substitution patterns: a review of possible underlying mutational or selective mechanisms. Gene **238**(1), 65–77 (1999)
3. S. Ghaemmaghami, W.K. Huh, K. Bower, R.W. Howson, A. Belle, N. Dephoure, E.K. O'shea, J.S. Weissman, Global analysis of protein expression in yeast. Nature **425**(6959), 737 (2003)
4. M. Gouy, C. Gautier, Codon usage in bacteria: correlation with gene expressivity. Nucleic Acids Res. **10**(22), 7055–7074 (1982)
5. R. Grantham, C. Gautier, M. Gouy, M. Jacobzone, R. Mercier, Codon catalog usage is a genome strategy modulated for gene expressivity. Nucleic Acids Res. **9**(1), 213–213 (1981)
6. M.J. Greenacre, *Correspondence analysis* (Academic Press, London, 1984)
7. R. Hershberg, D.A. Petrov, General rules for optimal codon choice. PLoS Genet. **5**(7), e1000556 (2009)
8. Y. Ishihama, T. Schmidt, J. Rappsilber, M. Mann, F.U. Hartl, M.J. Kerner, D. Frishman, Protein abundance profiling of the escherichia coli cytosol. BMC Genomics **9**(1), 102 (2008)
9. J.R. Lobry, Asymmetric substitution patterns in the two dna strands of bacteria. Mol. Biol. Evol. **13**(5), 660–665 (1996)

10. J.O. McInerney, Replicational and transcriptional selection on codon usage in borrelia burgdorferi, in *Proceedings of the National Academy of Sciences*, vol. 95(18) (1998), pp. 10698–10703
11. A. Muto, S. Osawa, The guanine and cytosine content of genomic dna and bacterial evolution, in *Proceedings of the National Academy of Sciences*, vol. 84(1) (1987), pp. 166–169
12. S. Osawa, T.H. Jukes, K. Watanabe, A. Muto, Recent evidence for evolution of the genetic code. Microbiol. Mol. Biol. Rev. **56**(1), 229–264 (1992)
13. J.F. Paden, CodonW. (University of Nottingham, 1999)
14. J.B. Plotkin, H. Robins, A.J. Levine, Tissue-specific codon usage and the expression of human genes, in *Proceedings of the National Academy of Sciences*, vol. 101(34) (2004), pp. 12588–12591
15. B.R. Powdel, M. Borah, S.K. Ray, Strand-specific mutational bias influences codon usage of weakly expressed genes in escherichia coli. Genes to Cells **15**(7), 773–782 (2010)
16. P. Puigbò, I.G. Bravo, S. Garcia-Vallve, Caical: a combined set of tools to assess codon usage adaptation. Biol. Direct **3**(1), 38 (2008)
17. R. Raghavan, Y.D. Kelkar, H. Ochman, A selective force favoring increased g+ c content in bacterial genes, in *Proceedings of the National Academy of Sciences*, vol. 109, issue 36 (2012), pp. 14504–14507
18. P. Rice, I. Longden, A. Bleasby, Emboss: the european molecular biology open software suite. Trends Genet. **16**(6), 276–277 (2000)
19. A. Roth, M. Anisimova, G.M. Cannarozzi, Measuring Codon Usage Bias. Codon Evolution: Mechanisms and Models (Oxford University Press Inc, New York, 2012), pp. 189–217
20. S.S. Satapathy, B.R. Powdel, A.K. Buragohain, S.K. Ray, Discrepancy among the synonymous codons with respect to their selection as optimal codon in bacteria. DNA Res. **23**(5), 441–449 (2016)
21. S.S. Satapathy, B.R. Powdel, M. Dutta, A.K. Buragohain, S.K. Ray, Selection on ggu and cgu codons in the high expression genes in bacteria. J. Mol. Evol. **78**(1), 13–23 (2014)
22. S. Satapathy, S. Ray, A. Sahoo, T. Begum, T. Ghosh, Codon usage bias is not significantly different between the high and the low expression genes in human. Int. J. Mol. Genet. Gene Ther **1**(1) (2015)
23. P.M. Sharp, E. Bailes, R.J. Grocock, J.F. Peden, R.E. Sockett, Variation in the strength of selected codon usage bias among bacteria. Nucleic Acids Res. **33**(4), 1141–1153 (2005)
24. P.M. Sharp, W.H. Li, The codon adaptation index-a measure of directional synonymous codon usage bias, and its potential applications. Nucleic Acids Res. **15**(3), 1281–1295 (1987)
25. F. Supek, N. Škunca, J. Repar, K. Vlahoviček, T. Šmuc, Translational selection is ubiquitous in prokaryotes. PLoS Genet. **6**(6), e1001004 (2010)
26. F. Supek, K. Vlahoviček, Inca: synonymous codon usage analysis and clustering by means of self-organizing map. Bioinformatics **20**(14), 2329–2330 (2004)
27. G. Wu, D.E. Culley, W. Zhang, Predicted highly expressed genes in the genomes of streptomyces coelicolor and streptomyces avermitilis and the implications for their metabolism. Microbiology **151**(7), 2175–2187 (2005)
28. X. Xia, An improved implementation of codon adaptation index. Evol. Bioinform. **3** (2007) (117693430700300028)
29. X. Xia, Dambe7: new and improved tools for data analysis in molecular biology and evolution. Mol. Biol. Evol. **35**(6), 1550–1552 (2018)
30. CodonW: http://www.codonw.sourceforge.net
31. National Center for Biotechnology Information: http://www.ncbi.nlm.nih.gov

Blockchain-Based Transparent and Secure Decentralized Algorithm



Shreya Sudhakaran, Sunil Kumar, Priya Ranjan and Malay Ranjan Tripathy

1 Introduction

Blockchain-based currency makes a cheap, quick, and secure record. It has many usages apart from financial tasks such as building a decentralized app that can help to cast our votes, proving the actual authenticity of material, i.e., tracing back material to its origin. Blockchains take away control from the central authority and give them back to the users. In a blockchain network, users have control over what they want to share and how much of the data that they want to share. We no longer will have to depend on a central agent [1]. It uses a hash function and SHA256 algorithm which is responsible for the Avalanche effect, i.e., even for a minute change in the block, the hash function will change completely. This will, in turn, result in changing the block of the next block and so on. So, even if an attacker tries to change the configuration of a block it will alert the participants in the network and if **51% of the participant** vote against any such change it will be restored back to the normal [2, 3]. Therefore, it is a transparent and secure system. It has been noncontroversial and has worked flawlessly over the years (Fig. 1).

Current system that we are familiar with is totally dependent upon trust. For example, when we take the taxi, we need to trust the driver that he will charge a fair price. That was the reason that meters were installed, i.e., we are no longer trusting

S. Sudhakaran · S. Kumar (✉) · P. Ranjan · M. R. Tripathy
Amity School of Engineering and Technology, Amity University, Noida, Uttar Pradesh, India
e-mail: skumar58@amity.edu; sk4sunilkumar@gmail.com

S. Sudhakaran
e-mail: shreyasudhakaran5@gmail.com

P. Ranjan
e-mail: pranjan@amity.edu

M. R. Tripathy
e-mail: mrtripathy@amity.edu

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_32

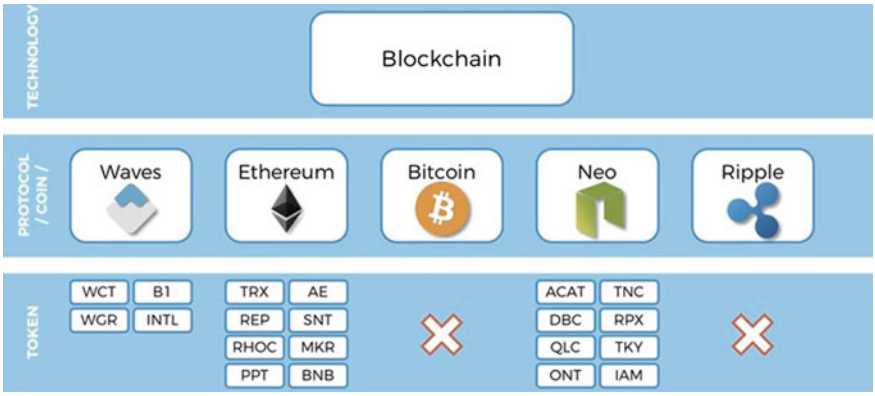


Fig. 1 Architecture of blockchain technology [2]

a human being. We have shifted our trust from a human to a machine. Another example can be that, you send a mail to someone we need to rely on that software to tell us the truth about the delivery of our mail or it can be a bank telling us that that particular transaction has been successful and has reached your family. Also, there have been instances that people have claimed or have put allegation that the EVM used to conduct polls were hacked [4]. With the implementation of blockchains, all these problems can be solved very smoothly. The main idea is to cut out the middleman completely. It can provide secure and verifiable transactions that can reduce operational and counter-party risks (Fig. 2).

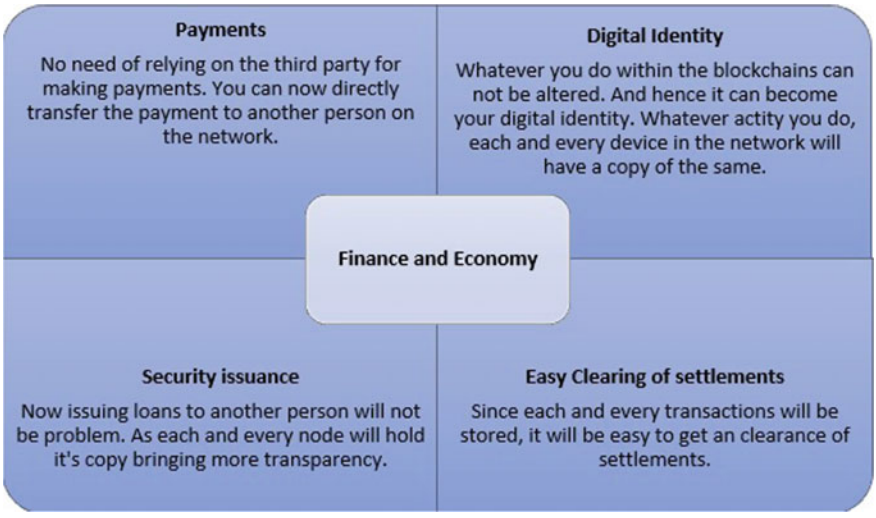


Fig. 2 Finance and economy model [4]

Transaction in a blockchain network is very much different from that of a bank that we are used to. There is no account in a blockchain where we can deposit all our money. There are only transactions which are called UTXOs, i.e., Unspended Transactional Output. In order to perform a transaction, the UTXOs should be killed off and the only way it can be done is by feeding it to new transactions. Now, as we have seen that the miners will pick only those transactions which have a higher fee attached to it. Although, it is said that the transactions are free and no one is taking money from you but we need to assign a certain fee to it so that the transaction gets picked up and is put in a mempool. Higher the transaction fee, more likely it will be picked up. There is no coin, no paper nor any number that exists in the transaction. It is very different from the conventional banking approach [5–7].

1.1 Web 3.0: The Future

Interested persons buy their own server and it started the decentralization of the website. Now bigger and bigger technologies started using a central server [4]. People who want access to the services should be connected to the server, and it is really unfortunate in the following sense:

1. It is shrinking the economy.
2. Too much centralized power.
3. We have no say in how our data is used.
4. We should get paid for our data.
5. It can be shut down (Fig. 3).

In a decentralized system even if one node gets shut down due to any reason, the rest of the system will work just fine. There is no central node and every node is termed equal in terms of authority. The results and failures are better handled in these types of systems. Web 3.0 is the future of our Internet, where HTTP will be replaced with IPFS. A web can be thought of like a pendulum. It started off as decentralized then it became centralized and again it is going to be decentralized [15] (Fig. 4).

2 Related Work

Bitcoin was built by Nakamoto in 2008 and has since enjoyed a rapid growth both in value as in transaction volume [8–10]. However, the design of Bitcoin intrinsically limits the rate it can process transactions. Recent work considers how Bitcoin's network can delay or prevent block propagation [11–13]. It has also focused on how blockchain technology can alter the current economy of the world. There are many research papers also written upon how the Bitcoin Mining will affect the world as it requires a lot of electricity and can only be used if the currency has cheap electricity

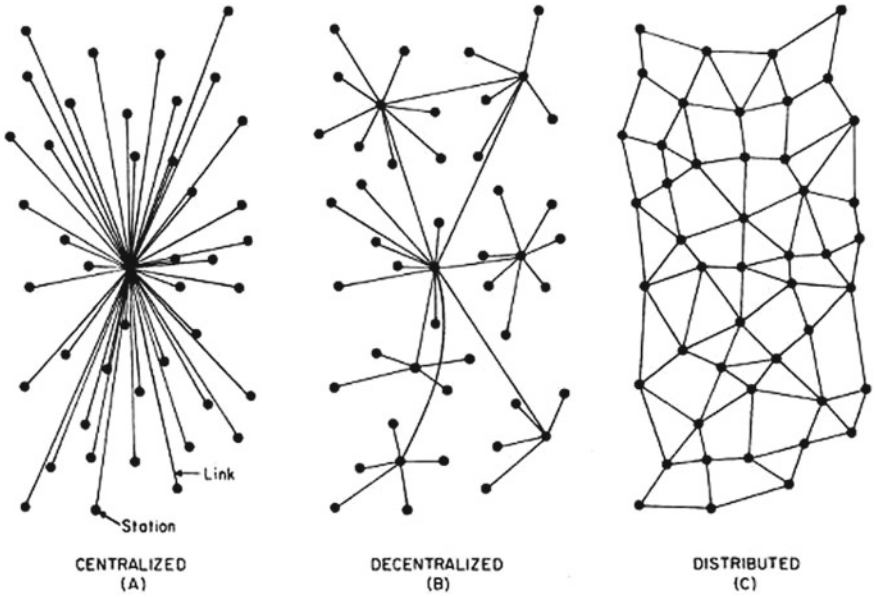


Fig. 3 Techniques of data sharing [13]

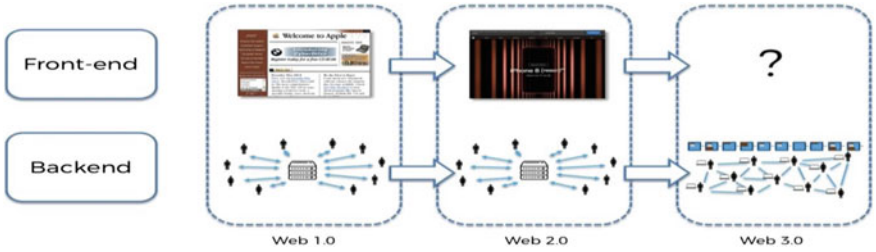


Fig. 4 In Web 1.0 the front end was not very interactive or beautiful. In Web 2.0 the front end became extremely interactive and user-friendly but is still a centralized system. In the future, we will see Web 3.0 which will be a decentralized system and every user will have equal roles and participation in the decision-making. But we still do not know how the front end of Web 3.0 will look like

like China [14]. Most of the research is concentrating on uncovering and improving restrictions of blockchain from protection and security points of view. However, huge numbers of the proposed arrangements lack concrete evaluation on their effectiveness. Anonymity, data integrity, and security attributes set a lot of intriguing difficulties and questions that should be discovered and surveyed with excellent research. Scalability is likewise an issue that should be illuminated for future needs. The blockchain is for the most part known for its connection to Bitcoin digital money. Bitcoin utilizes blockchain innovation in cash exchanges. In any case, Bitcoin digital

currency is not the main arrangement that utilizes blockchain innovation. Therefore, it is essential to locate the present applications created by utilizing blockchain innovation. Identifying different applications can comprehend different directions and approaches to utilize blockchain.

3 Problems in Blockchain

DAO was the first decentralized app. It was based on Ethereum. It was stateless. DAO was crowd-funded and it raised $150,000,000. Unfortunately, there was an error in the code of the smart contract. It was attacked and hacked in 2016 for $50,000,000. Interestingly, the attacker did not do anything illegal. He just saw the flaw and used it to transfer the money. And nobody could do anything about it because the contract is immutable and once it is on blockchain you cannot change it. It gives rise to a question Is Code law? The problem was in the coding of DAO and not in Ethereum. Some of the major disadvantages of blockchain are discussed in Fig. 5.

3.1 Problem Statement

How can we bring transparency and security in voting that is trusted by all? We go to Facebook that has a centralized server. It acts as a middleman. They have control over our data and privacy. They are a regulator of the ads that we are seeing and they can even charge us for the services that we are taking. There is a lot of control that we are giving to them. Similarly, while casting votes in a democratic country, we use EVM machines that are hackable. Which will give one party an unfair chance of winning the elections. Which actually defeats the purpose of democracy.

3.2 Technique that Can Be Used to Solve This Problem

The smart contract is collection of simple codes that are written in a blockchain network which are agreeable to all parties involved. Once the smart contract is implemented in the system it just cannot be changed. These are of extreme importance as the codes are law here. They contain the code as to how to behave when a new node joins the network or an existing node leaves the network. They are all predefined. It is written in Solidity which is a Turing complete language. The smart contract has the following features:

1. It can actually store the history of all smart contracts that have been there on the bockchain.

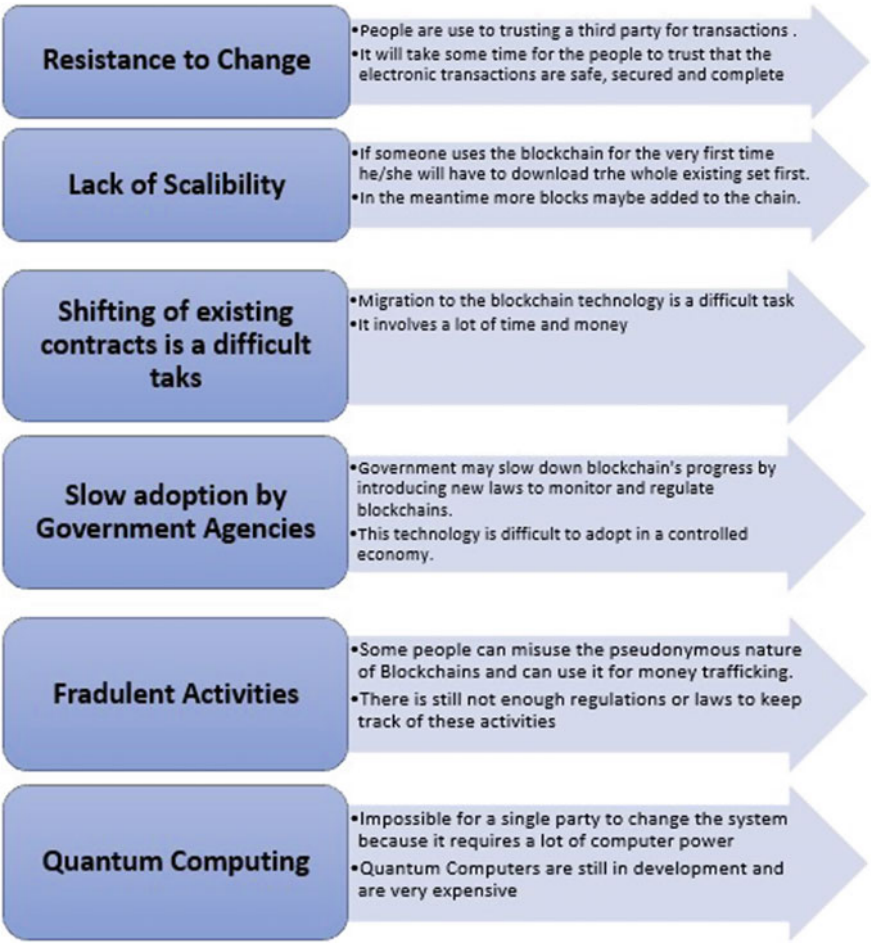


Fig. 5 Disadvantages of Blockchain [18]

2. History of all transactions that have taken place. It can be of extreme importance especially in a banking center where all the records of loan transfers can be maintained for a transparent and fair system.
3. It also shows the current state of all smart contracts.

3.3 Problem Solution

We all are interconnected! In Web 3.0 the people will actually have control over the data. Knowledge is power, and data can be transformed into knowledge. Data

is the natural resource of our time. In Web 3.0 there will not be any middlemen. There will be blockchains and every member in the chain will have a copy of them in their system. There will be predefined codes as to how the system will work when a new user joins in and when an existing user quits. It becomes so much easier as we completely cut out the middlemen. We will have control over the data that we have generated. I. Steemit is a blockchain version of Twitter. The control over the data will be stored using a password on the blockchain and only you will have control over the data. And you will have complete control over the fact that who gets to see your data and how much of your data you want to reveal. You get the control back of your data. We are leveraging our own machines instead of a central server. We can have a social network that does not belong to anyone. It just runs and it runs by the people who are using it. It is crowdsourced. People will have votes on how it will work.

3.3.1 Decentralized Applications (Dapps)

It contains interface for people to connect with an application. It contains both frontend and backend. The front end of these applications can be made using Node.js or JavaScript as in Web 2.0 but the real power lies in the backend. In the backend, these will not be connected to any server or central authority. Instead, they will be connected to a blockchain that has end-to-end connectivity. The participant in that network will have a copy of that particular blockchain in their system. It is said that the world's next supercomputer will not be one giant piece of computers. Rather, it will be a collection of a group of computers that interact in a network. Such a decentralized application or in simple words Dapp will be controlled by the people and not by any other agency. There will be transparency in the system. There will not be centralization of power and we will have a proper say in how our data will be used and we could actually get paid for the natural resource that we are generating.

3.3.2 Program Code—Making a Decentralized Voting Application

The big difference between this decentralized application and the normal voting system that exists today lies in the fact that this voting app is unhackable. But it is hackable in the fact that you will have to have access to 500 supercomputers in the world combined. So, it is basically unhackable unless you have access to that much computing power in the world.

```
Pragma solidity 0.4.11;
Contract EVote{  Mapping(bytes32 => uint8) public votesReceived;
  bytes32[] public candidatesInElection;
   function EVote(byte32[] candidates){
            candidatesInElection=candidates;}

   function total(bytes32 candidate) returns (uint8){
   if(validCandidate(candidate)==false) throw;
   else
   return votesReceived[candidate];
   function voteForCandidate(bytes32 candidate){
   if(validCandidate(candidate)==false) throw;
            votesReceived[candidate]=votesReceived[candidate]+1;}
function validCandidate(bytes32 candidate) returns(bool)
while(i<candidatesInElection){
     if(candidatesInElection[i]==candidate){
     i++;
     return true;}}
return false;
} }
```

3.3.3 Comparison of Present Centralized System to that of a Decentralized System for Voting

S. no.	Centralized system	Decentralized system
1.	It is not 100% secure. It can be changed or manipulated	It is extremely secure. Data once stored in a blockchain can never be manipulated or changed
2.	EVMs can be breached if the chip is manipulated or the Bluetooth is inserted	Since blockchains are spread across a large number of computers even if one system is breached other 51% will be secured and force will take the decision
3.	Trojans can be inserted	The data can be corrupted only if the smart contract is corrupted. That is the smart contract is created by some unfair means
4.	EVMs can go out of order and due to that time is wasted to replace the machine	A secure application using blockchains can never go out of order and can never be shut down

Consensus Algorithm

Algorithm based on consensus is a set primarily for achieving settlement randomly in the networks. Compared to most people balloting, it emphasizes that the complete organization should benefit by attaining a consensus. Such coordinated consensus may be tampered in the presence of malicious actors and defective processes. For example, a terrible actor might also secretly create conflicting messages to make institution members fail to act in unison, which breaks down the effectiveness of the

institution to coordinate its actions. This technique is also known as the "Byzantine Generals Problem" (BGP). The failure of accomplishing consensus because of faulty actors is referred to as Byzantine fault. Leslie Lamport, Marshall Pease, and Robert Shostak confirmed in 1982 that Byzantine fault tolerance may be executed handiest, if a majority agreement can be reached by means of the sincere generals on their strategy. The consensus algorithms popularly utilized in contemporary blockchain structures offer a probabilistic solution to BGP.

4 Conclusion

Even if we build secure and safe voting applications using blockchains, it will give rise to many other questions. Who gets the authority to make decisions? How can the system ensure that the validators are who they say they are? A person may feel more secure by using blockchains but it actually only gives the user more power to change whatever they want even without the permission of the fellow users and that very much contradicts the very purpose of the blockchains. "Secure" is very hard to define in the context of blockchains as secure from whom? secure for what? It just totally depends on the perspective.

References

1. M. Conti, E.S. Kumar, C. Lal, S. Ruj, A survey on security and privacy issues of bitcoin. IEEE
2. G. Wood, ETHEREUM: a secure decentralized generalised transaction ledger. Tech. Rep. (2014). http://gavwood.com/Paper.pdf
3. T. Moore, N. Christin, Beware the middleman: Empirical analysis of bitcoin-exchange risk, in *Financial Cryptography and Data Security: 17th International Conference* (Springer, Berlin, Heidelberg, 2013), pp. 25–33
4. S. Goldfeder, J. Bonneau, E.W. Felten, J.A. Kroll, A. Narayanan, Securing bitcoin wallets via threshold signatures (2014). http://www.cs.princeton.edu/stevenag/bitcointhresholdsignatures.pdf
5. M. Vasek, M. Thornton, T. Moore, Empirical analysis of denial-of-service attacks in the bitcoin ecosystem, in *Financial Cryptography and Data Security: FC 2014 Workshops, BITCOIN and WAHC 2014* (Springer, Berlin, Heidelberg, 2014), pp. 57–71
6. Block Chain Info, Total Bitcoins in circulation. https://markets.blockchain.info/
7. S. Nakamoto, Bitcoin: A peer-to-peer electronic cash system. https://bitcoin.org/bitcoin.pdf
8. K. Croman, C. Decker, I. Eyal, A.E. Gencer, A. Juels, A. Kosba, A. Miller, P. Saxena, E. Shi, E.G. Sirer, D. Song, R. Wattenhofer, On scaling decentralized blockchains, in *2016 International Conference on Financial Cryptography* (2016)
9. A. Ghosh, M. Mahdian, D.M. Reeves, D.M. Pennock, R. Fugger, Mechanism design on trust networks, in *Internet and Network Economics* (2007)
10. A. Miller, M. Moser, K. Lee, A. Narayanan, An Empirical analysis of linkability in the Monero blockchain. CoRR, abs/1704.04299 (2017)
11. S. Meiklejohn, M. Pomarole, G. Jordan, K. Levchenko, D. McCoy, G.M. Voelker, S. Savage, A fistful of bitcoins: characterizing payments among men with no names, in *Proceedings of the 2013 Internet Measurement Conference* (2013)

12. B. Viswanath, M. Mondal, K.P. Gummadi, A. Mislove, A. Post, Canal: scaling social network-based sybil tolerance schemes, in *Proceedings of the 2012 European Conference on Computer Systems* (2012)
13. https://www.udemy.com/blockchain-101-beginners-free-course-bootcamp-cryptocurrency/
14. https://www.udemy.com/your-first-decentralized-app/
15. M. Milutinovic, W. He, H. Wu, M. Kanwal, Proof of luck: an efficient blockchain consensus protocol, in *Proceedings of the 1st Workshop on System Software for Trusted Execution, ser. SysTEX 16. ACM, 2016.* = *Evolution Programs*, 3rd edn. (Springer, Berlin, Heidelberg, New York, 1996)

Prediction of Cancer Diagnosis Patients from Fine-Needle Aspirates Using Machine Learning



Deepak Mehta and Chaman Verma

1 Introduction

From 1989 to 1995, University of Wisconsin Hospital collected data of cancer patients [1]. Instead use of statistical methods [2, 3], the use of machine learning ML in trending in the prediction task is trending [4–7]. In this paper, the author used ML to predict the malignant diagnosis of a patient using SVM and neural network machine learning classifier models to solve binary classification problems [8–10]. The reason behind the selection of machine learning research is to learn the ability to recognize the complex pattern that exists in data and try to make an intelligent decision [11]. The supervised learning works on the assumption that given training data are labeled by class and forecast is made on those labels [9]. The classification model is popular from the past and their studies can be found in prediction based on label present in terrorism, weather, finance, medical, and many more fields. Binary or multiclass classification [12, 13] is considered in logical regression, in which binary classification can be described as classifying the two groups based on classification rule [14]. The author used the model of the binary classifier to predict the diagnosis of positive and negative output. Machine Learning (ML) is the process of estimating unknown dependencies or structures in a system using a limited number of observations and it is used in data mining applications to retrieve hidden information and used in decision-making [1]. In ML, for classification, and regression problem various classifiers can be used for learning decision trees, rules, Bayes networks, artificial neural networks and support vector machines, and different knowledge representation models can be used to support decision-making methods [11]. Nowadays, in

D. Mehta (✉)
Bharat Group of College, Punjab, India
e-mail: dr.d.k.mehta81@gmail.com

C. Verma
Eötvös Loránd University, Budapest, Hungary
e-mail: chaman@inf.elte.hu

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_33

addition to statistical methods [2, 15], the use of ML in trending in the educational domain in the prediction of various targets in educational datasets [3–6]. The Support Vector Machine (SVM) is a supervised learning model presented for binary classification in both linear and nonlinear versions [16] and it performs classification by constructing an *N*-dimensional hyperplane that optimally separates the data into two categories [17]. Random Forest (RF) is a combination of tree predictors such that each tree depends on the values of a random vector sampled independently and with the same distribution for all trees in the forest [18]. Artificial Neural Network (ANN) is simple mathematical model defining a function f: X→Y or distribution over X or both X and Y, but sometimes models are also intimately associated with a learning algorithm or learning rule [8]. The binary Logistic Regression (LR) is much suitable for a binary classification problem [9]. Recently, students' demographic features such as residence state and gender were predicted using ML classifiers [10, 19, 20]. Dimension Reduction (DR) is a way to a reduced number of features in the given dataset. Feature extraction is the type of DR can be implemented using Principal Component Analysis (PCA) which transforms or projects a space composing of many dimensions into a space of fewer dimensions [13].

2 Research Design and Methodology

2.1 Dataset

To predict the readmit status of patient, the breast cancer dataset features are computed from the digitized image of fine-needle aspirates donated Olvi Mangasarian, Computer Science Department, University of Wisconsin, WI [1]. The donated medical dataset consists of 569 instances of 32 attributes of potential risk factors with diagnosis label that indicate the patient is malignant suffering or benign. The distribution percentage of records based on readmission for Malignant is 63% and Benign is 37%.

2.2 Dataset Preprocessing

After the data cleaning process, only 31 variables were retained for analysis and prediction of malignant diagnosis is shown in Table 1.

As more features exist there is no use of those features like id and one-nil feature in the dataset and there can be the possibility of redundant feature which correlates among them. This phenomenon leads to a multicollinearity problem; therefore, PCA is recommended in this case. After reducing the dimension prediction is made with support vector machine and neural network model.

Table 1 Independent factors for predicting readmit rate

Radius_mean	Symmetry_mean	smoothness_se
Texture_mean	Fractaldimensionmean	concavity_se
Perimeter_mean	Diagnosis	radius_worst
Area_mean	radius_se	perimeter_worst
Smoothness_mean	texture_se	smoothness_worst
Compactness_mean	perimeter_se	concavity_worst
Concavity_mean	area_se	symmetry_worst
Concavepoints_mean	compactness_se	texture_worst
concavepoints_worst	fractal_dimension_se	compactness_worst
symmetry_se	concave points_se	area_worst
fractaldimension_worst		

2.3 Support Vector Machine and Artificial Neural Network

A brief description of the SVM and ANN classification algorithm is considered in this article.

2.3.1 Support Vector Machine

The Support Vector Machine (SVM) algorithm used for regression and classification problems and it can be considered as a family of linear classification. The SVM has a property of minimizing the empirical classification to enhance the prediction accuracy. For this reason, it is also called maximum margin classifier. It maps the given vector as input to space of higher dimensional where construction is made according to a maximum hyperplane which has the property of separable. In this, two parallel hyperplanes are designed which separates the data on each side. The SVM provides better generalization error if there would be large distance between these two parallel hyper-planes.

2.3.2 Artificial Neural Networks

It can be used to model the complex relationship between input and output or can be used to find the pattern in data. It is one of the best machine learning classifiers that can be considered as powerful as it has established the best place in speech processing and machine learning prediction in many more areas. In this article, we have used Multilayer Perceptron (MLP) that is fed forward to the artificial neural network model. It maps multiple inputs to outputs.

2.4 Feature Extraction Using Principal Component Analysis

Whenever multivariate correlated dataset is under analysis, we feel the requirement of some techniques to reduce the dimension of available features in the dataset. PCA is replying to those questions or needs, which does some combination of available variables into a limited correlated group of uncorrelated variables.

From Table 2, there are k-random variables (X) of n instances. PCA explores most variation available in the data, i.e., using the k-random variables $X_1, X_2, X_3, \ldots, X_k$ to produce uncorrelated component $Z_1, Z_2, \ldots, Z_k$.

Z_1 represents a linear combination of ordered k variables with maximum variance and Z_2 combination of second largest variance and so on. This can be understood with Eqs. (1)–(4) with as a principal component coefficient.

$$Z_1 = a_{11}X_1 + a_{12}X_2 + a_{13}X_3 + \ldots A_{1k}X_k \tag{1}$$

$$Z_2 = a_{21}X_1 + a_{22}X_2 + a_{23}X_3 + \ldots A_{2k}X_k \tag{2}$$

$$\vdots \tag{3}$$

$$Z_k = a_{k1}X_1 + a_{k2}X_2 + a_{k3}X_3 + \ldots A_{kk}X_k \tag{4}$$

Our targeted data after PCA is represented in Table 3, which tells 90% concave point mean is loaded in component 1 (Z_0) and 1% in component 2 (Z_2). If we remove Z_3, Z_4, and Z_5 then we lose a maximum 10% of the information, this will address the multicollinearity problem.

PCA Algorithm:
Input: Multi-dimension matrix X (m, n): m—number of samples, n = number of variables
Outputs: PCA Components

Step 1. Compute the covariance of the given matrix between different variables.
Step 2. Calculate eigenvalues λi ($e_1, e_2, \ldots, e_n$) value of the covariance matrix).
[Eigen value represents the variance of principal component Zi]
Step 3. Calculate the eigenvector associated with each eigenvalue
[Coefficient of principal components $a_1, a_2, a_3 \ldots a_n$ is also called loadings].

Table 2 An arrangement of data for PCA

Instances	X_1	X_2	…	X_k
1	X_{11}	X_{12}	…	X_{ik}
2	X_{21}	X_{22}	…	X_{2k}
:	:	:	…	:
N	X_{n1}	X_{2n}	…	X_{n2}

Table 3 Uncorrelated component of correlated features

S. no.	Variables	Z_1	Z_2	Z_3	Z_4	Z_5
		%				
1	concave.points_mean	90	1	0	1	0
2	concavity_mean	89	2	0	0	1
3	concave.points_worst	84	0	8	0	0
4	compactness_mean	76	13	2	0	0
5	perimeter_worst	74	23	1	0	0
6	concavity_worst	70	5	8	1	6
7	perimeter_mean	69	26	0	0	0
8	radius_worst	69	28	1	0	0
9	area_mean	65	30	0	1	0
10	radius_mean	64	31	0	0	0
:	:	:	:	:	:	:
30	smoothness_se	0	24	27	0	9

After applying PCA to the data matrix, then carefully select the number of components which will remain in the analysis. The main goal behind this PCA process is to summarize the data and reduce the dimension.

Some tips can be followed for selection of important components

- Select those components whose eigenvalues are greater than 1 when using a correlation matrix to extract the components
- Scree plot can also be used for selection of components by finding a break between the large and small eigenvalues with eigenvalues λi (Variance in y-axis) in a scree plot as reflected in Fig. 1.

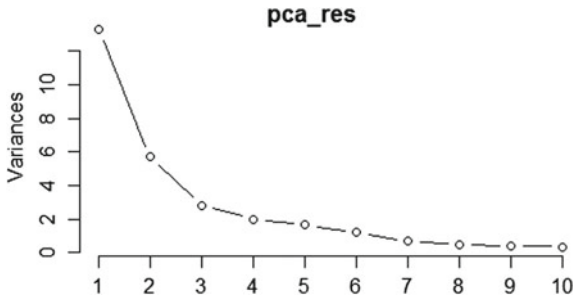


Fig. 1 Scree plot

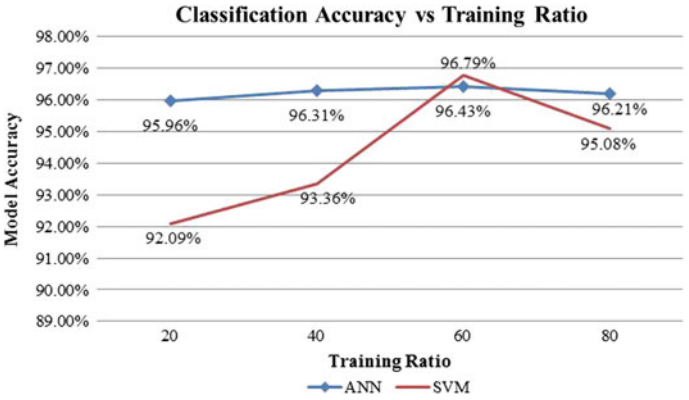


Fig. 2 Classification accuracy using training ratio

3 Experiments and Result Discussions

3.1 Experiment-I

Usually, a machine learning model is developed with partitioning the whole dataset into training and testing datasets. Partitioning involves selection of optimized splitting ratio for training and testing with the aim of developing a highly accurate model. It is basically a holdout method of cross-validation.

In this article, we use various training–testing ratios to verify the accuracy like 20:80, 40:60, 60:40, and 80:20. Afterward, we applied both PCA-ANN and PCA-SVM classifiers at various training ratios with 90% variance of PCA (only 10 important features) to forecast better accuracy. As it is reflected in Fig. 2 the better accuracy of PCA-ANN (96.43%) is achieved at 60–40 and got reduced (96.21%) after the next level on 80–20. At the initial stage, the minimum split 20–80 ratio, the accuracy can be observed (95.96%).

In the case of PCA-SVM classifier at the initial stage of split 20–80, the accuracy is 92.1 and maximum accuracy at 60–40 split is 96.8 and next level little bit difference is achieved in the decimal notion which seems negligible. Hence, better split ratio for both classifiers is 60–40.

3.2 Experiment-II

Another popular validation approach is cross-validation, in which the dataset is sampled into k subsets. Among those subsets, $(k - 1)$ are test sets and k-rest sets are train sets. During the classification process, this approach evaluates the classifier in a more meaningful manner. In our article, we use the k-fold method by varying values

by 2, 4, 5, 6 for the dataset and applied classifier with 90% variance of PCA (only 10 important features) accordingly. By every method, accuracy result is noted down to achieve maximum among of them. The k-fold algorithm works in the following manner.

Algorithm 1. K-Fold Cross-Validation Method

```
Step 1. Let K=2 (Split the entire data into the training
set into 2 equal subsets)
Step 2. For I=1 to K
Step3.    Let Fi as testing sets and retain all remaining
          k-1 fold as a training sets in CV
Step 4.  Apply ANN and SVM to train sets and compute the
          accuracy of the classifier.
Step 5. Select among classifier of higher accuracy.
```

With cross-validation, Fig. 3 clearly implies classification accuracy at various k-fold numbers. ANN Classification accuracy is increasing up to fivefold and afterward, it gets down by percentage. So from chart, maximum accuracy achieved is (98.74%) but in SVM, accuracy is not affected which remains same as 96.71%.

3.3 Experiment-III

As we know, PCA explains the total variation in the data with a few linear combinations of original variables which can be positively or negatively correlated. As Fig. 4 exhibits the effect of variation threshold on accuracy of model, 90% threshold

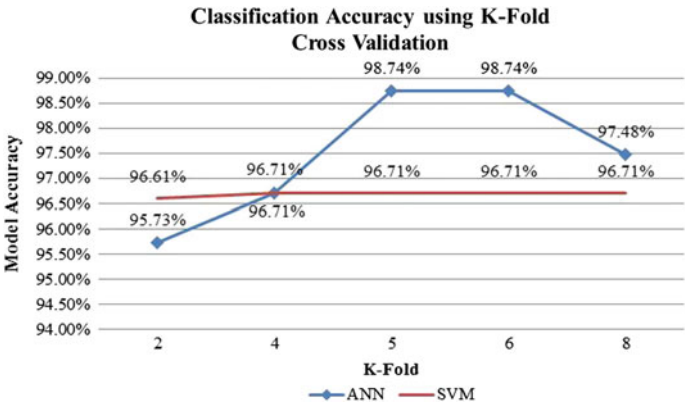


Fig. 3 Classification accuracy using K-fold

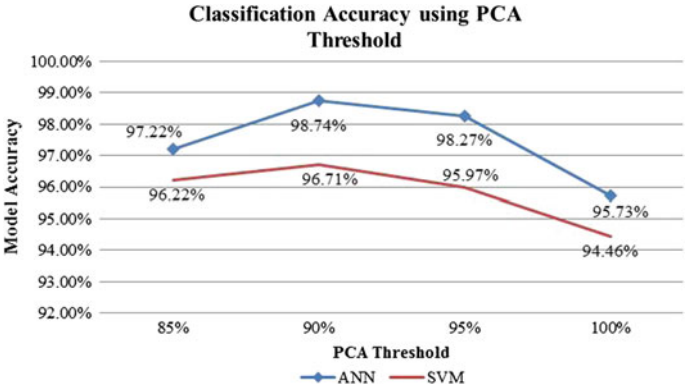


Fig. 4 Classification accuracy with respect to PCA threshold ratio

covers all important features which express all important components. In this experiment, we use fivefold cross-validation with PCA threshold ratios. Higher accuracy (98.74%) in case of ANN and also (96.71%) in case of SVM is achieved with 90% principal components. If we use all features (100%) which can be redundant and become a victim of the multicollinearity problem and effectively the accuracy gets reduced to 95.73% (ANN) and 94.64 (SVM).

4 Performance Evaluation

The result of an experimental study of classifier models are evaluated using the following four major performance matrices: (a) Accuracy: It is the percentage of correct number of predicted diagnosis from the overall diagnosis in the dataset. (b) Error: the total number of misclassification from prediction. (c) Sensitivity: It is the ability of a test to correctly identify those with the disease (true positive rate). (d) Specificity: It is the ability of the test to correctly identify those without the disease (true negative rate).

5 Results Evaluation and Discussion

Based on the classification, the prediction counts of ANN and SVM with and without principal component analysis, the result finding can be seen in tabular forms. Table 4 represents the confusion matrix in model prediction with actual results are compared and with the help of the confusion matrix table, accuracy is computed from the data. It clearly states that the accuracy of SVM and ANN without using feature reduction techniques is lower than with PCA-SVM and PCA-ANN.

Table 4 Confusion matrix

Models Prediction		SVM Prediction		ANN Prediction		PCA-SVM Prediction		PCA-ANN Prediction	
Gender		B	M	B	M	B	M	B	M
Actual	B	103	4	106	1	104	3	106	1
	M	2	61	3	60	2	61	1	62

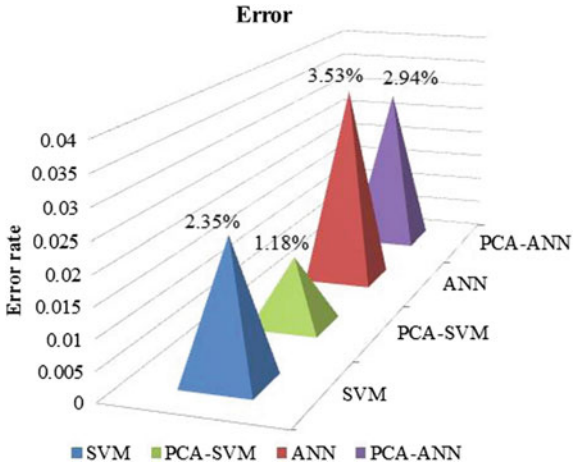


Fig. 5 Error rate of classification models

Table 5 Performance metrics

Classifiers	Accuracy (%)	Sensitivity	Specificity	Error
ANN	97.6	0.98	0.97	2.4
SVM	96.5	0.94	0.98	3.5
PCA-ANN	98.8	0.98	0.99	1.2
PCA-SVM	97.1	0.95	0.98	2.9

Figure 5 also clearly implies that the error rate of PCA-ANN is very low (1.18%) among all calculated classification models and sensitivity, i.e., true positive rate, is also higher. If we are interested in calculating false positive rate, i.e., specificity, then PCA-SVM can be considered.

Table 5 represents the performance metrics of prediction models. The PCA-ANN classifier outperformed others in prediction accuracy and model is highly sensitive at 0.98. It can be seen that PCA enhanced the prediction accuracy by 1.2% of ANN classifier. Also, SVM model sensitivity is also increased by 0.01 using PCA. The highest prediction error 3.5% is found by SVM without PCA.

6 Conclusion

The present experimental study is conducted to predict the diagnostic outcome. Breast cancer sample dataset is used to train the model for predicting the outcome of whether the patient is malignant or benign. The feature extraction method named PCA selected 10 important high variance features to enhance the accuracy of both classifiers. The

study favors the ANN model in each case and proclaimed the best model with high accuracy and low error rate. The accuracy of ANN is maximum and constant using K-fold method with k = 5 and k − 6. Also, using training ratio we found the highest accuracy of ANN at 60:40. It is also found that PCA enhanced the accuracy of ANN and SVM by 1.2% and 0.6%, respectively. Further, the future work includes more wrapper methods or filter methods to be applied to enhance prediction accuracy on the same dataset. Also, the authors recommend 90% significant features to be used in real-time website of Hospital to predict cancer-prone patients for readmission.

Acknowledgements The corresponding author thank the UCI website to provide significant datasets to pursue this research. The authors' institutions, Bharat Group of College and Eötvös Loránd, did not require ethical committee approval to be granted for this study.

References

1. UCI. https://archive.ics.uci.edu/ml/datasets/breast+cancer+wisconsin+%28Diagnostic%29 (2019)
2. D. Kabakchieva, Student performance prediction by using data mining classification algorithms. Int. J. Comput. Sci. Manag. Res. **1**(4), 686–690 (2012)
3. C. Maria Teresa, Noel R. Maria, Prediction of university student's academic achievement by linear and logistic models. Spanish J. Psychol. **2**(1), 275–288 (2015)
4. D. Kolo, A decision tree approach for predicting students' academic performance. Int. J. Educ. Manag. Eng. **5**, 12–19 (2015)
5. C. Verma, Educational data mining to examine mindset of educators towards ICT knowledge. Int. J. Data Min. Emerg. Technol. **7**, 53–60 (2017)
6. B. Deshmukh, A. Patil, B. Pawar, Comparison of classification algorithms using weka on various datasets. Int. J. Comput. Sci. Inf. Technol. **4**(2), 85–90 (2011)
7. R.L. Cheu, D. Srinivasan, E. Tian, Support vector machine models for freeway incident detection. In: *Proceedings of the Intelligent Transportation Systems*, vol. 1 (IEEE, 2003), pp. 238–243
8. C. Verma, S. Ahmad, V. Stoffová, Z. Illés, Forecasting residence state of indian student based on responses towards information and communication technology awareness: a primarily outcomes using machine learning. In: *International Conference on Innovations in Engineering, Technology and Sciences* (IEEE, India, 2018), In Press
9. C. Verma, V. Stoffová, Z. Illés, S. Dahiya, Binary logistic regression classifying the gender of student towards Computer Learning in European schools. In: *The 11th Conference of Ph. D Students in Computer Science* (Szeged University, Hungary, 2018), p. 45
10. C. Verma, V. Stoffová, Z. Illés, An ensemble approach to identifying the student gender towards information and communication technology awareness in European schools using machine learning, Int. J. Eng. Technol. **7**, 3392–3396 (2018)
11. C. Verma, S. Dahiya, Gender difference towards information and communication technology awareness in indian universities. SpringerPlus **5**, 1–7 (2016)
12. C. Verma, Z. Illés, V. Stoffová, Attitude prediction towards ICT and mobile technology for the real-time: an experimental study using machine learning. In: *The 15th International Scientific Conference eLearning and Software for Education* (University Politehnica of Bucharest, Romania, 2019), In Press
13. C. Verma, Z. Illés, V. Stoffová, Real-time prediction of development and availability of ICT and mobile technology in Indian and Hungarian University. In: *2nd International Conference on Recent Innovations in Computing* (J & K University, India, 2019), In Press

14. C. Verma, S. Ahmad, V. Stoffová, Z. Illés, M. Singh, National identity predictive models for the real time prediction of european schools students: preliminary results. In: *International Conference on Automation, Computational and Technology Management* (IEEE, London, 2019), In Press
15. C. Verma, S. Dahiya, D. Mehta, An analytical approach to investigate state diversity towards ICT: a study of six universities of Punjab and Haryana: Indian. J. Sci. Technol. **9**, 1–5 (2016)
16. S.R. Kalmegh, Comparative analysis of weka data mining algorithm random forest, random tree and lad tree for classification of indigenous news data, Int. J. Emerg. Technol. Adv. Eng. **5**(1), 507–517 (2015)
17. M. Minsky, S. Papert, *Perceptrons: An Introduction to Computational Geometry*, (MIT Press, 2017)
18. C. Verma, S. Ahmad, V. Stoffová, Z. Illés, S. Dahiya, Gender prediction of the european school's teachers using machine learning: preliminary results. In: *International Advance Computing Conference* (IEEE, India, 2018), pp. 213–220
19. C. Verma, V. Stoffová, Z. Illés, Age group predictive models for the real-time prediction of the university students using machine learning: preliminary results. In: *International Conference on Electrical, Computer and Communication* (IEEE, India, 2019), In Press
20. C. Verma, V. Stoffová, Z. Illés, Rate-Monotonic vs early deadline first scheduling: a review. In: *International Conference on Education Technology and Computer Science in Building Better Future* (University of Technology and Humanities, Poland, 2018), pp. 188–193

Recognition of Facial Expression Based on the Position of Hands Surrounding the Face Through Median Filter



Samta Jain Goyal, Arvind Kumar Upadhyay and Rakesh Singh Jadon

1 Introduction

There is a famous researcher Ekman, who proposed a FACS (Facial Action Coded System) which is used to describe and recognize human expressions through the basic coded emotions code such as joy, anger, surprise, disgust, fear, and sadness. There are many researchers who proposed many algorithms and systems for perceived human emotions against many default challenges such as lightning conditions, background, position of body language, and so on. The MIT media lab provides a particular data set for the facial expressions which contain natural and spontaneous images. The currently developed HCI systems should be able to understand the intention of human activities during interaction through the machine. FER system based on hand gesture first detects and then tracks the facial emotions through CV and Machine Learning techniques. These developed systems are used in Medical Science, Education, Virtual Learning, Surveillance systems, daily interaction which machine, interactive gaming system and so on. There we develop new approach to recognize facial emotions through the position of hands and its gesture which is very useful application of HCI Systems [1–5]. These HCI systems use body gestures and facial expressions for interaction through computer systems so that they can react. The objective of this proposed work is to find better results with better accuracy also with less time during recognition process. There we have taken 8 subjects with 469 images of different emotions and body gestures in terms of hand. First, we detect a required object and pass them into a median filter from where root mean square features are extracted.

S. J. Goyal (✉) · A. K. Upadhyay
Department of Computer Science & Engineering, Amity University,
Gwalior, Gwalior 474005, MP, India
e-mail: samtajaingoyal@gmail.com

R. S. Jadon
Department of Computer Applications, MITS, Gwalior 474001, MP, India

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_34

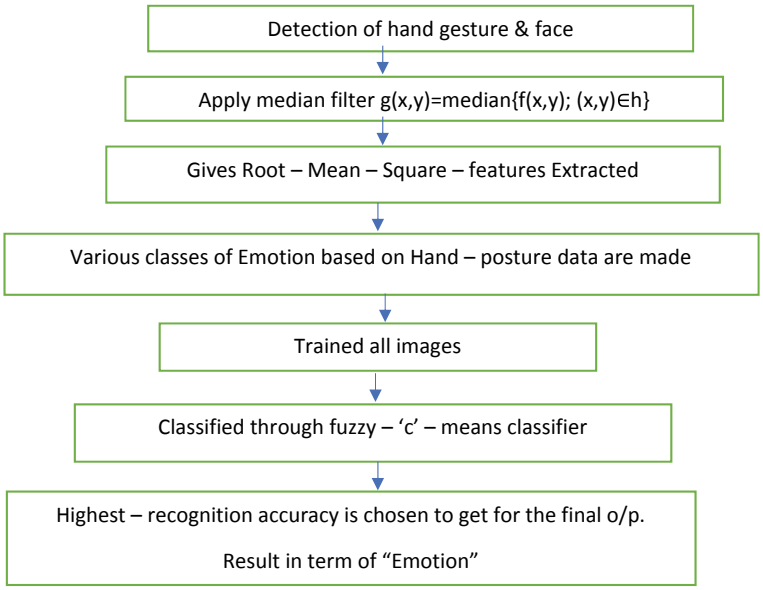


Fig. 1 Framework of the study

Average accuracy is more than 90% which proves the better refinement from the existing solutions [6–10].

There are many real time applications where human and machine interaction place grade important such as in media in medical science, mechanical engineering, neuroscience, computer science, and education. But more importantly a place better where the people are handicapped and afraid to say anything only they can show his or her intention through facial expressions and body language (Fig. 1). The field of HCI systems places an interface where features are face and body gesture acts as information's for the developed systems and then this system reacts accordingly. To design such systems is quite challenging in terms of accuracy, efficiency, and effectiveness. Although there is a great challenge to design in the form of logic, nonunderstanding capabilities in terms of functionalities to use. In the present era, many remarkable interfaces in terms of recognition systems have been developed to recognize human facial expressions in terms of emotions. There we use eye gaze tracking method to get exact positions of eyes on the detected image of human face. To design such systems, a camera is used to capture images on which we apply image processing methods for analysis of the input image thoroughly. During this process many circumstances such as computational cost, development cost, image size, complexity of design, quality of the developed systems occur. During designing such systems, analysis of an input image in deep, effectiveness of that system, accuracy is the most important factor. High speed, best recognition rates in accuracy, and low-cost computation are actually a vital factor. All the specified features must be targeted to achieve best performance of the newly designed systems [11–22].

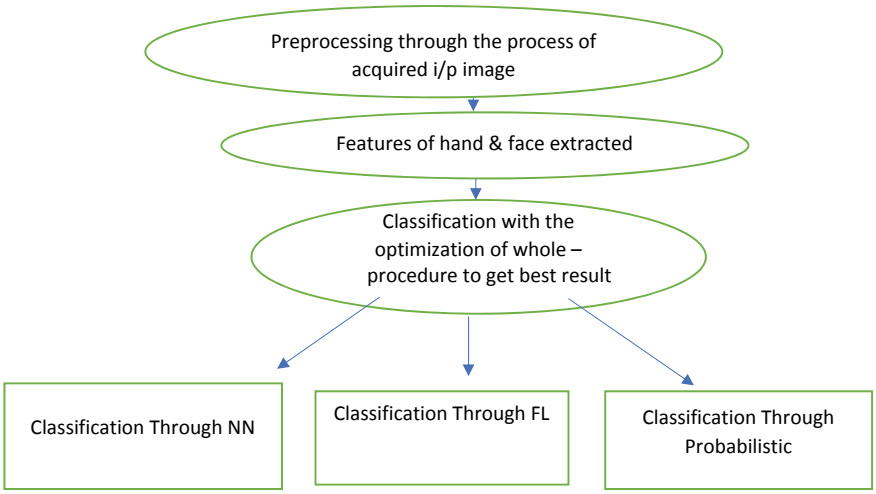


Fig. 2 Ways of the classification techniques

The role of classifier in the classification stage place a major role in the recognition system. The optimize classifier must be accurate and speedy. Optimized classifier depends on many factors such as pattern of the training, type of objects, application type, etc.

There are two main issues that occurred with the existing developed system. One is fixed number of facial and hand gesture which does not gives flexibility during designing and second recognition rate decreases if we apply any other type of body gestures. In this work, trying to design such systems which would be more flexible and accurate than the existing [23–29]. So that newly developed system becomes more reliable, flexible, accurate to recognize human emotions based on hand gesture, and its positions surrounding the face (Fig. 2).

2 Proposed Work

The procedure of our proposed work is shown in Fig. 3. Eight subjects have participated as the main dataset for this work. Figure 4 shows all possible emotions through gestures are performed by each subject. There are some major expressions and hand gestures such as happy, sad, anger, surprise and pointing someone, thinking, victory, and so on, respectively.

Preprocessing is a significant task for getting better input for further processing, preprocessed image after many subphases are saved for the further process of the recognition system. This procedure repeated for all input images to make them better so that output result also becomes better and of course more accurate, each processed

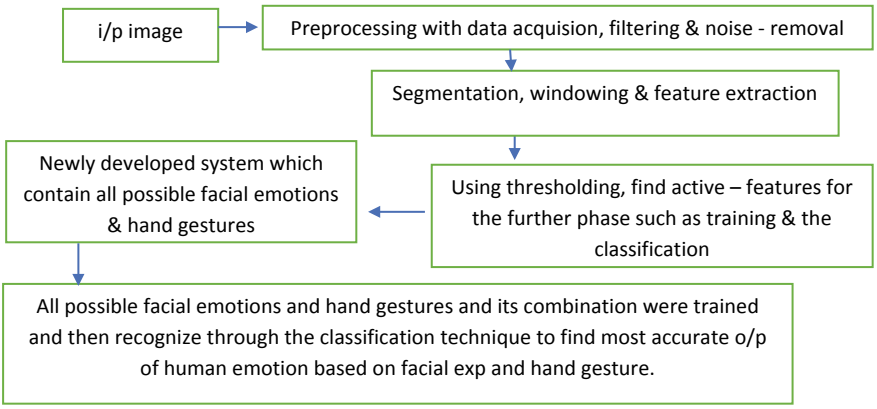


Fig. 3 Procedure of our proposed work

image has a different emotion, gestures to represent his or her view or intensions. But it is obvious that every subject has different ways to presentation of a particular emotion. All images that belong to the same category of emotion class are placed for the classification stage. Even all the input images are passed through a median filter which includes some important information.

3 Segmentation and Feature Extraction and Selection

All processed images are segmented to get the feature for the identification of the human emotion which offers the extreme probability approximation. After feature extraction, these features are sieved to collect the prompter feature set through thresholding. The main objective of this work is to design HCI based systems to control and react accordingly to human emotions. As mentioned earlier, hand gestures are considered as input addon parameters to get their intentions. The number of input information in terms of features is varied, based on the subject's nature, mood, and expressions. The proposed design is for the various real time applications. This system is used for all types of combinations of facial expressions and hand gestures and then assign them to different categories of emotion classes. In out = r system, images are trained and then classified with the highest recognition rate of accuracy for each emotional combination of facial expressions and hand gesture with different number of facial emotions are made using the following equation

$$\text{fehg} = \text{n}!\,\text{k}!(\text{n} - \text{k})! \tag{1}$$

where fe is a facial expression, hg is hand gesture, n = number of facial emotions and k is no. of hand gestures.

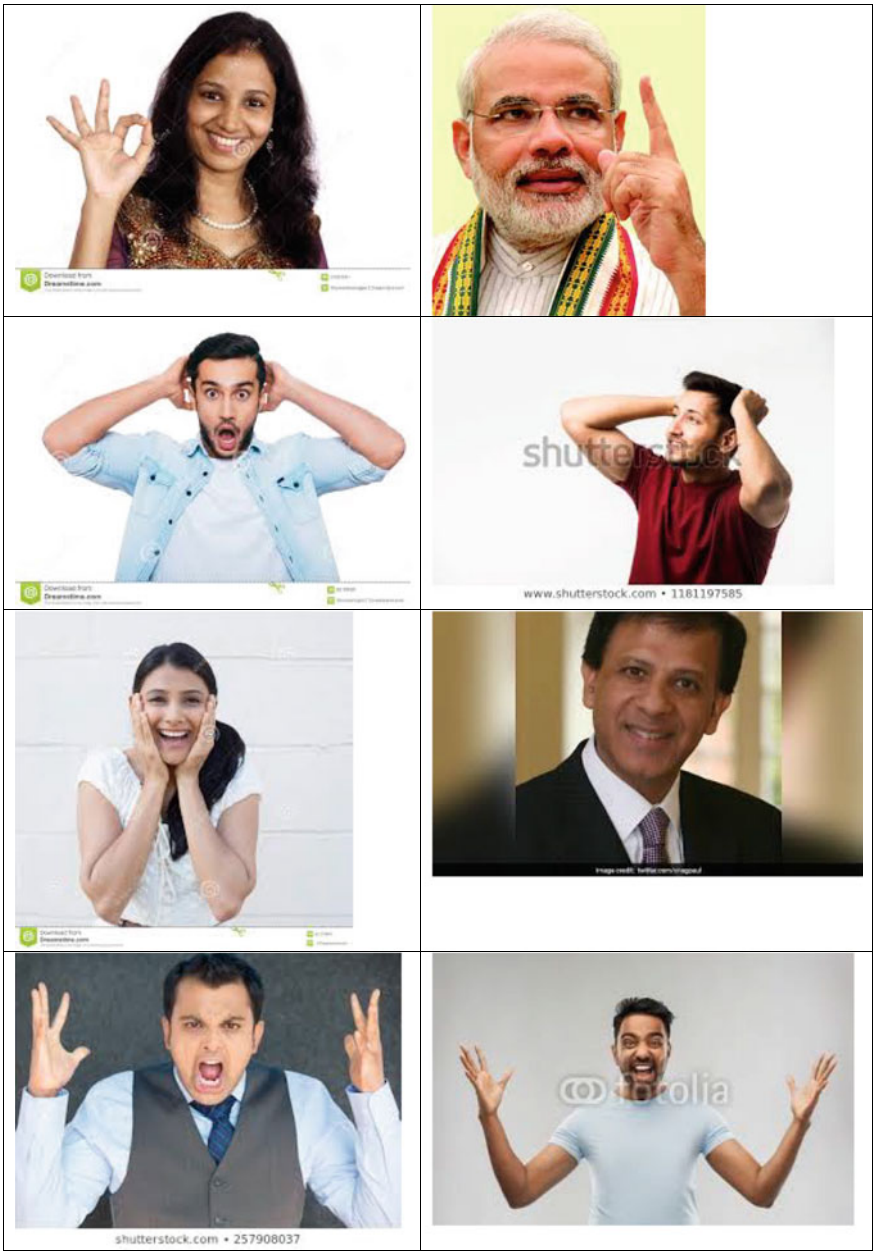


Fig. 4 Different facial expressions with hand gesture [44]

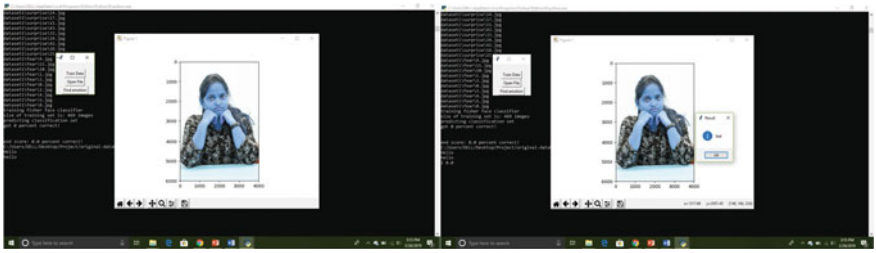
Based on Eq. (1), number of human emotions are calculated and puts them into a particular category of emotion. Lots of emotion with hand gestures such as anger with pointing someone is used to blame someone. Same anger with holding own head shows the ere great and so on [30–40].

4 Classification

In this phase of our recognition system, all combinations of images are trained and classifies = d for different human emotions such as anger, happy surprise, sad, and disgust. Each emotion class contains facial expressions and hand gesture. The selected features are then used to classify and recognize human emotion based on classifier. This work uses fuzzy-c-man's clustering (FCM) because of its simplicity, efficiency also it permits input image to belong two or more clusters. This process makes system more flexible to use. During training procedures, class labels use to direct images conveniently. This is basically a theme of supervised learning so overall procedure on facial expressions and hand gesture with the idea of supervised FCM where number of clusters is already given before train the input images in the form of data. All the obtained feature set with some additional parameters are used to compute the location of each cluster and also membership value for each data towards the cluster. We get optimized results when this procedure repeats until we get last and optimized output. This technique is very effective in the recognition processes used to get better output in terms of human emotion.

5 Result

For designing such recognition system, a flexible, trained set of input are required. So that we can apply them in various real time applications. The best image of facial expression and hand gesture are assigned as an input. All 469 images of 8 subjects are trained and classified them through improved FCM to get highest recognition accuracy in the output. There we have taken two criteria to find best input image. First its classification and recognition performance and the other is the way of the distribution and discrimination. Facial emotion can be recognized more accurately in terms of intensions through the hand gesture.



6 Distribution

The main objective of this paper is to design a human emotion recognition system based on hand gestures in HCI based application systems. It is accomplished through the help of human's basic emotions and hand gestures. The result of experiments shows that the used approach gives better results in terms of emotions. There are many hand gestures, positions of hand especially surrounding the face, facial expressions offer the difference between the existing developed system and newly designed system. The comparison of the results shows that the current system gives more accurate and better results than the existing ones. Also, this current system is easy in the implementation, simple to use, high speed, more compatible, more flexible, low-cost, and user-friendly [41–43].

7 Conclusion

The human emotion recognition system based on hand gesture is designed and applied to the HCI based applications. Where we have taken 8 subjects. A total 469 images have been taken which consist of facial expression and hand gestures. This system identifies all input images with the highest recognition rate. Best feature extraction, feature classification approaches are used to get better output in terms of standard emotions, better accuracy, better recognition rate, simple to use, and high speed during the processing of an input image for the recognition purpose.

References

1. A. Mehrabian, J.A. Russell, *An Approach to Environmental Psychology* (MIT Press, Cambridge, MA, USA, 1974)
2. P. Ekman, W.V. Friesen, *Facial Action Coding System: A Technique for the Measurement of Facial Movement* (Consulting Psychologists Press, Palo Alto, CA, USA, 1978), pp. 271–302

3. P. Lucey, J.F. Cohn, T. Kanade, J. Saragih, Z. Ambadar, I. Matthews, The extended Cohn-Kanade dataset (CK+): a complete dataset for action unit and emotion-specified expression, in *Proceedings of the IEEE Computer Society Conference on Computer Vision Pattern Recognition Workshops* (2010), pp. 94–101
4. M. Turk, A. Pentland, Eigenfaces for recognition. J. Cognit. Neurosci. **3**(1), 71–86 (1991)
5. P.N. Belhumeur, J.P. Hespanha, D.J. Kriegman, Eigenfaces vs. fisherfaces: recognition using class specific linear projection. IEEE Trans. Pattern Anal. Mach. Intell. **19**(7), 711–720 (1997)
6. K. Etemad, R. Chellappa, Discriminant analysis for recognition of human face images, J. Opt. Soc. Amer. A, Opt. Image Sci. **14**(8), 1724–1733 (1997)
7. L. Wiskott, J.-M. Fellous, N. Kuiger, C. von der Malsburg, Face recognition by elastic bunch graph matching. IEEE Trans. Pattern Anal. Mach. Intell. **19**(7), 775–779 (1997)
8. M.S. Bartlett, J.R. Movellan, T.J. Sejnowski, Face recognition by independent component analysis. IEEE Trans. Neural Netw. **13**(6), 1450–1464 (2002)
9. J. Yang, D. Zhang, A.F. Frangi, J. Yang, Two-dimensional PCA: a new approach to appearance-based face representation and recognition. IEEE Trans. Pattern Anal. Mach. Intell. **26**(1), 131–137 (2004)
10. D.-J. Kim, K.-W. Chung, K.-S. Hong, Person authentication using face, teeth and voice modalities for mobile device security. IEEE Trans. Consum. Electron. **56**(4), 2678–2685 (2010)
11. M.S. Bartlett, G. Littlewort, I. Fasel, J.R. Movellan, Real time face detection and facial expression recognition: development and applications to human computer interaction, in *Proceedings of the Computer Vision and Pattern Recognition Workshop* (2003), p. 53
12. M. Yeasin, B. Bullot, R. Sharma, From facial expression to level of interest: a spatio-temporal approach, in *Proceedings of the IEEE Computer Society Conference on Computer Vision Pattern Recognition* (2004), pp. 922–927
13. C. Shan, S. Gong, P.W. McOwan, Robust facial expression recognition using local binary patterns, in *Proceedings of the IEEE International Conference on Image Processing* (2005), pp. 370–373
14. P. Ekman, Universals and cultural differences in facial expressions of emotion, in *Proceedings of the Nebraska Symposium on Motivation* (1987), p. 712
15. S.S. Tomkins, *Affect, Imagery, Consciousness: Cognition Duplication and Transformation of Information* (Springer, New York, NY, USA, 1963)
16. J.A. Russell, A circumplex model of affect. J. Pers. Social Psychol. **39**(6), 1161–1178 (1980)
17. L.F. Barrett, J.A. Russell, Independence and bipolarity in the structure of current affect. J. Pers. Soc. Psychol. **74**(4), 967–984 (1998)
18. T. Ahonen, A. Hadid, M. Pietikäinen, Face description with local binary patterns: application to face recognition. IEEE Trans. Pattern Anal. Mach. Intell. **28**(12), 2037–2041 (2006)
19. I. Borg, P.J. Groenen, *Modern Multidimensional Scaling: Theory and Applications* (Springer, New York, NY, USA, 2005)
20. G.A. Seber, *Multivariate Observations*, vol. 252, no. 1 (Wiley, Hoboken, NJ, USA, 2009)
21. T.F. Cox, M.A. Cox, *Multidimensional Scaling* (Chapman & Hall, London, U.K., 2001)
22. J.Y. Davis, B. Kulis, P. Jain, S. Sra, I.S. Dhillon, Information theoretic metric learning, in *Proceedings of the International Conference on Machine Learning* (2007), pp. 209–216
23. R.A. Fisher, The use of multiple measurements in taxonomic problems. Ann. Hum. Genet. **7**(2), 179–188 (1936)
24. Y. Rubner, C. Tomasi, L.J. Guibas, The earth mover's distance as a metric for image retrieval. Int. J. Comput. Vis. **40**(2), 99–121 (2000)
25. S.D. Cohen, L.J. Guibas, The earth mover's distance: Lower bounds and invariance under translation. Technical Report TR-97-1597 (Department of Computer Science, DTIC, Fort Belvoir, VA, USA, 1997)
26. M. Werman, S. Peleg, A. Rosenfeld, A distance metric for multidimensional histograms, Comput. Vis. Graph. Image Process. **32**(3), 328–336 (1985)
27. G. Levi, T. Hassner, Emotion recognition in the wild via convolutional neural networks and mapped binary patterns, in *Proceedings of the ACM International Conference on Multimodal Interaction* (2015), pp. 503–510

28. H. Zhang, S. Luo, O. Yoshie, Facial expression recognition based upon human cognitive regions. IEEJ Trans. Electron. Inf. Syst. **134**(8), 1148–1156 (2014)
29. M.H. Bindu, P. Gupta, U.S. Tiwary, Cognitive model—based emotion recognition from facial expressions for live human computer interaction, in *Proceedings of the Computer Intelligence Image Signal Processing* (2007), pp. 351–356
30. R.W. Picard, E. Vyzas, J. Healey, Toward machine emotional intelligence: analysis of affective physiological state. IEEE Trans. Pattern Anal. Mach. Intell. **23**(10), 1175–1191 (2003)
31. C. Tiwari, M. Hanmandlu, S. Vasikarla, Suspicious face detection based on eye and other facial features movement monitoring, in *Proceedings of the IEEE Application Image Pattern Recognition Workshop* (2015), pp. 1–8
32. D. Patel, X. Hong, G. Zhao, Selective deep features for microexpression recognition, in *Proceedings of the International Conference on Pattern Recognition* (2017), pp. 2258–2263
33. D. Huang, M. Ardabilian, Y. Wang, L. Chen, Asymmetric 3D/2D face recognition based on LBP facial representation and canonical correlation analysis, in *Proceedings of IEEE International Conference on Image Processing* (2009), pp. 3325–3328
34. B. Zhang, G. Liu, G. Xie, Facial expression recognition using LBP and LPQ based on Gabor wavelet transform, in *Proceedings of IEEE International Conference on Computer Communication* (2016), pp. 365–369
35. Z. Shokoohi, R. Bahmanjeh, K. Faez, Expression recognition using directional gradient local pattern and gradient-based ternary texture patterns, in *Proceedings of the International Conference on Pattern Recognition and Image Analysis* (2015), pp. 1–11
36. D. Shi, X. Chen, J. Wei, R. Yang, User emotion recognition based on multi-class sensors of smartphone, in *Proceedings of IEEE International Conference on Smart City* (2015), pp. 478–485
37. J.B. Tenenbaum, V. de Silva, J.C. Langford, A global geometric framework for nonlinear dimensionality reduction. Science **290**(5500), 2319–2323 (2000)
38. W. Ou, X. You, D. Tao, P. Zhang, Y. Tang, Z. Zhu, Robust face recognition via occlusion dictionary learning. Pattern Recognit. **47**(4), 1559–1572 (2014)
39. X. You, Q. Li, D. Tao, W. Ou, M. Gong, Local metric learning for exemplar-based object detection. IEEE Trans. Circuits Syst. Video Technol. **24**(8), 1265–1276 (2014)
40. Z. Zhu et al., An adaptive hybrid pattern for noise-robust texture analysis. Pattern Recognit. **48**(8), 2592–2608 (2015)
41. M. Pantic, L. Rothkrantz, Automatic analysis of facial expressions: the state of the art. IEEE Trans. Pattern Anal. Mach. Intell. **22**(12), 1424–1445 (2000)
42. T. Ahonen, A. Hadid, Face recognition with local binary patterns, in *Proceedings of the European Conference on Computer Vision* (2004), pp. 469–481
43. B. Fasel, J. Luettin, Automatic facial expression analysis: a survey. Pattern Recognit. **36**(1), 259–275 (2003)
44. https://www.ft.com/__origami/service/image/v2/images/raw/http%3A%2F%2Fcom.ft.imagepublish.upp-prod-us.s3.amazonaws.com%2F6638c874-5a9d-11e9-840c-530737425559?fit=scale-down&source=next&width=700

Secure Sharing of Location Data Using Elliptic Curve Cryptography

Nikhil B. Khandare and Narendra S. Chaudhari

1 Introduction

Elliptic key cryptography is a very important contribution to the field of public key cryptography, most of the cryptographic systems in today's era are based on ECC. Solving the discrete logarithm problem based on ECC is hard. Difficulty of solving this problem makes the cryptosystems more secure. The level of security provided by RSA algorithm with key size of 1024 bit is same as the level of security provided by 160 bit ECC. Elliptic curve over the finite field is represented by equation

$$y^2 mod\ p = \left(x^3 + ax + b\right) mod\ p \tag{1}$$

Recommendations for key management were given by NIST in [1]. Diffie–Hellman Key Exchange Algorithm was originally proposed in [2] for secure exchange of keys between two parties. Generator is known to both the parties, X_A and X_B are the private keys of A and B, respectively, each user gets the shared secret key without actually sharing the key over insecure channel (Fig. 1).

ECC can be combined with Diffie–Hellman algorithm to make it nearly impossible to break. Security of elliptic curve cryptography relies on difficulty of discrete logarithm problem. Other hard problem in ECC which do not have solution in the polynomial time are

N. B. Khandare (✉)
Department of Computer Science and Engineering, Visvesvaraya National Institute of Technology, Nagpur, India
e-mail: nikhil.khandare@students.vnit.ac.in

N. S. Chaudhari
Department of Computer Science and Engineering, Indian Institute of Technology, Indore, India
e-mail: nsc0183@yahoo.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_35

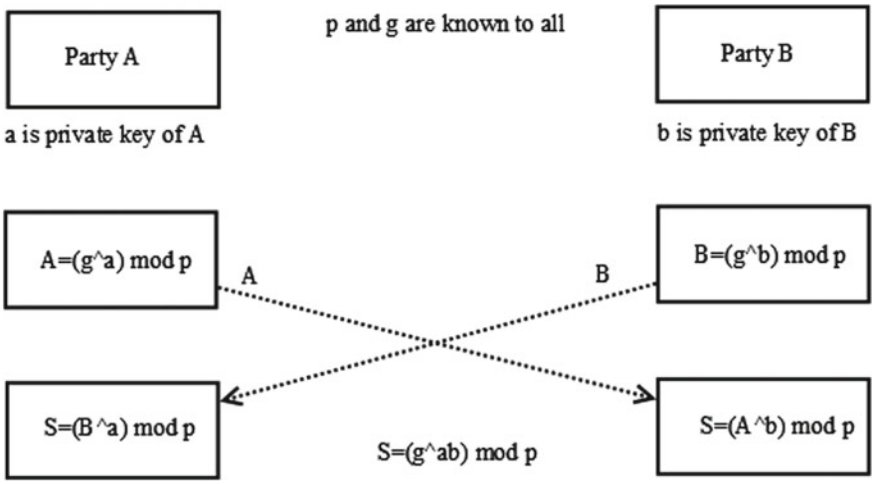


Fig. 1 Diffie–Hellman key exchange algorithm

Elliptic Curve Discrete Logarithm Problem (ECDLP): Given two elements A and B, then it nearly impossible to solve the equation for the value of k, Given that $A = kB$, where B is group generator.
Elliptic Curve Factorization Problem (ECFP): Given that $Z = aX + bY$, then it is impossible to find aX and bY such that $Z = aX + bY$.
Computational Diffie–Hellman Problem (CDHP): Given X, aX, bX, then it is impossible to find abX.
Decisional Diffie–Hellman Problem (DDHP): Given (X, aX, bX, cX), then it is impossible to check whether $c = ab$ or not. not $cP = abP$, i.e. to decide $c = ab$.

2 Related Works

Taking into consideration the security requirements of Internet of Things and continuous support that it requires, comparisons of solutions for Internet of Things that are applicable to security and privacy in connected mobile Internet of Things network were done in [3]. The attack linkage disposal decision-making method for edge computing network systems is done in [4] which is based on attribute attack graphs. To solve the tradeoff between privacy and power consumption level, a customized model for power consumption specifically for location-based service was proposed in [5]. An attack was developed to prove mathematically that k-anonymity does not satisfy the privacy level needed by location-based services [6]. A new encryption notion, called Order-Retrievable Encryption (ORE) for social networking applications was proposed in [7]. Various other works were done for protecting privacy and enhancing security of location-based services [8–10]. To address the privacy issue in

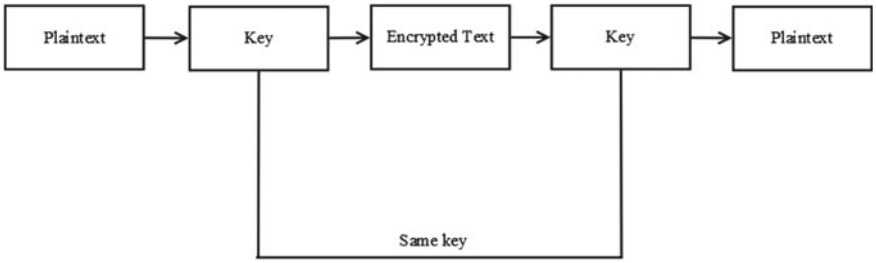


Fig. 2 Symmetric key cryptography

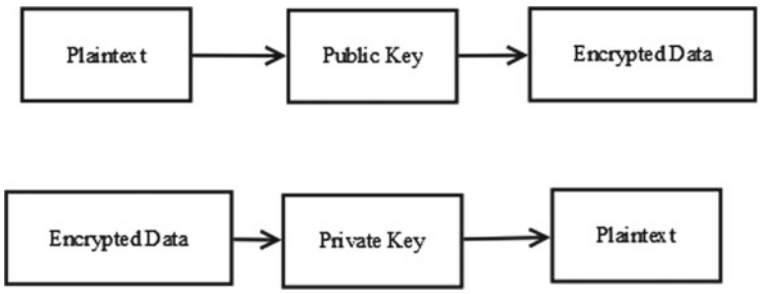


Fig. 3 Asymmetric key cryptography

location-based services Dummy Location Selection (DLS) algorithm was proposed in [11] to achieve k-anonymity for users in LBS. Privacy-preserving solutions in which location-based queries can be answered by data owners without sharing their data with other businesses and without accessing sensitive information such as the customer list was proposed in [12]. A new Elliptic Curve Cryptography (ECC)-based mobile-banking application tool called m-BAT that runs in client–server environment was proposed in [13]. A framework for preserving the privacy of nodes and event in wireless sensor network was given in [14]. This paper proposed the combination of elliptic curve cryptography for secure sharing of geographic location over the insecure channel.

3 Research Methodology

Symmetric Key Cryptography is shown in Fig. 2, in this cryptographic technique, same key is used for encryption as well as decryption of the message, this mechanism is fast but less secure.

Asymmetric Key Cryptography is shown in Fig. 3, here different keys are used for encryption and decryption, here encryption is done using public key and decryption is done using private key.

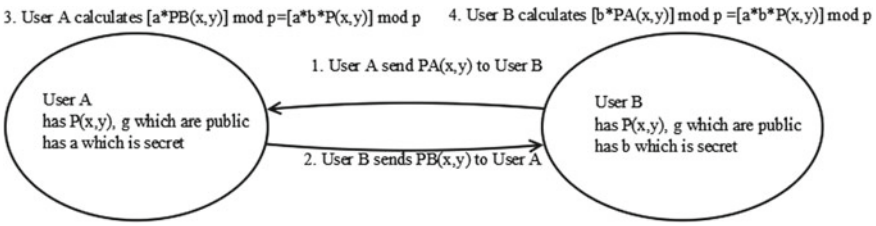


Fig. 4 Elliptic curve Diffie–Hellman key exchange algorithm

In this paper, asymmetric key cryptography is used wherein elliptic curve cryptography is combined with Diffie–Hellman algorithm.

3.1 Elliptic Curve Diffie–Hellman Key Exchange Algorithm

Figure 4 shows the elliptic curve Diffie–Hellman key exchange, wherein both parties agree on the elliptic curve over finite field given in equation A. a and b are the secret keys of User A and User B, respectively. Shared secret key is established as follows:

1. User A calculates P_A = [a*P(x,y)] mod p and sends it to B.
2. User B calculates P_B = [b*P(x,y)] mod p and sends it to A.
3. User A multiplies received P_B by a, i.e. it calculates [a*P_B(x,y)] mod p = [a*P*b] mod p.
4. User B multiplies received P_A by b, i.e. it calculates [b*P_A(x,y)] mod p = [b*P*a] mod p.

Thus, both users establish a shared secret key, i.e. [a*b*P(x,y)] mod p.

3.2 Proposed Work

The proposed work is focused on privacy in location data. Location data is provided to location service provider at the time of accessing the LSP. LSP can be search-based or share-based LSP. In order to deal with handling and querying huge data LSP go for Cloud Service Providers (CSPs), thus the location data needs to be secured from CSPs and LSPs. The proposed work is divided into four modules as follows:

1. Secure agreement of geographic location.
2. Sharing of geographic data between user and LSP.
3. Sharing of geographic data between LSP and Cloud.
4. Sharing of geographic data between users.

3.2.1 Secure Sharing or Secure Agreement of Geographic Location

Figure 5 illustrates the secure agreement of geographical location between two parties, here the situation wherein the two parties who want to agree on the location or who want to decide the place of meeting securely will use this protocol. In this protocol, the security of both the users from the intruder, LSP and CSP is provided. At the end of the protocol both users have same latitude, longitude coordinates without sharing the actual coordinates over the network. Here the new user will have to login using his IP address and password then he can login into the system.
Stepwise working of the proposed module of the protocol is given as follows:

1. User who wants to login should be registered first and then registered user can login using the IP address and Password.
2. User A and User B will agree on the elliptic curve (E/Fq) over the finite field Fq, where E is of the format given in (A).
3. User A and B choose point P(x,y) on elliptic curve in (A), this point is called group generator.
4. User A will multiply this group generator point using his own secret key a, i.e. A will get say $P_a(x,y) = [a*P(x,y)] \bmod p$, which is another point. User B will multiply this group generator point using his own secret key b, i.e. B will get say $P_b(x,y) = [b*P(x,y)] \bmod p$, which is another point.
5. A will send this point to B over insecure channel, B will send this point to A over insecure channel.
6. User A will multiply this received point P_b using a and get $[a*P_b(x,y)] \bmod p = [a*b*P(x,y)] \bmod p$ and User B will multiply this received point P_a using b and get $[b*P_a(x,y)] \bmod p = [a*b*P(x,y)] \bmod p$.

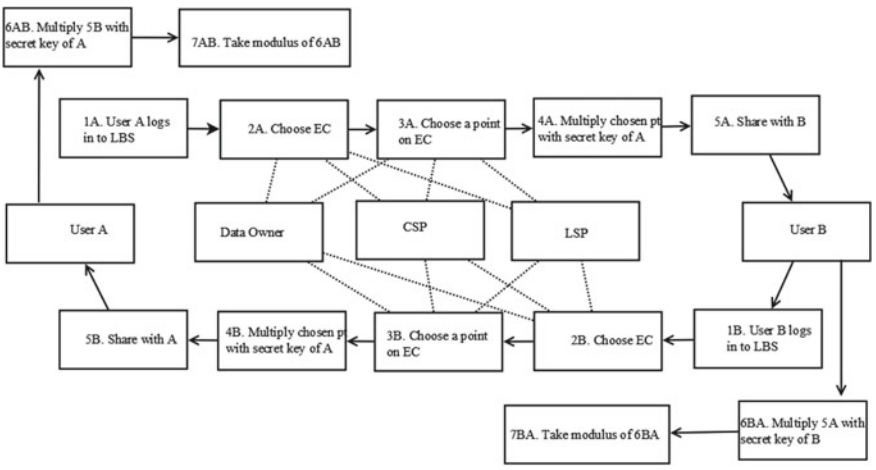


Fig. 5 Secure agreement of geographic location using ECC

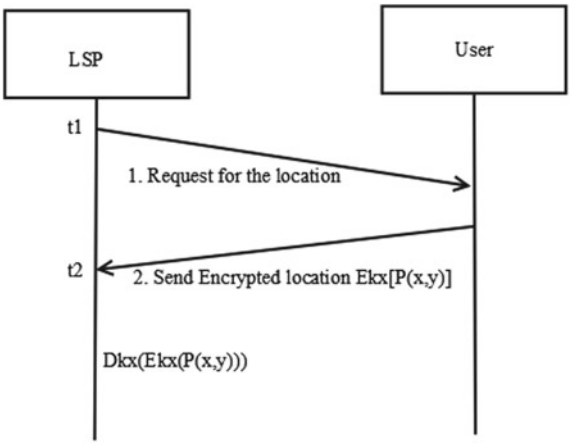


Fig. 6 Secure sharing of geographic location between user and LSP

7. This received point a*b*P(x,y) is operated by modulus operator p and thus agreed location is a*b*P(x,y) mod p.

Thus, two parties can securely agree on the location of meeting using above protocol.

3.2.2 Sharing of Geographic Data Between User and LSP

Whenever location-based service (search-based or share-based) is used, it is required to share location with location service provider. In such case attacks like man in middle and other attacks in insecure channel and protection the location data from LSP should be considered. Key agreement and authentication is done using ECDH-key exchange. The secure sharing of geographical data between user and LSP (Fig. 6) works as follows:

1. LSP requests User A for Location data, i.e. latitude and longitude data.
2. User then sends the location data, i.e. latitude and longitude data expressed as point P(x,y) on elliptic curve over finite field encrypted using x-coordinate of the key.
3. Authenticated LSP will only be able to decrypt (using x-coordinate of key) the encrypted location data E_{Kx} P(x,y). LSP will perform D_{Kx} (E_{Kx}P(x,y)).

3.2.3 Sharing of Geographic Data Between LSP and Cloud

In order to efficiently handle and querying the huge (geographic) data coming from user, location-based service provider goes for cloud service provider. This solves the issue of handling huge data but it unfolds the privacy issue that the data needs to be protected from CSP as well. Here, key agreement and authentication is done

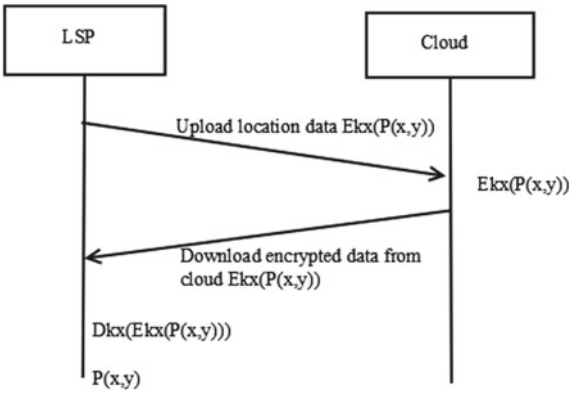


Fig. 7 Secure sharing of geographic location between LSP and cloud

between LSP and Cloud using ECDH-key exchange. Secure sharing of geographic data between LSP and cloud (Fig. 7) works as follows:

1. LSP will upload the location data on cloud encrypted using x-coordinate of the key and the encrypted data will only be stored on cloud, i.e. ($E_{Kx}P(x,y)$)
2. When LSP needs the geographical data it will download and decrypt the location data, i.e. D_{Kx} ($E_{Kx}P(x,y)$).

3.2.4 Sharing of Geographic Data Between Users

Location sharing amongst friends and family members is very common, people share their location on WhatsApp, Facebook, Google to name the few. However, the sharing of location between the two users is prone to many attacks and to ensure privacy and confidentiality of the data the protocol for secure exchange of location between users (Fig. 8) works as follows:

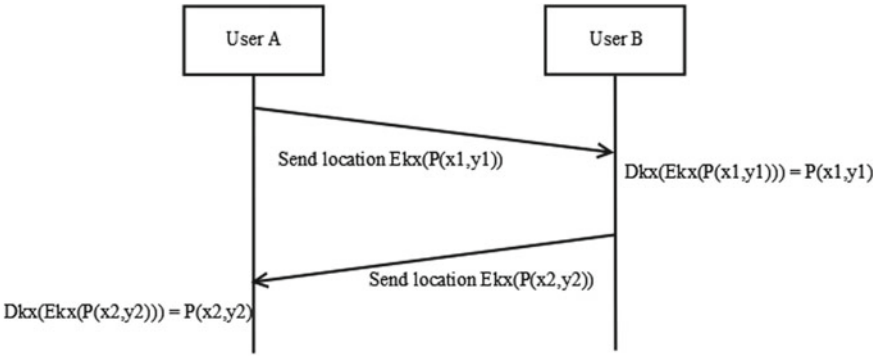


Fig. 8 Secure exchange of location between users using ECC

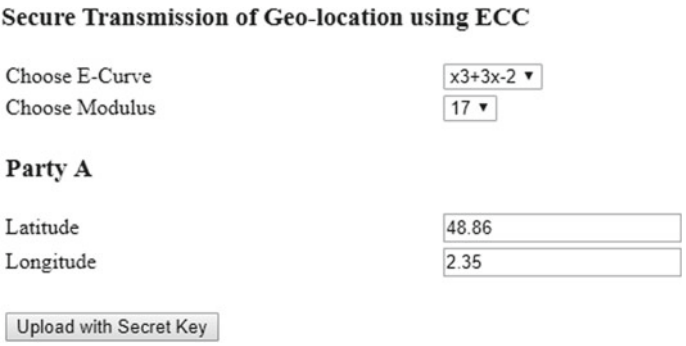


Fig. 9 Secure transmission of geolocation using ECC, party A

1. Let (X_1,Y_1) and (X_2,Y_2) be the location of two users, User A will send the location to User B by encrypting using x-coordinate of the key $(E_{Kx}P(x_1,y_1))$ and User B will decrypt the same using the mutually agreed key, i.e. D_{Kx} $(E_{Kx}P(x_1,y_1))$
2. User B will send the location to User A by encrypting using x-coordinate of the key $(E_{Kx}P(x_2,y_2))$ and User A will decrypt the same using the mutually agreed key, i.e. D_{Kx} $(E_{Kx}P(x_2,y_2))$.

This protocol is also used for continuous location sharing for the stipulated time.

4 Results and Discussion

4.1 Secure Agreement of Geographic Location Using ECC

This section discusses the results obtained from the implementation of secure agreement of geographic location using ECC which mainly consist of choosing an elliptic curve of the form $y^2 \bmod p = (x^3 + ax + b) \bmod p$, choosing modulus, entering latitude, longitude and upload by multiplying the data using secret key for party A and party B (Figs. 9 and 10).

4.2 Receiving the Intermediate Data

This section discusses the intermediate data generated when location P(x,y) is multiplied by the secret key of party A and party B. It mainly discusses the intermediate latitude, intermediate longitude (both are unreadable), uploading with secret key of A and B. After uploading for second time, directly shared secret key will be generated which is same for both users (Figs. 11 and 12).

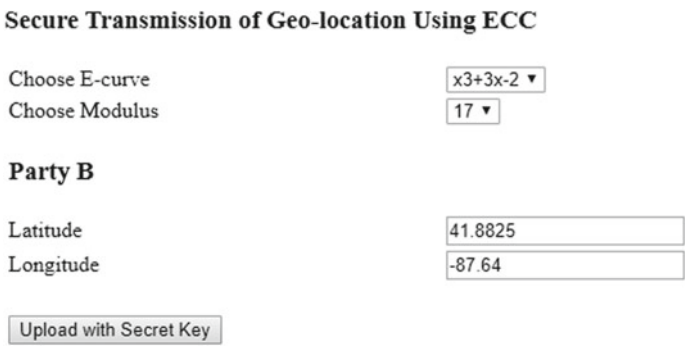


Fig. 10 Secure transmission of geolocation using ECC, party B



Fig. 11 Receiving intermediate location, party A



Fig. 12 Receiving intermediate data, party B

4.3 Computation Time and Space Requirement

Proposed system of sharing the location shows better results when compared with symmetric key cryptography algorithms in terms of speed of computation, space requirement and network bandwidth consumption since the size of key is 160 for elliptic key cryptography when compared with RSA which is 1024 bit. Figure 13 shows the computation time of RSA and ECC for the same key size, for 144 bit key size computation time of RSA is 13.5 s whereas for ECC it is 1.2 s.

Space requirement is lower as the size of key is 160 bit as compared to asymmetric RSA which is 1024 bit. If the size of key in RSA is reduced, the level of security is reduced. The proposed protocol is more secure and has less computation overhead.

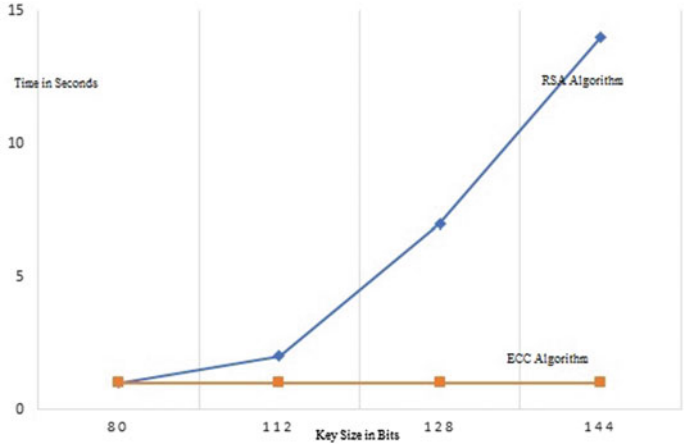


Fig. 13 Computation (encryption and decryption) time of RSA versus ECC

4.4 Security Analysis

As discussed earlier, the security of elliptic curve cryptography relies on difficulty of solving discrete logarithm problem. In the proposed secure location agreement protocol, even if someone gets the intermediate data, he/she may not be able to get the correct interpretation of location.

Consider an example to support our discussion, consider the point on elliptic curve as P(x,y) = (3,7) and modulus is 17, secret key of User A is 2 and secret key of User B is 3. Intermediate data is always sent over the insecure channel. In first case (User A) intermediate data is 2*(3,7) mod 17 = (6,14) mod 17 = (6,14). In second case (User B) intermediate data is 3*(3,7) mod 17 = (9,21) mod 17 = (9,4). Now even if intruder comes to know these intermediate values, i.e. (6,14) or (9,4) it is impossible for him/her to find out that (6,14) = 2*(3,7) or (9,4) = 3*(3,7). This is elliptic curve discrete logarithm problem, i.e. given two elements A and B, then it is nearly impossible to solve the equation for the value of k, given that A(x,y) = k*B(x,y), where B(x,y) is group generator. Also given that Z(x,y) = a*X(x,y) + b*Y(x,y), then it is impossible to find a*X(x,y) and b*Y(x,y) such that Z(x,y) = a*X(x,y) + b*Y(x,y).

Now (6,14) is received by B and (9,4) is received by A. A calculates shared secret key as 2*(9,4) mod 17 = (18,8) mod 17 = (1,8), B calculates shared secret key as 3*(6,14) mod 17 = (18,42) mod 17 = (1,8). Thus, both parties agree that the location of meeting is (1,8). Now given (3,7), (6,14), (9,4), i.e. given P, aP, bP intruder is unable to compute agreed location abP, i.e. (1,8). This is elliptic curve computation Diffie–Hellman problem, i.e. given X(x,y), a*X(x,y), b*X(x,y), then it is impossible to find a*b*X(x,y).

Given (3,7), (6,14), (9,4) and (1,8) still intruder is not able to take the decision that (1,8) is the agreed location of meeting. This is elliptic curve decision Diffie–Hellman

problem, i.e. given (X(x,y), a*X(x,y), b*X(x,y), c*X(x,y)), then it is impossible to check whether c = ab or not, or c*P(x,y) = a*b*P(x,y), i.e. to decide c = ab.

Few popular attacks are discussed and how the proposed system defends them

Man in the middle attack: The proposed system authenticates the user each time when he or she logs into the system, for each of the four modules in the proposed work authentication is done before initiating the module. MITM is the attack wherein user gets the key and decrypts the further messages. User A sends the message as [a*P(x,y)] mod p] but solving this equation for value of 'a' is discrete logarithm problem which is hard and impossible.

Non-repudiation: In case of any conflict between the users, any user cannot deny the ownership of message sent or received by him, the proposed system does the user registration and authentication (*username, IP address and password*) which would leave the footprint on the message and thus cannot deny the ownership of message.

Denial of Service attack: If the registered user enters a wrong password while logging into the system more than three times, then user is automatically blocked. Also the duration between sending of message (t_1) and receiving of message (t_2) is noted and if duration ($t_2 - t_1$) is greater than threshold (t) the request is dropped.

5 Conclusion

The secure protocol for agreement of geographic location was proposed wherein two parties are able to decide the location of meeting by using elliptic curve cryptography. Another issue of protection of location data from location service provider and cloud service provider was proposed, communication between user and LSP, communication between LSP and CSP and communication between users is secured using elliptic curve cryptography. Implementation of the secure location agreement protocol was discussed in result and discussion, cost (in terms of time) of breaking the cipher is far more as compared to RSA with same level of security. Decryption time is 13.5 s for 144-bit RSA, whereas it is 1.5 s for 144-bit ECC. Security analysis was done by using toy example to show that the system is secure by mathematical hard problems EC-DLP, EC-CDHP, EC-DDHP and robust to popular attacks.

Acknowledgements Authors are thankful to the anonymous reviewers for their comments on the previous version of the manuscript.

References

1. E. Barker, Q. Dang, NIST special publication 800-57 Part 1, Revision 4. Technical Report (NIST, 2016)
2. W. Diffie, M. Hellman, New directions in cryptography. IEEE Trans. Inf. Theory **22**(6), 644–654 (1976)
3. P. Gope, R. Amin, S.H. Islam, N. Kumar, V.K. Bhalla, Lightweight and privacy-preserving RFID authentication scheme for distributed IoT infrastructure with secure localization services for the smart city environment. Futur. Gener. Comput. Syst. **83**, 629–637 (2018)

4. S. Zhang, X. Li, Z. Tan, T. Peng, G. Wang, A caching and spatial K-anonymity driven privacy enhancement scheme in continuous location-based services. Futur. Gener. Comput. Syst. **94**, 40–50 (2019)
5. M.S. Alrahhal, M. Khemekhem, K. Jambi, Achieving load balancing between privacy protection level and power consumption in location-based services (2018)
6. J. Bou Abdo, T. Bourgeau, J. Demerjian, H. Chaouchi, Extended privacy in crowdsourced location-based services using mobile cloud computing. Mob. Inf. Syst. (2016)
7. R. Schlegel, C.Y. Chow, Q. Huang, D.S. Wong, Privacy-preserving location sharing services for social networks. IEEE Trans. Serv. Comput. **10**(5), 811–825 (2017)
8. Y. Sun, N. Wang, X.L. Shen, J.X. Zhang, Location information disclosure in location-based social network services: Privacy calculus, benefit structure, and gender differences. Comput. Hum. Behav. **52**, 278–292 (2015)
9. L. Kuang, Y. Wang, P. Ma, L. Yu, C. Li, L. Huang, M. Zhu, An improved privacy-preserving framework for location-based services based on double cloaking regions with supplementary information constraints. Secur. Commun. Netw. (2017)
10. P. Galdames, C. Gutierrez-Soto, A. Curiel, Batching location cloaking techniques for location privacy and safety protection. Mob. Inf. Syst. (2019)
11. C. Ma, Z. Yan, C.W. Chen, SSPA-LBS: Scalable and social-friendly privacy-aware location-based services. IEEE Trans. Multimed. (2019)
12. E. Yilmaz, H. Ferhatosmanoglu, E. Ayday, R.C. Aksoy, Privacy-preserving aggregate queries for optimal location selection. (2018)
13. S. Ray, G.P. Biswas, M. Dasgupta, Secure Multi-Purpose Mobile-Banking Using Elliptic Curve Cryptography. Wireless Pers. Commun. **90**(3), 1331–1354 (2016)
14. B. Chakraborty, S. Verma, K.P. Singh, Differentially Private Location Privacy Preservation in Wireless Sensor Networks. Wireless Pers. Commun. **104**(1), 387–406 (2019)

Cyberbullying Checker: Online Bully Content Detection Using Hybrid Supervised Learning



Akshi Kumar and Nitin Sachdeva

1 Introduction

Nowadays, Internet has drastically reformed the way people express their views-opinions-thoughts on social media. People rely more on the use of social forums like Twitter, Facebook, Formspring.me, MySpace, Ask.fm, etc. for sharing their views and opinions which result in producing unprecedented volume of user-generated online data that is available generally in the form of tweets, blog posts, reviews, question-answering forums, etc. [1–3]. Due to the heavy dependence of the mass on such multimedia content for appropriate opinions shows the increasing relevance of the utilization of Web 2.0 technologies and tools in our daily lives [4, 5]. Hence, we can say that social media [6] has global reach and has become widespread. Its pervasive reach has in return given some unpremeditated consequences as well where people are discovering illegal and unethical ways of using such communities. One of its most severe upshots is known as cyberbullying where individuals are searching new means to bully one another over the Internet.

Cyberbullying (CB) is often described as bullying that occurs utilizing the electronic technology like mobiles, computers, Internet, etc. where people are involved in sharing their interests and information through online social media platforms, messages or apps, etc. in order to harm or embarrass them [7, 8]. It has grown as a social menace that puts a negative effect on the minds of both the bully and victim. It is more persistent way of bullying a person before an entire online community especially when we talk in terms of social networking websites which can ultimately result in psychological and emotional breakdown for the CB victim developing the

A. Kumar · N. Sachdeva (✉)
Department of Computer Science & Engineering, Delhi Technological University, Delhi, India
e-mail: nits.usit@gmail.com

A. Kumar
e-mail: akshi.kumar@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_36

feeling of low self-confidence, depression, stress, anger, sadness, health degradation, loneliness, suicides, etc. [8–10]. It is typically a multistep phenomenon comprising many subtasks like Data collection, Data Preprocessing, Feature Extraction, Feature Selection, and Classification of messages. Including all social media, there is huge gamut of data available on Web and it is quite intractable to manually classify each abusive or offensive comment as either cyberbullying or non-bullying messages. Eventually a large number of features are also there, thus, *effective feature selection* is a computationally hard task [11–13] that affects the overall accuracy of the classification. This necessitates the constant requisite for exploring and analyzing new computational methodologies to find the most optimal set of features that enhance the mining performance of the classifier. This confluence of identifying or detecting CB across various social media like Twitter, Facebook, MySpace, Ask.fm, Formspring.me, etc. has completely revolutionized the way computing is done by using the power and collective intelligence of supervised learning techniques. CB is emerging as an upcoming area of research and is, however, facing issues pertaining to availability of well-known established benchmark datasets [7]. Although, various datasets are also accessible and appear to majorly constituent discussions among adults, therefore, we have developed our own-labeled corpus using the social media forums, namely, Formspring.me, MySpace, and Ask.fm for identifying bullying content in it.

Cyberbullying (CB) is primarily associated with the utilization of digital media in order to bully someone [9]. It has grown to a level where it is seriously affecting and damaging individual's lives where social media forums play a key role by providing a fecund means for bullies and the ones using such portals are more vulnerable to attacks. All this has gradually increased the fuzziness, dimensionality, and complexity related to the "user-generated unstructured online data" which further encourages the need to search for optimized and enhanced classification techniques that can cater to the identification or detection of CB on any social media. Motivated by this, we can infer that such online detrimental behavior has instigated a need for devising an automated mechanism using data-driven methodologies for critically analyzing, assessing, and detecting such unfavorable activities involving CB on social media. Studies are continuously being conducted to examine new supervised learning paradigms that handle fuzziness, uncertainty, partial truth, imprecision, approximation, and allow reproduction of human intelligence for tractable and personalized results. Supervised learning techniques fit well with this reasoning and can automatically classify messages under most optimal circumstances [2, 3]. CB uses various supervised learning-based techniques for identifying bullying content, whereas, feature level CB, particularly deals with the application of various swarm intelligence (SI-based) algorithms for identifying and selecting relevant features from the data sources, thus are majorly used for the feature selection subtask of CB detection mechanism. Motivated by the collective adaptive learning behavior of supervised learning techniques, the major contributions of our work are as follows:

- Automatic CB detection using hybrid supervised learning.

- Implement eight supervised techniques for cyberbullying content detection: Naïve Bayes (NB), Logistic Regression (LogR), Support Vector Machines (SVM), K-Nearest Neighbor (kNN), Projective Adaptive Resonance Theory (PART), JRIP is RIPPER Algorithm, Decision Tree (DT), and Artificial Neural Networks (ANN).
- The environment used for implementation: Waikato Environment for Knowledge Analysis (Weka) tool.
- Empirical evaluation to detect the best supervised textual cyberbullying content detection model using Accuracy (Ac), Recall (Re), F-Measure (F), and Precision (Pr) as efficacy criteria.

The rest of the paper is structured as follows. Section 2 focuses on the background work done in this area. Section 3 discusses the system architecture. Section 4 describes the implementation of the work including the analysis of results succeeded by conclusion and future work.

2 Related Work

This section briefs about the background work that has been already done in the pertinent literature on the selected problem statement. One of the author Ptaszynski et al. [14] in 2010 applied support vector machines for detection of CB of parent–teacher association of websites of Japanese schools. The author appreciating results in classifying CB entries in the specified dataset. Dinaker et al. [15] (2011) applied NB, DT, and SVM for detecting textual CB when analyzed for a corpus containing 4500 YouTube comments. Reynolds et al. [16] in 2011, implemented decision tree, instance-based learner for CB detection of Formspring.me. Among all, DT produced improved results. Dadvar et al. [17] (2012) focused on the implementation of SVM for training gender-specific language features in order to detect CB. The author did the implementation using MySpace corpus and obtained the highest precision for male-specific features. Kontostathis et al. [18] (2013) worked on the data obtained from Formspring.me website and applied a supervised machine learning approach named as essential dimensions of LSI. The author achieved average precision of around 91% at rank 100 for five of their queries. Potha and Maragoudakis [19] in 2014, applied singular value decomposition, support vector machines, and neural network for detection of CB. The results suggested that SVM yielded improved results. Huang et al. [20] (2014) investigated the impact of social network features for improving the accuracy of detection of CB. The author applied algorithms like bagging, DT, NB, zero regression, etc., on textual and social features using Twitter corpus from the CAW 2.0 dataset which contains 900,000 posts of 27,135 users. The results claimed that the social features enhance the result rate.

Hosseinmardi et al. [21] (2015) gathered dataset from Instagram consisting of images using human labelers. The author implemented SVM and NB for detecting incidents of CB automatically from the labeled dataset. NB showed improved recall rate of around 78%. Sarna and Bhatia [22] (2015) categorized the messages as either

indirect or direct CB messages and had proposed a solution in order to control CB through inspecting the user creditability. The author had implemented NB, KNN, DT, and SVM. The best results were obtained by KNN. Al-garadi et al. [23] (2016) implemented machine learning classifiers like NB, SVM, RF, and KNN and had proposed set of unique features derived from Twitter. Their proposed model had yielded improved accuracy with KNN technique for detecting CB. Ozel et al. [24] (2017) worked on the datasets derived from Instagram and Twitter messages in Turkish language. The author applied techniques like DT, SVM, NB, and KNN using feature selection methods like information gain and chi-square test. Among all, NB had produced enhanced results for CB detection. Zhang et al. [25] (2018) proposed the application of deep learning algorithms like convolutional and gated recurrent networks for detecting hate speeches on Twitter. The results showed that the proposed approach produced enhanced F1 score. Still, many other supervised learning techniques are yet to be applied to the field of CB in textual and non-textual forms and are part of open research areas as well.

3 System Architecture

This section illustrates the overall architecture and methodology of the process. Steps involved are data collection and data preprocessing, identification of bad words, addition of features to original dataset, and classification as shown in Fig. 1.

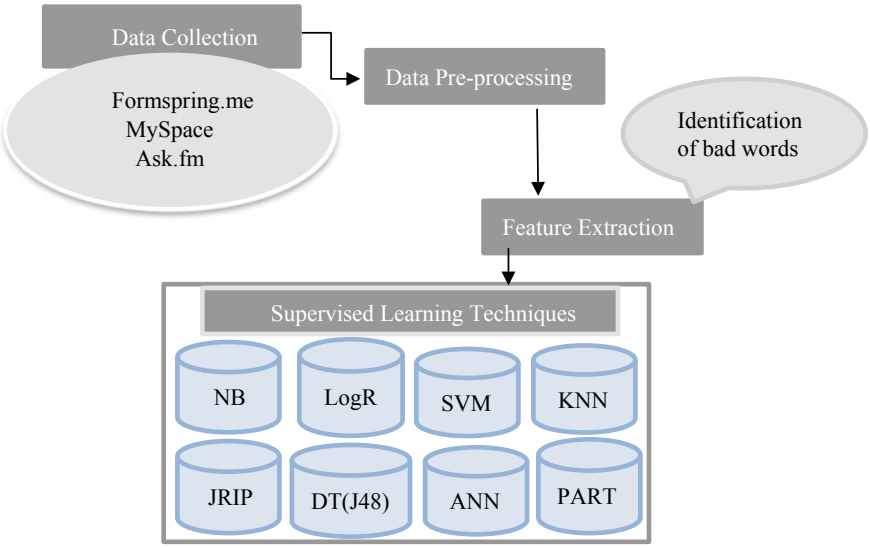


Fig. 1 System architecture

3.1 Data Collection and Preprocessing

The data was collected from the website Formspring.me (DS1), Myspace.com (DS2), and Ask.fm. (DS3). Formspring.me was peculiarly chosen as it is very famous website among all the teens and college-going students. It is heavily populated by the young minds and generates quality level answers to their queries or questions. Since its inception, it is gaining popularity based on social Q&A format (Question and Answering) allowing anyone to ask you the questions and providing you a platform for answering them thus simulating an interview-based layout. The most distinguishing feature of this website is that you invite others for questioning yourself over any topic that they want and the other person has two choices in demand for this situation. Either he can leave the question anonymously or leave his user information. It is principally this option of anonymity that makes this website highly prone to CB. It is expected to contain huge percentage of CB content that could be fruitful to be used for investigation in our study. For obtaining data from Formspring.me, we crawled a subset of this website and had mined information [26] from the sites of 19,000 randomly selected users. The opted range for selection of questions per user comprised of 1 post to 1000 posts.

MySpace is another social networking [27, 28] website that offers an interactive and user-submitted network of known friends [26]. A small subset of data from crawl of MySpace groups was also used. The dataset had 2800 posts from more than 10 separate chats.

Ask.fm is also one of the types of social networking websites that allows users to create their own profiles and later they may send each other question as well. It was earlier considered to be a form of anonymous social media as it allowed users to submit their questions anonymously. A large data was crawled from the website during summer of 2014 for 72 normal users and 38 common users with over 1 lakh posts per ID. We selected one public ID at random and used it for our work.

Tenfold cross-validation technique was used for choosing the randomly selected files for training and testing purposes. Data preprocessing involves cleaning of all the three datasets to deal with nuances. Labeling of data instances is done to facilitate supervised learning.

Cleaning: The first step in preprocessing of data is to clean the data. First some data points were removed which were not in English language. Thereafter, we removed highly repeated words like question and answer from all the posts. Some data points also consisted of HTML elements that were downloaded as text during crawling. Such data points were not relevant for our study and were hence removed.

Labeling: The Formspring.me dataset was prelabeled using Amazon's Mechanical Turk (AMT). However, the other two datasets required labeling. As CB is a type of subjective task, for this, three workers were employed that labeled each post as either "yes" for having CB content in it and "no" for not having any such content. The final labeling was selected with majority voting.



Fig. 2 List of bad words with their corresponding severity

3.2 Identification of "Bad" Words

During labeling, it was clear that the presence of certain bad words made it more likely for a post to be labeled as CB. Identification of insulted and swear words was done from the website www.noswearing.com and list of around 349 terms was made. Severity level was associated with every word such as 100 (for idiot)-200 (for trash)-300 (for asshole)-400 (for fuckass), and 500 (for buttfucker) as shown in Fig. 2.

3.3 Addition of Features to the Original Dataset

A new dataset was extracted out of the original dataset by the addition of new features. As the original dataset just had the answer and question data for every post, thus additional features were required to aid model development. Additional features were generated to measure the overall "wickedness" of a post. We call these features as SUM and TOTAL. SUM is computed by taking a weighted average of the "bad" words severity. TOTAL is calculated by taking the ratio of total severity of the sentence and the total number of words in the sentence.

Algorithm to Generate Extended Dataset: The pseudocode of the algorithm to generate NUM dataset from the original dataset by feature addition is given below:

Input:

- Dataset P which is a set of tuples containing the posts

- Dataset S which is a set of tuples containing swear words
- Map Severity containing the severity value for each swear word in S
- Function Sum, which returns the weighted average of bad words based on severity
- Function Total, which returns the weighted average of bad words based on total count.

Output:

- Dataset NUM containing Num, Sum, and Total values for each post with the predicting variable Bully

Algorithm:
1. Define a list posts_data that contain all the posts in Dataset P
2. Define a list swear_data that contain all the swear words in Dataset S
3. Declare five lists numi where i varies from 1 to 5 to keep count of swear words for each severity level
4. Declare lists sum_num and total_num to store values returned by Sum and Total function for num values
5. for each post in posts_data:
    for each word in swear_data:
        if word belongs to post:
            keep the counts of swear words for each severity level
6. Store the counts of swear words for each severity level in their respective lists numi, where i varies from 1 to 5
7. Store the sum_num and total_num for num values in num lists returned by Sum and Total functions respectively
8. Save the lists num1, num2, num3, num4, num5, sum_num and total_num in NUM data set with the predicting variable Bully
9. Return NUM dataset

3.4 Classification

Various SC techniques were applied on NUM to predict the presence of CB in a post and categorizing it into YES or NO. Classification algorithms were used to compare and contrast the results of both the datasets.

4 Result and Discussions

This section describes the experimentation results involved in this study. Weka tool (version-3.8.1) [12, 15, 16] was used for performing the empirical assessment of the aforesaid SC techniques when applied to the chosen datasets. The efficacy measures, namely, Ac, Pr, F, and Pr [29, 30] are used to evaluate the overall performance of CB classification tasks. All the values are expressed in percentages (%).

Results are presented in the following tables (expressed in percentages). Tables 1, 2, and 3 illustrate the application of the chosen techniques [31–35] on the three datasets, namely, DS1, DS2, and DS3.

From these tables, the following observations were made:

Table 1 Empirical comparison of various techniques using DS1

Techniques	Ac	Pr	Re	F
JRIP	95.45	60	75	66.67
LogR	90.4	73.2	59.4	65.5
PART	90	68.8	68.8	68.8
J48	89.5	65.7	71.9	68.7
SVM	89.2	70.8	53.1	60.7
NB	89	69.2	56.3	62.1
KNN	88.5	65.5	59.4	62.3
ANN	88	75	37.5	50

Table 2 Empirical comparison of various techniques using DS2

Techniques	Ac	Pr	Re	F
JRIP	90.7	72.2	74.3	73.2
LogR	90.5	71.1	77.1	74
PART	90	71.4	71.4	71.4
J48	89.9	70.6	68.6	69.6
SVM	89.6	71.9	65.7	68.7
KNN	89.5	73.3	62.9	67.7
NB	88	69	57.1	62.5
ANN	86.36	53.84	70	60.86

Table 3 Empirical comparison of various techniques using DS3

Techniques	Ac	Pr	Re	F
JRIP	95.5	91.7	57.9	71
LogR	94	81.8	47.4	60
PART	93.8	62.5	78.9	69.8
J48	93.7	66.8	63.4	65
SVM	93.6	66.7	63.1	64.7
NB	93.5	68.8	57.9	62.9
KNN	93.4	75	47.4	58.1
ANN	90.9	66.67	66.67	66.67

The results of our study suggest that the best accuracy is achieved by JRIP for all the three datasets, i.e., data fetched from Formspring.me, MySpace, and Ask.fm. JRIP outperformed all other supervised classification algorithms in terms of accuracy, followed by LogR, and PART techniques for all the three datasets. J48 and SVM also showed encouraging results. Among all, these four techniques showed almost quite close results with the accuracy ranging around 89% to 90% for DS1 and DS2, and 93% to 94% for DS3, respectively. NB had comparable accuracy quite akin to kNN for all the datasets. ANN demonstrated lower accuracy for all the three corpuses. From the results, it is deduced that more improved and optimized results were observed for DS1 and DS3 in contrast to DS2. The following Figs. 3 and 4 give the comparison of the eight supervised learning techniques for all the datasets based on performance accuracy (Ac) and F-score (F).

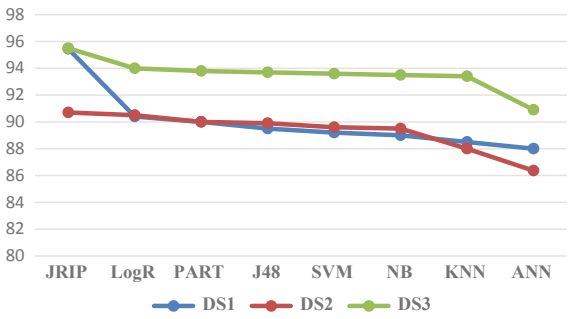


Fig. 3 Comparison of DS1, DS2, and DS3 based on Ac

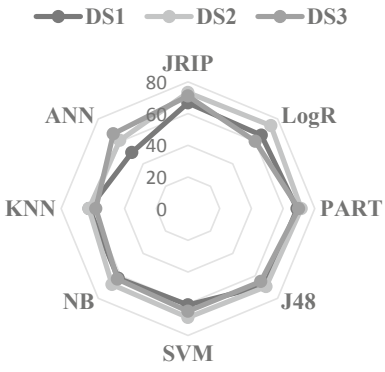


Fig. 4 Comparison of DS1, DS2, and DS3 based on F

5 Conclusion

This paper empirically contrasted the three social media portals, Formspring.me (DS1), Myspace.com (DS2), and Ask.fm. (DS3) for cyberbullying detection using eight supervised classification algorithms, namely, Naïve Bayes (NB), Logistic Regression (LogR), Support Vector Machines (SVM), K-Nearest Neighbor (kNN), Projective Adaptive Resonance Theory, JRIP is RIPPER Algorithm, Decision Tree (DT), and Artificial Neural Networks (ANN). The results were evaluated for the classifier performance, based on precision (Pr), recall (Re), accuracy (Ac), and F-score (F). The best accuracy and precision are achieved using JRIP for all the three datasets, followed by Logistic Regression, Projective Adaptive Resonance Theory, Decision Tree, Support Vector Machines, and Artificial Neural Networks which demonstrated the lowest accuracy.

This model's accuracy can be enhanced by including other more refined feature selection methods that could aid in better and improved modeling of the present system. It is a recent area of research and has not been explored substantially for all the domains including health, politics, entertainment, etc. There exists a vast scope of application of other learning techniques like fuzzy logic, swarm optimization, etc. that can be assessed for cyberbullying detection and prevention. Deep learning models including Convolutional Neural Network, etc., could also be used for detection of cyberbullying when applied to other datasets.

References

1. A. Kumar, P. Dogra, V. Dabas, Emotion analysis of Twitter using opinion mining. In: *2015 Eighth International Conference on Contemporary Computing (IC3)* (IEEE, 2015), pp. 285–290
2. A. Kumar, R. Khorwal, S. Chaudhary, A survey on sentiment analysis using swarm intelligence. Indian J. Sci. Technol. **9**(39), 1–7 (2016)
3. A. Kumar, R. Khorwal, Firefly algorithm for feature selection in sentiment analysis. In: *Computational Intelligence in Data Mining* (Springer, Singapore, 2017), pp. 693–703
4. A. Kumar, T.M. Sebastian, Sentiment analysis on Twitter. Int. J. Comput. Sci. Issues (IJCSI) **9**(4), 372–438 (2012)
5. A. Kumar, T.M. Sebastian, Sentiment analysis: a perspective on its past, present and future. Int. J. Intell. Syst. Appl. **4**(10), 1–14 (2012)
6. A. Kumar, A. Joshi, Ontology driven sentiment analysis on social web for government intelligence. In: *Proceedings of the Special Collection on eGovernment Innovations in India* (ACM, 2017), pp. 134–139
7. E. Raisi, B. Huang, Cyberbullying detection with weakly supervised machine learning. In: *Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining* (ACM, 2017), pp. 409–416
8. A. Kumar, N. Sachdeva, Cyberbullying detection on social multimedia using soft computing techniques: a meta-analysis. Multimed. Tools Appl., 1–38 (2019)
9. M. Foody, M. Samara, P. Carlbring, A review of cyberbullying and suggestions for online psychological therapy. Internet Interventions **2**(3), 235–242 (2015)

10. A. Kumar, S. Nayak, N. Chandra, Empirical analysis of supervised machine learning techniques for Cyberbullying detection. In: *International Conference on Innovative Computing and Communications* (Springer, Singapore, 2019), pp. 223–230
11. A. Kumar, A. Jaiswal, Systematic literature review of sentiment analysis on Twitter using soft computing techniques. Concurrency Comput. Practice Exp. 5107 (2019)
12. A. Kumar, A. Jaiswal, S. Garg, S. Verma, S. Kumar, Sentiment analysis using cuckoo search for optimized feature selection on Kaggle tweets. Int. J. Inf. Retr. Res. (IJIRR) **9**(1), 1–15 (2019)
13. A. Kumar, A. Jaiswal, Swarm intelligence based optimal feature selection for enhanced predictive sentiment accuracy on twitter. Multimed. Tools Appl. 1–25 (2019)
14. M. Ptaszynski, P. Dybala, T. Matsuba, F. Masui, R. Rzepka, K. Araki, Y. Momouchi, In the service of online order: tackling cyber-bullying with machine learning and affect analysis. Int. J. Comput. Linguist. Res. **1**(3), 135–154 (2010)
15. K. Dinakar, R. Reichart, H. Lieberman, Modeling the detection of textual cyberbullying. In: *Fifth International AAAI Conference on Weblogs and Social Media* (2011), pp. 11–17
16. K. Reynolds, A. Kontostathis, L. Edwards, Using machine learning to detect cyberbullying. In: *2011 10th International Conference on Machine Learning and Applications and Workshops* vol. 2 (IEEE, 2011), pp. 241–244
17. M. Dadvar, F.D. Jong, R. Ordelman, D. Trieschnigg, Improved cyberbullying detection using gender information. In: *Proceedings of the Twelfth Dutch-Belgian Information Retrieval Workshop (DIR 2012)* (University of Ghent, 2012)
18. A. Kontostathis, K. Reynolds, A. Garron, L. Edwards, Detecting cyberbullying: query terms and techniques. In: *Proceedings of the 5th Annual ACM Web Science Conference* (ACM, 2013), pp. 195–204
19. N. Potha, M. Maragoudakis, Cyberbullying detection using time series modeling. In: *2014 IEEE International Conference on Data Mining Workshop* (IEEE, 2014), pp. 373–382
20. Q. Huang, V.K. Singh,, P.K. Atrey, Cyber bullying detection using social and textual analysis. In: *Proceedings of the 3rd International Workshop on Socially-Aware Multimedia* (ACM, 2014), pp. 3–6
21. H. Hosseinmardi, S.A. Mattson, R.I. Rafiq, R. Han, Q. Lv, S. Mishra, Detection of cyberbullying incidents on the instagram social network. arXiv preprint arXiv:1503.03909 (2015)
22. G. Sarna, M.P.S. Bhatia, Content based approach to find the credibility of user in social networks: an application of cyberbullying. Int. J. Mach. Learn. Cybernet. **8**(2), 677–689 (2017)
23. M.A. Al-garadi, K.D. Varathan, S.D. Ravana, Cybercrime detection in online communications: the experimental case of cyberbullying detection in the Twitter network. Comput. Hum. Behav. **63**, 433–443 (2016)
24. S.A. Özel, E. Saraç, S. Akdemir, H. Aksu, Detection of cyberbullying on social media messages in Turkish. In: *2017 International Conference on Computer Science and Engineering (UBMK)* (IEEE, 2017), pp. 366–370
25. Z. Zhang, D. Robinson, J. Tepper, Detecting hate speech on Twitter using a convolution-GRU based deep neural network. In: *European Semantic Web Conference* (Springer, Cham, 2018), pp. 745–760
26. B.S. Nandhini, J.I. Sheeba, Online social network bullying detection using intelligence techniques. Proc. Comput. Sci. **45**, 485–492 (2015)
27. A. Kumar, N. Ahmad, ComEx miner: expert mining in virtual communities. Int. J. Adv. Comput. Sci. Appl. (IJACSA) **3**(6) (2012)
28. M.P.S. Bhatia, A. Kumar, Paradigm shifts: from pre-web information systems to recent web-based contextual information retrieval. Webology **7**(1) (2010)
29. A. Kumar, A. Jaiswal, Empirical study of Twitter and Tumblr for sentiment analysis using soft computing techniques. In: *Proceedings of the World Congress on Engineering and Computer Science*, vol. 1 (2017), pp. 1–5
30. N. Sachdeva, R. Dhir, A. Kumar, Empirical analysis of machine learning techniques for context aware recommender systems in the environment of IoT. In: *Proceedings of the International Conference on Advances in Information Communication Technology & Computing* (ACM, 2016), pp. 39

31. M.P.S. Bhatia, A. Kumar, Information retrieval and machine learning: supporting technologies for web mining research and practice. Webology **5**(2) (2008)
32. D.W. Aha, D. Kibler, M.K. Albert, Instance-based learning algorithms. Mach. Learn. **6**(1), 37–66 (1991)
33. R. Quinlan, *C4.5: Programs for Machine Learning* (Morgan Kauffman, San Mateo, CA, 1993)
34. G.H. John, P. Langley, Estimating continuous distributions in Bayesian classifiers. In: *Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence* (Morgan Kaufmann Publishers Inc., 1995), pp. 338–345
35. A. Kumar, N. Sachdeva, A. Garg, Analysis of GA optimized ANN for proactive context aware recommender system. In: *International Conference on Health Information Science* (Springer, Cham, 2017), pp. 92–102

Location-Wise News Headlines Classification and Sentiment Analysis: A Deep Learning Approach



Ashwin Kabra and Seema Shrawne

1 Introduction

News Data is one of the structured formatted data, which carries attributes like source, date, location, author, headline text, and detailed data, which is text formatted data. To extract features from the textual data, usually, the preference is for Natural Language Processing approach, which deals with all the language- and text-related aspects [1, 2].

Machine Learning is the new trending phenomenon observed recently, nearly, all applications use some type of intelligence and automation in their respective areas. Machine Learning is the technique by which the system or machines gets the ability to learn and perform operations by self-intelligence. The input to the machine learning algorithm so as to learn is a dataset or more specifically, features of the data. Deep learning approaches are recently being used for more accurate output on the textual data [3–5]. Deep Learning is a part of machine learning, where Artificial Neural Networks are being used with varying hidden layer units and number of hidden layers [6–8].

Deep Learning networks work best and gives more accurate output on the long sequence and large inputs such as sound, images, large text sequences, so most of the text classification systems use Deep Learning approaches. Recurrent Neural Networks (RNN) and LSTM are mostly used for the purpose of classification of text. News headlines are long sequence data, so normal machine learning algorithms such as TFIDF, SVM, and Multi-Layer Perceptron won't work that accurately, so we preferred the Deep Learning approach for this study [3].

A. Kabra · S. Shrawne (✉)
VJTI, Mumbai University, Mumbai, India
e-mail: scshrawne@ce.vjti.ac.in

A. Kabra
e-mail: kabraashwin123174@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_37

This study carries study of classification and sentiment analysis on the News articles with the enhancement of geospatial approach, where each news article associated with its location is analyzed.

This paper is divided into five sections. The first section gives an introduction to our research work, followed by the second section of literature survey, wherein previous research work performed in this field is mentioned. The third section focuses on the model analysis where various Machine Learning and Deep Learning techniques that are being used for this study is explained. The fourth section consists of experiment and analysis and the brief about the dataset used, where the actual experimental results of various models and techniques used is analyzed. The final and fifth section gives the conclusion and discussion over the research study.

2 Literature Survey

Various researchers such as Cecchini et al. [9], Shahi et al. [1], Mohiuddin et al. [10], Jirasirilerd et al. [2], Kaur et al. [11] have performed news classification using machine learning techniques like SVM, Naïve Bayes, and Neural Networks, to classify news into different categories. Deep Learning techniques like LSTM have been used by Li et al. [4], Zhang et al. [5], and have been used for the purpose of classifying the news articles. Rao et al. [12], Azzopardi et al. [13], Zhang et al. [14], Nagalavi et al. [15] have performed the classification of news articles depending on particular location using machine learning techniques like Random Forest, Naïve Bayes, and clustering using surface-based methods and graph models. UHL [16], Mohiuddin et al. [10], Mason et al. [17], Bobichev et al. [18], Ko et al. [19], Abali et al. [20] have used the news information over various real-time situations such as classification, sentiment analysis, text summarization, or clustering over streaming data of news, over the social media, identification of terrorist attacks using news articles, getting the current situation of market using the consumer behavior analysis over news data, using news information to detect the problems faced by citizens.

The past research work is focused on categorical classification or hybrid of similar approaches. Our research changes the paradigm of news text analysis with the inclusion of geospatial parameter. It enables us to use location as an additional parameter.

3 Model Analysis

The goal of the study is to do sentiment analysis over news data dependent on the location. The data which was needed for this study was obtained by scraping from various news websites of various locations. This data comprised of Headlines, Location, Latitude and Longitude values, and the News text. Since the data was comprised of collection of news from all locations, it had to be segregated into sets, so that each set consist of news of a particular location. This required to apply a

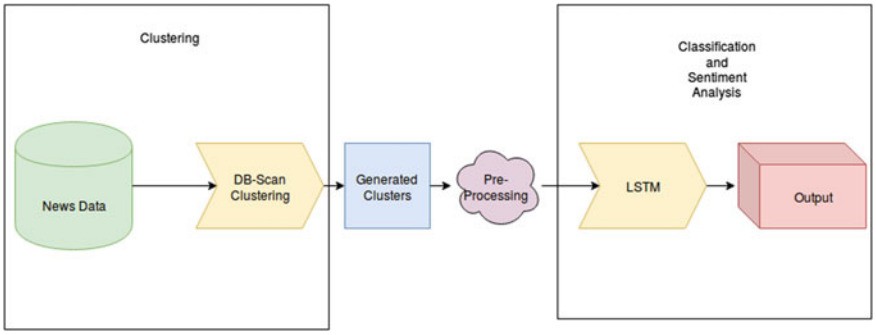


Fig. 1 Working model

clustering technique to perform location-wise clustering. After clustering, the next task was to perform the classification as we need to perform the sentiment analysis dependent on every category of news. Classification task segregates news headlines into different categories. The resultant data after classification along with relevant category can then be used for sentiment analysis. This analysis gives positive output if the words in particular news for a particular location and particular category is positive otherwise negative. For example, crime has happened and the criminal is free then it is negative but if he's been held then it is positive (Fig. 1).

3.1 Clustering

Clustering is the task of merging up the similar type of input elements into groups and different type into another group. In clustering approach, the main goal is to reduce the intra-cluster distance and increase the inter-cluster distance. Various algorithms are being used for the task of clustering such as K-means, Hierarchical Clustering, etc. [21].

The algorithm that we preferred for the purpose of the study is the DBSCAN algorithm.

DBSCAN algorithm:
The DBSCAN algorithm is a density-based and centroid-independent algorithm. The parameters that this algorithm carries are the only two which are minimum number of elements to form a cluster and the epsilon value so as to calculate the distance between two points. For the first part, we have to perform the clustering on the data, which carries the latitude and longitude values which we are going to perform the DBSCAN clustering algorithm. The DBSCAN clustering algorithm assigns the labels to each cluster which are nothing but the city.

3.2 Classification

Classification is the task of mapping the input to the respective output variables. The classification is divided into two types, i.e., binary classification (e.g., Yes/No) and multi-class classification [6]. The approach that we are using for the purpose of classification is a deep learning approach and we are using the LSTM model.

LSTM
Long Short-Term Memory is the advancement to the normal recurrent neural networks where Logic gates are being used for the purpose of storing the input till long time but which would be sequence dependent or group of sequence dependent. The LSTM takes input sequence and gives output after the sequence ends, output of every hidden node is given as an input to the other hidden node and not only that it is stored till the end using the logic gates.

3.3 Sentiment Analysis

Sentiment analysis is the study which leads us to give an opinion regarding various information. Sentiment analysis also known as opinion mining, is the process of making opinions regarding particular situation or we can say it as over particular input sequence or information (Fig. 2).

4 Experiment Analysis

The Working model for the research is mainly divided into two phases, first one is the clustering phase which is made by using the DBSCAN Algorithm and the second phase is classification and sentiment analysis which is made by using the LSTM network.

4.1 Dataset

The dataset used for the research study carries the data that depends on five cities {Mumbai, Delhi, Bangalore, Hyderabad, and Kolkata} and 19 class labeled data for the news are being used, which was scraped from different news sources. The dataset does not have any missing values or any outliers so the preprocessing for the data was not required, which have reduced the amount of time required for preprocessing (Fig. 3).

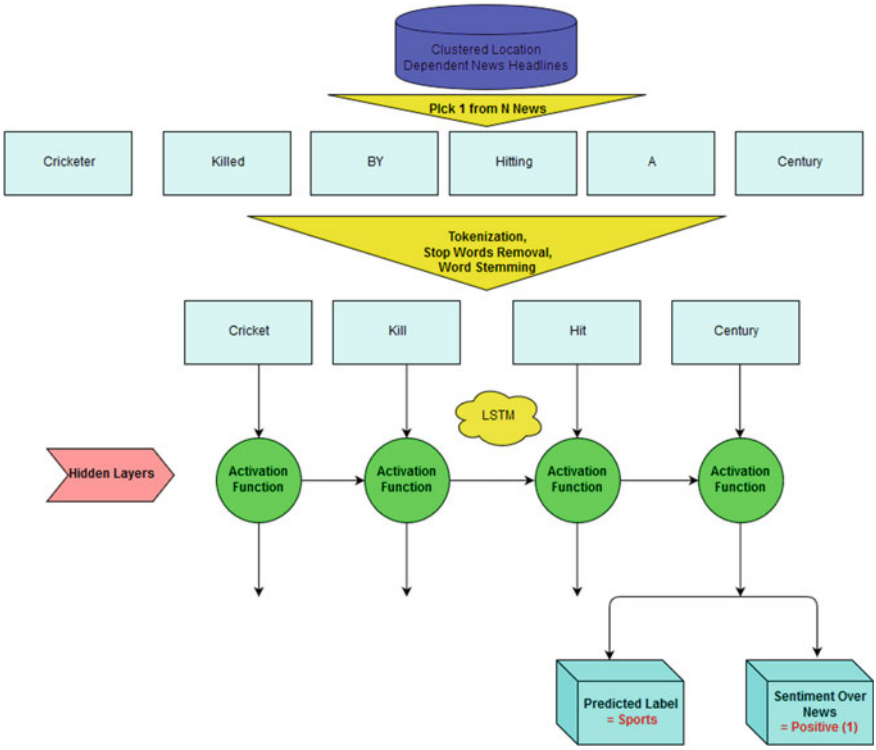


Fig. 2 Using LSTM for news classification and sentiment analysis

4.2 DBSCAN Clustering

The input to the DBSCAN clustering algorithm is the data which was scraped from the various news sites, and the clustering is performed on the basis of Latitude and Longitude assigned for the particular location with epsilon value as 1.5/6371.0088 where 6371.0088 is the radius of earth, and minimum number of samples to be considered as 5. The clusters formed from the DBSCAN were assigned to different labels which were after used for splitting that data, which are nothing but location-wise news sets (Fig. 4).

4.3 Preprocessing

For the task of preprocessing, the NLP operations such as tokenization where each input sequence is split into small tokens of words, stop words removal {e.g., is, was, it, they} where unnecessary words were removed, word stemming where only the

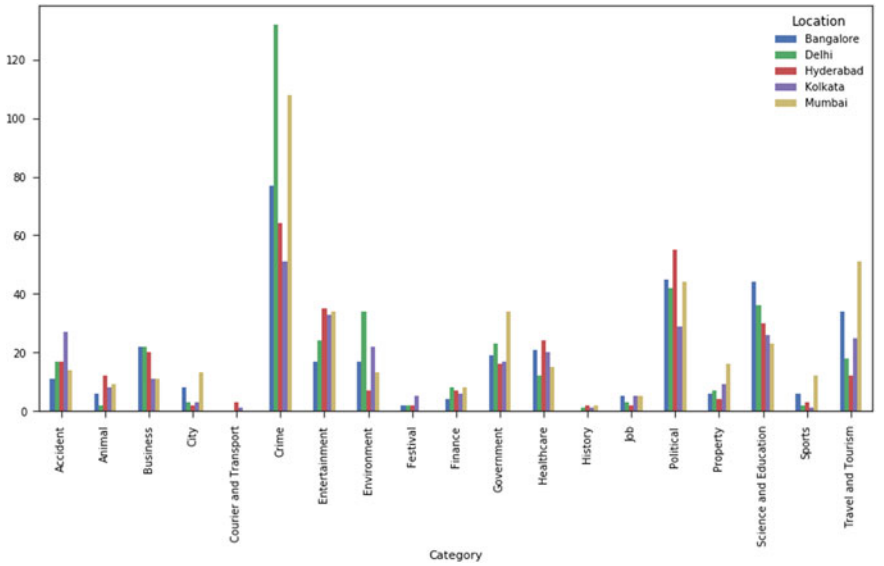


Fig. 3 Location-wise news headlines

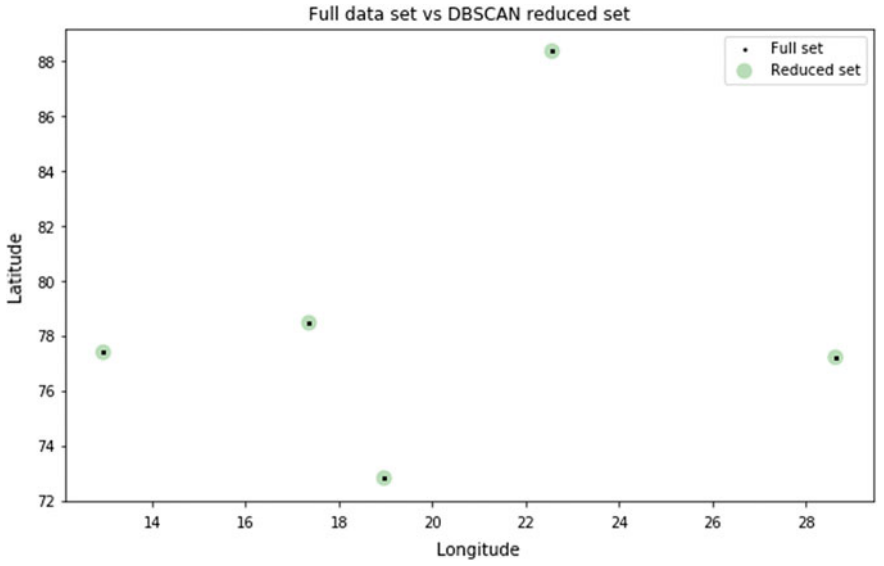


Fig. 4 Clusters according to location using LAT and LONG values

initial form of word is considered {e.g., (Hire, Hiring, Hires, Hired)→Hire} and then it is transferred to the LSTM as input.

4.4 LSTM for Classification and Sentiment Analysis

First, the words or tokens are taken into the vectors and dictionary was made using the word to vector model where each word is vectored using the related surrounding words. These words are used for creating the dictionary of words and for word embedding and optimization, model used is ADAM. This word of dictionary is used for the generation of training rules. The study has used Anaconda and TENSORFLOW as backend.

The numbers of dimension that were used for the purpose of study are 32 and the number of layers for the study were 5 each of size 256 with learning rate of 0.001.

4.5 Experimental Analysis

The accuracy score for the training set were observed as 0.54 and accuracy for testing set was observed as 0.58 with the loss rate of 0.68. In this research, we also have used the Precision, Recall, and F1-score values for the purpose of accuracy calculation (Fig. 5).

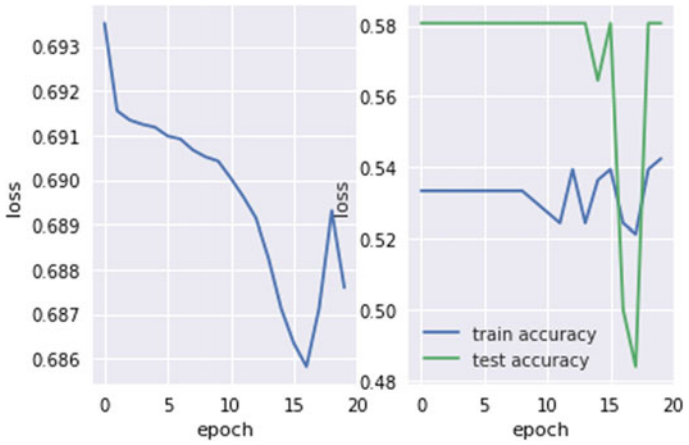


Fig. 5 Accuracy and loss graph

	Precision	Recall	F1-Score	Support
Negative	0.33	0.03	0.06	30
Positive	0.51	0.94	0.66	32
Avg/total	0.42	0.50	0.37	62

5 Conclusion and Discussion

In this paper, we have used DBSCAN clustering algorithm, which was used for the purpose of location-wise clustering and LSTM used for the purpose of classification and sentiment analysis, simultaneously the word to vector model and word embedding is performed for the creation of dictionary. Thus, we can say that we can perform the operations of sentiment analysis and classification with respect to the location-dependent news articles. This study can be used to find out that locations safety, growth or failures with respect to Education, Sports, Finance, Business, Health, Property, etc.

We got the accuracy of 58% which will increase with more data. The accuracy calculated shows us the sentiment of positive and negative articles over each category as per the locations. Our future research would involve increasing the accuracy by adding data and using different models for the purpose of classification and sentiment analysis and simultaneously optimize the amount of time required for the execution. Also, temporal analysis of the data can be done to find a change in sentiments with respect to time. Our future study will also involve visualization of the sentiments of different categories of news on geographic maps.

References

1. T.B. Shahi, A.K. Pant, Nepali news classification using Naïve-Bayes, support vector machines and neural networks. in *2018 International Conference on Communication, Information Computing Technology (ICCICT), Feb. 2–3, Mumbai, India*
2. W. Jirasirilerd, P. Tangtisanon, Automatic labeling for Thai news articles based on vector representation of documents, in *2018 International Conference on Engineering, Applied Sciences, and Technology (ICEAST)* (IEEE, 2018)
3. P. Kaushik, A.R. Sharma, Literature survey of statistical, deep and reinforcement learning in natural language processing. in *International Conference on Computing, Communication and Automation (ICCCA 2017)* (IEEE, 2017)
4. C. Li, G. Zhan, Z. Li, News text classification based on improved Bi-LSTM-CNN. in *2018 9th International Conference on Information Technology in Medicine and Education (ITME)*
5. J. Zhang, Y. Li, J. Tian, T. Li, LSTM-CNN hybrid model for text classification. in *2018 IEEE 3rd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)*
6. S. Sharma, J. Agrawal, S. Agrawal, S. Sharma, *Machine Learning Techniques for Data Mining: A Survey* (IEEE, 2013)

7. Z. Madhoushi, A.R. Hamdan, S. Zainudin, Sentiment analysis techniques in recent works, in *Science and Information Conference 2015* (IEEE, 2015)
8. T.D. Bui, D.K. Nguyen, T.D. Ngo, Supervising an unsupervised neural network. in *First Asian Conference on Intelligent Information and Database System* (IEEE, 2009)
9. D. Cecchini, L. Na, Chinese news classification. in *2018 IEEE International Conference on Big Data and Smart Computing*
10. U. Mohiuddin, H. Ahmed, M.A. Ismail, NEWSD: a real time news classification engine for web streaming data. in *International Conference on Recent Advances in Computer Systems* (RACS, 2015)
11. K. Gurmeet, B. Karan, News classification and its techniques: a review. IOSR J. Comput. Eng. (IOSR-JCE) **18**(1) Ver. III (2016)
12. V. Rao, J. Sachdev, A machine learning approach to classify news articles based on location. in *Proceedings of the International Conference on Intelligent Sustainable Systems (ICISS 2017)*
13. J. Azzopardi, C. Staff, Fusion of news reports using surface-based methods. in *26th International Conference on Advanced Information Networking and Applications Workshops* (2012)
14. J. Zhang, C.-T. Lu, M. Zhou, S. Xie, Y. Chang, P.S. Yu, HEER: heterogeneous graph embedding for emerging relation detection from news. in *IEEE International Conference on Big Data (Big Data)* (2016)
15. D. Nagalavi, M. Hanumanthappa, A new graph based sequence clustering approach for news article retrieval system. in *IEEE International Conference on Power, Control, Signals and Instrumentation Engineering (ICPCSI-2017)*
16. M.W. UHL, Explaining U.S. consumer behavior with news sentiment. ACM Trans. Manag. Inf. Syst. **2**, Article 9 (2011)
17. R. Mason, B. McInnis, S. Dalal, Machine learning for the automatic identification of terrorist incidents in Worldwide News Media. in *2012 IEEE, ISI 2012 June, 11–14, Washington D.C., USA*
18. V. Bobichev, O. Kanishcheva, Sentiment analysis in the Ukrainian and Russian news. in *2017 IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON)*
19. B.S. Ko, C. Park, D. Lee, J. Kim, H.-J. Choi, D. Han, Finding news articles related to posts in social media: the need to consider emotion as a feature. in *2018 IEEE International conference on Big Data*
20. G. Abali, E. Karaarslan, A. Hurriyetoglu, F. Dalkilic, Detecting Citizen problems and their locations using Twitter data. in *6th International Istanbul Smart Grids and Cities Congress and Fair (ICSG)* (IEEE, 2018)
21. Nisha, P.J. Kaur, *A Survey Of Clustering Techniques and Algorithm* (IEEE, 2015)

Review of Plagiarism Detection Technique in Source Code



Anala A. Pandit and Gaurav Toksha

1 Introduction

Stanford defines plagiarism as the "use, without giving reasonable and appropriate credit to or acknowledging the author or source, of another person's original work, whether such work is made up of code, formulas, ideas, language, research, strategies, writing, or other form" [1]. With recent access to easy and low-cost Internet facilities, information is available on the click of a finger. However, this benefit is usually utilized in an undesirable manner. The enormous access to web content has turned plagiarism into a serious problem be it for researchers, publishers, or educational institutions.

In an academic environment, the most common forms of plagiarism are textual plagiarism which is the document level where documents are essays, reports, scientific papers, and the other is the source code plagiarism in programming assignments.

The students resort to such malpractices since these are easily available on the Internet. With an explosion of Massive Open Online Courses, Github repositories to showcase work, "stackoverflow" where people post their assignments and receive suggestions from community and educational websites such as "geeksforgeeks" makes plagiarism an easy way out for many. This is just an example sources and by no means exhaustive.

This makes the problem of assignment plagiarism detection an important task. One of the effective ways is *manual inspection,* but it is laborious and time consuming.

A. A. Pandit (✉) · G. Toksha
Veermata Jijabai Technological Institute, H.R Mahajan Marg, Matunga (E), Mumbai 400091, India
e-mail: aapandit@mc.vjti.ac.in

G. Toksha
e-mail: gotoksha_16@mc.vjti.ac.in

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_38

It is hard to manually inspect and detect (decide whether a submission is genuine or plagiarized) plagiarized assignments in a large class. This necessitates the need for identifying the tools, techniques available that could help detect plagiarism in source code. There are excellent survey papers available for plagiarism detection [2, 3]. However, none of them have focused specifically on source code plagiarism. In this paper, the focus is only for source code plagiarism detection.

The rest of the paper follows the given organization. Currently available techniques are discussed in Sect. 2 where details of the various algorithms/techniques along with their advantages and disadvantages are discussed, Sect. 3 discusses the similarity measures and proposes the appropriate technique to choose. Section 4 discusses future scope and Sect. 5 concludes the paper.

2 Literature Review

Traditional systems include the use of partial check-sum [4] and partial code checking using basic string comparison algorithms [5]. However, the main disadvantage of such systems is that they are not robust to even the simple kind of code obfuscation [6].

Code obfuscation is the process of making source code more difficult to understand by introducing various dummy functions and variables in the code. These introductions do not make any significant contribution to main logic or the functionality of the program.

A simple arrangement of code block to some other place would immediately fail in detecting the plagiarized code [5].

2.1 Algorithm Architecture

An algorithm architecture gives a framework around which an algorithm should be made. It is also considered as the baseline, which ensures that we won't miss any important factors while deriving a new algorithm (Fig. 1).

Traditional Plagiarism detection techniques adhere to the architecture shown above. It has the following components:

- *Input text document*: This is the input documents provided to the algorithm to analyze plagiarized pair of documents.
- *Collection of Documents*: This is more of a corpus concept where in the algorithm will look for document originality. The algorithm would compare the input document with this corpus to determine the similarity between documents.
- *Threshold*: This is one of the critical concepts in plagiarism detection techniques. Threshold would help the algorithm identify which documents are plagiarized and

which aren't. For example, if the similarity score is greater than a specified number then that pair of document are considered as plagiarized.

- *Plagiarized Segments with References*: The output of such algorithm is to give snippets of the code where the code similarity is above the threshold and considered as plagiarism.

2.2 Algorithm Properties

Saul Schleimer in his paper [7] about MOSS has given some fundamental properties every algorithm should satisfy.

- *Whitespace Insensitivity*: It is a human tendency to add white space in between source code components to enhance its readability. However, a user may add extra white space according to his convenience. This makes it essential for the algorithm to be whitespace insensitive. For documents other than source code while performing data cleaning and before comparison of text, redundant part that is generally removed is white spaces, punctuations, capitalizations, etc.
- *Noise suppression*: Matching variable names and types of loop structures would be redundant and could be considered as noise. For other types of document, noise would mean stopwords such as "a" or "the", etc. Inclusion of such noise elements would hamper the results of similarity. Hence, the algorithm should be robust enough to handle noise..
- *Position independence*: Many times, it is found that rearranging the words or sentences is resorted to while plagiarizing. Re-arrangement of the set of matches (permutation of content) in the plagiarized documents should not affect the result of similarity. This is true for every type of document comparison whether it is a source code or a textual document.

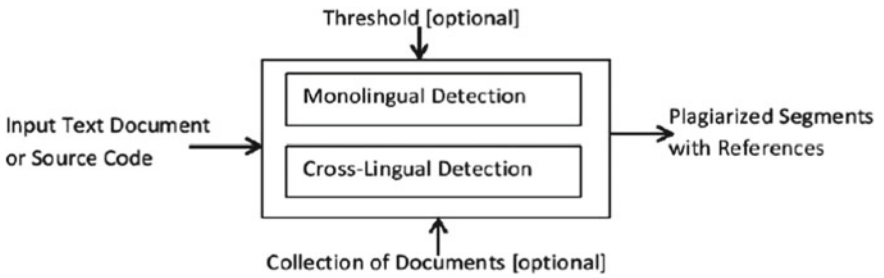


Fig. 1 Algorithm architecture [2]

2.3 Algorithms Used for Plagiarism Detection

N-Grams: It is also known as k-gram. It is a famous technique used in text mining for finding similar documents [8]. The concept is simple, first, the N-grams are extracted from the words of the documents and then, the similarity is computed by identifying the n-grams that are common.

Document 1: Mumbai is a city;
Document 2: Mumbai is not a city;
Considering the value of N is 5;
5-grams of document 1: mumba umbai mbaii baiis aiisa iisac isaci sacit acity;
5-grams of document 2: mumba umbai mbaii baiis aiisn iisno isnot snota notac otaci tacit acity.

$$\text{Similarity} : \text{common n} - \text{grams/total n} - \text{grams} = 5/18 = 0.277$$

Since the number of N-grams of reasonably big paragraph would be huge, it is necessary to only a subset of such N-grams which represent the document well enough.

Advantage of such a system is easy implementation and provides interpretable results. However, it has poor performance in terms of precision and recall [8].

Document Fingerprinting: After generating N-grams for such a document, it is advisable to transform/hash them into some numbers to improve the space efficiency and for easy representation. This enables the user to select some representational subset of the document for fingerprinting [7].

For example,

(a) *Input*: The quick brown fox jumps over the lazy dog.
(b) *Text with redundant features removed*: The quick brown fox jumps over the lazy dog.
(c) *Array of 5-grams derived from the text*: thequ hequi equic quick uickb ickbr ckbro kbrow brown rownf ownfo wnfox nfoxj foxju oxjum xjump jumps umpso mpsov psove sover overt verth erthe rthel thela helaz elazylazyd azydo zydog.
(d) *Hypothetical sequence of hashes of 5-grams*: 56 67 29 60 37 80 76 36 45 20 12 63 64 8 97 5 79 61 92 9 3 65 26 48 69 95 73 17 74 50
(e) *The array of hashes selected using 0 mod 4*: 12 64 8 92 48

The logic for selection of fingerprints for this particular example is 0 mod 4. However, this logic should be formed in such a way where the chances of missing any relevant information is minimized.

Since each document will have different positional relevance in the structure of the text, it is necessary to form a more generalized way to select fingerprints which

represents the entire text in an uniform way. Winnowing algorithm [7] maximizes the representation of the document by using the concept of moving window. For each moving window, Winnowing selects the minimum value and if there are more than one minimum value it will select the rightmost one. Example below shows moving window of hashes of length 4.

For example,

(a) *Window of hashes*:

(77, 74, 42, 17) (74, 42, 17, 98)
(42, 17, 98, 50) (17, 98, 50, 17)
(98, 50, 17, 98) (50, 17, 98, 8)
(17, 98, 8, 88) (98, 8, 88, 67)
(8, 88, 67, 39) (88, 67, 39, 77)
(67, 39, 77, 74) (39, 77, 74, 42)
(77, 74, 42, 17) (74, 42, 17, 98)

(b) *Fingerprints selected by the algorithm*: 17 17 8 39 17

Note: The value of the window size and the "N" of N-Gram should be provided by the user. However, based on empirical study [7], the size should satisfy the following condition for optimal performance:

$$T = W + N - 1 \tag{1}$$

where T is the threshold, W is window size, and N is the size of N-Gram.

Moss algorithm is developed by MIT as a way to detect plagiarized assignments on the student's submission [7]. Even though it is an old detection system, it is still relevant and gives reasonably good results for detecting plagiarism.

It uses the techniques mentioned above, i.e., N-grams, document fingerprinting, Winnowing to extract relevant fingerprint and uses various similarity measures to present the similarity score.

Advantages of applying Moss is that it works on source code as well as on language documents. This is more of a generalized approach to detect similarity between documents. It doesn't use any syntactic information of the language or the source code.

Disadvantage of Moss is that it doesn't work well enough in the cases of code obfuscation.

In [9], an open-source plagiarism detection tool is discussed which is integrated into Submitty course management platform. It creates tokens of source code and generates fingerprints for comparison.

Parse Tree Based Detection: The program source code is converted as a parse tree where each node represents variables used in the code, reserved words, operators, etc. [10].

The leaf nodes are the actual values of the parent nodes. The Parse tree for a program can be generated from antlr.org website. ANTLR [11] proposed by Parr

and Quong, is a language tool that gives a framework to construct recognizers, interpreters, compilers, and translators from grammatical descriptions. The parse tree provides syntactic structural information that necessitates the creation of a performance measure for the parse tree that reflects the entire structural information. The parse tree kernel is one of such metrics. It compares parse trees without manually designed structural features.

Advantage of such technique is that it uses semantic information of the code rather than working directly on text (as in MOSS). If the code is not much complex, then comparing the parse trees would give better results.

Disadvantage is obvious that the parse trees will be unwieldy when the program structure is extremely complex with multiple recursions.

Graph Based Detection: A program dependence graph (PDG) [12] represents the source code of a procedure in the form of a *graph*. The program vertices in a PDG are mandatory statements like variable declarations, assignments, and procedure calls. The edges between program vertices are data and control dependencies between statements.

PDG detection technique stores certain information about the relations of vertices and the similarity is determined by comparing these vertices. Advantage of such an approach is that complex source code can be broken down and represented in easy-to-interpret structure. It could also help in identifying dummy functions which are made in an attempt for code obfuscation. In case of extremely complex code (deep recursive codes), however, the disadvantage could be that the trees may appear very complex and may not be interpretable in some cases (Fig. 2).

Attribute-Based Detection: This is a source code plagiarism detection tool based on attribute counting [13]. This framework has the speed advantage. It also works on

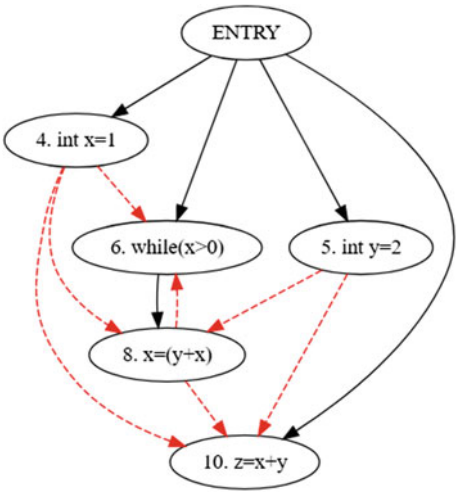


Fig. 2 Program dependence graph [24]

large quantum of data. The process creates a token list for each source code and thus eliminates the need to process the complete code each time. In this algorithm, the first stage is selection of attributes from the source code, the second stage is assigning a unique token to each attribute. Finally, the similarity is calculated by counting the common tokens.

In [14], an attribute-counting system was proposed by Halstead where in a program A is treated as tokens classified into operators and operands, further, these tokens are grouped into variables of $N1$ number of operator occurrences, $N2$ number of operands occurrences, $n1$ the number of distinct operators, and $n2$ the number of distinct operands.

These variables are later grouped in a tuple and are used for comparisons. In due course, more attributes are added such as number of variables, loops statements. In [13], more complex counts were added for better comparison.

Design-Based Detection: A detection technique [15] in which the design pattern of the source code is extracted/generated and similarity is measured on the design pattern of the source code.

Style-Based Detection: In style-based technique [16], an attempt is made to identify the author of the source code by maintaining the features of each author. The coding style of the author of the code is used to identify whether the submission is plagiarized or not.

Machine Learning-Based Detection: There are mainly two types of machine learning approaches to detect similar codes:
Unsupervised Learning: Relevant features in the source code are extracted from the source codes and are clustered using distance metrics such as Euclidean distance. Small clusters formed after this procedure are considered similar [17].
Supervised learning: From each source code, semantic features are extracted from the compiler as well as other types of features are extracted such as difference in length of submission, Similarity of Comments, Similarity in String literals that are used as features. Since this is a supervised approach, a dataset is already prepared with the given features and a final class value as Cheated or Not Cheated (could be hot encoded). Deep learning models such as RNN with LSTM support are trained on the dataset [18].

Another CNN based plagiarism detection algorithm is proposed in [19]. The algorithm uses TFIDF features of each source code token. It is then represented as word embedding which is fed into CNN.

String Matching-Based Detection: One of the popular String Matching techniques based detection system is JPlag [20]. It is available to access for free on the Internet and has an open source repository. JPlag is capable to evaluate plagiarism of source code written in C, C++, Scheme, and Java. It adopts the following process: the source code is initially parsed and later transformed into token of strings. The tool then matches these array of strings by using *Running Rabin Karp Greedy String Tiling algorithm*. The output of the process is then displayed in a browser.

Hybrid Detection Techniques: Hybrid Techniques uses two or more methods of plagiarism detection techniques to get better similarity scores.
Combination of Parse Trees and PDG: Since Parse Trees and PDG works best on different types of source codes, it would be better to create a composite of both of these approaches into one model leveraging the advantages of both of the models [21]. The model uses the semantic benefits of both the features of source code. However, to determine the weights of usage of each technique, cyclomatic complexity is used.

Advantage of such a model is that it leverages two semantically evaluated information, hence, the quality of the result is improved. It does help in the issue of code obfuscation to some extent.

Disadvantage is that depending upon the complexity of the source code, it might be possible that the time complexity of the model is not reasonable since creating graph for a deep recursive program would be time consuming and not scalable.
PDetect: It is a hybrid detection technique where in the similarity detection is done in two phases [22]:

In Phase one, a set of program files is supplied as input. For each of the programs supplied as input, a representation based on the keywords contained in the program is produced. A similarity metric is determined for each pair of programs. This pair-wise result is stored in a flat file with its corresponding similarity value, in a descending order. In practice, this would mean first the documents could be tokenized into words and the set of words for each document would be generated.

Then their similarity is calculated using Jaccard similarity measure, However, there are several ways to represent a program, one of them is using attribute based in which functions, variables, Boolean values, and loops information are preserved.

In Phase two, the entire set of program pairs in the flat file are represented as a weighted non-directed graph. The vertices will represent the programs and the similarities between the programs is represented as edges. An appropriate graph-clustering algorithm is then used to cluster these graphs.
ES-Plag: It uses the novel concept of *Running Karp Rabin Greedy String Tiling Algorithm (RKRGST)* to calculate plagiarized documents [23]. RKRGST is an effective technique where in the documents would be converted into tokens first. Then, the corresponding pair of tokens will be matched by comparing their hashes. This part is done by RKR and the comparison of the subsequences are done by GST. RKRGST is very time costly. RKR itself takes linear time complexity and GST is of quadratic complexity. Together they would have Cubic time complexity. This is very costly considering the scalability feature.

ES-Plag reduces this complexity by reducing the number of pairs of program given to RKRGST by using Cosine Similarity. The user of the detection technique would give a cutoff value to the program. The values above the cutoff criterion would be inserted into the RKRGST algorithm for further evaluation. The pairs below the cutoff criterion would be ignored. But there is a disadvantage for using RKRGST for comparing subsequences.

This algorithm is insensitive to subsequence rearrangements. Due to this, the value of similarity would increase and would give a false impression about plagiarism.

ES-Plag uses penalties to maintain the error caused by this insensitivity.

3 Proposed Evaluation Measures

After studying in detail all of the detection techniques mentioned above and some additional research papers, it was observed that determination of best technique in comparison to others under different conditions is difficult to identify.

The issues observed while researching were: lack of implementation details of the concepts and unavailability of open-source code that support implementation of these techniques. Absence of a working software backing these research papers, performing any experimentation to know which algorithm works better became a challenging task.

3.1 Similarity Metrics

Following are the similarity metrics which are used selectively in many of the algorithms discussed previously.

- *Jaccard Similarity*: No. of elements in intersection of sets/No. of elements in union of sets.
- *Cosine Similarity*: No. of elements in intersection of sets/product of No. of elements in the sets.
- *Dice Similarity*: 2 × No. of elements in intersection of sets/summation of No. of elements in sets.

Dice and Cosine similarity usually show approximately the same similarity values, Jaccard shows a significantly less similarity value

Similarity value are considered to be in the range of 0–1, where 0 is considered as no similarity and 1 is considered as completely identical source code.

To determine the similarity measure to use, it is dependent on the use case of the system. For example, if the requirement is that there is no tolerance for plagiarism in the system, Dice or Cosine would be better similarity measures. After experimenting several times with different plagiarized documentations, it has been observed that these measures would overestimate the similarity values and would even mark those pairs whose similarity values are closer to 0.

Figure 3 shows the comparison performed on a sample source code which includes the basic function of iterating through a list and swapping two variables. The comparison is done using MOSS.

Figure 4 shows a comparison of two completely different programs.

Following are the observation from experiments:

- Cosine similarity value is highest amongst all the three similarity measures and hence it is a better similarity measure for plagiarism intolerant systems like program codes that have been submitted for patent granting.
- Dice similarity value is in between Jaccard similarity and Cosine similarity but it is much closer to Cosine similarity as compared Jaccard similarity.

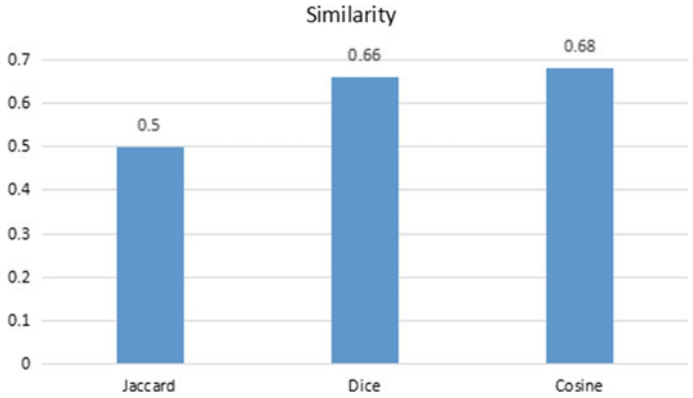


Fig. 3 Comparison of similarity measures on sample source code

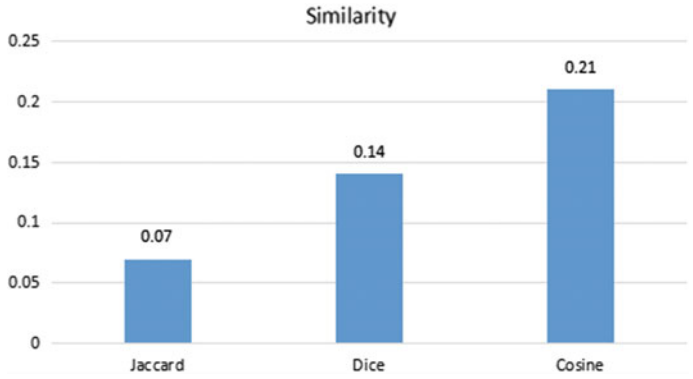


Fig. 4 Comparison of similarity measures on two completely different source codes

- Jaccard similarity value is the lowest similarity value and hence it is advisable when the system is supposed to be very lenient.

If one is not sure about the nature of the system, Dice similarity seems to be a better option to choose.

Thus, if the system has a very strict policy for plagiarism, it is advisable to use Cosine similarity, if it is lenient then Jaccard similarity. If the policy is moderate or in between moderate to high then Dice similarity is a better measure.

3.2 Model Performance Measures

Following are the performance measures which helps to evaluate the efficiency of the model.

- Precision is defined as

$$\mathrm{TP}/\mathrm{TP} + \mathrm{FP} \tag{2}$$

- Recall is defined as

$$\mathrm{TP}/\mathrm{TP} + \mathrm{FN} \tag{3}$$

- F-measure is defined as

$$(2 * \text{Precision x Recall})/\text{Precision} + \text{Recall} \tag{4}$$

Here, TP refers to True Positive, FP refers to False Positive (wrongly identified as positive), and FN refers to False Negative (wrongly identified as negative). In general cases, F-measure is sufficient to compare performances and it is not necessary to worry about the individual metrics. However, in specific cases where the systems output is very critical, recall is given more importance.

4 Future Scope

In recent times, there have been major improvements in the machine learning area for plagiarism detection it is very foreseeable in future to have models which work better than the traditional ones. With proper cross-validation techniques and a good quality training datasets deep learning models such as RNN would potentially outperform the traditional techniques and they would be robust against code obfuscation practices.

The existing techniques have utilized only the semantic information for detecting plagiarism. The methods have ignored other parameters such as the difference between submission timings, patterns in comments, patterns in variable names can be utilized as feature vectors in machine learning algorithms to train better models, however, whether there is a relation between such semantics and the code being plagiarized or not is still remains to be explored.

5 Conclusion

This paper has presented a reasonably complete, but not exhaustive study of techniques of plagiarism detection available at current stage. It has presented a taxonomy of various source code plagiarism detection techniques currently being used. It has given a fair idea of advantages and disadvantages of some of the traditional algorithms. Although in the past two decades, there are many detection tools introduced,

and there are still several challenges and issues present in the current available systems.

References

1. Plagiarism-Wikipedia, https://en.wikipedia.org/wiki/Plagiarism#cite_note-22. Accessed 04 Nov 2019
2. H. Chowdhury, D. Bhattacharyya, Plagiarism: taxonomy, tools and detection techniques. *arXiv preprint* arXiv:1801.06323 (2018)
3. S.M. Alzahrani, N. Salim, A. Abraham, Understanding plagiarism linguistic patterns, textual features and detection methods, in *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 42, no. 2 (IEEE, New York, 2012), pp. 133–149
4. V. Kelly Adam, Method for detection plagiarism, Patent No. US6976170
5. M. Wise, String similarity via greedy string tiling and running Karp – Rabin matching, Unpublished Basser Department of Computer Science Report (1993)
6. Wikipedia, http://en.wikipedia.org/wiki/Obfuscation_(software). Accessed 12 Nov 2019
7. S. Schleimer, D. Wilkerson, A. Aiken, Winnowing: local algorithms for document fingerprinting, in *Proceedings of the 2003 ACM SIGMOD International Conference on Management of Data* (ACM, 2003), pp. 76–85
8. E. Stamatatos, Intrinsic plagiarism detection using character n-gram profiles, Threshold **2**(1), 500 (2009)
9. M. Peveler, T. Gurjar, E. Maicus, A. Aikens, A. Christoforides, B. Cutler, Lichen: customizable, open source plagiarism detection in submitty, in *50th ACM Technical Symposium on Computer Science Education, Minneapolis, USA* (2019)
10. J. Son, S. Park, S. Park, Program plagiarism detection using parse tree kernels, in *Pacific Rim International Conference on Artificial Intelligence* (Springer, Berlin, Heidelberg, 2006), pp. 1000–1004
11. ANTLR Homepage, https://www.antlr.org/. Accessed 04 Nov 2019
12. C. Liu, C. Chen, J. Han, P.S. Yu, GPLAG: detection of software plagiarism by program dependence graph analysis, in *Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (ACM, Philadelphia, USA, 2006), pp. 872–881
13. J.A. Faidhi, S.K. Robinson, An empirical approach for detecting program similarity and plagiarism within a university programming environment. Comput. Educ. **11**(1), pp. 11–19 (1987)
14. M. Halstead, *Elements of Software Science* (Elsevier, New York, 1977)
15. A. Asadullah, M. Basavaraju, I. Stern, V. Bhat, Design patterns based pre-processing of source code for plagiarism detection, in *2012 19th Asia-Pacific Software Engineering Conference* vol. 2 (IEEE, Hongkong, China, 2012), pp. 128–135
16. O.M. Mirza, M. Joy, G. Cosma, Style analysis for source code plagiarism detection—an analysis of a dataset of student coursework, in *IEEE 17th International Conference on Advanced Learning Technologies (ICALT)* (Timisoara, Romania, 2017), pp. 296–297
17. J. Yasawi, B. Katta, G. Srikailash, A. Chilupuri, S. Purini, C. Jawahar, Unsupervised learning based approach for plagiarism detection in programming assignments, in *ISEC. 2017, Jaipur, India* (2017)
18. J. Yasaswi, S. Purini, C.V. Jawahar, Plagiarism detection in programming assignments using deep features, in *4th Asian Conference on Pattern Recognition (ACPR 2017), Nanjing, China* (2017)
19. M. Abuhamad, J. Rhim, T. AbuHmed, S. Ullah, D. Nyang, Code authorship identification using convolutional neural networks. Future Gen. Comput. Syst. 104–115 (2018)
20. L. Prechelt, G. Malpohl, Finding plagiarisms among a set of programs with JPlag. J. Univ. Comput. Sci. **8**(11), 1016–1038 (2003)

21. H. Song, S. Park, S. Young Park, Computation of program source code similarity by composition of parse tree and call graph, in *Mathematical Problems in Engineering*, vol. 2015 (Hindawi, United Kingdom, 2015)
22. L. Moussiades, A. Vakali, PDetect: a clustering approach for detecting plagiarism in source code datasets. Comput. J. **48**(6), 651–661 (2005)
23. L. Sulistiani, O. Karnalim, ES-Plag: efficient and sensitive source code plagiarism detection tool for academic environment. Comput. Appl. Eng. Educ. **27**(1), 166–182 (2019)
24. StackOverflow, https://stackoverflow.com/questions/46872521/draw-a-program-dependence-graph-with-graphviz. Accessed 04 Nov 2019

Study on the Future of Enterprise Communication by Cloud Session Border Controllers (SBC)



Siddarth Kaul and Anuj Jain

1 Introduction

Cloud network Virtualization has nowadays gained momentum and is used to address the solutions of the exploding high volume of network unified communication traffic the service providers can provide the new cloud Virtualized solution with the benefits of Agility, Fasserservice for new markets, Capacitive Expenditure (CAPEX) and Operating Expenditure (OPEX) savings through Infrastructural flexibility and operational efficiencies [1]. Network Function Virtualization (NFV) architecture can use various communication technologies like Skype for Business, Asterisk, Free PBX, and GENBAND. The Network Function Virtualization function was recently presented by GENBAND in their Cloud Virtualization infrastructure orchestration for Unified Communication [2]. The paper covers the opportunities and Challenges in virtualizing the on premises Session Border Controller (SBC) implementation to Network Function Virtualization Cloud for better Cloud Communication scalability and automation in communication [3].

S. Kaul · A. Jain (✉)
Bhagwant University, Ajmer, Rajasthan, India
e-mail: anuj.22631@lpu.co.in; a1978jain@gmail.com

S. Kaul
e-mail: kaulsiddarth2013@gmail.com

A. Jain
SEEE, Lovely Professional University, Jalandhar, India

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_39

2 Cloud Session Border Controller Virtualization (SBC—Virtualization)

The session Border Controller (SBC) is a vital communication element and it provides secure Voice Communication with better session security, internetworking, and advanced control for real-time voice, messaging, and video communication. The SBC virtualization redefines communication and CSP and enterprises are, nowadays, started to realize how cost-effective and efficient the solution is for enterprise usage [3] (Fig. 1).

3 Cloud Session Border Controller (SBC) Virtualization Benefits

- Service Agility and Operational Simplicity.
- Better Cost Efficiency.
- New Revenue Opportunity.
- Better efficiency usage via automation and elastic utilization of resources.

4 Cloud Session Border Controller (SBC) Virtualization View in the Cloud

Cloud management in one of the scopes in which SBC are virtualized in a single management view in the cloud network, and it essentially underlines the complexity to

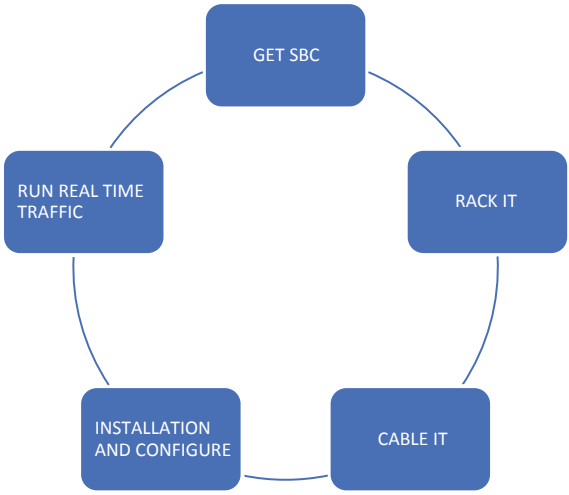


Fig. 1 Diagrammatic representation of loop of operational and cost inefficiencies

cloud application and the deployment resources through automation and complexity in a single cloud view, the virtualization of cloud Session Border Controller (SBC) works in tandem to high scaling and complexity including the logical representation of Sonus cloud application through management console, through cluster view of applications and distribution of licenses in cloud-based configuration [3]. Another view of Virtualization in the cloud is through the automatic registration of nodes in cloud and the open stack implementation in NFV domain within the cloud for fault management and reporting [3] (Fig. 2).

Virtualization of SBC help in service agility and Operational simplicity for new markets and new services with better removing life cycle hardware dependencies and better elastic scalability. It is very important to indeed understand the virtualization of Session Border Controller (SBC) is the need of the hour and helps in a solution which is highly agile in today's-centric environment [4].

5 Cloud Session Border Controller (SBC) Virtualization Test Setup in the Cloud

The test can be performed in a small computer incorporated Lab having voice traffic generator or analyzer are required the System under Test (SUT) requires virtualized instances of signaling session controllers responsible for SIP traffic generation instances in a Media Session Controller (MSC) which process media stream. Each of the instances are configured with 8 CPU and 4 NIC with 16 GB of RAM was allocated to SSC on MSC the test bed consists of 2 SIPp server generators and 2 cisco catalyst switch and a router with an open stack cluster and a met switch. In most of the cases, the SIP signaling is processed in such a way that the call processing happens quickly without any issues and the flow of calls in the Cloud environment is very quick the Met switch controls the orchestration of transmitting the signaling

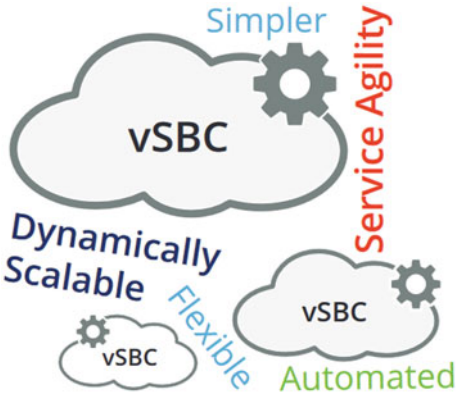


Fig. 2 Virtual SBC benefits

information and media streams in a direction for analyzing the call flow for packet loss and delay in call flow [4] (Fig. 3).

The example of call signaling performance rate on such a test bed has already been demonstrated by EANTC CORPORATION in its white paper titled for virtualization of Perimeter Session Border Controller (SBC) where the calls on different solutions on Cloud Session Border Controller (SBC) has been demonstrated in the table (Tables 1, 2, 3, and 4).

Figure 4 describes the call flow session in a Cloud SBC and it includes the call communication between a subscriber and a SIP proxy and SBC as gateway, there are various mechanisms of call flow as shown in call flow test bed diagram.

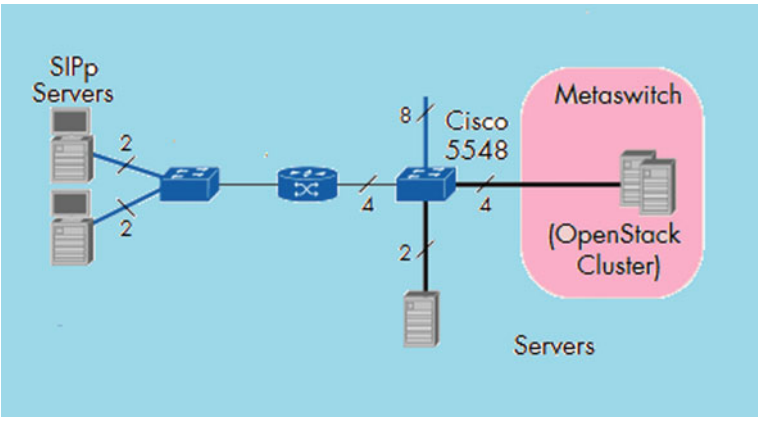


Fig. 3 Test bed setup in cloud

Table 1 Call signaling performance rate on test bed

Scenario	Max. Call Rate [calls/s]
SIP	700
SIP + TLS	600
VoLTE	240
VoLTE + IMS-AKA	228

Table 2 Registration performance rate on test bed

Scenario	Reg. Rate [msg/s]
Plain	2600
TLS-Encrypted	720
Fast registration (Plain)	14000
Fast registration (TLS)	4400
Authenticated VoTLE IMS-AKA	810

Table 3 SIP session call capacity and ram usage on test bed

Scenario	Concurrent Calls	RAM usage [%]
SIP	85000	56
SIP + TLS	85000	78
VoLTE	85000	52
VoLTE + IMS-AKA	50000	37

Table 4 RTP session calls capacity usage on test bed

Scenario	Concurrent Calls
Plain RTP	78000
sRTP	24000
sRTP w/silence suppression	35000
Plain RTP (VMware)	70000

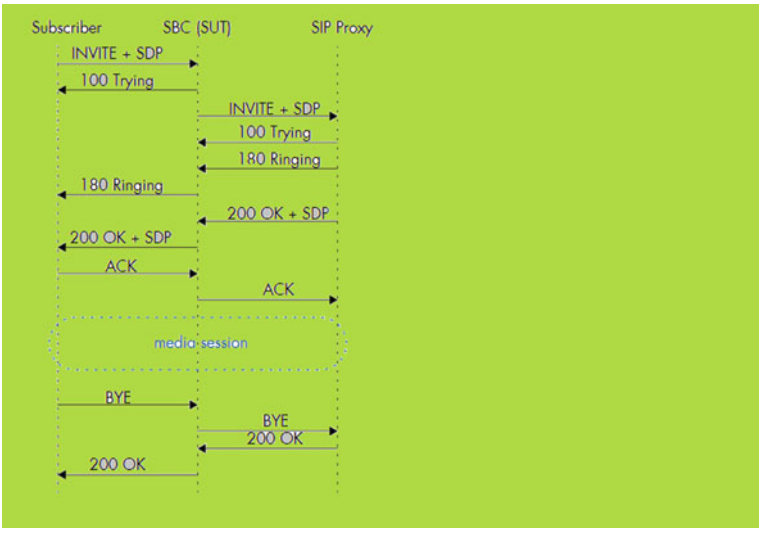


Fig. 4 SIP call flow in the test bed

Though the model can be implemented on a VM ware platform for the test on VM ware we may require one Signaling Session Controller (SSC) and Media Session Controller (MSC) units in a high availability cluster on a VM ware platform [5].

A test conducted by EANTAC Corporation reveals that such type of test conducted on a vm platform can generate almost 70000 concurrent calls as per the test revealed by EANTAC Corporation [1].

There are many more Cloud Session Border Controller (SBC) manufacturers like Oracle, Sangoma, Sonus, and Avaya, which are also working in the same direction toward a Cloud SBC solution on a VM platform.

Another example for the Cloud VM approach was recently performed by GenBand and is given out in the whitepaper titled the SBC NFV Virtualization Approach, and the advanced software capabilities demonstrated by Genband in the white paper do illustrate various software capabilities like VOIP DOS attacks and fraud. Advanced routing options including ENUM, DNS, and SIP redirect these all are some of the best illustrations which describe the Cloud Session Border Controller (SBC) in VM-centric environment [1, 6].

6 Cloud Session Border Controller (SBC) Key Differentiators

- Cloud Session Border Controller (SBC) optimizes the data signaling resources and media.
- Cloud Session Border controllers provide enhanced Session Border Controller (SBC) redundancy and failover.
- Unique licensing and wide flexibility to deploy multiple Survival Branch Appliances (SBA).

Figure 5 illustrates Huawei simple Cloud-Based Desktop solution, and the solution itself is an example of how remote Cloud desktop solution works and opens up a new area for understanding whether this solution can be used in scope of how Cloud Session Border Controller (SBC) can be used with these remote Cloud-based desktop solution [7].

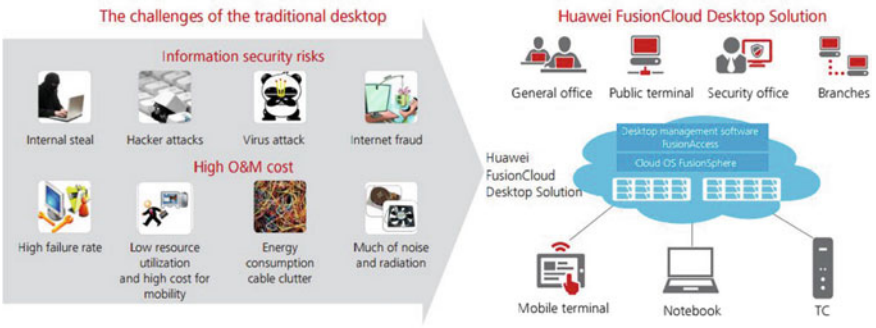


Fig. 5 Huawei cloud SBC environment

7 Cloud Session Border Controller (SBC) Virtualization Big Data Analytics

Big data analytics can be used for report generation of data from these Cloud Session Border Controllers (SBC) in a simple way like data can be used with Dashboards which have widgets, Charts, tables, and Graph for Call quality Reports.

The Big data analytics can also be used for network data performance analysis and Call Detailed Records (CDR). Signaling across the media events can be demonstrated with HTML for call Pie data chart analytics or protocol packet Flow Information.

The data from these tables can be integrated with Power BI for a reporting dash board from where reports can be fetched anytime and from anywhere despite any restriction of access from anywhere anytime whether from home or from Office [8].

8 Difference Between Cloud Session Border Controller (SBC) and Cloud PBX (Private Branch Exchange)

Cloud Session Border Controller (SBC) involves the Virtual implementation on VM ware like EXSI host integration with the help of cloud connector to the host client whereas Cloud PBX (Private Branch Exchange) involves implementation on a Physical server or VM hosted on a Physical server at some remote location.

Cloud Session Border Controller (SBC) involves complete transcoding on the VM ware hosting the virtual implementation whereas Cloud PBX (Private Branch Exchange) involves channel route patterns on Server on which Cloud PBX resides [9].

9 Cloud Session Border Controller (SBC) Virtualization Conclusion

The evolution of Cloud SBC represents a fundamental shift in the type approach of communication service providers there steps to mitigate the best solutions for their endusers and also the benefits of deploying VM-based Cloud SBC solutions with good scalability and high automation for less time consuming way to enabling a faster service rollout, better service agility, and improved quality of service.

Virtualization offers a better approach to multiple vendor Interoperability and rapid transformation of network infrastructure resources for better and rapid advantages to healthy business across the globe.

References

1. Whitepaper EANTC Corporation EANTC-Metaswitch-Perimeta-SBC-Marketing Report http://www.eantc.de/fileadmin/eantc/downloads/News/2016 (Figure 3) (Table 1, 2, 3, 4)
2. K. Harper, A. Whitefield, *SBC FAQ A Ready Reference* (2018), pp. 111–113
3. M. Berg, *Managing Microsoft Hybrid Cloud* (2015), pp. 25–26 and 28
4. Whitepaper Genband WP_NFV_SBC_v10616.pdf http://www.exl-technologies.com (Figure 1,2) http://www.exl-technologies.com/wp-content/uploads/2017/08/WP_NFV_SBC_v10616.pdf
5. K. Inamdar, et al., *Understanding Session Border Controller CCIE Material CISCO (Figure 4)* (2018), pp. 3–50. CCIE Material for SBC Workflows
6. K. Mochizuki, H. Yamazaki, A. Misawa, Bandwidth guaranteed method of relocation of virtual machines, in *Paper Presented in IEEE Conference* (2013), https://ieeexplore.ieee.org/document/6665287/
7. Whitepaper 2018 Huaweicarrier.huawei.com/~/media/CNBG/Downloads/Product/IT (Figure 5) Huawei carrier services
8. Whitepaper 2018 SBC_v10616.pdf http://www.exl-technologies.com/wpcontent/uploads
9. S. Kaul, et al., Cloud PBX (private branch exchange) security: future enterprise communication, http://www.jetir.org/papers/JETIR1901335.pdf

Task Scheduling Based on Hybrid Algorithm for Cloud Computing



A. Vijaya Krishna, Somula Ramasubbareddy and K. Govinda

1 Introduction

Cloud computing has significantly evolved over the future. As a theory around 10 years ago it has become a practical approach to deploy our net resources and data into the internet. A virtualized resource pool and capability of large distributed paradigm, cloud computing has surpassed the other types of computing methods namely, grid computing, utility computing, web services, parallel, and distributed computing. With the inclusion of pay-per-use model and great variety of cloud services the cloud computing has also managed to cope up and create with green computing.

With low carbon future and economy, cloud computing's highly flexible and saleable resource pool management and terminal relaxation can cut costs of the user as well as data center and conserve significant amount of energy too. An application operation in Inter-Cloud has town main parameter time and cost. Tough the response time of a jib cannot be clearly estimated but this is primarily due to higher processing time than its original estimation. This is mainly because of delays occurring from the provider's side. So in order to counter this, most of the services are estimated with respect to time constraints, which helps in curbing the overtime head and achieving greater speed at which the job can be executed. The main goal here is to achieve the best thorough output and highest performance. According to classification of jobs in scheduling algorithm the main types are Batch mode heuristic algorithm and online mode heuristic. In BMHA, the jobs are queued into a line for a specific amount of time and then the scheduling algorithms are applied. The main examples are

A. Vijaya Krishna · K. Govinda
Scope, VIT University, Vellore, Tamilnadu, India

S. Ramasubbareddy (✉)
Information Technology, VNRVJIET, Hyderabad, Telangana, India
e-mail: svramasubbareddy1219@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_40

Round Robin, Shortest Job First, Min-Max algorithm, First Come First Served, and Priority Based algorithm. On the other hand, online heuristic scheduling algorithm schedules jobs as soon as they arrive in the system. But since the cloud environment is heterogeneous the online heuristic algorithms are more appropriate [1].

The Features of Task Scheduling In The Cloud Computing Environment
Virtualization is a technology provider and the solution to all the issues in cloud computing. The resource management and the scheduling of tasks are attained through this technology. Some of the other features of this are as follows:

The definition of job scheduling is to allocate jobs. The main concern is to allocate is to correct resources at a particular time. Thus, cloud computing has a lot of jobs to be scheduled and excited, this is the main issue coming with its many of the uses and to make cloud in the future the most conventional use of our resources. This is the only reason the job scheduling should be dynamic in nature [2]. The factors that need improvement are bandwidth, reduction in time completion, and memory. The efficient utilization of jobs in the VM is used to focus more on these issues [3]. The yielding of least response time to make sure that the execution of submitted jobs occurs at the lowest possible time and the validity of occurrence of intima reallocated resources. Some of the guidelines to be kept in mind while developing the algorithm.

2 Background

The proposed job shop scheduling algorithm that they have proposed is eligible for cloud environments. The name that they have given is PJSC. In the above section, we have already discussed the various terms and the conditions as well as the functions that come up with the requirement of the jib scheduling. And by the analysis we can say that the proposed algorithm needs more efficiency for the starvation of jobs problem. An appropriate job scheduling must be taken into account and to be developed. It is a necessity.

Mayank Mishra in his paper has stated the pay as you go model that is to be used when cloud computing is concerned. This model is kind of completely different from earlier infrastructure models, wherever enterprises would invest large amounts of money in building their own computing infrastructure. Also the wastage of resources during the non-peak periods can be taken into concentration. Thus, traditional data centers can be very costly and inefficient to use by small enterprises for their computational jobs.

Venkatesa Kumar. V and S. Palaniswami,'s paper has proposed the concepts of low utilization power and overall resource management and use. This paper satisfies all the above factors. It not only prefers the higher priority tasked to be utilized but also takes care of the low priority tasks to come into scheduling. Also, the unique turnaround utility programming approach is effective in this regard.

Zhongni Zheng, Rui Wang did the analysis of a genetic algorithm for scheduling programs in the cloud. The idea of using GA for optimization and sub-optimization in

cloud scheduling was highlighted in that paper. We tend to use and exploit the same. Unbalanced Assignment Problem is considered to be used as scheduling problem, mathematically. Future work can embody a lot of complete characterization of the constraints for programming in a cloud computing field and enhancements for the convergence with a lot of advanced issues.

Liang Luo et al. used CloudSim and Java programming to construct a new VM based Load Balancing Algorithm. The algorithm has been quite effective which made us use CloudSim to be implemented in this hybrid algorithm also.

3 Proposed Method

Job Scheduling is a process and technique of scheduling jobs that request the broker. The key components in this are the data centers, the virtual machines, and the user requests. The CPU utilization in this scenario is high. It provides high complexity with regard to optimization and low time complexity. In fact, it reduces the time complexity significantly. The existing techniques or the algorithms used have a disadvantage of starvation. This can be applied to First Come First Serve as well as the Shortest Job First. The FCFS uses the scene where the job that comes first is utilized and resolved while in SJF the job with shortest burst time is to be evaluated and processed. But this causes starvation and takes more time.

Priorit-based scheduling being used in this scene uses the jobs with less waiting time. However, this programing technique additionally suffers from the matter, where jobs are scheduled on the premise of priority of jobs in cloud, but can't perform once priority is greater also if burst time more that leads to a less computer processing utilization and takes a lot of waiting time as compared to the traditional execution.

This paper purposes a hybrid model of Shortest Job First and Priority scheduling to be incorporated in cloud computing for job scheduling.

Cloudlet selection inspired honey bee behavior algorithm.

The Fig. 1 represents the initial basic task scheduling. A number of servers are into play. The main point is it can accept or deny the job values from the customer. The jobs that are selected are to be queued into the scheduling pool. If the job is allotted the time the accept function is called or else, it is denied. The below mentions are individual methodologies of the scheduling algorithms along with hybrid one.

3.1 Shortest Job First

Initialize the cloud Simulation environment.

Set the number of VMs, cloudlets, datacenters along with users and resources to access.

Schedule the jobs according to resources with SJF algorithm.

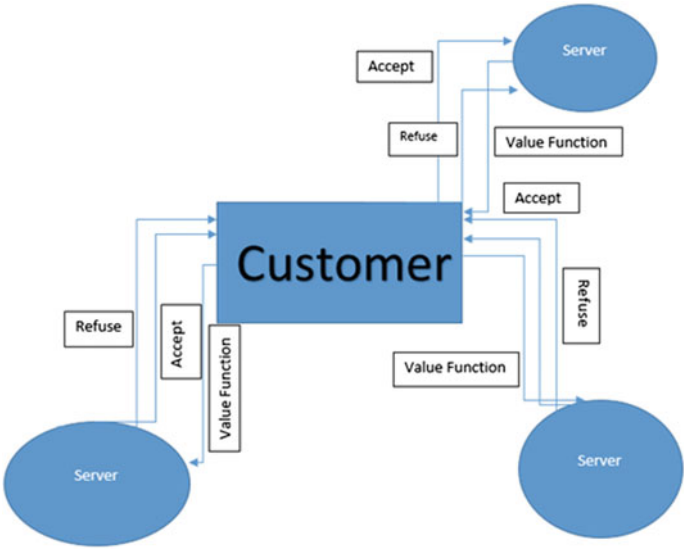


Fig. 1 Task scheduling model

3.2 Priority Based Scheduling

Initialize the cloud Simulation environment.

Set the number of VMs, cloudlets, datacenters along with users, and resources to access.

Schedule the jobs according to resources with a Priority algorithm.

3.3 Proposed Algorithm

The proposed hybrid constitutes of two algorithms of the shortest job first and priority-based scheduling. This is not restricted to a public cloud but can be also applied to hybrid Cloud.

1. Set N as number of requests to be sent from U users of the cloud.
2. Each User U as a specific burst time B. The burst time is sent to data centers DC through brokers.
3. Let T1,T2…..Tn be the time bursts of various users U and the requests of jobs J
4. Assign P1,P2…Pn to be the priorities to various jobs sent to data centers and brokers.
5. Check and find out the highest priority in the job pool.
6. Now find the burst time of the specific job.

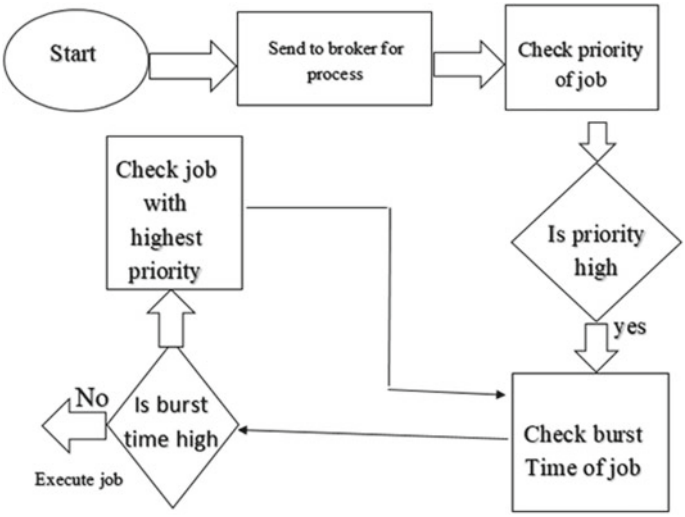


Fig. 2 Proposed architecture

7. If the burst time Ti is very less, execute the job or else the job with the shortest burst time T_i is executed (Fig. 2).

The equations are an exception to the prescribed specifications of this template. You will need to determine whether or not your equation should be typed using either the Times New Roman or the Symbol font (please no other font). To create multileveled equations, it may be necessary to treat the equation as a graphic and insert it into the text after your paper is styled.

4 Experimental Results

For the purpose of this algorithm, a Cloud Simulator called CloudSim [4]. CloudSim is a simulator tool providing virtual machines creation, data centers, and broker centers [5–7]. It also helps in simulation of cloud services and its infrastructure, the simulation is done with 40 Cloudlets (Figs. 3 and 4).

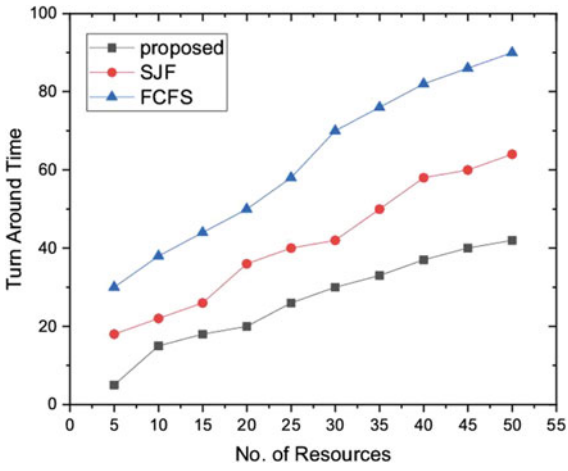


Fig. 3 Processing time versus no of resources

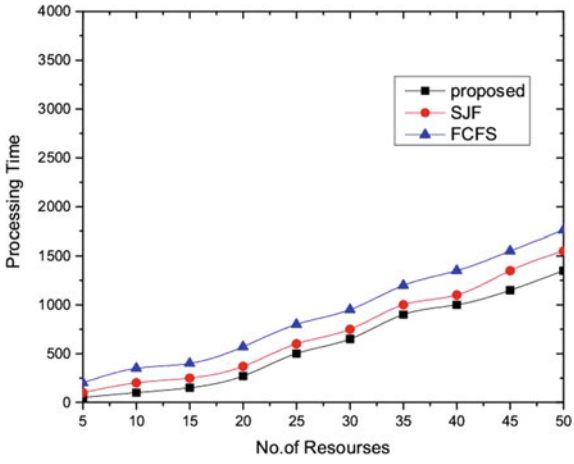


Fig. 4 Turn around time versus no of resources

5 Conclusion

It is observed that the proposed algorithm actually works in less time than the individual scheduling algorithm done separately. It provides high computation and less waiting time. The turnaround time of the jobs to be scheduled is also reduced significantly. The projected methodology enforced on the cloud simulation setting provides solely costs only 315.05 average response time in the simulation of cloud environment. On the other hand the cost of computation of running the simulation is only $0.57.Thus, to conclude the proposed algorithm does work in hybrid model and it can be effectively efficient to use in real time purposes. The main advantage is that the degree of starvation in this algorithm is significantly reduced and the jobs with greater priority get more preference. The preferred spelling of the word

“acknowledgment” in America is without an “e” after the “g.” Avoid the stilted expression “one of us (R. B. G.) thanks…”. Instead, try “R. B. G. thanks…”. Put sponsor acknowledgments in the unnumbered footnote on the first page.

References

1. Z. Zheng, R. Wang, H. Zhong, X. Zhang, *An Approach for Cloud Resource Scheduling Based on Parallel Genetic Algorithm*, 978-1-61284-840-2/11 (IEEE, 2011), pp. 444–447
2. V. Venkatesa Kumar, S. Palaniswami, A dynamic resource allocation method for parallel data processing in cloud computing, J. Comput. Sci. **8**(5), ISSN 1549–3636, Science Publications, pp. 780–788 (2012)
3. M. Mishra, A.0 Das, P. Kulkarni, A. Sahoo, *Dynamic Resource Management Using Virtual Machine Migrations*, 0163-6804/12, IEEE Communications Magazine (2012), pp. 34–40
4. Cloudsim.com/packages
5. TerrySimTutorials/youtube/SJF
6. Open Nebula. An open source tool kit for data center virtualization, http://opennebula.org/
7. Open Stack. Open source software for building private and public clouds, http://openstack.org/

An Integrated Approach for Botnet Detection and Prediction Using Honeynet and Socialnet Data



Mahesh Banerjee, Bhavna Agarwal and S. D. Samantaray

1 Introduction

Today, every organization around the globe gather, process, and store information on digital devices such as computers, smartphones, etc. A significant segment of that information can be sensitive and confidential data, breach of such data could have negative outcomes. Since the breach of confidential data has created a lot of negative impact in the recent past resulting in losses of billions of dollars, it has become a major area of concern and needs to be introspected. To counter such security threats such as malware, ransomware, botnets, etc. which cause such data breaches, we need to develop frameworks for the effective detection of the threats and also a framework for prevention of such threats.

1.1 Botnets

Botnets have become a menace to the Internet as the cybercriminals are not only using the botnets to conduct DOS attacks but are now using botnets as a mechanism for capturing information, distributing many other kinds of malicious software such as malware, rootkits, keyloggers, etc. A bot is a compromised machine whose resources

M. Banerjee (✉) · B. Agarwal · S. D. Samantaray
Department of Computer Engineering, GBPUAT, Pantnagar, India
e-mail: maheshbanerjee92@gmail.com

B. Agarwal
e-mail: bhavnaagarwal04@gmail.com

S. D. Samantaray
e-mail: sdsray@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_41

are exploited remotely by the attacker by the means of malicious codes. Botnets are generally detected using passive monitoring techniques such as honeypots.

Following are the most common operations of botnets:

Scanning: The attacker scans networks for vulnerable and unpatched systems.
Spreading: The infected machine searches for other machines that can be compromised and infects them to make them a part of the botnet.
Sign-On: The infected machine connects to the CnC channel for receiving further instructions.

1.2 Honeypot

Honeypots are generally computer system running unpatched version of operating systems making them an easy target for the attacker, they act like a legitimate system. It stores the attack-related information such as the IP and port of the attacker, the vulnerabilities exploited by the attacker and finally the post-exploitation action taken by the attacker. Post-exploit events are stored by the honeypot in the form of system event logs for which a dedicated service is executed embedded in the kernel of honeypot. It also captures packets and payloads involved in the attack. This information is important for analyzing the attacker's activities. A honeynet is a network of honeypots coupled with other tools such as a honeywall, SNORT, etc.

We have proposed a new and integrated approach for botnet detection using the data obtained from honeynets and social networks. Our work focuses on proactive botnet detection based on the honeynet data. For implementation, local honeynets have been deployed. The attack data is preprocessed for feature extraction. For each attack location, time components such as country, state, city, date, day, and time information are derived and added to the feature set. Further, for a given duration and location, the socialnet is explored and key information, events, activities, and news details are extracted. These are appended as new attributes, thus producing an augmented Transaction database. The data set obtained from the honeynets is used for the detection of botnets by means of correlation and similarity while association rule mining techniques help in predicting a botnet attack.

2 Related Work

BotHunter is an intrusion detection system developed by Gu et al. [1]. It uses the snort-IDS rules for detecting the botnets. The BotHunter scans the network, captures the payload, and does analysis on the payload to detect common malware intrusions by correlating the payload traffic with Snort rules and triggers an alarm for any anomaly behavior detected. To evade detections, botmaster uses encrypted traffic so that the

BotHunter cannot detect the bot. Analyzing the payload is too costly as payloads are heavier and contradict the principle of user privacy. BotSniffer developed by Gu et al. [2] uses anomaly detection algorithms for building custom ruleset inside Snort-IDS for detecting botnets. BotSniffer is deployed on a real-world network and was able to detect botnets with high accuracy and low false positives. BotMiner developed by Gu et al. [3] uses two plane clustering methods for detecting botnets, namely, A-Plane and C-plane. The C-plane logs traffic flows and A-plane detects suspicious activities. The information obtained from the two planes is used to detect the botnets by the means of cross-plane correlation. Li et al. [4] have proposed a honeynet-based botnet scan traffic analysis method where they developed a general technique which has a minimum monitoring overhead for observing botnet behavior, and hard to evade by botnets. They have aggregated the measurements for getting an accurate picture of the botnets. Analysis of the behavioral list is done and it is found that the botnet scanning behavior is grained to the botnet because this is the most effective way for them to recruit new bots. Moreover, monitoring scanning is relatively easy. With a honeynet installed, people can easily get the botnet scanning traffic. With this motivation, they designed a framework to extract botnet-related scanning events and analyzing methods. Moon et al. [5] proposed a method for detection of botnets before activation through an enhanced honeypot system for intentional infection and behavioral observation of malware. They proposed a system consisting of two major modules:

Module 1 is the malware collection module.
Module 2 is the module for analyzing behaviors of the collected malware and for identifying suspicious malware for botnets.

Karasaridis et al. [6] proposed a wide-scale botnet detection and characterization method. An anomaly-based passive analysis algorithm has been used to detect IRC botnet controllers achieving less than 2% false positive rate. The algorithm is able to detect IRC botnet controllers running on any random port without the need for known signatures or captured binaries analysis.

3 Methodology

For our research, we have used the captured data from our local honeynet. The captured data is a network traffic dump containing all traces of the malicious traffic generated by various cyber threats. The honeynet provides us data in three formats.

- Malicious network traffic dump (PCAP files),
- Binaries of the malware (binary files), and
- Log files.

	A	B	C	D	E	F	G	H
1	Date	Time	Source_IP	Source_port	Target_IP	Target_port	Country	
2	1/22/2019	14:40.0	213.21.56.81	48634	14.139.229.44	80	Russia	
3	1/22/2019	34:23.0	90.150.137.41	59217	14.139.229.44	23	Russia	

Fig. 1 Data set attributes

date	time	source_ip	source_port	target_ip	target_por	country	state	city	longitude	lattitude
1/22/2019	14:40.0	213.21.56.81	48634	14.139.229.44	80	Russia	Saint Peter	Saint Peter	59.9321	30.1968
1/22/2019	34:23.0	90.150.137.41	59217	14.139.229.44	23	Russia	Sverdlovsk	Reftinskiy	57.0925	61.680801

Fig. 2 Sample of the spatial–temporal database

3.1 Phase-I

In this, the log files are extracted which are stored in the honeynet in a folder named system logs. It is a text file with no formatting with a (.txt) format. The text file is converted to a (.csv) file using a Python module that we have developed with the use of some open-source Python libraries such as the tabula Python library. From the CSV file, the attribute which is no longer required is dropped such as the tools or the agents the attacker used since they are currently not being used for the analysis. The resulting CSV file contains attributes such as (Fig. 1).

The CSV file is then imported to the database for further feature extraction. The source IP field is used to extract the location details such as the city, state, and coordinates corresponding to the IP (using the GeoIP tool). The location details are thus appended to the file. This whole operation is done in an automated fashion and the resulting file is then used to create a spatial–temporal database with attributes such as date and time of the attack, origin of the attack, i.e., source IP and port, country, state, city, longitude, and latitude and also target IP and port which can used to analyze the spatial–temporal details of the attack data which can be used for detecting and analyzing new types of threats (Figs. 2 and 3).

3.2 Phase-II

The spatial–temporal database obtained in Phase-I can be used for the detection and prediction of a botnet. Similarity and correlation are generally used to detect botnet traffic. The proposed approach filters the botnet traffic on the basis of protocol matching. Once the traffic is filtered the transaction of the traffic is grouped on the basis of the source IP, port, and then is checked if they belong to any pre-established CnC channel. The traffic is analyzed for checking the communication patterns such as most of the communication do not exhibit a high measure of synchronization as the

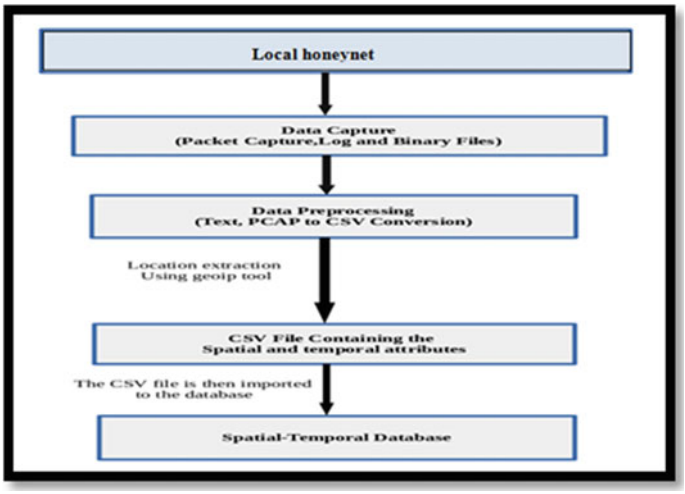


Fig. 3 Schematic of the workflow

communication of the bot's exhibit. Two types of checks are used for differentiating the communication patterns.

Response-Crowd-Density-Check

The network traffic is analyzed to filter the client activity/message responses. Then the density of the responses is checked and the clients who are having frequent activity are added to the crowd. Hence, forming a dense crowd for this a variable Y_i is used for the density check, i.e., $Y_i = 0$ and $Y_i = 1$ where the values 1, 0 define whether the crowd is dense or not, respectively. The presence of a botnet is defined by calculating probability on the basis of the hypotheses. H_1 being the hypothesis for "botnet" and H_2 being the hypothesis for "benign or not botnet". Finally, the probabilities are calculated as $Pr(Y_i|H_0) = \theta_0$ and $Pr(Y_i|H_1) = \theta_1$ where θ_1 represents a botnet and θ_0 represents benign.

Response-Crowd-Homogeneity-Check

The response crowds are checked on the basis of the similarity of the responses of the clients because here also bots are distinguishable from the rest since they respond to the botmaster or the CnC channel in a pre-programed way. In this proposed approach, DICE distance is used to find the homogeneity of the crowd for this n-gram analysis used and then DICE distance is used to calculate DICE coefficient which ratio of n-grams shared by strings to the total number of n-grams in both strings. The n-gram analysis uses a sliding window with a span of n characters to extract sub-strings from a string being checked. For given string X with length l, the number of n-gram will be N where $N = l - n + 1$.

$$Dice(X, Y) = \frac{2|ngrams(X) \cap ngrams(Y)|}{|ngrams(X)| + |ngrams(Y)|} \quad (1)$$

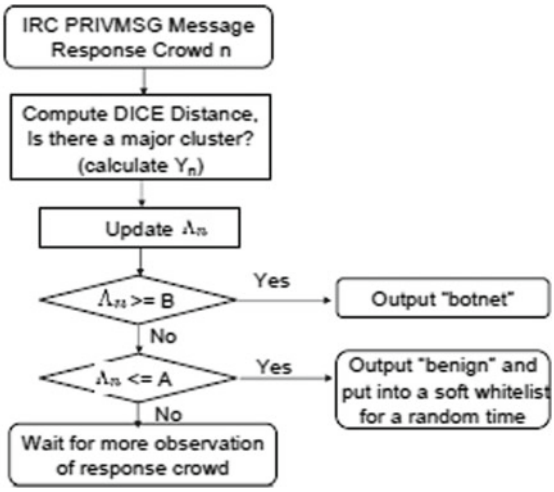


Fig. 4 IRC message correlation

To check the homogeneity of a crowd Y_i, first, clustering methods are used to find the largest cluster of similar messages in the crowd and calculate the ratio of the size of the cluster to the size of the crowd. A crowd is said to be homogeneous if the ratio is greater than a given threshold and not homogeneous if it is less than the given threshold. The similarity of the messages is calculated for every client in the crowd, every unique pair is compared using DICE distance greater than a given threshold, i.e., the similarity of the messages in percentage (Fig. 4).

4 Results and Discussion

Phase-I: The spatial–temporal database provides us with many interesting and useful attack patterns that can be used further analysis for finding various attack vectors. The results presented in this section are based on the attack captured from November 23, 2018 to January 23, 2019 from our local honeynets (Fig. 5).

Phase-II: The network data so obtained from the honeynets is used to monitor the intrusion from malicious sources and detecting the botnets using the network traffic alert logs and corresponding tcpdumps (Figs. 6 and 7).

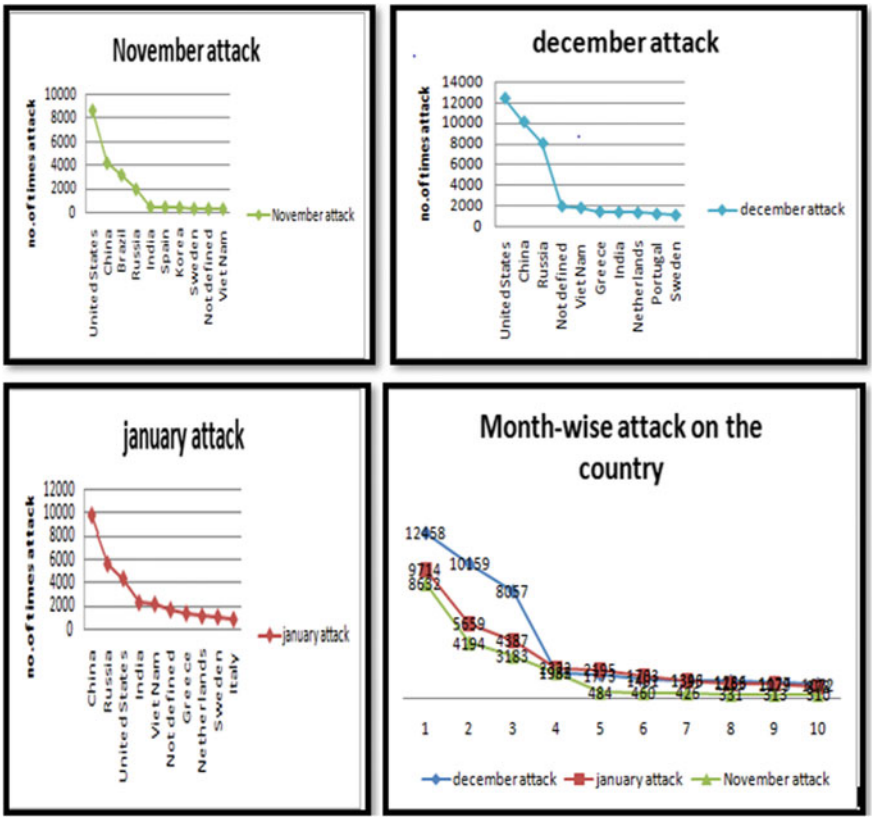


Fig. 5 Observations from the data

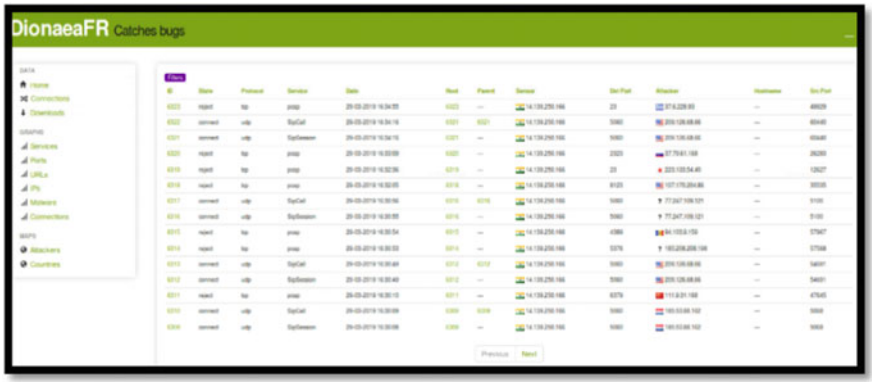


Fig. 6 DionaeaFR dashboard

Bin ID	Binary MD5	Captured Date	AV Name	AV Classification	Malware Family
5500	c752208ee42ad60dd54152cd413ddabd	2018-09-02 08:21:51.0	Qihoo-360	Win32/Backdoor.6dc	Backdoor
5501	84c73696edd32db9951485825 6b75155	2018-09-02 08:22:09.0	Rising	Backdoor.Overie!1.648D (CLOUD)	Backdoor
5587	6d5e0392145c87dd0b12c1207d9dcf6f	2018-09-05 05:07:49.0	McAfee	BackDoor-EXZ	Backdoor
5780	b34b2a65bf00f533ab33c725a50bbcf2	2018-09-08 02:58:17.0	Microsoft	Backdoor:Win32/Zegost	Backdoor
5809	812a848e6d13527c871fc3048aef97b5	2018-09-11 15:35:55.0	Rising	Backdoor.Overie!1.648D (CLOUD)	Backdoor
21	770d0caa24d9644ea7c04ff5daf290f08	2018-08-10 11:56:40.0	AhnLab-V3	HackTool/Win32.Shadowbrokers.C1925198	Botnet
22	649b368c52de83e52474a20ce4f83425	2018-08-10 11:56:40.0	AhnLab-V3	HackTool/Win32.Shadowbrokers.C1925362	Botnet
23	4803a7863da607333378b773b8a1714c	2018-08-10 11:56:40.0	ESET-NOD32	Win32/HackTool.Equation.AA	Botnet
24	43aac72a9602ef53c57696f04e1be7386	2018-08-10 11:56:40.0	Malwarebytes	HackTool.Agent	Botnet
25	f01f09fe60d08f10c44dce4e94785227	2018-08-10 11:56:40.0	ESET-NOD32	a variant of Win32/HackTool.Equation.AB	Botnet
26	9a5cec05e9c158cbc51cdc972693363d	2018-08-16 11:54:26.0	ESET-NOD32	a variant of Win32/HackTool.Equation.AI	Botnet
28	5e8ecdc3e70e2ecb0893cbda2c189061	2018-08-16 11:54:26.0	ESET-NOD32	Win32/HackTool.Equation.AO	Botnet
29	8c80dd97c37525927c1e549cb59bcbf3	2018-08-10 11:56:40.0	Fortinet	HackTool.SHADOWBROKERS!tr	Botnet
31	01d5adbfee39c5807ee46f7990ff5fda7	2018-08-16 11:54:26.0	AhnLab-V3	HackTool/Win32.Shadowbrokers.C1921149	Botnet
32	d1aae806243cc0bedb83a22919a3a660	2018-08-16 11:54:26.0	AhnLab-V3	HackTool/Win32.Shadowbrokers.C1921173	Botnet
33	1fa609bc0d252ca0915d6aed2df7ccc2	2018-08-16 11:54:26.0	AhnLab-V3	HackTool/Win32.Shadowbrokers.C1921219	Botnet
34	6b7276e4aa7a1e50735d2f6923b40de4	2018-08-16 11:54:26.0	AhnLab-V3	HackTool/Win32.Shadowbrokers.C1921174	Botnet
38	e4ad4df4e412405 87b4fe8bcb32d815	2018-08-16 11:54:26.0	Ikarus	HackTool.Shadowbrokers	Botnet

Fig. 7 Some of the identified botnets

5 Conclusion and Future Works

Our work is focused on proactive botnet detection using the honeynet data. For implementation, we deployed two honeypots at Pantnagar, namely, the CDAC CTMS Honeynet and PantHoneynet. And the data has been collected for a period of 3 months from November 23, 2018 to January 23, 2019. For each attack location, time components such as country, state, city, date, day, and time information are derived and added to the feature set. The data set obtained from the honeynets is used for the detection of botnets by the means of correlation and similarity while association rules mining techniques help in predicting a botnet attack. We have used data from PantHoneynet for experimentation that is used to detect the botnets and the result is verified using the data provided by the CDAC CTMS Honeynet. Our implementation detects botnets with low false positives while operating on the Internet.

For future advancement of this work, social network information can also be included for better analysis and prediction of botnets. Further, for a given duration and location, the socialnet can be explored and key information, events, activities, and news details are extracted. These can be appended as new attributes, thus producing an augmented transaction database which can provide a diverse data set for the prediction of botnet attacks.

References

1. G. Gu, P. Porras, V. Yegneswaran, M. Fong, W. Lee, BotHunter: Detecting Malware Infection Through IDS-Driven Dialog Correlation (2007)
2. G. Gu, J. Zhang, W. Lee, BotSniffer: Detecting Botnet Command and Control Channels in network traffic (2008)

3. G. Gu, Ro. Perdisct, J. Zhang, W. Lee, BotMiner: Clustering Analysis of Network Traffic for Protocol-and Structure-Independent Botnet Detection (2008)
4. Z. Li, A. Goyal, Y. Chen, Honeynet-Based Botnet Scan Traffic Analysis (2008)
5. Y.H. Moon, E. Kim, S.M. Hur, H.K. Kim, Detection of botnets before activation: an enhanced honeypot system for intentional infection and behavioral observation of malware (2012)
6. A. Karasaridis, B. Rexroad, D. Hoeflin, Wide-Scale Botnet Detection and Characterization (2016)
7. K. Anagnostakis, S. Sidiroglou, P. Akritidis, K. Xinidis, E. Markatos, A. Keromytis, Detecting targeted attacks using shadow honeypots, in *Proceedings of 14th USENIX Security Symposium*, August (2005)
8. M. Bailey, E. Cooke, D. Watson, F. Jahanian, N. Provos, A hybrid honeypot architecture for scalable network monitoring. Technical Report CSE-TR-499-04, U. Michigan, October (2004)
9. D. Dagon, X. Qin, G. Gu, W. Lee, J. Grizzard, J. Levin, H. Owen, Honeystat: local worm detection using honeypots, in *Proceedings of the 7th International Symposium on Recent Advances in Intrusion Detection (RAID)* (2004)
10. X. Jiang, D. Xu, Collapsar: a vm-based architecture for network attack detention center, in *Proceedings of 13th USENIX Security Symposium*, August (2004)
11. N. Provos, A virtual honeypot framework, in *Proceedings of 13th USENIX Security Symposium*, August (2004)

Teaching–Learning-Based Functional Link Artificial Neural Network for Short-Term Electrical Load Forecasting



Rudra Narayan Pandey, Sarat Mishra and Sudhansu Kumar Mishra

1 Introduction

Load forecasting has always been a zone of concentrate for scholarly research as well as for business applications because of its significant role in the viable and financial task of intensity utilities [1]. Load forecasting has dependably been a challenging task as the data collected may not be the genuine data, and may change abruptly. These vulnerabilities in the data are due to different reasons, such as the sudden breakdown of the power plant, natural calamity, erroneous instrument for noting down the load consumption reading, human errors, etc. Building a good model for forecasting load is a challenging task, and it might function admirably with certain power systems while they fail for some other systems. Most of these models are unpredictable, exceptionally nonlinear, and relies upon climate, occasional, and social elements [2, 3]. Over the past decades, numerous techniques have been proposed for load forecasting, such as statistical, regression, artificial intelligence [4], Artificial Neural Network (ANN) [5], etc. Recently, a variant of ANN has been proposed by Pao, which eliminates the need of hidden layers, thus making it simple and computationally inexpensive, and since then it is being extensively used in load forecasting problems [6].

R. N. Pandey · S. Mishra · S. K. Mishra (✉)
Department of Electrical and Electronics Engineering, Birla Institute of Technology, Mesra, Ranchi, India
e-mail: sudhansumishra@bitmesra.ac.in

R. N. Pandey
e-mail: rudrakumar118@gmail.com

S. Mishra
e-mail: mishra.sarat@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_42

A great deal of research has gone into the advancement of ANN models based on derivative-based training algorithm such as Least Mean Squares (LMS), Normalize Least Mean Squares (N-LMS), Backpropagation (BP), Recursive Least Squares (RLS), etc. However, these derivative-based methodologies may trap in neighborhood optimal point. To avoid the bottleneck of these techniques, a range of intelligent soft computing techniques such as Genetic Algorithms (GA), Teaching–Learning-Based Optimization (TLBO) [7, 8], JAYA [9], and Particle Swarm Optimization (PSO) [10] have been introduced over the last two decades. Numerous researchers have used GA and PSO techniques for load forecasting, in [11]. The PSO is considered for optimizing the network for transient load forecasting in [12]. Similarly, modified TLBO is used for unit commitment issue in [13]. In [14], the authors compared the results of different statistical models, such as ARMA, ARIMA with the artificial intelligence based models by implementing in forecasting of load, wind power and electricity prices.

In this paper, TLBO is utilized to update the weights of the FLANN model for short-term load forecasting. FLANN model is chosen because of its simpler yet effective design, and TLBO optimization techniques are considered because of its dependencies on fewer numbers of parameters. The data to be forecasted are collected from BIT, Mesra, Ranchi substation.

2 Structure of the Artificial Neural Network Filters

In this section, different variants of artificial neural networks are explained briefly and the basic idea of different functional expansions is also illustrated.

2.1 Multilayer Perceptron

MLP is a parallel distributed computing network for solving different nonlinear problems. MLP with one hidden layer is considered here. At first, the weighted sum of the input layer is calculated by using Eq. (1); further, this weighted sum is made to penetrate through a nonlinear activation function which represents the output of the hidden layer (Eq. (2)). Similarly, the same steps are followed for obtaining the outputs of the output layer. The Mean Square Error (MSE) is calculated by applying Eq. (5). By employing the BP algorithm, the weights of the MLP is updated

$$h = \sum x * w^T \tag{1}$$

$$y_h = \varphi(h) \tag{2}$$

$$h1 = \sum yh * wh^T \tag{3}$$

$$yp = \varphi(h1) \tag{4}$$

$$MSE = \frac{1 * (\sum_{i=0}^{N} (Yp_i - Yd_i)^2}{N} \tag{5}$$

where the total number of samples is represented by N, MSE represents mean square error, predicted output is represented as Y_p, and Y_d is taken as the desired output.

$$w_{\text{new}} = w_{\text{old}} + \mu \frac{dMSE}{dw} \tag{6}$$

where μ represents the learning rate.

2.2 Functional Link ANN

The FLANN is a single-layer ANN structure as represented in Fig. 1. It is capable of generating nonlinear decision boundaries for performing complex decision regions. In the FLANN model, the input layer is expanded into a higher dimension using some functional expansions, such as algebraic polynomial expansion, trigonometric

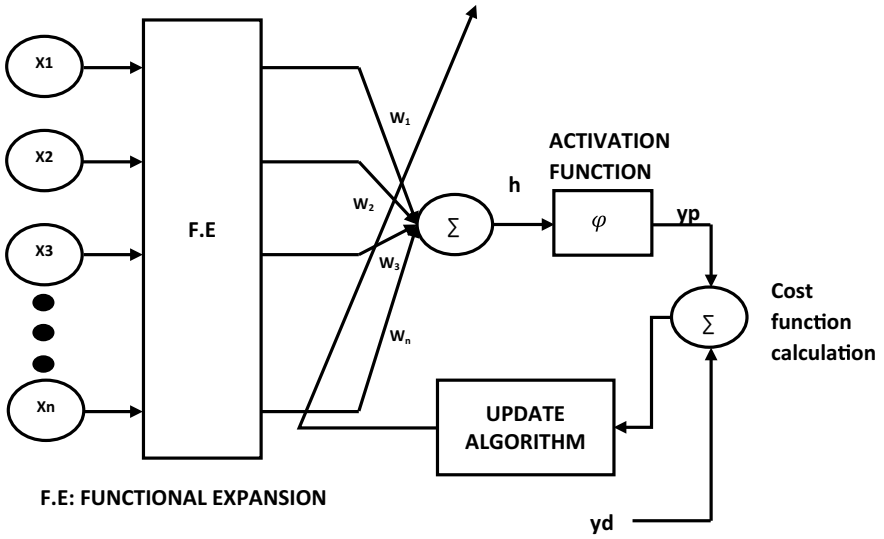


Fig. 1 A FLANN structure

expansion, Chebyshev expansion, etc. This process eliminates the need for hidden layers making it easier to compute as compared to MLP. The FLANN model is trained by first calculating the weighted sum of the expanded input by applying Eq. (8), and then it is made to penetrate through a nonlinear activation function as shown in Eq. (9), which represents the output layer. The weights of the network are updated by using error between the desired pattern and the output layer node. The MSE is considered as a cost function that needs to minimize.

$$X = F(x) \tag{7}$$

$$h = \sum X * w^T \tag{8}$$

$$y_o = \varphi(h) \tag{9}$$

where *F()* represents an expansion function, X is the enhanced input, x is the original input, and φ is the nonlinear activation function.

2.3 Different Functional Expansions

Here, any of the underneath mentioned expansion techniques can be utilized by functional expansion block. For instance, the expanded input sample X^l by considering a two-dimensional input sample $X = [x1x2]^T$ is obtained by using algebraic polynomial expansion,

$$X^l = \left[x1\, x2\, x1^2\, x2^2\, x1 * x2\right]^T \tag{10}$$

If the input pattern is $X = [x1\, x2\, \mathrm{x3}]^T$, then the enhanced pattern obtained will be

$$X^l = \left[x1\, x2\, x3\, x1^2\, x2^2\, x3^2 x1 * x2\, x1 * x3\, x2 * x3\right]^T \tag{11}$$

By using trigonometric expansion

$$X^l = \left[x1\cos\left(\Pi x1'\right)\sin(\Pi x1')\ldots\, x2\cos(\Pi x2')\sin(\Pi x2')\ldots\, x1\, x2\right]^T \tag{12}$$

Similarly, by using exponential expansion, it will be

$$X^l = [x1\exp^{x1}\exp^{2x1} x2\exp^{x2}\exp^{2x2}]^T \tag{13}$$

Correspondingly, by using Chebyshev polynomial expansion, the higher Chebyshev polynomials for 1 < x may be generated by applying Eq. (14)

$$T_{n+1} = 2 * x * T_n(x) - T_{n-1}(x) \tag{14}$$

The first few Chebyshev polynomials are given by

$$\left.\begin{aligned} T_0(x) &= 1 \\ T_1(x) &= x \\ T_2(x) &= 2 * x^2 - 1 \end{aligned}\right\} \tag{15}$$

3 Different Optimization Techniques

Optimization techniques are used to minimize the cost function of the model. Optimization techniques can be categorized into two categories, i.e., derivative-based (LMS, BP, etc.) and derivative-free optimization (PSO, GA, TLBO, etc.). The main drawback of derivative-based optimization techniques is that they get stuck in local minima problem. To vanquish the shortcomings of derivative-based optimization techniques, numerous researchers have come with derivative-free optimization techniques.

All the optimization techniques have some pros and cons. In this section, different optimization techniques which are used in this paper for comparative analysis are elucidated in brief.

3.1 Least Mean Square (LMS)

The Least Mean Squares (LMS) algorithm is the simplest of all the other optimization techniques. This algorithm is mostly used for the single-layer ANN. The FLANN is a single-layer ANN model, and hence this simple algorithm can be applied for updating the weights. It is worth comparing FLANN-LMS model with other optimization techniques because of its simple structure and less computational complexity. The weight updating formula for the LMS algorithm is as follows:

$$w_{\text{new}} = w_{\text{old}} + 2 * \mu * x * error \tag{16}$$

where μ is the learning rate that controls the convergence speed. A small μ leads to slow convergence speed, and it may occur that within the prescribed maximum epochs model fails to converge, a high μ leads to large misadjustments, thus an adaptive learning rate (μ) is chosen [6]. In LMS model, weights are adjusted by using Eq. (16).

3.2 Particle Swarm Optimization (PSO)

The Particle Swarm Optimization (PSO) is a heuristic methodology, first proposed by Kennedy and Eberhart in 1995 [10] as a swarm intelligence algorithm, developed for dealing with the optimization of continuous and discontinuous functions. The PSO algorithm is based on the biological and sociological conduct of animals, such as schools of fish and flocks of birds searching for their nourishment. The PSO copies this conduct by creating a population with an arbitrary search solution, and each potential solution is represented as a particle in a population (called a swarm.). Each particle is flown through the multidimensional pursuit space with arbitrary and versatile velocity so as to discover the lower function values (global minimum).

The updating of the position of the particle is done by the following equation:

$$V_i(n+1) = wt * V_i(n) + cp * rd1 * (pbest_i - X_i(n)) + sp * rd2 * (globest - X_i(n)) \tag{17}$$

$$X_i(n+1) = X_i(n) + V_i(n+1) \tag{18}$$

where X_i and V_i are the position and velocity of ith particle, respectively, wt is inertia weight, cp represents cognitive parameter, sp is taken as a social parameter, $rd1$ and $rd2$ are a random number between 0 and 1, $pbest_i$ is the particle best of ith particle, and $globest$ is the global best of entire particle.

3.3 JAYA

JAYA is a heuristic-based algorithm which is introduced in [9]. This is a population-based optimization technique. The algorithm always endeavors to get closer to success (i.e., achieving the best arrangement) and endeavors to avoid failure (i.e., moving away from the most exceedingly terrible arrangement). The algorithm strives to become victorious by reaching the best solution, and hence it is named as Jaya (a Sanskrit word meaning victory) [9].

$$z = w_i^n + rd * \left(best_{value} - abs\left(w_i^n\right)\right) - rd * \left(worst_{value} - abs\left(w_i^n\right)\right) \tag{19}$$

$$w_i^{n+1} = z \tag{20}$$

where rd is the random number between 0 and 1, the weight which has the lowest cost work esteem (i.e., cost function value) is represented as $best_{value}$ and the weight which has the highest cost work esteem (i.e., cost function value) is represented as $worst_{value}$.

In JAYA, at first, random set of weights are generated, and after that $best_{value}$ and $worst_{value}$ are chosen depending upon the cost function value, subsequently using the

obtained $best_{value}$ and $worst_{value}$ weights are updated by using Eq. (19) and if the new weights give effective result than the existing one then the new weight is considered for the next iteration as shown in Eq. (20), else weights are left unchanged.

3.4 Teaching–Learning-Based Optimization (TLBO)

Teaching–Learning-Based Optimization (TLBO) is a heuristic algorithm which is presented in [7]. This algorithm has been effectively applied in various engineering fields, e.g., mechanical engineering, to solve the optimization problem [7]. It has been additionally utilized to unravel substantial scale non-direct improvement issues [8]. This method is based on the population.

$$z = w_i^n + rd * mea * (teacher_{value} - round(1 + rd)) \tag{21}$$

$$w_i^n = z \tag{22}$$

$$z1 = w_i^n + rd * \left(w_i^n - w_j^n\right) \tag{23}$$

$$w_i^{n+1} = z1 \tag{24}$$

where rd is the random number between 0 and 1, mea is the mean value of each column of the populations, and $teacher_{value}$ represents the best students who have the lowest cost function value.

The population is fundamentally the same as the population of the class. Surely, every student in the class can improve their insight dependent on two steps: the initial step involves gaining from the educator and improving the knowledge, and this progression is defined as teacher phase. The second step includes learning by trading information with companions, and this step is defined as the learner phase. In this study, TLBO is utilized for transient load forecasting. In TLBO, at first, random set of weights are generated which constitute the population of the class. The mean value of the population is calculated, and after that brightest student of the class is chosen as the teacher for the next iteration, i.e., that weight is chosen as $teacher_{value}$ which gives the lowest cost function value (MSE). Further, by utilizing the chosen $teacher_{value}$ weights are updated by applying Eq. (21), and if the new weights give a better result than the existing one, then the new weight is considered for the next iteration as shown in Eq. (22), else weights are left unchanged.

After teacher phase, any random weight is chosen, and subsequently weights are adjusted by applying Eq. (23). If the new weights give a better result than the existing one, then the new weight is considered for the next iteration as shown in Eq. (24), else weights are left unchanged, and with this learner phase comes to an end. This process repeats as long as the desired response is not obtained.

4 Proposed Technique

In this paper, comparative analysis between TLBO and other optimization techniques for one day ahead load forecasting is carried out. Figure 2 illustrates how TLBO optimization technique is applied for minimizing the cost function. While using TLBO optimization technique, in spite of using the modification equation mentioned in [13], a new modification equation is used. All the FLANN models are trained with

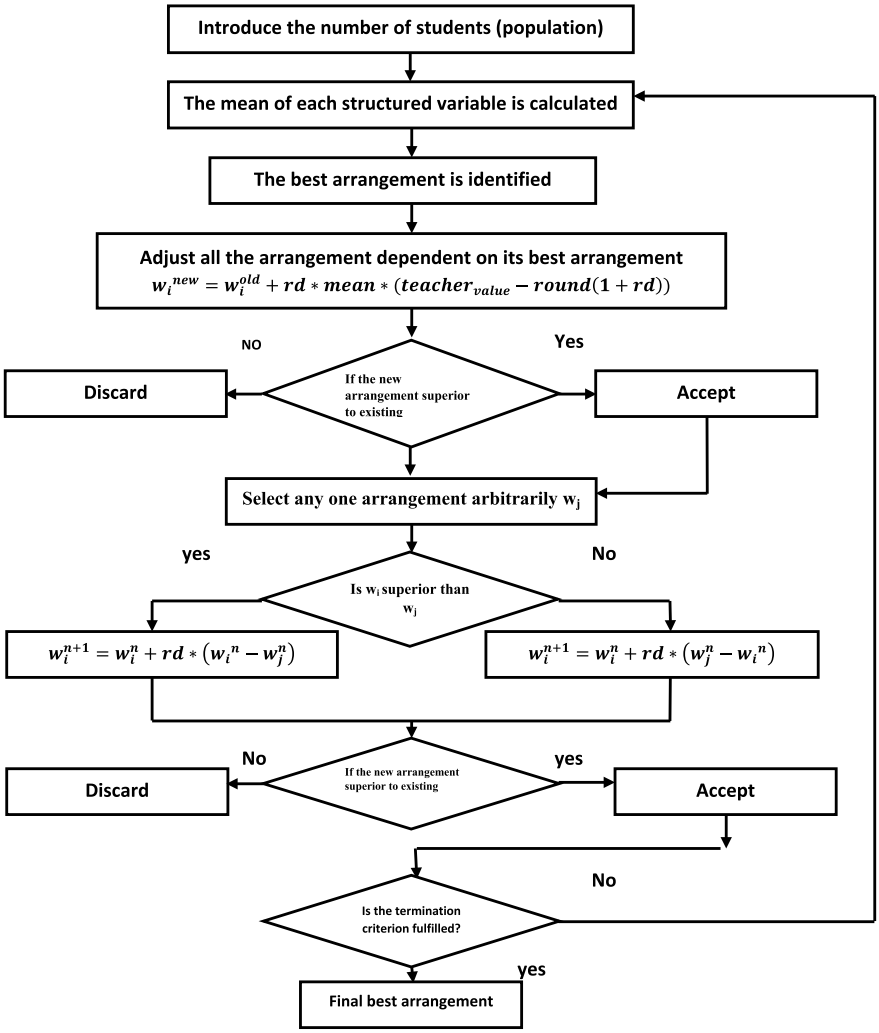


Fig. 2 Algorithm flowchart for TLBO

the load consumption value of first, second, and third day and predicting the load consumption value of the fourth day.

5 Simulation Study

The present study explains the supremacy of the TLBO optimization technique in short-term load forecasting as compared to the other cutting-edge methods. To highlight the supremacy of the TLBO over others, various simulations have been carried out on the data obtained from the SLDC, Jharkhand, and different optimization techniques, such as TLBO, JAYA, PSO, and LMS with the same number of epochs and activation function have been considered. Data obtained from the SLDC, Jharkhand is first cleaved into two sets, i.e., train set and test set in the ratio 8.5:1.5. FLANN model with various optimization techniques is trained on the train set, and afterward the trained model is tested utilizing the test set. The comparative study is performed in terms of an objective analysis using Root Mean Square Error (RMSE) metric.

5.1 Performance Indices

The formulation for Root Mean Square Error (RMSE) is calculated by applying Eq. (25)

$$\text{RMSE} = \left(\frac{1 * (\sum_{i=0}^{N} (Yp_i - Yd_i)^2}{N} \right)^{1/2} \tag{25}$$

where N represents the total number of tests, predicted output is represented as yp, and Y_d is taken as the desired output. Low RMSE score indicates high accuracy, and thus the model which has the lowest RMSE score has the highest accuracy.

5.2 Simulation Result

Specification of the system on which tests were executed is Lenovo G500s having Windows 10 operating system, Intel® Core™ i3-3120 CPU @ 2.50 GHz processor, RAM of 4 GB, and in MATLAB platform. One day ahead load forecasting taking the load consumption of day 1, 2, and 3 for training the FLANN model and predicting the load consumption of fourth day, using epochs of 30, activation function of tanh, and expansion of algebraic polynomial is applied. It is observed from Table 1 that TLBO

optimization technique outperforms all the other optimization techniques taken into consideration in almost all the test runs.

The Friedman statistical test is additionally performed by using the MATLAB prebuild function (Friedman) to confirm the impact and repeatability of the obtained outcomes. Table 2 exhibits the average ranking of the different tested optimization techniques dependent on Friedman's test; lower the ranking, higher is the accuracy and performance, and from Table 4, which is obtained by using MATLAB prebuild function Friedman and passing display out="on" as one of the parameters, it is clear that the critical value obtained from the Friedman test is 1.29E−10. As the critical value under $\alpha = 0.05$, it signifies that the null hypothesis can be rejected. The result obtained by running the Friedman test reveals the supremacy of the TLBO over the other (Tables 3 and 4).

For a pairwise comparison of the TLBO optimization technique with any other optimization techniques under consideration, the nonparametric statistical tests, such as the sign test and the Wilcoxon signed test are illustrated. The critical number of wins needed to achieve both levels of significance is shown in Table 6.

Table 1 Comparative results in RMSE of various optimization techniques for 20 random test runs

Sr. no.	RMSE			
	JAYA	TLBO	PSO	LMS
1	1.527349	1.458615	1.494501	5.965024
2	1.470038	1.478917	1.699533	7.534004
3	1.464111	1.489194	1.544837	8.171045
4	1.472327	1.463365	1.655971	9.434481
5	1.67629	1.479367	1.706719	8.125089
6	1.536734	1.538177	1.612444	8.028031
7	1.594185	1.530984	1.54415	8.801587
8	1.659558	1.438783	1.679895	7.973747
9	1.588524	1.501106	1.842998	8.009808
10	1.446379	1.50571	1.618987	6.88606
11	1.510583	1.438243	1.744695	7.899063
12	1.444617	1.484451	1.558766	6.497116
13	1.51325	1.462623	1.869691	2.864388
14	1.536237	1.479741	1.636468	7.501838
15	1.550368	1.452381	1.593119	6.92254
16	1.476808	1.553244	1.501512	8.079399
17	1.462223	1.428074	1.594569	8.465191
18	1.576989	1.496393	1.57122	3.892762
19	1.695496	1.424899	1.749084	8.271448
20	1.480499	1.455011	1.608788	5.448685

Table 2 Ranks obtained according to Friedman test

Methods	JAYA	TLBO	PSO	LMS
Mean ranks	9.25	6.75	14	20

Table 3 Comparison of computational time

CPU time (s)	TLBO	JAYA	PSO	LMS
	77.63	34.64	13.63	3.84

Table 4 Friedman test values

Source	Sum of square (SS)	Degree of freedom (DoF)	Mean square (MS)	Chi-square	Critical value (p)
Columns	81.7	3	27.2333	49.02	1.29E–10
Error	18.3	57	0.3211		
Total	100	79			

Table 5 Critical values obtained for the two-tailed sign tests at $\alpha = 0.05$ and $\alpha = 0.01$ utilizing RMSE metric as a triumphant parameter

TLBO	JAYA	PSO	LMS
Wins (+)	14	19	20
Loss (−)	6	1	0
Detected difference	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.05$

From Table 5, it can be concluded that TLBO beats all the other techniques as in every pairwise comparison, the number of wins obtained is greater than the critical value, which proves its supremacy (Table 6).

The Wilcoxon signed test is performed using MATLAB prebuilt function signrank, and the h-value obtained by using this function specifies whether the null hypothesis can be rejected or not. Null hypothesis depicts that the data under comparison is the same and there is no significant difference between them, and h-value of 1 signifies that the null hypothesis can be rejected. So, the h-value obtained from the Wilcoxon test supports the supremacy of the TLBO over the other as shown in Table 7.

Table 6 Minimum wins needed for the two-tailed sign test at $\alpha = 0.05$ and $\alpha = 0.01$

No. of cases	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
$\alpha = 0.05$	5	6	7	7	8	9	9	10	10	11	12	12	13	13	14	15	15	16	17	18	18
$\alpha = 0.01$	5	6	6	7	7	8	9	9	10	10	11	12	12	13	13	14	14	15	16	16	17

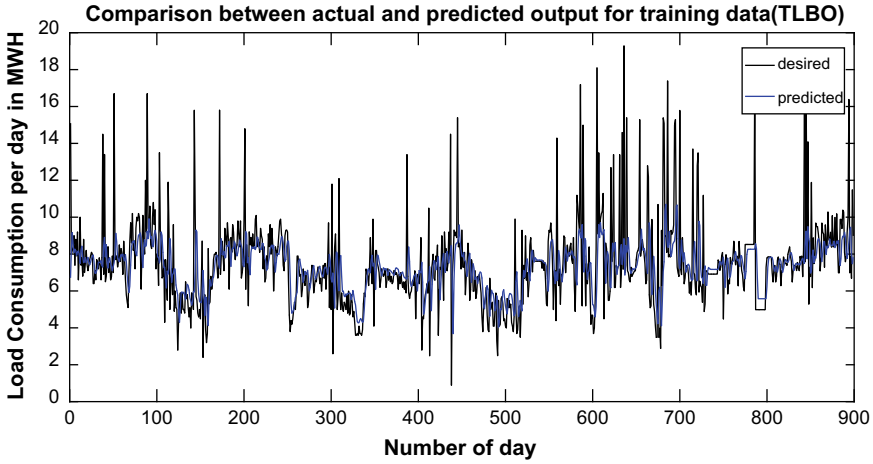
Table 7 Wilcoxon signed test utilizing RMSE metric as a triumphant parameter

Comparison	p-value	h-value
TLBO with JAYA	0.012374	1
TLBO with PSO	0.00014	1
TLBO with LMS	8.86E−05	1

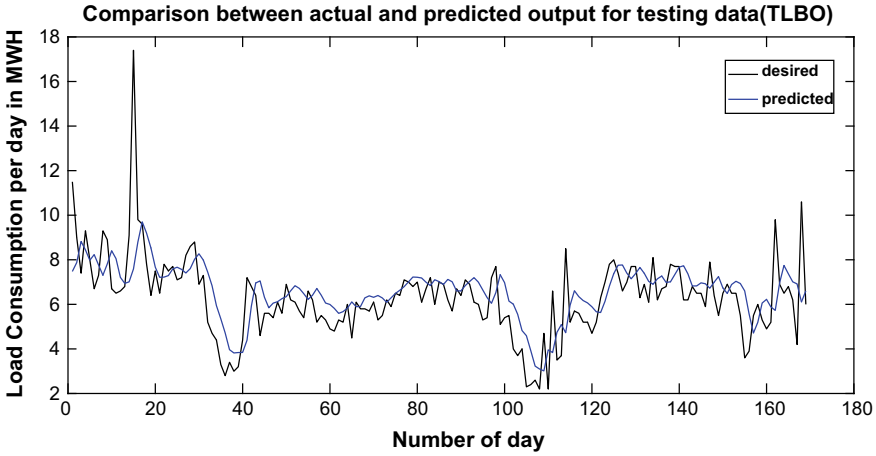


Fig. 3 Comparison of RMSE value of different optimization techniques in various test runs

The visual performance measure can also be carried out by observing the plots. Load consumption per day versus the number of day plot is plotted for all the optimization techniques taken into consideration. For every optimization technique, two plots are plotted one on the train set containing 900 data and another on the test set containing 170 data. On close visual inspection of Figs. 1, 2, 3, and 4, it can be concluded that TLBO performs better as compared to other optimization techniques for short-term load forecasting (Figs. 5, 6 and 7).

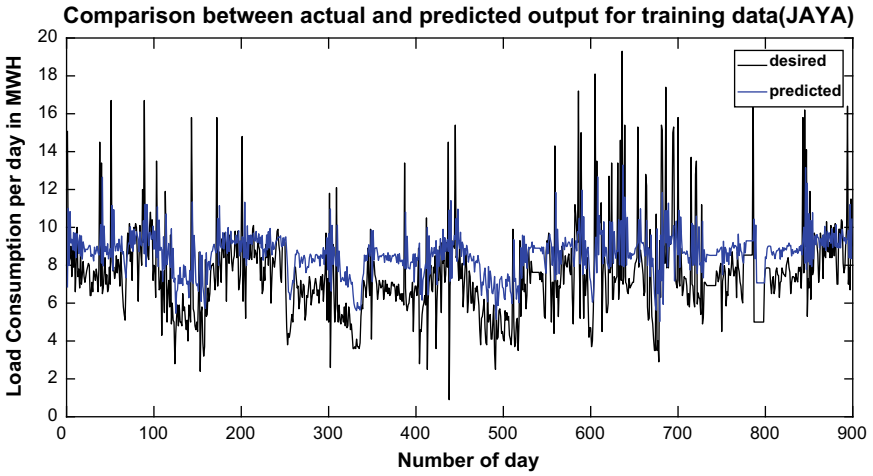


(a) Prediction Plot on training set

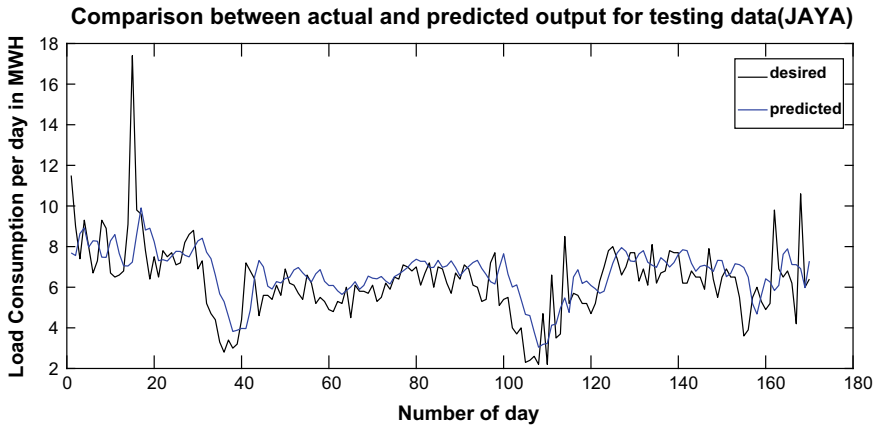


(b) Prediction Plot on the testing set

Fig. 4 One day ahead prediction plot of FLANN-TLBO with the tanh activation function

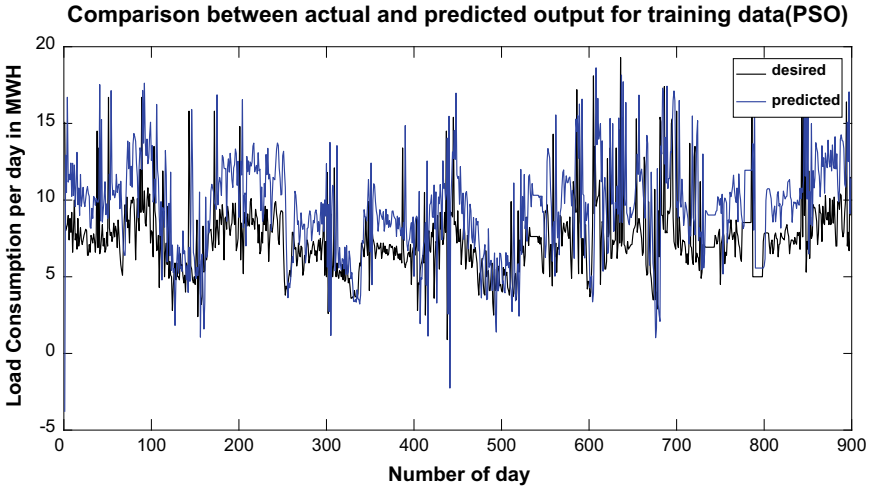


(a) Prediction Plot on training set

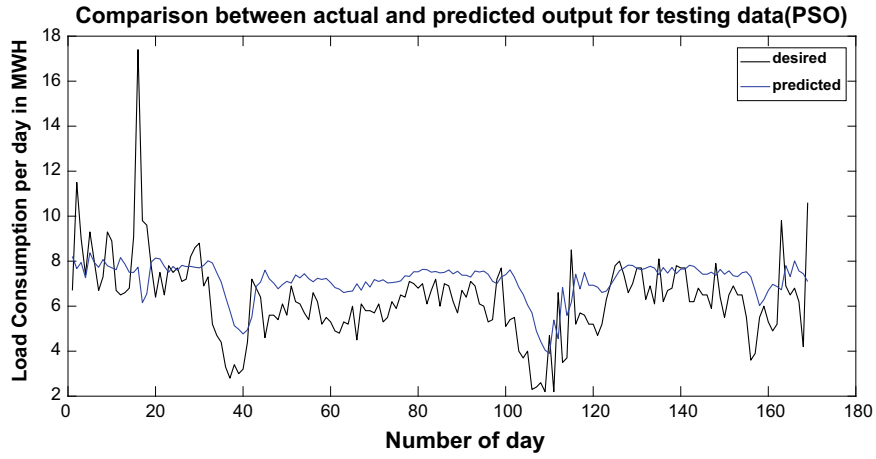


(b) Prediction Plot on the testing set

Fig. 5 One day ahead prediction plot of FLANN-JAYA with the tanh activation function

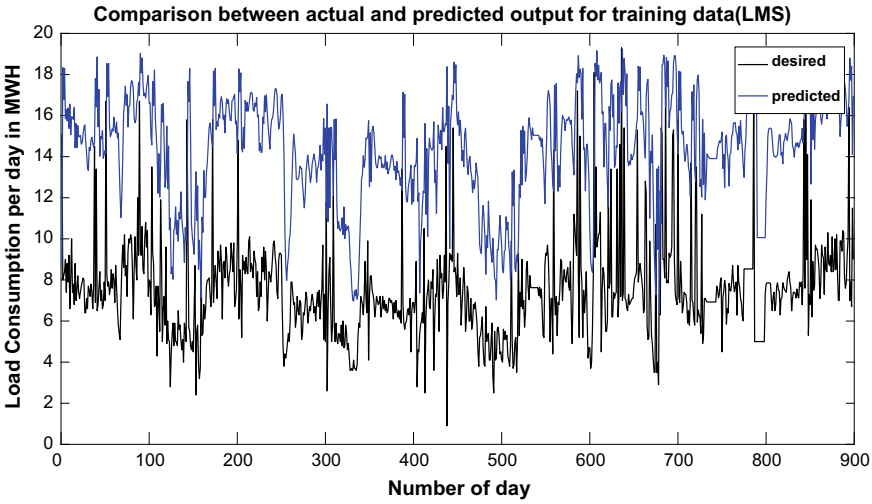


(a) Prediction Plot on training set

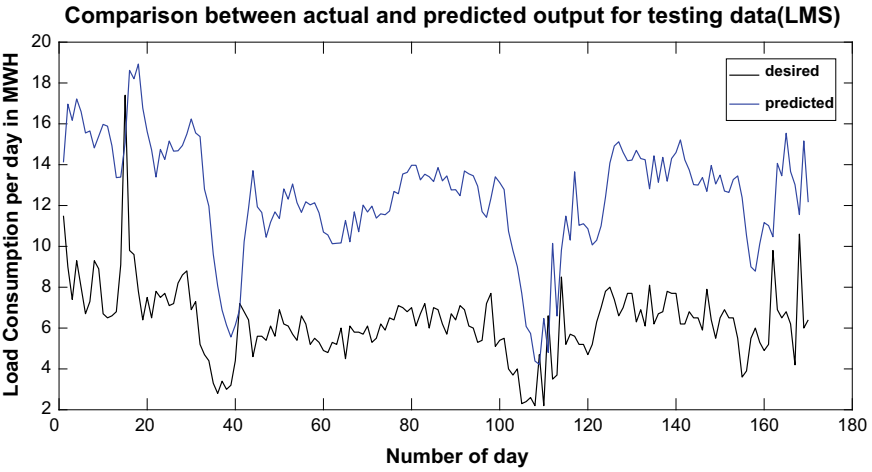


(b) Prediction Plot on the testing set

Fig. 6 One day ahead prediction plot of FLANN-PSO with the tanh activation function



(a) Prediction Plot on training set



(b) Prediction Plot on the testing set

Fig. 7 One day ahead prediction plot of FLANN-LMS with the tanh activation function

6 Conclusion

In this paper, the FLANN model with TLBO optimization technique is introduced for one day ahead load forecasting. The detailed comparison between TLBO and the other optimization techniques, such as JAYA, PSO, and LMS, has been carried out in this work. All the tests are conducted on the obtained data from the SLDC, Jharkhand for the duration of January 2016 to March 2019. The data is further cleaved into the train set and the test set. All the models are prepared utilizing the train set and evaluated using the same test set. The RMSE metric is used as a performance metric. After testing the FLANN-TLBO model for 20 times, it can be concluded that FLANN-TLBO model is superior to the other cutting-edge models. To test the impact and repeatability of the obtained results, the nonparametric statistical tests, such as the Friedman test, Sign test, Wilcoxon test, are additionally performed. The plot between the actual load consumption and the predicted load consumption by the different models is also plotted for visual inspection. The proposed technique can also be used for time series prediction problems, such as stock price prediction, weather forecasting, etc.

References

1. S. Tzafestas, E. Tzafestas, Computational intelligence techniques for short-term electric load forecasting. J. Intell. Rob. Syst. **31**, 7–68 (2001)
2. E. Kyriakides, M. Polycarpou, Short term electric load forecasting: a tutorial, in *Trends in Neural Computation*, ed. by K. Chen, L. Wang. Studies in Computational Intelligence, vol. 35 (Springer, Berlin, Heidelberg, 2007)
3. Dudek G., Artificial immune system for short-term electric load forecasting, in *Artificial Intelligence and Soft Computing—ICAISC 2008*, ed. by L. Rutkowski, R. Tadeusiewicz, L.A. Zadeh, J.M. Zurada. ICAISC 2008. Lecture Notes in Computer Science, vol. 5097 (Springer, Berlin, Heidelberg, 2008)
4. E.A. Feinberg, D. Genethliou, Load forecasting, in *Applied Mathematics for Restructured Electric Power Systems*, ed. by J.H. Chow, F.F. Wu, J. Momoh. Power Electronics and Power Systems (Springer, Boston)
5. H.S. Hippert, C.E. Pedreira, R.C. Souza, Neural networks for short-term load forecasting: a review and evaluation. IEEE Trans. Power Syst. **16**(1), 44–55 (2001)
6. R. Majhi, G. Panda, G. Sahoo, Development and performance evaluation of FLANN based model for forecasting of stock markets. Expert. Syst. Appl. **36**, 6800–6808 (2009)
7. R.V. Rao, V.J. Savsani, D.P. Vakharia, Teaching–learning-based optimization: a novel method for constrained mechanical design optimization problems. Comput.-Aided Des. **43**, 303–315 (2011)
8. R.V. Rao, V.J. Savsani, D.P. Vakharia, Teaching–learning-based optimization: an optimization method for continuous non-linear large scale problems. Inf. Sci. **183**, 1–15 (2012)
9. Rao, R., Jaya: a simple and new optimization algorithm for solving constrained and unconstrained optimization problems. Int. J. Ind. Eng. Comput. **7**(1), 19–34 (2016)
10. Kennedy, R.C. Eberhart, Particle swarm optimization, in *Proceedings of the IEEE International Conference on Neural Networks*, vol. 4 (1995), pp. 1942–1948
11. A. Baliyan, K. Gaurav, S.K. Mishra, A review of short term load forecasting using artificial neural network models. Procedia Comput. Sci. **48**, 121–125 (2015)

12. N. Zeng, H. Zhang, W. Liu, J. Liang, F.E. Alsaadi, A switching delayed PSO optimized extreme learning machine for short-term load forecasting. Neurocomputing **240**, 175–182 (2017)
13. P. Khazaei, M. Dabbaghjamanesh, A. Kalantarzadeh, H. Mousavi, Applying the modified TLBO algorithm to solve the unit commitment problem, in *World Automation Congress (WAC)*, Rio Grande (2016), pp. 1–6
14. A. Muñoz, E.F. Sánchez-Úbeda, A. Cruz, J. Marín, Short-term forecasting in power systems: a guided tour, in *Handbook of Power Systems II*, ed. by S. Rebennack, P. Pardalos, M. Pereira, N. Iliadis. Energy Systems (Springer, Berlin, Heidelberg, 2010)

An Enhanced K-Means MSOINN Based Clustering Over Neo4j with an Application to Weather Analysis



K. Lavanya, Rani Kashyap, S. Anjana and Sumaiya Thasneen

1 Introduction

Rapid climatic behavior entangles the normal livelihood of a living being. The evaluation of level risk associated with the climatic changes needs to be done in order to repudiate the effect of prevalent condition.

The national disaster management authority of India released a natural disaster list, which clearly states that the rate of natural calamities is alarmingly increasing year by year [1].

In the year 2014, heavy torrential rainfall in Kashmir led to a massive flood which caused the death of around 500 peoples. The Uttarakhand Flash flood in 2014 due to huge and heavy cloudbursts caused floods in river Ganga. The death toll was around 5700 [2]. Likewise, there are so many other events that endanger the life of the living being that put mankind in a thinkable situation about the prevailing causes and prevention mechanism.

Clustering is the technique, which can be used for different types of predictions, which helps in weather analysis. The clustering algorithm creates a cluster with different data values. Incremental K-means is one of the popular clustering algorithms,

K. Lavanya (✉) · R. Kashyap · S. Anjana · S. Thasneen
School of Computer Science and Engineering,
Vellore Institute of Technology, Vellore, Tamil Nadu, India
e-mail: lavanya.k@vit.ac.in

R. Kashyap
e-mail: rani.kashyap2017@vitstudent.ac.in

S. Anjana
e-mail: anjana.s2017@vitstudent.ac.in

S. Thasneen
e-mail: isumaiyathasneen@vit.ac.in

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_43

which can be used to create the clusters. It creates the cluster by calculating the distance between each data object and all in the center after each iteration.

Processing and representing data are major challenges in the current computing environment because data grows exponentially day by day. Social networking, online shopping, and online storage facilities are the major contributors to the mass data. These connected datasets can be represented and accessed more efficiently if they are represented in the graph. Graph database contains a set of nodes which are the entities connected with each other using edges which represent the relationship between the nodes. Graph databases are more suitable when the data is complex and connected. Neo4j, one of the most popular graph databases, is used to store the dataset. The scalability and simplicity of neo4j make it more suitable for data processing and representation. The cipher query language used in neo4j is easy to understand and implement which gives the user optimized storage and access. This index-free database helps the user to access data by different inbuilt functions as well as by traversals. Neo4j creates nodes for each entity in the dataset with appropriate attributes given in the weather report.

The K-means clustering algorithm uses Euclidean's distance formula which calculates the distance between each data object. The squared Euclidean distance of two n-dimensional vectors is determined as

$$\|x - y\|2 = (x1 - y1)2 + (x2 - y2)2 + (x3 - y3)2 + \cdots + (x\mathrm{n} - y\mathrm{n})2 \quad \text{where } x = [x1, \ldots, x\mathrm{n}] \text{ and } y = [y1, \ldots, y\mathrm{n}] \tag{1}$$

After calculating the distance, the numbers of clusters are selected and the number of centroids is assigned for each cluster. The centroid is referred to as a center for each cluster. The efficiency of this method is not that much high because of static selection of clusters [3].

In this paper, we are going to use K-means MSOINN based on the SOINN algorithm. The MSOINN distance measure calculates the distance between two values for a defined attribute, the number of distinct values for every categorical attribute and the amount of every value in the dataset is calculated [4]. Once the distance is calculated between the data objects, the K-means algorithm is used for assigning the data objects into different clusters by finding the closest centroid using the neo4j database.

2 Literature Survey

Weather analysis has a major impact on various day-to-day activities. This section briefs about other techniques used in weather analysis. Neo4j is used in multiple domains as it outperforms other NoSQL databases in terms of accuracy and response time. Petri Kivikangas and Mitsuru Ishizuka describe how Neo4j helps in improving semantic queries. The connected nature of the database helps in representing the semantic relationships in a better way [5].

The query performance of the neo4j graph database depends on data size, query complexity, query numbers, and the cost of development or maintenance. Sometimes the performance depends on the application scenario [6].

Rohit Kumar Kaliyar says about the evolution of the graph database and their significance in the current computing environment. It also essays the details of the current graph databases and their comparisons which give the idea of current scenarios happening in data modeling using graphs [7].

Roberto De Virgilio introduces a new methodology to convert rdf to a graph database. Networking is enabled to form a large amount of connected datasets. Graph databases are an effective and efficient solution to represent these datasets. Hence, many complex and connected datasets were converted to graph database to improve performance [8].

Kosei Ueta et al. look into the issues such as distribution and security in implementing graph database for data management in IoT. Data models of IoT systems use graph because these models can specify relationships in a straightforward manner among entities such as devices, users, and information that constructs IoT systems [9].

Xianlong Lv et al. propose the traditional technique for data storage and manages in power network data system is very worst. With the development of technology, it is getting more complex. For that, a new reliable solution is introduced. The GRAPH-CIM model which is based on Hadoop architecture and process of big data energy Internet is proposed. This paper gives an application development method for power network topology processing, which will make data robust and parallel usable [10].

G. Drakopoulos explains about the PubMed which is the largest online open-source database which includes document related to life science and biomedical research under the authorization of NIH. It contains near 40 million abstracts so far. For that reason, there is a need for a perfect method which can execute document analysis. For that, text-mining methodology introduced with PubMed which executes traditional document-term matrix representation, an architecture for content-based retrieval, whose core is a document-term author third-order tender. This paper implements text mining with Python and neo4j [11].

G. Drakopoulos proposes harmonic centrality with the structural ranking presentation, which is executed easily sort time. A directed graph ranking is necessary for vertices to get a result. And if the graph contains a large number of vertices and nodes then it is a big challenge to design algorithm and implement for vertices ranking. In the Twitter network overcome on network oblivious Tensor fusion methodology is implemented with Neo4j and Tensor toolbox [12].

Since neo4j is strongly based on nodes and relations, it is best suitable for Social Networking. As Social Networking is booming all around the world, Asham Virk and Rinkle Rani explain how Neo4j is efficient approach for social recommendations [13].

3 Methodology

Clustering analysis is a prominent exploration concern in data mining expected its diversity of accomplishment. The aim of our paper is to analyze the weather changes by processing the weather report of Delhi in 2016 using neo4j. Weather dataset has three weather-type samples, i.e., Fog, Clear, and Smoke. The K-means MSOINN algorithm has been applied to the dataset to form the clusters.

3.1 The K-Means MSOINN Clustering Algorithm

K-means MSOINN clustering algorithm forms the cluster based on the MSOINN distance measure. It considers a dataset with some sample values and each sample has data for n properties, and each sample can be considered as a real n-dimensional vector. The MSOINN distance measure calculates the distance between two values for a defined attribute, the number of distinct values for every categorical attribute, and the amount of every value in the dataset is calculated. Once the distance is calculated between the data objects, the K-means algorithm is used for assigning the data objects into different clusters by finding the closest centroid. The centroid is referred to as a center for each cluster. Centers are computed for each cluster by selecting k trial from dataset.

3.2 Distance Measure

The categorical distance of two instances I_i and $I_j \in$ D is given by Eq. (2)

$$\text{dist}_{\text{cat}}(I_i, I_j) = \sum [\delta(I_i^{a_k}, I_j^{a_k}) \times w^{a_k}\left(I_i^{a_k}, I_j^{a_k}\right)]^2 \tag{2}$$

where a_k is the attribute in dataset D exploiting supervised and unsupervised information.

The numeric distance of two instances I_i and I_j is given by Eq. (3)

$$\text{dist}_{\text{num}}(I_i, I_j) = \sum_{\substack{a \\ k}\in \text{NUM}} (I_i^{a_k} - I_j^{a_k})^2 \tag{3}$$

Finally, the mixed distance of two instances I_i and I_j is calculated by Eq. (4)

$$\text{dist}_{\text{mix}}(I_i, I_j) = \sqrt{\text{dist}_{\text{num}}(I_i, I_j) + \text{dist}_{\text{cat}}(I_i, I_j)} \tag{4}$$

3.3 Cluster Assignment Step

The sample value w is allocated to the cluster respective to their centroids, i.e., the centroids closest to w. It is allocated to one of the centroid set.

$$c_i : \left\{ \sqrt{\sum (c_i - w)^2} = \min_i \sqrt{\sum (c_j - w)^2} \right\} \tag{5}$$

3.4 Centroid Update Step

In this step, the average of all allocated samples is calculated and then the new centroid value is determined when all the samples are allocated to a centroid. So for allocated sample value $w_1, \ldots, w_n$, the vector for the new centroid c_i will be

$$c_i = \frac{1}{n} \sum_{i=1}^{n} w^i \tag{6}$$

The above two Eqs. (5) and (6) are used to iterate until the algorithm coincides. So best practice would be to repeat the algorithm many times with distinct allocation of the centroids and select the best among them.

4 Implementation

4.1 The Dataset

For clustering, we selected weather dataset of Delhi from Kaggle. The dataset consists of 416 samples of weather type, i.e., Clear, Fog, and Smoke with 53 clear, 167 fog, and 196 smoke values for each of them. The dataset has seven attributes such as dew, humidity, pressure, temperature, visibility, wind_direction, and wind_speed. It can be considered as a seven-dimensional vector. The labeling of the dataset is done to identify the weather type. This is done to see the results different from original labeling.

4.2 Implementation in Neo4j

Two different kinds of nodes have been created. The first node is weather and another one is centroids. Both of them contain the seven attribute values from the dataset. Weather nodes have been labeled as 1, 2, and 3 for identification of the weather types. Centroids have attribute "index" which provides information of the cluster and "repetition" attribute which tells about the centroid's iteration. The centroids can be selected by picking up the weather data from one of the labels.

The nodes can be created easily. The identifier "label" is added before the colon to identify the type of node. The attribute values of the nodes can be specified after the colon.

```
CREATE(:weather{dew:9,humidity:82,pressure:1012,temperatu
re:12,visibility:2,wind_direction:20,wind_speed:5.6,label
:1})
CREATE(:Centroid{
dew:6,humidity:93,pressure:1015,temperature:7,visibility:
1,wind_direction:0,wind_speed:0, index:1, repetition:1})
```

Figure 1 shows the number of weather nodes created of each type by using CREATE feature.

Cluster Assignment. The feature which we have used is MATCH which can be used for matching the nodes, paths, and relationships. For calculating the distances for each of the three centroids, we need all the nodes with an identifier as the weather. After calculating the distance, we will find the closest centroid.

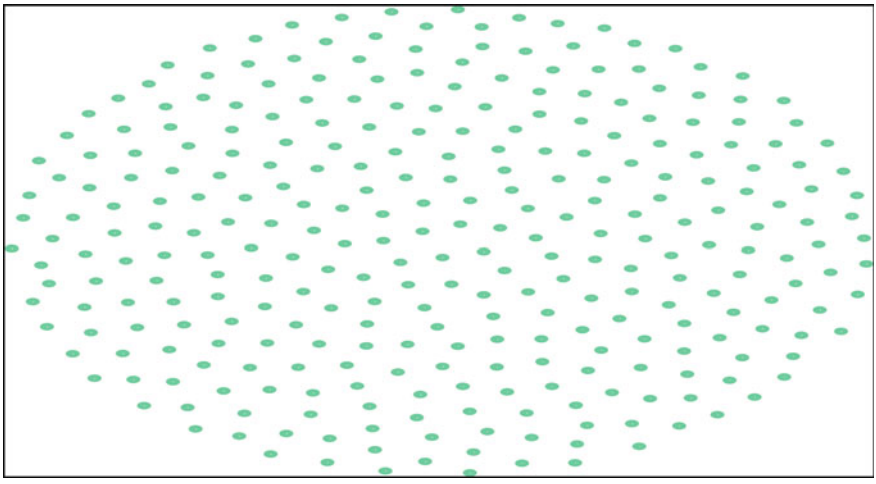


Fig. 1 The number of nodes for each weather data type

The next feature used is SET which adds new attributes distanceC1, distanceC2, and distanceC3 to each weather, containing the distance from weather to the centroid. WITH clause is used for binding the variables which can be used for matching later. Each weather node is kept as w and nearest centroid as minC, and lastly we create IN_CLUSTER relationship from weather w to the centroid minC. Succeeding the CREATE one could have bound up the weather and removed the other three distance attributes.

1. Select the weather node and centroid {C1, C2, and C3}.
2. Calculate and set the w.distance for C1,C2,C3 by using MSOINN:

```
SET w.distanceC1 = sqrt(w.dew-c1.dew)^2+
(w.humidity-c1.humidity)^2 +(w.pressure- c1.pressure)^2
+(w.temperature-c1.temperature)^2+(w.visibility-
c1.visibility)^2 + (w.wind_direction-c1.wind_direction)^2
+ (w.wind_speed-c1.wind_speed)^2)
```

3. Categorize the node using distanceC1, distanceC2, distanceC3 value.
4. Apply when and else condition.
5. End up the minC.
6. Create relation between node and centroid named as [in_cluster].
7. Return the centroid.

```
{return*}
```

Figure 2 shows the three clusters formed in the first iteration based on the specified centroid value.

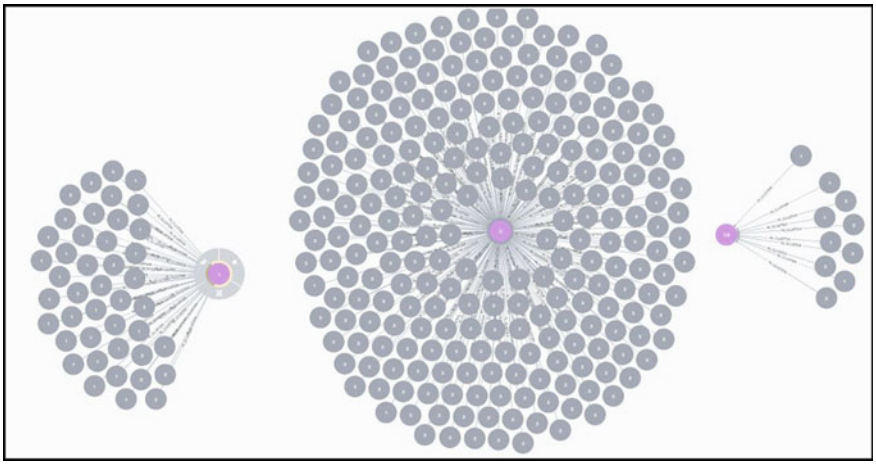


Fig. 2 The clusters formed in the first iteration

Centroid Update. The next step is to find the new centroids for the next allocation step, by calculating the mean of the seven attributes for the weather allocated to each cluster. If the MATCH statement is like MATCH (w1:Weather)-[:IN_CLUSTER]-> (c1:Centroid {index: 1, repetition: 1}), (w2:Weather)-[:IN_CLUSTER]-> (c2:Centroid {index: 2, repetition: 1}), (w3:Weather)-[:IN_CLUSTER]-> (c3:Centroid {index: 3, repetition: 1}), we would get the Cartesian product over w1, w2, and w3.

This will return the correct numbers when using the average function as computing the average over multiple copies gives back the original value.

Calculate the node again with their average values and create the new centroid accordingly using the following steps:

1. Selecting all the nodes property from previous centroid.
2. Calculate the average of the nodes using repetition: 1.

```
MATCH (w2:weather)-[:IN_CLUSTER]->(c2:Centroid {index: 2,
repetition: 1})
WITH w1Dew, w1Humidity, w1Pressure, w1Temperature,
w1Visibility, w1 Wind_direction, w1Wind_speed,
avg(w2.dew) as w2Dew, avg(w2.humidity) as w2Humidity,
avg(w2.pressure) as w2Pressure, avg(w2.temperature) as
w2Temperature, avg(w2.visibility) as w2Visibility,
avg(w2.wind_direction) as w2 Wind_direction,
avg(w2.wind_speed) as w2Wind_speed
```

3. New properties centroid creating (repetition 2).

```
CREATE (:Centroid {dew: w2Dew, humidity: w2Humidity,
pressure: w2Pressure, temperature: w2Temperature, visi-
bility: w2Visibility, wind_direction: w2 Wind_direction,
wind_speed: w2Wind_speed, index: 2, repetition: 2})
```

4. Return to dataset.

```
{return*}
```

Next Cluster Assignment. The same statement as the last does the next cluster assignment, only with the iteration number in the centroids increased to two, and some weather is moved to a new cluster.

Calculate the node again with their average values and create the new centroid accordingly using the following steps:

1. Selecting all the nodes property from previous centroid.
2. Calculate the average of the nodes using repetition: 1.

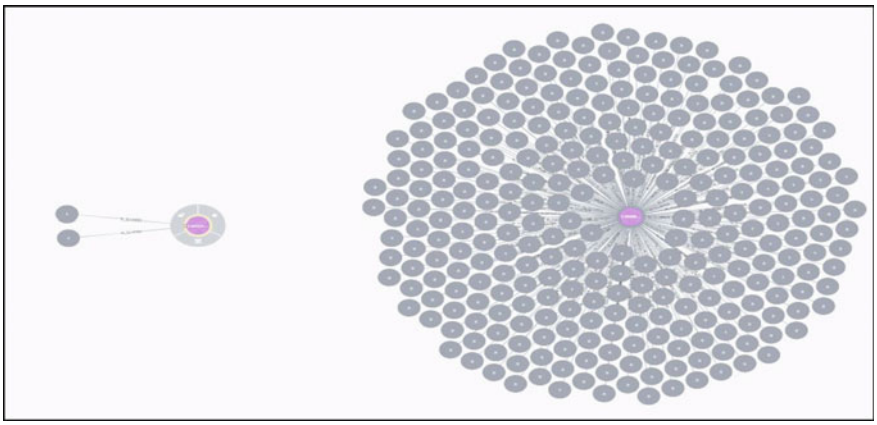


Fig. 3 The clusters formed in the second iteration

```
MATCH (w2:weather)-[:IN_CLUSTER]->(c2:Centroid {index: 2,
repetition: 1})
WITH w1Dew, w1Humidity, w1Pressure, w1Temperature,
w1Visibility, w1 Wind_direction, w1Wind_speed,
avg(w2.dew) as w2Dew, avg(w2.humidity) as w2Humidity,
avg(w2.pressure) as w2Pressure, avg(w2.temperature) as
w2Temperature, avg(w2.visibility) as w2Visibility,
avg(w2.wind_direction) as w2 Wind_direction,
avg(w2.wind_speed) as w2Wind_speed
```

3. New properties centroid creating (repetition 2).

```
CREATE (:Centroid {dew: w2Dew, humidity: w2Humidity,
pressure: w2Pressure, temperature: w2Temperature, visi-
bility: w2Visibility, wind_direction: w2 Wind_direction,
wind_speed: w2Wind_speed, index: 2, repetition: 2})
```

4. Return to dataset.

```
{return*}
```

Figure 3 shows the two clusters formed in second iteration based on the specified centroid value.

5 Results

Table 1 compares the result between the K-means with Euclidean's distance and K-means MSOINN algorithm.

Table 1 Comparison between K-means and K-means MSOINN

	K-means		K-means MSOINN	
Iteration	I	II	I	II
No. of clusters	3	2	3	2
No. of nodes in CLUSTER: C1	221	294	243	296
No. of nodes in CLUSTER: C2	61	4	43	2
No. of nodes in CLUSTER: C3	15	–	10	–

6 Conclusion

The graph database Neo4j is best suitable for representing the connected and semi-structured dataset. The creation and processing of the mass weather dataset are made easy by using the clustering algorithm, i.e., K-means algorithm MSOINN. The weather data is increasing exponentially and growing dynamically day by day. Neo4j is the best suitable database to represent such complex and multi-connected dataset. The K-means clustering algorithm plays a vital role in analyzing and managing weather dataset. The similar dataset is grouped using the algorithm and is analyzed to understand the behavior of the weather, which helps to predict the weather pattern. The K-means MSOINN algorithm also forms the clusters with closest nodes. The number of nodes in each cluster tends to show better accuracy by using K-means MSOINN compared to K-means with Euclidean's distance formula.

References

1. Some Major Disasters in India. http://www.ndma.gov.in/en/disaster-data-statistics.html
2. Top 10 Natural Disasters in the History of India. https://www.mapsofindia.com/my-india/travel/top-10-natural-disasters-in-the-history-of-india
3. N. Shi, X. Liu, Y. Guan, Research on k-means clustering algorithm: an improved k-means clustering algorithm, in *3rd International Symposium on Intelligent Information Technology and Security Informatics*, IITSI 2010 (2010), pp. 63–67. https://doi.org/10.1109/iitsi.2010.74
4. F. Noorbehbahani, S.R. Mousavi, A. Mirzaei, An incremental mixed data clustering method using a new distance measure. Soft. Comput. **19**, 731–743 (2014). https://doi.org/10.1007/s00500-014-1296-7
5. P. Kivikangas, M. Ishizuka, Improving semantic queries by utilizing UNL ontology and a graph database, in *Proceedings—IEEE 6th International Conference on Semantic Computing ICSC 2012* (2012), pp. 83–86. https://doi.org/10.1109/icsc.2012.50
6. H. Huang, Z. Dong, Research on architecture and query performance based on distributed graph database Neo4j, in *2013 3rd International Conference on Consumer Electronics, Communications and Networks, CECNet 2013—Proceedings* (2013), pp. 533–536. https://doi.org/10.1109/cecnet.2013.6703387
7. R. Kumar Kaliyar, Graph databases: a survey. Int. Conf. Comput. Commun. Autom. ICCCA **2015**, 785–790 (2015). https://doi.org/10.1109/CCAA.2015.7148480
8. R. De Virgilio, Smart RDF data storage in graph databases, in *Proceedings—2017 17th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing CCGRID 2017* (2017), pp. 872–881. https://doi.org/10.1109/ccgrid.2017.108

9. K. Ueta, X. Xue, Y. Nakamoto, S. Murakami, A distributed graph database for the data management of IoT systems, in *Proceedings—2016 IEEE International Conference on Internet of Things; IEEE Green Computing and Communications; IEEE Cyber, Physical and Social Computing; IEEE Smart Data, iThings-GreenCom-CPSCom-Smart Data 2016* (2017), pp. 299–304. https://doi.org/10.1109/ithings-greencom-cpscom-smartdata.2016.74
10. X. Lv, X. Cheng, S. Tian, C. Fang, Storage and parallel topology processing of the power network based on Neo4j, in *Chinese Control Conference CCC* (2017), pp. 10703–10707. https://doi.org/10.23919/chicc.2017.8029061
11. G. Drakopoulos, A. Kanavos, Tensor-based document retrieval over Neo4j with an application to PubMed mining, in *IISA 2016—7th International Conference on Information, Intelligence, Systems and Applications* (2016), pp. 1–6. https://doi.org/10.1109/iisa.2016.7785366
12. G. Drakopoulos, Tensor fusion of social structural and functional analytics over Neo4j, in *IISA 2016—7th International Conference on Information, Intelligence, Systems and Applications* (2016), pp. 1–6. https://doi.org/10.1109/iisa.2016.7785365
13. A. Virk, R. Rani, recommendations using graphs on Neo4j, in 2018 International Conference on Inventive Research in Computing Applications (2018), pp. 133–138

Proactive Preventive and Evidence-Based Artificial Intelligene Models: Future Healthcare



Kamal Kr. Sharma, Shivaji D. Pawar and Bandana Bali

1 Introduction

This study emphasizes on different AI models on which researchers are working from the last 2 years and implementing their thought process for future development. In this paper, we have studied recent inventions in the field of AI in health care and for the better understanding, we have divided this study into three different categories like proactive AI domain, preventive AI domain, and evidence AI domain shown in Fig. 1.

In proactive AI model, many researchers provide unique solutions to health care by using a noninvasive approach. A proactive approach is always better for future health care; hence, in this model, we can monitor human and geographical activity to predict diseases or illness before it happens. The hypertriglyceridemic waist (HW) is a better predictive tool for type 2 diabetic patients as compared to the triglyceride [1]. Facial abnormalities detection can be the best model for childcare, mental illness, and old age people [2]. Environmental and geographical sensor data can be useful to predict malaria and asthma-related diagnosis [3, 4]. Atrial fibrillation detection by smartphone is another noninvasive diagnosis tool [5]. Factor dimensions (factor D) can be a predictive tool for brain disease and age detection [6].

In prevention, AI model works on the concept of health care, may be in the form of electronic health care or creating collaborative healthcare environments especially

K. Kr. Sharma
Electronics and Electrical Engineering, Lovely Professional University, Jalandhar, Punjab, India
e-mail: Kamal.23342@lpu.co.in

S. D. Pawar (✉)
Computer Science and Engineering, Lovely Professional University, Jalandhar, Punjab, India
e-mail: paras_pawar08@rediffmail.com

B. Bali
Uttarakhand Technical University, Dehradun, India

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_44

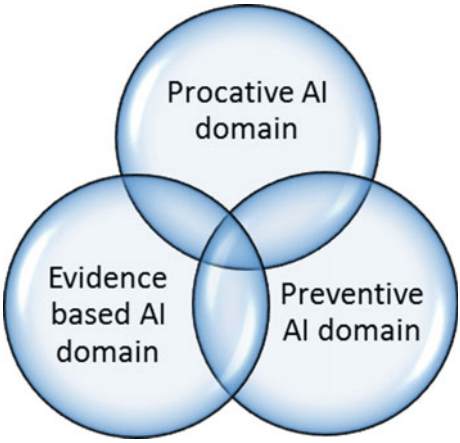


Fig. 1 Summary of AI model

for aged people and childcare purpose. Many researchers have discovered electronic E-Health options with enhanced security [7–9]. Automatic food recognition system for a diabetic patient, especially old age people for diet monitoring and artificial pancreas, can be a future of diabetic prevention technique. Recent research also highlights co-relation between Vitamin D and diabetics is also acting as a proactive tool [10–12]. Deep learning approach on node sensor data analysis can increase the accuracy of smartphone and wearable devices [13]. In the prevention model, data accuracy and integration with run time diagnosis of AI model are essential features.

Evidence-based AI diagnosis model helps the radiologist for fast and accurate diagnosis. This model is a supportive tool in radiology for predictive analysis during diagnosis. Many researches exist in recent times in terms of how to enhance medical image sensitivity and specificity in the complex domain. The medical field always consists of data from various illnesses in complex form having different intensities, shapes, textures, and wavelength. Area of radiology covers Magnetic Resonance Image (MRI), Computed Tomography (CT), Ultrasound (US), Radiography, Mammography, Breast Tomosynthesis, and Positron Emission Tomography (PET) which provides a higher version of the image sensor.

Automatic lesion detention by using ultrasound imaging with the neural network [13], automatic age detection from MRI data, and machine learning algorithms [14] can act as a better run time diagnosis tool for breast cancer and medical diagnosis. Anatomically constrained neural network (ACNNS) can be another option for cardiac image enhancement and segmentation [15]. Automatic localization of anatomical landmarks with decision forest can act as the best choice for medical imaging in cardiac detention [16]. Nonnegative canonical polyadic decomposition for tissue-type differentiation in gliomas is a better method than the previous matrix-based decomposition for identification of brain tumor [17]. Infrared imaging technology can be another alternative for breast cancer detection [18]. AI offers a new and promising set of methods for analyzing image data in radiology [19–21]. The latest

advancement in deep learning technologies is a new and effective paradigm to obtain end-to-end learning models from complex data [22].

2 Literature Review

Health care should be cost-effective, fast in diagnosis, preventive, and patient oriented. This all attributes of health care may be possible with AI, and hence the world is excited to see the new era of health care in 2019 and beyond. In this section, we will see the main recent literature (26 papers) related to the application of deep and machine learning in medical imaging, EHR, smartphone, and wearable devices. For better understanding, we divide this literature survey into three AI models.

2.1 Proactive AI Model

Key challenges of health care are the variation in data of each patient depending upon their genetic makeup, economical and geographical background. Hence, it is very difficult to develop a common proactive tool for health diagnosis in early days.

But due to smartphone and wearable devices, it is easy to develop proactive model for health care with the support of deep learning algorithms. In all the proactive models, we can generate personalized data of each patient with the help of sensor and by applying deep learning algorithm. In recent literature, many researchers have developed attributes of human activity, facial abnormalities, environmental and geographical sensors, radio genomics, atrial fibrillation, and factor dimension for developing proactive models for differential diagnosis.

Identification of Type 2 diabetic patients is carried out by using anthropometry and triglyceride with Machine Learning Algorithms (NB/NR), and this model identifies a strong association between HW and diabetic2 as compared to TG [1]. A survey shows the correlation between face abnormalities with AI can be the best proactive model for childcare, old age people, and mental illness [2]. Asthma-related emergency visits, Twitter data, and SVM can develop asthma proactive model [3]. Another multitier application for malaria detection known as malaria analytics can handle interoperability on a distributed database [4]. Inbuilt microelectromechanical accelerometer and gyroscope sensor as an internal measurement unit can detect whether the patient is suffering from the A-fib or not with 97% accuracy as compared to previous camera flash methods [5]. Cognitive changes in aging can be a proactive tool for brain diseases [6]. Fatal D is another innovation which works on a change in geometry lines as per aging [23]. A recent survey suggests, Vitamin D is one of the important diagnostic tools for breast and diabetic detection and also more research is required in this area to develop automated detection tool with the help of artificial neural networks in the future [11]. These models will provide valuable tools for

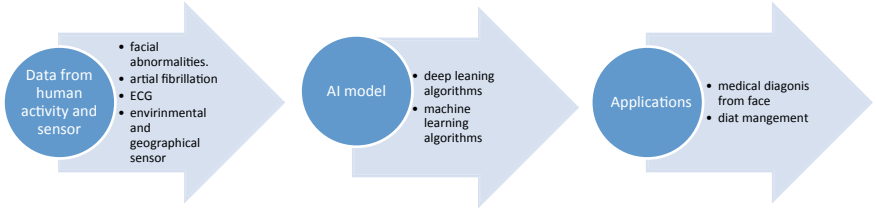


Fig. 2 Summary of proactive model

Table 1 Reviewed trends of artificial intelligence in proactive health care

Parameter	AI algorithms	Diseases
Anthropometry	NB/NR	Diabetic prediction
Face abnormalities	Cosine similarity and K-nearest neighbor	Child and old age care
Emergency visit and Twitter data	ANN	Asthma attack prediction
Data from medical reports	Service management agents (SMA)	Malaria prediction
ECG clinical data	A-fib classification	Atrial fibrillation detection
Cognitive numerical analysis	UWB-RTLS system	Brain disease
Vitamin D	Future research gap	For diabetic and breast cancer detection

practitioners by eliminating subjective bias and reducing diagnosis time and costs (Fig. 2 and Table 1).

Future challenge for this model is to handle heterogeneous data in terms of images, videos, and texts, and hence advanced sensor technology and data acquisition systems should be sufficient enough to provide on time processing of data. Also, a deep learning algorithm should be smart enough to provide an accurate prediction. Also, this modeling requires a larger data warehouse to construct accurate decision.

2.2 Prevention AI Model

The basic objective behind this model is to provide prevention technique to the patient against the illness. Health care of child and old age people is the major concern across the world. Recently, many researchers have proposed privacy-enhanced electronic health record system which consists of a database of different community people which is integrated at the common platform. Security is the major challenge in this area, XML based, cloud computing based, big data analytics, and adaptive Merkle

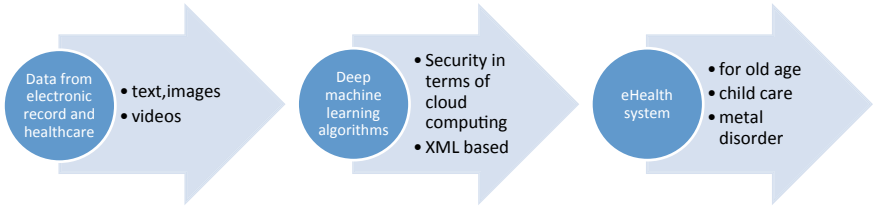


Fig. 3 Summary of prevention model

tree based security mechanisms are proposed in all the above studies [7–10]. Biosignal sensor with ECG and unsupervised Bayesian interface can act as the best tool for personalized health care [8].

When an adult's age grows more than 60, then near about two-third population of the world will be under illness of diabetic, hypertension, and cardiovascular. Proper nutrition management can avoid evidence-based complication. Hence, activity management and diet management by artificial intelligence with a smartphone can be a future area of research development [11, 12]. In order to enhance the life of the diabetic patient, continuous blood sugar monitoring and proper medicine are options.

The future need is to develop new noninvasive sensor technology with the help of smart algorithms to monitor blood glucose [13] (Fig. 3 and Table 2).

Table 2 Reviewed trends of artificial intelligence in preventive health care

Data input	Security/New technique	Application
Heterogeneous data	Deep learning algorithms	To increase on node processing power of smartphone and wearable devices
Korotkoff stethoscope sound	CNN	For blood pressure measurements
Heterogeneous data	Electrocardiogram identification method	To enhance security of health care
Heterogeneous data	XML-based security	To enhance security of health care
Heterogeneous data	Cloud computing-based security	To enhance security of health care
Heterogeneous data	Biosignal sensor with ECG and unsupervised Bayesian interface	To enhance security of health care
Human activity	Future scope	To enhance personal care
Diet monitoring	Future scope	To increase health care of old age people
Blood glucose monitoring	New invasive sensor technology	For diabetic patients

Variation of the Korotkoff stethoscope sounds during blood pressure measurement with the help of convolutional neural networks can act as an important tool for personalizing care [24]. The basic limitation of wearable devices is the inability to process on time data for personalized care. This drawback can be overcome with the help of deep learning algorithms with better accuracy and sensitivity. Multi-scale autoregressive security model on the base of electrocardiogram identification method is the best future alternative to enhance the security of health care [14].

Future challenges of this model are to integrate world patient data into a single domain with increased security. Active participation of patients by overcoming economic, geographical, social standards, and essential support from a physician, clinical staff, and government policies is also important for accurate implementation of preventive-based AI models in health care.

2.3 Evidence-Based AI Model

Basic role behind this model is to provide accurate and cost-effective diagnosis to the patients. This model works to improve medical image sensitivity and specificity by utilizing the ability of artificial intelligence to provide a fast and accurate diagnosis. Sometimes it is very difficult for the radiologist to identify small inpoints or smaller changes of images, which is possible by deep learning algorithms. Observer fatigue can be also overcome by artificial intelligence.

Digital mammography is one of the most applied tools for breast cancer detection but these methods consist of many drawbacks like dense breasts where lesions have similar attenuation as they are surrounded by other tissues. Many kinds of research are undergoing in this area where a combination of ultrasonic imaging and CNN can provide a better diagnostic tool for breast cancer detection as compared to the digital mammography [25]. Automatic age detection from MRI data of hand, clavicle, and teeth and deep convolutional neural networks can be the new revolution in forensic and medical diagnosis for age 13–19 years [14]. Multichannel SCG using a novel seismocardiogram spectrum system can be used to overcome position dependency, time delay, and signal attenuation of single-channel ECG [16]. Blind source separation technique from MRI data known as nonnegative canonical polyadic decomposition (NCPD) is used to separate tissue types for glioma patients [17]. Novel stratified decision forests can be the better method for accurate landmark location in cardiac images as compared to the previous classifier and regression models [26]. Early detection of breast cancer always increases the rate of survival as many practices are there for an early prediction like mammography, ultrasound, and MRI but these methods get fail in case of dense breast dynamic infrared thermography as it can be the best alternative to improve breast cancer detection [18] (Fig. 4 and Table 3).

Future challenges in this area are to provide the cheapest, fastest, and safest medical diagnosis. Ultrasonic imaging is one of the areas where we can provide real-time guidance with the support of AI in the area of vessel detection in breast cancer known as an AI breast. In MRI, the AI model can increase diagnosis speed by

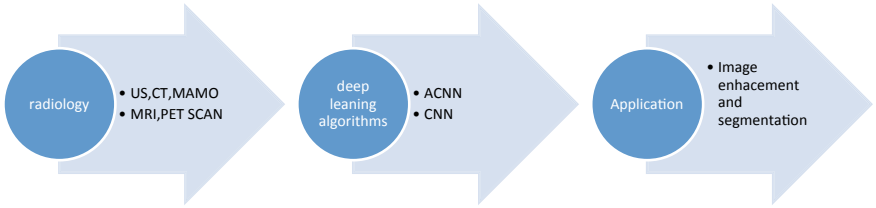


Fig. 4 Summary of evidence-based model

Table 3 Reviewed trends of artificial intelligence in evidence-based health care

Data input source	New algorithms/Technique	Application
US	CNN	Breast cancer detection
Infrared imaging	Pennes model	Dense breast cancer detection
MRI	CNN	Automatic age detection
Medical image	ACNN	Image enhancement and separation
MRI	Blind source separation technique	To separate tissue types in Glioma patients
Cardiac image	Decision forest	Accurate localization of landmarks

10 times to reduce high doses of gadolinium. MRI with AI can be used to detect liver fat for weight loss programs. AI accelerates PET process which can provide faster scans with reduced doses. Chain of handcrafted signal processing models can be replaced with the image of reconstruction algorithms with deep learning. AI imaging research would be benefited, if national and international image sharing network takes place and generating a standard protocol for standardization and optimization of medical imaging. Generating reference dataset of proven cases for further analysis is the essential requirement of this area. Safety, failure and judicial transparency, and privacy are major challenges for the implementation of artificial intelligence in radiology [22].

3 Discussion

In this section, we will focus on some key points of this study for discussion and take away points (Table 4).

Table 4 Summary of all models

AI models	Area of application	Area of future scope
Proactive models	Activity monitoring, age detection, face abnormalities detection, cardiac monitoring, diet monitoring, malaria, and asthma detection. Breast cancer and diabetic detection	Data collection and storage technique Enhancement in sensor technology Run time processing power of the sensor Higher accuracy deep learning predictive algorithms Power handling capacity of the model
Preventive models	Blood glucose monitoring Healthcare systems Health monitoring systems	Collection of the heterogeneous database Cybersecurity Legal and ethical issues Sentimental and emotional analysis Training session of patient and clinical staff
	MRI CT Mammography Radiography PET scan	Increases the speed of diagnosis Reduces diagnosis time Reduces gadolinium doses Image enhancement Image segmentation Image reconstruction by using deep learning algorithms Sharing of national and international images Generation of reference dataset of proven cases The new protocol for standardization and optimization of images Dense breast cancer detection

4 Conclusion

Due to artificial intelligence, there is futuristic hope in health care that can deliver more responsive, cost-effective, and higher accuracy models to the patient. But there are a lot of technical, social, economic, legal, and ethical issues for the development of real-time models. Hence, combined efforts in the technical and medical field are required to achieve the goal. This study gives an overview of different AI-based models in medical application and their future scope and challenges.

From technical point of view, there is a need for technological advancement in automated algorithms in terms of accuracy and speed of diagnosis with the help of clinical trials. Despite tremendous growth in health care, there are still untangle challenges glimpsing the adoption of machine learning specifically and AI in general. Even though AI cannot replace the need of a physician but can act as an advanced

technical tool to help physician become more efficient, and provides better health care and decreases health care cost by early detection and better treatment. As an attitude of gratitude, AI can surely act as the best patient care model in the near future.

References

1. B.J. Lee, J.Y. Kim, Identification of type 2 diabetes risk factors using phenotypes consisting of anthropometry and triglycerides based on machine learning. IEEE J. Biomed. Health Inform. **20**(1) (2016)
2. J. Thevenot, M.B. Lopez, A. Hadid, A *survey on computer vision for assistive medical diagnosis from faces*. IEEE J. Biomed. Health Inform. **22**(5), 1497–1511 (2018). https://doi.org/10.1109/JBHI.2017.2754861
3. S. Ram, W. Zhang, M. Williams, Y. Pengetnze, Predicting asthma-related emergency department visits using Big Data. IEEE J. Biomed. Health Inform. **19**(4), 1216–1223 (2015). https://doi.org/10.1109/JBHI.2015.2404829
4. J.H. Brenas, M.S. Al-Manir, C.J.O. Baker, A. Shaban-Nejad, A malaria analytics framework to support evolution and interoperability of global health surveillance systems. IEEE Access **5**, 21605–21619 (2017)
5. O. Lahdenoja, T. Hurnanen, Z. Iftikhar, S. Nieminen, T. Knuutila, A. Saraste, T. Koivisto et al., Atrial fibrillation detection via accelerometer and gyroscope of a smartphone. IEEE J. Biomed. Health Inform. **22**(1), 108–118 (2018). https://doi.org/10.1109/JBHI.2017.2688473
6. W.D. Kearns, J.L. Fozard, V.O. Nams, Movement path tortuosity in free ambulation: relationships to age and brain disease. IEEE J. Biomed. Health Inform. **21**(2), 539–548 (2017). https://doi.org/10.1109/JBHI.2016.2517332
7. K. Seol, Y.G. Kim, E. Lee, Y.D. Seo, D.K. Baik, Privacy-preserving attribute-based access control model for the XML-based electronic health record system. IEEE Access **6**, 9114–9128 (2018). https://doi.org/10.1109/ACCESS.2018.2800288
8. L. Yu, W.M. Chan, Y. Zhao, K.L. Tsui, The personalized health monitoring system of elderly wellness at the community level in Hong Kong. IEEE Access **6**, 35558–35567 (2018). https://doi.org/10.1109/ACCESS.2018.2848936
9. R. Sanchez-Guerrero, F.A. Mendoza, D. Diaz-Sanchez, P.A. Cabarcos, A.M. Lopez, Collaborative eHealth meets security: privacy-enhancing patient profile management. IEEE J. Biomed. Health Inform. **21**(6), 1741–1749 (2017). https://doi.org/10.1109/JBHI.2017.2655419
10. M.M. Anthimopoulos, L. Gianola, L. Scarnato, P. Diem, S.G. Mougiakakou, A food recognition system for diabetic patients based on an optimized bag-of-features model. IEEE J. Biomed. Health Inform. **18**(4), 1261–1271 (2014). https://doi.org/10.1109/JBHI.2014.2308928
11. A.A. Rizvi, Nutritional challenges in the elderly with diabetes. Int. J. Diabetes Mellit. (2009). https://doi.org/10.1016/j.ijdm.2009.05.002
12. H.C. Wang, A.R. Lee, Recent developments in blood glucose sensors. J. Food Drug Anal. (2015). Elsevier Taiwan LLC. https://doi.org/10.1016/j.jfda.2014.12.001
13. D. Ravi, C. Wong, B. Lo, G.Z. Yang, A deep learning approach to on-node sensor data analytics for mobile or wearable devices. IEEE J. Biomed. Health Inform. **21**(1), 56–64 (2017). https://doi.org/10.1109/JBHI.2016.2633287
14. D. Stern, C. Payer, N. Giuliani, M. Urschler, Automatic age estimation and majority age classification from multi-factorial MRI data. IEEE J. Biomed. HealthInformatics (2018). Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/JBHI.2018.2869606
15. E. Shelhamer, J. Long, T. Darrell, Fully convolutional networks for semantic segmentation. IEEE Trans. Pattern Anal. Mach. Intell. **39**(4), 640–651 (2017). https://doi.org/10.1109/TPAMI.2016.2572683

16. W.Y. Lin, W.C. Chou, P.C. Chang, C.C. Chou, M.S. Wen, M.Y. Ho, M.Y. Lee et al., Identification of location specific feature points in a cardiac cycle using a novel seismocardiogram spectrum system. IEEE J. Biomed. Health Inform. **22**(2), 442–449 (2018)
17. H.N. Bharath, D.M. Sima, N. Sauwen, U. Himmelreich, L. De Lathauwer, S. Van Huffel, Non-negative canonical polyadic decomposition for tissue-type differentiation in gliomas. IEEE J. Biomed. Health Inform. **21**(4), 1124–1132 (2017). https://doi.org/10.1109/JBHI.2016.2583539
18. S.G. Kandlikar, I. Perez-Raya, P.A. Raghupathi, J.L. Gonzalez-Hernandez, D. Dabydeen, L. Medeiros, P. Phatak, Infrared imaging technology for breast cancer detection—current status, protocols, and new directions. Int. J. Heat Mass Transf. (2017). Elsevier Ltd. https://doi.org/10.1016/j.ijheatmasstransfer.2017.01.086
19. R. Miotto, F. Wang, S. Wang, X. Jiang, J.T. Dudley, Deep learning for healthcare: review, opportunities, and challenges. Brief. Bioinform. (2017). https://doi.org/10.1093/bib/bbx044
20. M. Pavel, H.B. Jimison, H.D. Wactlar, T.L. Hayes, W. Barkis, J. Skapik, J. Kaye, The role of technology and engineering models in transforming healthcare. IEEE Rev. Biomed. Eng. **6**, 156–177 (2013). https://doi.org/10.1109/RBME.2012.2222636
21. F. Pesapane, M. Codari, F. Sardanelli, Artificial intelligence in medical imaging: threat or opportunity? Radiologists again at the forefront of innovation in medicine. Eur. Radiol. Exp. **2**(1), 35 (2018). https://doi.org/10.1186/s41747-018-0061-6
22. C. Liew, The future of radiology augmented with artificial intelligence: a strategy for success. Eur. J. Radiol. (2018). Elsevier Ireland Ltd. https://doi.org/10.1016/j.ejrad.2018.03.019
23. M. De La Puente-Yagüe, M.A. Cuadrado-Cenzual, M.J. Ciudad-Cabañas, M. Hernández-Cabria, L. Collado-Yurrita, Vitamin D: and its role in breast cancer. Kaohsiung J. Med. Sci. (2018). Elsevier (Singapore) Pte Ltd. https://doi.org/10.1016/j.kjms.2018.03.004
24. F. Pan, P. He, C. Liu, T. Li, A. Murray, D. Zheng, Variation of the Korotkoff stethoscope sounds during blood pressure measurement: Analysis using a convolutional neural network. IEEE J. Biomed. Health Inform. **21**(6), 1593–1598 (2017). https://doi.org/10.1109/JBHI.2017.2703115
25. M.H. Yap, G. Pons, J. Martí, S. Ganau, M. Sentís, R. Zwiggelaar, R. Martí et al., Automated Breast ultrasound lesions detection using convolutional neural networks. IEEE J. Biomed. Health Inform. **22**(4), 1218–1226 (2018). https://doi.org/10.1109/JBHI.2017.2731873
26. O. Oktay, W. Bai, R. Guerrero, M. Rajchl, A. De Marvao, D.P. O'Regan, D. Rueckert et al., Stratified decision forests for accurate anatomical landmark localization in cardiac images. IEEE Trans. Med. Imaging **36**(1), 332–342 (2017). https://doi.org/10.1109/TMI.2016.2597270

Utilization of Artificial Neural Network for the Protection of Power Transformer



Mudita Banerjee and Anita Khosla

1 Introduction

With increased requirement for power and rise in industrialization, the call for a consistent supply of power for today's world has expanded significantly. It further requires zero fault operation of the electrical network. The protection of power transformers from various faults is essential as they are large and expensive devices. The aim is to minimize the occurrence and time duration of undesirable outages of transformer.

The differential protection scheme for the protection of power transformer, sometimes maloperate because of the occurrence of magnetization inrush current. As inrush is a transient condition, mainly happening during the energization of power transformer, the differential relay is unable to identify whether the condition is inrush or internal fault. This is because the magnitude of inrush current is as large as that of internal fault current and hence causes maloperation in the circuit breaker. The second harmonic component is predominant in inrush current of the transformer [1].

Differential relay works on the principle of measurement of primary and secondary current does not avoid the false tripping when inrush condition takes place. Hence, it is essential to stop the false tripping of the differential relay during inrush condition.

Much research is going on in the field of transformer protection due to the presence of magnetizing inrush current. The researchers have explored and developed various algorithms to protect power transformer such as microcontroller-based system for transformer protection [2], differentiation based on symmetrical components

M. Banerjee · A. Khosla (✉)
Manav Rachna International Institute of Research and Studies, Faridabad, Haryana, India
e-mail: anitakhosla.fet@mriu.edu

M. Banerjee
e-mail: mmmudita@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_45

[3], morphological scheme based on wave shape for the identification of inrush current [4], autocorrelation method [5], equivalent instantaneous inductance (EII)—inductance-based method [6], and wavelet transform-based technique [7, 8] to classify inrush condition and internal fault condition. Recently, many methods based on artificial intelligence like fuzzy logic [9, 10], artificial neural network (ANN) [11], probabilistic neural network (PNN) [12], and fuzzy-neuro hybrid intelligent system [13] are developed for power transformer differential protection which increases the speed and robustness of present digital relays.

In spite of the various proposed algorithms, there still are certain restrictions. Some problems are the large calculation burden, less speed, large memory, harmonic pollution, and CT saturation depending upon the transformer parameters. Hence, more reliable, dependable, fast, and proficient approach for the utilization of artificial neural network to distinguish inrush current from fault current is presented in this paper.

2 Simulation of Power System

This segment demonstrates the simulation of electrical power system for the generation of magnetizing inrush current and internal fault current signals in a saturable core power transformer. Power system specification is given in Table 1.

Simulink outline of electrical power system is shown in Fig. 1. It consists of a three-phase AC generator connected to a three-phase transformer of specification given in Table 1. The system is feeding a connected load considering no-load condition. Any type of fault can be applied in the given system. The trained neural network is connected at the primary side of a transformer. Following three conditions can appear as an input to ANN during the energization of transformer:

- Normal condition,
- Magnetizing inrush, and
- Internal fault.

ANN is trained such that it will identify any internal fault condition as a trip and system will remain stable for inrush current to avoid malfunctioning of the differential relay. Modeling of ANN is performed using double-layer feedforward Levenberg–Marquardt backpropagation technique to discriminate between internal faults and magnetizing inrush condition in power transformer.

Table 1 Power system specifications

	Specification
Generator	220 kV, 50 Hz
Transformer	Three-phase, core type, two-winding, 50 MVA, 50 Hz, 220 kV/66 kV
Load	No-load condition

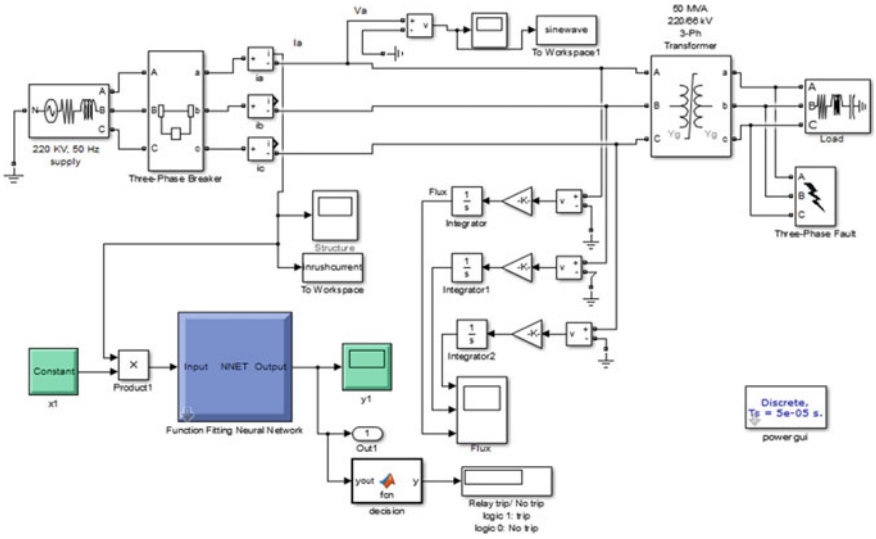


Fig. 1 Simulink diagram of electrical power system

For this transformer, input patterns of inrush and internal faults are generated. Figure 2 shows a resulted waveform of inrush current.

The magnitude of first peak of inrush current is 820 Amps, which is thrice the rated current of transformer, and therefore the relay gives false trip signal. The aim is to stop the false tripping of relay during inrush condition, and hence an ideal system should give No Trip/Logic 0 under inrush condition. It is further observed that after 2 s, inrush current decays to 270 Amps which is a rated current of 50 MVA power transformer.

Figure 3 shows the fault current waveform, when fault occurs within the protected zone at the load side. The relay should issue Trip/Logic 1 instantaneously for this large value of current.

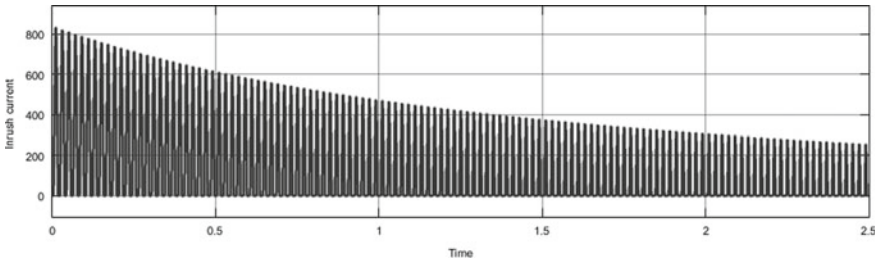


Fig. 2 Magnetizing inrush current waveform

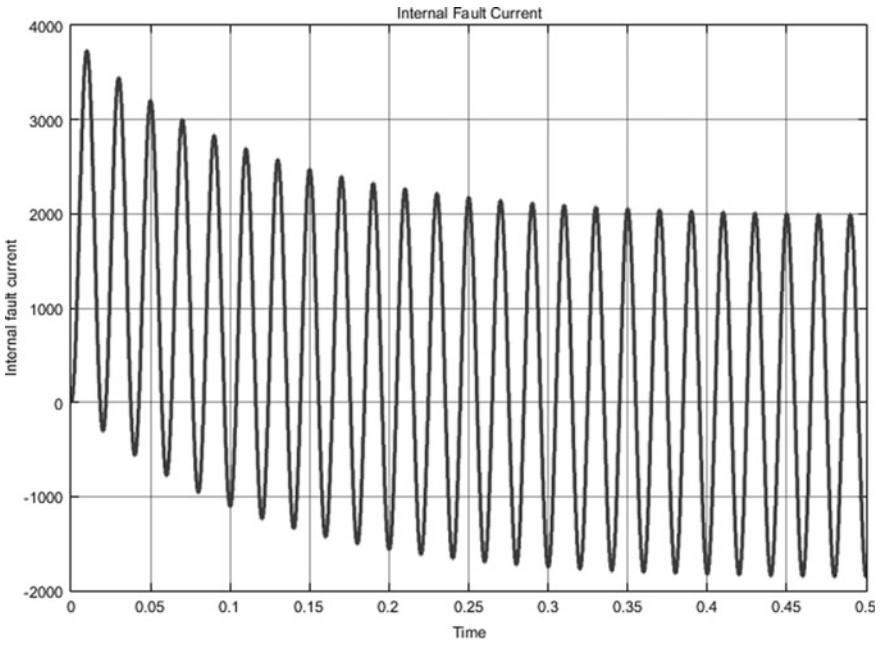


Fig. 3 Internal fault current waveform

Initially, this LLLG fault has a current magnitude of 3600 A and after 0.45 s, 2000 A flows through the unprotected system. Samples for both the conditions are kept for the training of artificial neural network.

2.1 Training of Neural Network

Neural network fitting toolbox which is obtainable in MATLAB is used for the training of ANN, for the system shown in Fig. 1. The ANN is trained for internal fault current waveform shown in Fig. 4. One cycle is required to take a decision for trip signal.

For the training of ANN, real-time instantaneous values are needed. A total of 801 samples from two cycles of training pattern have been collected and ANN is trained for internal fault current taken as a reference waveform. It consists of two-layer feedforward neural network and utilizes mean square error and regression analysis for computation of performance of network. The training is done with Levenberg–Marquardt backpropagation algorithm. The trained neural network is equated with the patterns for the actual internal fault and inrush current as input during the testing phase.

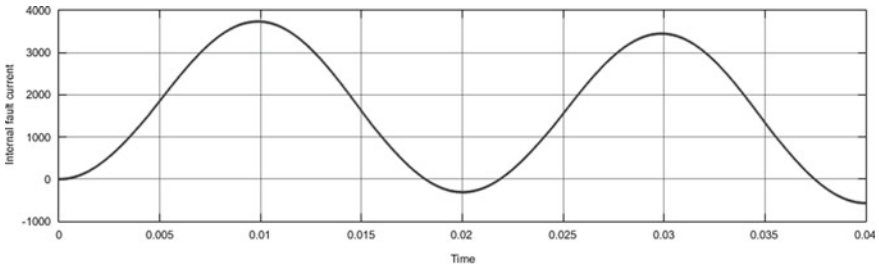


Fig. 4 Training pattern

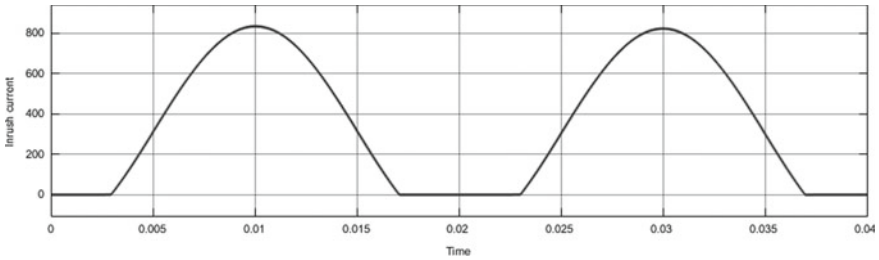


Fig. 5 Test pattern

2.2 Testing of Neural Network

The neural network has been tested with all the three conditions distinctly, i.e., normal, magnetizing inrush, and internal fault. When inrush occurs during energization, the ANN can accurately distinguish this condition with trained fault condition and accordingly system remains stable. The test pattern is shown in Fig. 5 up to two cycles of inrush current.

2.3 Result

TRIP signal has been observed for internal fault condition while testing the circuit in Fig. 1. For inrush current, the relay is not showing trip signal. Hence, the desired results have been found for the 50 MVA transformer.

3 Conclusion

This paper shows ANN-based approach for power transformer protection and an improved performance has been observed over other developed techniques. The

backpropagation algorithm reduces an average sum squared error by performing a gradient descent in the error space. Thus, the proposed ANN-based technique can accurately classify all conditions of inrush, normal, and internal fault for the protection power transformer. Hence, it results in promising security (ability not to trip during the occurrence of inrush current) and dependability (ability to trip during internal faults).

References

1. M. Banerjee, A. Khosla, Comparison and analysis of magnetizing inrush and fault condition for power transformer. Int. J. Eng. Technol. (2018)
2. A. Rafa, S. Mahmod, N. Mariun, W.Z. Wan Hassan, N.F. Mailah, Protection of power transformer using microcontroller-based relay, in *Conference on Research and Development Proceedings*, Shah Alam, Malaysia (2002)
3. H. Abniki, A. Majzoobi, H. Monsef, H. Dashti, H. Ahmadi, P. Khajavi, Identifying inrush currents from internal faults using symmetrical components in power transformers, in *Modern Electric Power Systems 2010*, Wroclaw, Poland (2010)
4. Z. Lu, W.H. Tang, T.Y. Ji, Q.H. Wu, A morphological scheme for inrush identification in transformer protection. IEEE Trans. Power Deliv. **24**(2) (2009)
5. H. Samet, T. Ghanbari, M. Ahmadi, An auto-correlation function based technique for discrimination of internal fault and magnetizing inrush current in power transformers. Electric Power Componen. Syst. (2015)
6. G. Baoming, A.T. de Almeida, Z. Qionglin, W. Xiangheng, An equivalent instantaneous inductance-based technique for discrimination between inrush current and internal faults in power transformers. IEEE Trans. Power Deliv. **20**(4) (2005)
7. R.A. Ghunem, R. El-Shatshat, O. Ozgonenel, A novel selection algorithm of a wavelet-based transformer differential current features. IEEE Trans. Power Deliv. **29**(3) (2014)
8. J. Faiz, S. Lotfi-Fard, A novel wavelet-based algorithm for discrimination of internal faults from magnetizing inrush currents in power transformers. IEEE Trans. Power Deliv. **21**(4) (2006)
9. A.A. Aziz, A.H. Abbas, A. Ali, Power transformer protection by using fuzzy logic. Iraq J. Electr. Electron. Eng. **5**(1) (2009)
10. I.S. Rad, M. Alinezhad, S.E. Naghibi, M.A. Kamarposhti, Detection of internal fault in differential transformer protection based on fuzzy method. Am. J. Sci. Res. (32), 17–25 (2011). ISSN 1450-223X
11. H. Khomhadi-Zadeh, Power transformer differential protection scheme based on symmetrical component and artificial neural network, in *IEEE 7th Seminar on Neural Network Applications in Electrical Engineering* (2004)
12. M. Tripathy, R.P. Maheshwari, H.K. Verma, Improved transformer protection using probabilistic neural network and power differential method. Int. J. Eng. Sci. Technol. **2** (2010)
13. H. Khorashadi Zadeh, M.R. Aghaebrahimi, A neuro-fuzzy technique for discrimination between internal faults and magnetizing inrush currents in transformer. Iran. J. Fuzzy Syst. **2**(2) (2005)

Analysing Tweets for Text and Image Features to Detect Fake News Using Ensemble Learning



Priyanka Meel, Harsh Agrawal, Mansi Agrawal and Archit Goyal

1 Introduction

Fake news is a part of our life since the birth of daily press. There are mainly three types of fake news as follows:

– The first type is deliberate fake news, which refers to spreading of fake news intentionally by someone in order to gain personal benefits. The person knows that the news is fake but still he/she spreads the news knowingly. This can be seen during elections where fake news about political parties or leaders are deliberately spread by the opponents so as to influence the voters. Main focus is to change the outcome of a particular situation by misleading others.
– The second type is not so much fake. In this type of fake news, the news that is spread is actually a modification or exaggeration of real news. Not all of the news is fake in this case, some part of the news is true. Though, this type of news also misguide others but not to that extent.
– The third type is unintentional fake news. In this case, the person spreads the fake news unknowingly. Actually, most of the fake news nowadays are of this type only. The person reads some news and believes that it is actually true, and share it with friends and family. Thus, in this way, the person is spreading the fake news but

P. Meel · H. Agrawal (✉) · M. Agrawal · A. Goyal
Delhi Technological University, New Delhi, India
e-mail: agharsh11@gmail.com

P. Meel
e-mail: priyankameel@dtu.ac.in

M. Agrawal
e-mail: mansiagrawal2806@gmail.com

A. Goyal
e-mail: architgoyal29@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_46

he/she does not know that. This is not done for any personal gain rather due to curiosity.

The main source of fake news is social media. Within seconds the information is spread among millions of users through various social networking sites like Facebook, Instagram, Twitter, etc. They are also doing their part to stop the spread of fake news. Facebook has third-party fact-checking, user ratings, option to report a particular post or user in case anyone finds it not suitable or fake. Twitter, too, has been removing fake profiles and bots at a faster rate than ever. It has also provided more transparency over political content. But still the spread of fake news is not getting affected by this. According to the report by Knight Foundation,[1] Twitter still contains huge amount of fake news activity. The group analysed around 10 million tweets from 700,000 Twitter accounts which had linked to more than 600 fake and conspiracy news outlets. They found that in the lead-up to the 2016 US Presidential election, more than 6.6 million tweets are linked to fake news and conspiracy news publishers. So, other means must be developed to stop the spreading of fake news and also people must be trained so that they can easily use these fake news detection frameworks.

In this paper, we have decided to propose a system for analysing the tweets for text and image features to detect fake news using ensemble learning. Previous papers have made some significant contributions to the field of fake news detection. They have assigned hard-coded weights to the features and are only limited to text tweets and do not take into account, images present in the tweets. We tend to improve on this by taking into consideration the implicit and explicit features of both text and images in the tweets. Also, the prediction algorithm in most research as of now gives a Boolean output, i.e. it tells whether the news is fake or not, which is unreasonable due to the fact that most of the fake news is based on certain real news or are derived from real news only by exaggerating the fact. Hence, it is important to predict the percentage reality in the news article or tweet rather than classifying it to real or fake. Also, most of the systems use static features and rankings, so a user can study the system and can spread the fake news in such a way that it is not detected by the system. Here, we are not assigning weights to the features statically, rather we will train the system continuously to get the best weights. Then user cannot study the system and create a way to spread the fake news such that it is not detected by the system. Sentiment analysis for explicit features of text, resolution of image, number of faces in the image for explicit features of image and CNN for implicit image features are used.

[1] https://www.knightfoundation.org/reports/disinformation-fake-news-and-influence-campaigns-on-twitter.

2 Related Works

With the increase in use of social networks, the detection of fake news spreading through them has gained significant importance. The solutions to this problem can be divided into two major parts, namely, machine learning based and non-machine learning based, our approach falls under the condition of machine learning-based models. Significant works have been done in this department by a lot of researchers.

In the paper TI-CNN [1] by Yang et al., they have proposed a system with separate features with a combination of both implicit and explicit features for both images and text which are combined later to predict the output. This typically just gives a Boolean value of whether the news is fake or real as output.

In the paper, 'A Credibility Analysis System for Assessing Information on Twitter' [2] by Majed et al., they have a similar approach with respect to text features but the impact of features of image on the final output is not taken into account. Also, they have assigned hard-coded weights to the features in order to get the output.

In the paper, 'One millisecond face alignment with an ensemble of regression trees' [3], Kazemi and Sullivan have shown how regression trees can be efficiently used to characterize certain landmarks on the face achieving an extremely reliable and real-time estimate of such facial landmarks, which can in turn be used to detect faces.

The methods discussed earlier gave different results due to the difference in the approaches used by them. The literature survey also shows that it is possible to build a system to predict whether the news is fake or real.

3 Dataset

The dataset[2] (also used in TI-CNN [1]), used here 20,015 news, out of which 11,941 is fake news and 8,074 is real news. Real news are crawled from famous news websites like the New York Times, Washington Post, etc. In this dataset, every news item contains multiple information, such as website, author, title, label information, image, text, language, country, likes, comments, etc. In our project, we are using only id, text, image and label information. Dataset is preprocessed to remove all the tweets that are not in English language. Also, the dataset contains some URLs to forbidden images, so all the tweets that contain link to forbidden images are removed before we train the model on this dataset. We have used the same dataset both for training and testing, and therefore the whole dataset is splitted to use 60% of the data for training the model and the other 40% of dataset is used for testing purpose. The main reason that we have selected this dataset is that it is an open-source dataset that is readily available. Also, this dataset contains tweets that have information about both text and the images relevant to that particular news. The dataset has been shown relevant by authors solving similar problems.

[2] https://drive.google.com/file/d/0B3e3qZpPtccsMFo5bk9Ib3VCc2c/view.

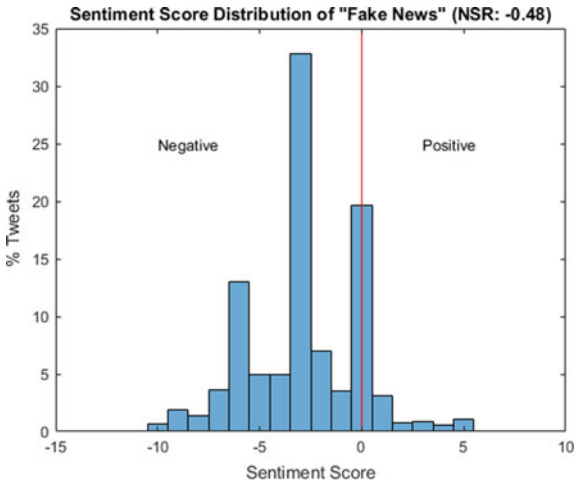


Fig. 1 Sentiment score distribution of fake news

4 Feature Extraction

4.1 Text Features

4.1.1 Explicit Features of Text

– Sentiment Analysis: The basis of sentiment analysis is to determine the polarity of the news in question to know how positive or negative the news is. Sentiment analysis can yield a multitude of feelings as a feature vector, and hence provide a more comprehensive view of the credibility. The sentiment analysis as shown in the graph[3] below, it is found that most of the fake tweets lie on the negative side of the sentiment polarity (Fig. 1).

4.2 Image Features

4.2.1 Explicit Features of Image

– Resolution: The resolution of an image, as shown in below figure,[4] is basically the details as covered by the image, i.e. it is nothing but the knowledge about the quality of the image, the more the resolution of the image, the more the details covered by the image, and the better is the quality of the image, the higher the

[3] https://blogs.mathworks.com/loren/2017/02/07/analyzing-fake-news-with-twitter/.

[4] https://blogs.mathworks.com/loren/2017/02/07/analyzing-fake-news-with-twitter/.

resolution, the more chances it is that the image is not a fake one, as seen in historical data.
– Number of faces: The number of faces refers to the count of the human faces in an image shared on Twitter. The greater count of human faces in an image implied that the photo was less likely to be fake, based on historical data.

4.2.2 Implicit Features of Image

While all the features explained up until now have been shown to have a large effect on determining fake news, a lot of features are implicit in nature to images and cannot be found explicitly, and to incorporate such features in the final result of the image, we decided on using a convolutional neural network classifying the image into fake or not fake taking the image as an input. We choose to use CNN due to its superior performance for images and easy implementation when compared to other deep learning networks (Fig. 2).

5 Proposed Approach

In this paper, we are proposing a system for analysing tweets for both text and image features to detect fake news using ensemble learning, which is basically an improvement of previous papers that consider only textual tweets but images in the tweets are not taken into account. They assign hard-coded weights to the features to predict whether the news is real or fake. Also, they only give result in Boolean value, whether the news is real or fake. We are training the system continuously, then user cannot study the system and create a way to spread the fake news such that it cannot be detected by the system. Here, we are training different machine learning models on implicit and explicit feature set of both text and images in the tweets.

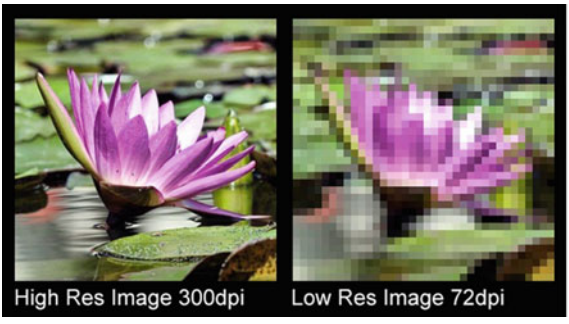


Fig. 2 Resolution of image

5.1 What is Ensemble Learning?

In ensemble learning, we combine different models like classifiers so as to filter the noise and overcome the shortcomings of a single model. This is based on the principle that a number of weak learners can be combined to form a single strong learner. Different ensemble learning techniques are

- Bootstrapping: Bootstrapping is random sampling of data allowing replacing picked up units, it is typically found helpful in understanding the bias and the variance of the dataset alongside being able to better recognize the contribution of mean and standard deviation, an integral component is that it is equally likely for an example to be selected in this kind of ensemble learning.
- Bagging: In bagging instead of just caring about the samples, we first get a subset from the dataset like we did in the bootstrapping explained followed by selecting the features which are particularly inclined towards a perfect split, now repeating this we create many different models which are aggregated to get the results.
- Boosting: In boosting, we exploit the concept of weighting, i.e. instead of independent feature selection as in the case of bagging, the features are weighted based on misclassification of previous models, thus trying to cover the wrong classifications in the past.

5.2 Dynamic Weighting of Feature Outputs

Weighting of features yields a better result of classification since some features had more impact on the result than others, but other researchers were using hard-coded weights for everything, and this made their system vulnerable to attacks to attackers who had cracked their system; in our case, we have improved this aspect and instead used a dynamically trained weighting module, and this would ensure that no attacker can actually guess the weights and also facilitate a more relevant weighting scheme.

5.3 Design Architecture

See Fig. 3.

$$\%Real = \frac{100 * (w1 * o1 + w2 * o2 + w3 * o3 + w4 * o4)}{(w1 + w2 + w3 + w4)} \tag{1}$$

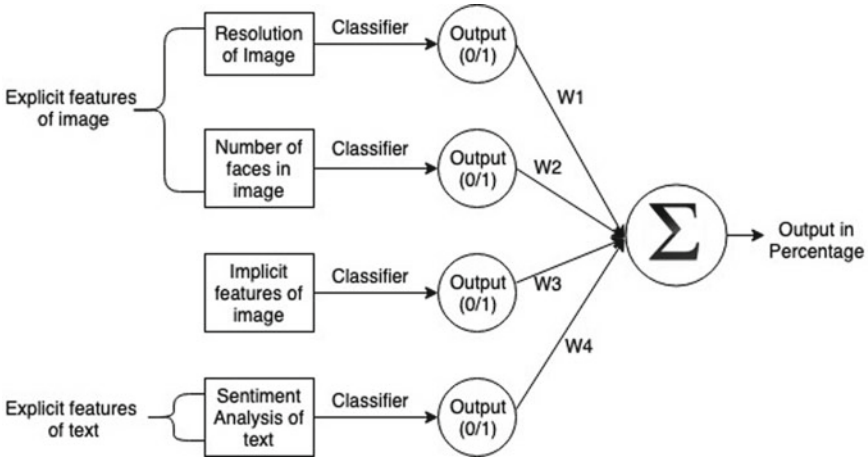


Fig. 3 Design architecture flow chart

6 Experiments And Results

6.1 Sentiment Analysis

To extract the sentiment from the text extraction module, we use the text blob API to get the polarity followed by AdaBoost classifier to generate the result, whether a particular text article is classified as fake or not, individually. Testing and training accuracies were measured, and listed in the table, using naive Bayes, SVM, gradient boost and AdaBoost classifiers when sentiment analysis on the text is done. We are using AdaBoost classifier for sentiment analysis of text (Fig. 4).



Fig. 4 Sentiment analysis flow chart

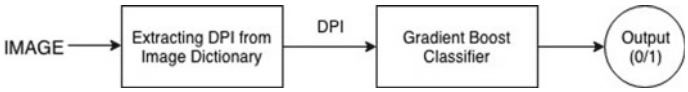


Fig. 5 Resolution of image flow chart



Fig. 6 Number of faces in image flow chart

6.2 Resolution of Image

To extract the resolution of the image, we use BytesIO to first read the image and then use the corresponding image's information dictionary to extract the resolution and then follow it by gradient boosting classifier to detect whether a news is fake or true. Testing and training accuracies for the resolution of image using naive Bayes, SVM, gradient boost and AdaBoost classifiers were compared, and classifier giving maximum accuracy was used. We are using gradient boost classifier for resolution of image because, as shown in the table, it has highest testing accuracy (Fig. 5).

6.3 Number of Faces

To detect number of faces, we have used DLIBs an ensemble of regression trees one millisecond face recognition to make face contours around all faces in the image followed by counting these contours, and then finally using a naive Bayes classifier, we classified it as fake or not fake. Testing and training accuracies for number of faces in image using naive Bayes, SVM, gradient boost and AdaBoost classifiers were compared, and classifier giving maximum accuracy was used. As shown in the table, naive Bayes classifier has the highest testing accuracy, and we are using naive Bayes for this model (Fig. 6).

6.4 Convolutional Neural Network

We decided to implement a sequential CNN model with batch size as 32 and 30 epochs. The CNN was then optimized using AdaDelta optimizer, calculating the loss from categorical cross-entropy loss and results are listed in the table.

6.5 Ensemble Learning Results

Results from all the above feature vectors are combined and trained using a neural network with a softmax classifying layer, and the weights obtained training the same are used in order to ascertain as to how much percentage news is real.

The training and testing accuracies for the neural network are listed in Table 1.

7 Conclusion and Future Works

In this paper, we have predicted by what percentage the news is credible. We have assigned dynamic weights to the implicit and explicit features of the text and the image. Various models, namely, sentiment analysis for text, number of faces in image, resolution of image and CNN for images, which analyse implicit and explicit features of both text and images in the tweets, were trained on a dataset of around 20,000 tweets. This approach differs from previous models which give Boolean value for the given instance of news and assigns static weights to the features in the model. We would like to highlight the issues we faced while developing the model, which when solved can improve the model:

- More labelled datasets especially for images.
- Along with the news posted, the information about the user who posted it like details of the user, tweeting history and usual behaviour pattern to check whether user is credible or not or if account has been compromised.
- Profile details like number of followers to following ratio, history of tweets by the user, number of retweets, how old the account is or demographic of users.

Table 1 Accuracies of all the models

Model	Accuracies	
	Training	Testing
Sentiment analysis		
naive Bayes	0.69285	0.70650
SVM	0.69285	0.70650
Gradient Boost	0.69285	0.70650
AdaBoost	0.69500	0.70789
Resolution of images		
naive Bayes	0.68000	0.72069
SVM	0.76000	0.76059
Gradient Boost	0.75667	0.77556
AdaBoost	0.75833	0.77057
Number of faces		
naive Bayes	0.68666	0.71820
SVM	0.71500	0.69077
Gradient Boost	0.71500	0.69077
AdaBoost	0.71500	0.693266
CNN for images		
Sequential CNN	0.50000	0.50000
Neural network results		
SoftMax	0.972	0.96

References

1. Y. Yang, L. Zheng, J. Zhang, Q. Cui, Z. Li, P.S. Yu, TI-CNN: convolutional neural networks for fake news detection. arXiv preprint arXiv:1806.00749 (2018)
2. M. Alrubaian, M. Al-Qurishi, M. Hassan, A. Alamri, A credibility analysis system for assessing information on Twitter. IEEE Trans. Dependable Secur. Comput. **PP**(99), 1–1
3. V. Kazemi, J. Sullivan, One millisecond face alignment with an ensemble of regression trees, in *2014 IEEE Conference on Computer Vision and Pattern Recognition*, Columbus, OH (2014), pp. 1867–1874. https://doi.org/10.1109/CVPR.2014.241
4. A. Campan, A. Cuzzocrea, T.M. Truta, Fighting fake news spread in online social networks: actual trends and future research directions, in *2017 IEEE International Conference on Big Data (Big Data)* (2017), pp. 4371–4375
5. Z. Jin, J. Cao, Y. Zhang et al., Novel visual and statistical image features for microblogs news verification. J. IEEE Trans. Multimed. **19**(3), 598–608 (2017)
6. K. Shu, A. Sliva, S. Wang, J. Tang, H. Liu, Fake news detection on social media: a data mining perspective. ACM SIGKDD Explor. Newsl. **19**(1), 22–36 (2017)
7. A. Figueira, L. Oliveira, The current state of fake news: challenges and opportunities. Procedia Comput. Sci. **121**, 817–825 (2017). https://doi.org/10.1016/j.procs.2017.11.106
8. E. Elmurngi, A. Gherbi, Detecting fake reviews through sentiment analysis using machine learning techniques, in *IARIA/DATA ANALYTICS 2017 the Sixth International Conference on Data Analytics*, November, Barcelona, Spain (2017), pp. 65–72. ISBN: 978-1-61208-603-3
9. S. Singhania, N. Fernandez, S. Rao 3HAN: A Deep Neural Network for Fake News Detection (2017). https://doi.org/10.1007/978-3-319-70096-0_59
10. V. Prez-Rosas, B. Kleinberg, A. Lefevre, R. Mihalcea, Automatic Detection of Fake News (2017)
11. A. Roy, B. Kingshuk, A. Ekbal, P. Bhattacharyya, A Deep Ensemble Framework for Fake News Detection and Classification (2018). CoRR abs/1811.04670
12. H. Thakur, A. Gupta, A. Bhardwaj, D. Verma, Rumor detection on twitter using a supervised machine learning framework. Int. J. Inf. Retr. Res. **8**, 1–13 (2018). https://doi.org/10.4018/IJIRR.2018070101
13. J. Fontanarava, G. Pasi, M. Viviani, An ensemble method for the credibility assessment of user-generated content, in *Proceedings of the International Conference on Web Intelligence (WI '17)* (ACM, New York, NY, USA, 2017), pp. 863–868. https://doi.org/10.1145/3106426.3106464

A Patchy Ground Antenna for Wide Band Transmission in S-Band Application



Anurag Saxena, Vinod Kumar Singh and Ashutosh Kumar Singh

1 Introduction

For improving the bandwidth of textile microstrip antenna, there are various methods such as using low dielectric material, impedance matching, feeding techniques, making partial ground, etc. This type of antenna can further be used in rectenna circuit for the measurement of power of RF energy presented in the environment. Different types of energy sources such as thermal, wind, solar, and RF energy are available in the environment [1–3].

Due to very less weight and size of presented antenna, it can easily be foldable or wearable. This antenna requires very less space to install. The antenna can be simulated in three steps. The first step is to make partial ground of copper material which is also known as patchy ground [4–7]. For this, copper tape is used at the backside of substrate. Second, textile material is required in which patch will be designed. In the last step, patch with copper tape is designed on the front surface of textile. For establishment of antenna, the apparatus required exceptionally less space. Body remote correspondence and Personal Area Network (PAN) structure are fundamentally found on antenna apparatus [8–12].

A. Saxena · V. K. Singh (✉)
Department of Electrical Engineering, SR Group of Institutions, Jhansi, UP, India
e-mail: singhvinod34@gmail.com

A. Saxena
e-mail: anurag.saxena.anurag@gmail.com

A. K. Singh
Indian Institute of Information Technology, Allahabad, Allahabad, India
e-mail: ashutosh_singh@iiita.ac.in

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_47

2 Designing of Antenna

For Wide Band (WB) communication, presented antenna is simulated in CST software in which front and back view of presented antenna is shown in Fig. 1. The dimension of partial ground which is also known as patchy ground is 53.3 * 14.9 mm. For a newly designed shape, the outer ring is made using copper tape having radius of 18 mm. A portion of ring has been removed by another circle having radius of 14 mm. Then a diamond shape is designed having radius of 10 mm. For feeding the antenna, a microstrip feed strip line is made having dimension of 2 * 28.9 mm. The details of parameters used in the designing of antenna are described in Table 1.

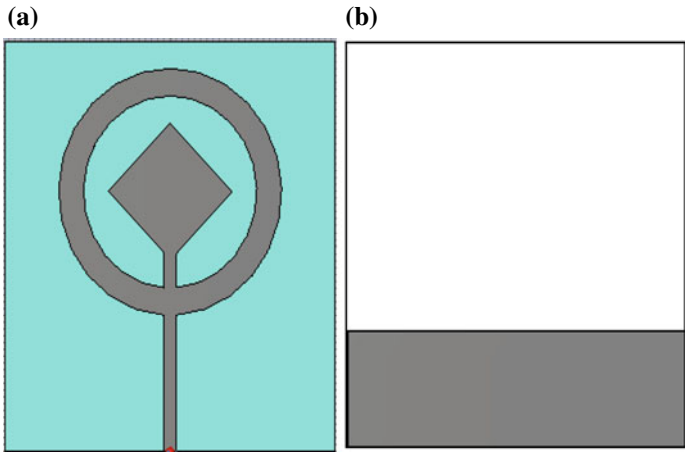


Fig. 1 Antenna shape **a** patch front **b** ground plane

Table 1 Parameters of proposed textile antenna

Antenna parameter	Values
Thickness of textile (mm)	1
Dielectric constant	1.7
Loss tangent	0.025
Patchy ground (mm)	53.3 × 14.9
Substrate dimension (mm)	53.3 × 59.8
Outer patch radius (mm)	18
Center circular cut radius (mm)	14
Microstrip feed line dimension (mm)	2 × 28.9
Center inner diamond radius (mm)	10

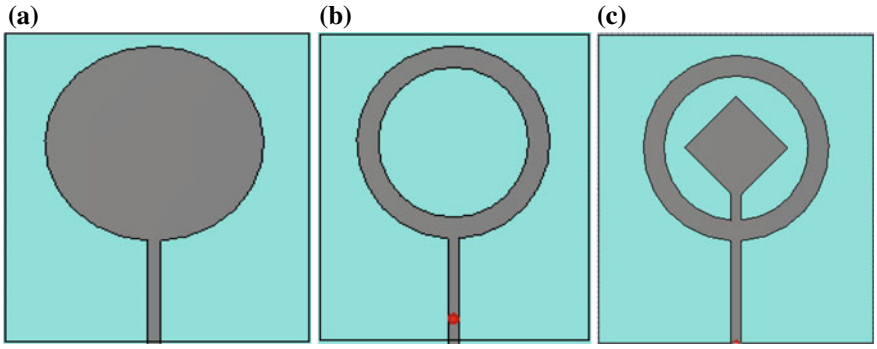


Fig. 2 Different design geometries **a** antenna-1, **b** antenna-2, and **c** presented antenna-3

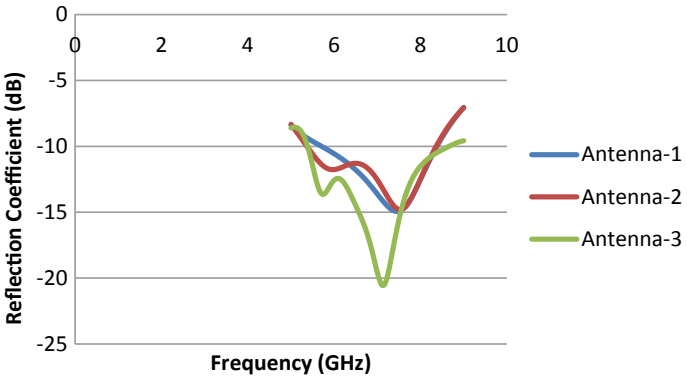


Fig. 3 Reflection coefficient versus frequency of different geometries

3 Optimization of Various Antenna Geometries

Antenna-1 and Antenna-2 are the initial designs and after modification, the antenna-3 (proposed antenna) has been simulated in CST software that gives desired results which is shown in Fig. 2. In the anticipated paper, antenna-3 is best suitable for S-Band and defense applications because it has large bandwidth as compared to antenna-1 and antenna-2 which are shown in Fig. 3 which indicates the graph between return loss versus frequency. Therefore, for various applications, antenna-3 was selected and designed in CST software which is suitable for wireless transmission.

4 Radiation Pattern

Polar plot radiation pattern which is generated by CST simulation software of proposed antenna at 6.17 GHz is shown in Fig. 4. This indicates the main lobe direction

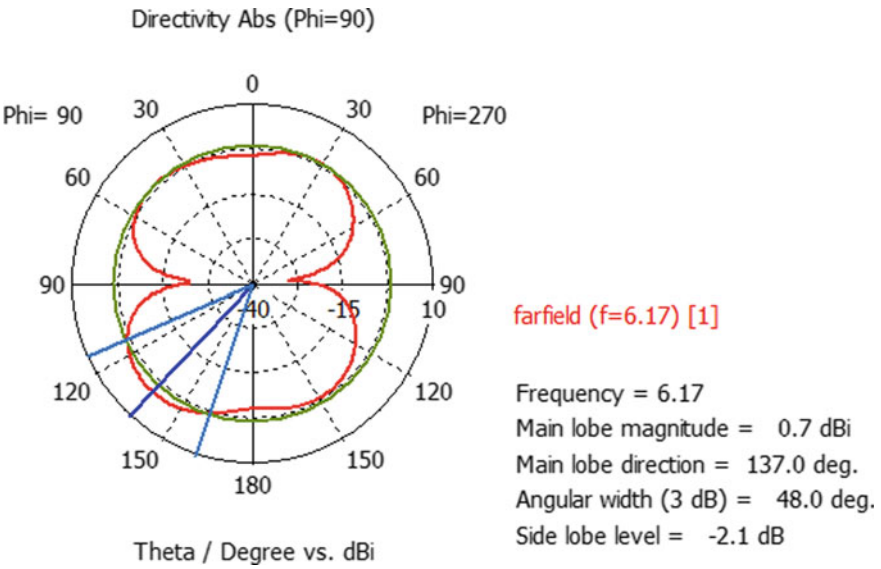


Fig. 4 Polar plot radiation pattern of anticipated textile microstrip antenna at 7.5 GHz

$= 137.0°$ and angular width (3 dB) $= 48°$ and main lobe magnitude $= 0.7$ dBi at $\varphi = 90$. Also, 3-D radiation pattern of proposed antenna at 6.17 GHz is shown in Fig. 5. The anticipated antenna gives excellent radiation efficiency of about -4.269 dB and also the directivity of 3.406 dBi.

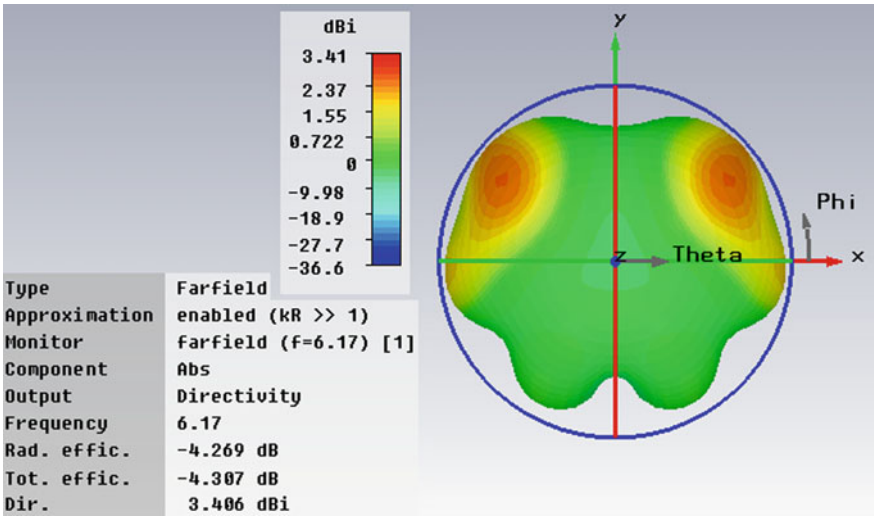


Fig. 5 3-D radiation pattern at resonant frequency 7.5 GHz

5 Conclusion

This anticipated antenna is remarkably helpful for the utilization of body-driven remote and satellite communication which is directional in nature. This patchy ground antenna used textile material for designing of substrate that resonates at 6.17 GHz. The overall bandwidth of this antenna is 66.6% from 4 to 8 GHz range which is useful in S-Band applications. Further, it can be used for various wireless transmissions.

References

1. S. Lemey, F. Declercq, H. Rogier, Textile antennas as hybrid energy-harvesting platforms. Proc. IEEE **102**(11) (2014)
2. K.-L. Wong, Compact and broadband microstrip antennas (Wiley, 2002). ISBNs 0- 471-41717-3 (Hardback)
3. A. Saxena, V.K. Singh, S.B. Mohini, G.-S. Chae, A. Sharma, A.K. Bhoi, Rectenna circuit at 6.13 GHz to operate the sensors devices. Int. J. Eng. Technol. **7**(2.33), 644–646 (2018)
4. V.K. Singh, S. Dhupkariya, N. Bangari, Wearable ultra wide dual band flexible textile antenna for WiMax/WLAN application. Int. J. Wirel. Pers. Commun. **95**(2), 1075–1086 (2017). Springer. ISSN 0929-6212
5. A. Saxena, V.K. Singh, A moon-strip line antenna for multi-band applications at 5.44 GHz resonant frequency, in *4th International Conference on Advances in Electrical, Electronics, Information, Communication and Bio-Informatics (AEEICB-18)*, August (2018)
6. V. Raghupatruni, R. Krishna, R. Kumar, Design of temple shape slot antenna for ultra wideband applications. Prog. Electromagn. Res. B **47**, 405–421 (2013)
7. B. Naresh, V.K. Singh, V, Bhargavi, Low power circularly polarized wearable rectenna for RF energy harvesting, in *Advances in Power Systems and Energy Management. Lecture Notes in Electrical Engineering*, ed. by A. Garg, A. Bhoi, P. Sanjeevikumar, K. Kamani, vol. 436 (Springer, Singapore, 2018), pp. 131–138
8. A. Saxena, V.K. Singh, A watch E-cut shape textile antenna for WB applications. J. Microwav. Technol. **5**(1), 29–32 (2018)
9. V.K. Singh, A. Saxena, Two parabolic shape microstrip patch antenna for single band application. J. Microwav. Technol. **5**(1), 33–36 (2018)
10. P. Van Torre, L. Vallozzi, H. Rogier, J. Verhaevert, Diversity textile antenna systems for firefighters, in *2010 Proceedings of the Fourth European Conference on Antennas and Propagation (EuCAP)*, 12–16 April (2010)
11. N. Singh, A.K. Singh, V.K. Singh, Design & performance of wearable ultra wide band textile antenna for medical applications. Microwav. Opt. Technol. Lett. **57**(7), 1553–1557 (2015). Wiley Publications, USA (ISSN: 0895-2477)
12. B. Naresh, V.K. Singh, V. Bhargavi, A. Garg, A.K. Bhoi, Dual-band wearable rectenna for low-power RF energy harvesting, in *Advances in Power Systems and Energy Management* ed. by A. Garg, A. Bhoi, P. Sanjeevikumar, K. Kamani. Lecture Notes in Electrical Engineering, vol. 436 (Springer, Singapore, 2018), pp. 13–21

Electromagnetic Scattering From Two Crossed Ferromagnetic Microwires



Tarun Kumar, Rajeev Kamal and Abhinav Sharma

Electromagnetic scattering from a ferromagnetic microwire has renewed interest and fervor among the researchers due to its applications metamaterials (MTMs) [1–7]. Ferrites are known for their application in nonreciprocal microwave devices. Recently, many authors have reported the application of ferrites in the designing of double negative (DNG) metamaterials due to their ability to produce negative permeability in the frequency range above ferromagnetic resonance (FMR) (see Fig. 2) [8]. Ferromagnetic resonance occurs in ferrites when **H** vector of a uniform plane wave is normal to the applied external magnetization $\mathbf{H_0}$ (i.e., TM_z polarization). Consequently, the permeability of ferrite medium turns out to be a tensor and an extraordinary wave propagates in the ferrite. The real part of effective permeability $Re[\mu_e]$ turns out to be negative for frequencies above the FMR frequency (see Fig. 2) [8]. When the **H** vector lies in other than the orthogonal plane w.r.t the direction of H_0 (i.e., TE_z polarization), no interaction between **H** and H_0 takes place. Hence, ordinary wave propagation takes place in the ferrite and the medium acts like a normal lossy dielectric medium [8] (Figs. 1 and 2).

Recently, Kumar et al. have reported in [9] that a ferromagnetic microwire grid produces negative values of effective permittivity and permeability but only for the case of TM_z polarization. The design proposed in this paper is shown in Fig. 1 whichconsists of two crossed ferromagnetic microwires kept along the y- and z-axes,

T. Kumar (✉) · A. Sharma
University of Petroleum and Energy Studies (UPES), Dehradun 248007, Uttarakhand, India
e-mail: tkumar@ddn.upes.ac.in

R. Kamal · A. Sharma
Dayananda Sagar University, Shavige Malleshwara Hills, Kumaraswamy Layout, Bengaluru 560078, Karnataka, India

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_48

respectively. Both microwires are assumed to be magnetized axially by a constant magnetization **H** and H_0. The boundary value-type solution of the scattering problem for normal and oblique incidence cases is reported by many authors [8, 10–12]. The analysis has been carried out in this paper by using tangential boundary conditions in order to calculate the unknown expansion coefficients of the scattered fields. As the result for single microwire positioned along the z-axis is already reported, only the solution for microwire placed along the y-axis is derived in this paper. After obtaining the results for microwire along the y-axis, total scattered field is obtained by combining the results of both microwires. Numerical results are obtained through MATLAB for TM_z, TE_z and at polarization angle $\alpha_0 = 45°$. Numerical results show that ferromagnetic resonance takes place in the structure at any arbitrary value of polarization angle.

$$\overline{\overline{\mu}} = \begin{pmatrix} \mu & j\kappa & 0 \\ -j\kappa & \mu & 0 \\ 0 & 0 & \mu_0 \end{pmatrix}.$$

where

$$\mu = \mu_0 \left(1 + \chi_p - j\chi_s\right), \tag{1}$$

$$\kappa = \mu_0 \left(K_p - jK_s\right), \tag{2}$$

$$\chi_p = \frac{\omega_0\omega_m\left(\omega_0^2 - \omega^2\right) + \omega_0\omega_m\omega^2\alpha^2}{\left[\omega_0^2 - \omega^2\left(1 + \alpha^2\right)\right]^2 + 4\omega_0^2\omega^2\alpha^2}, \tag{3}$$

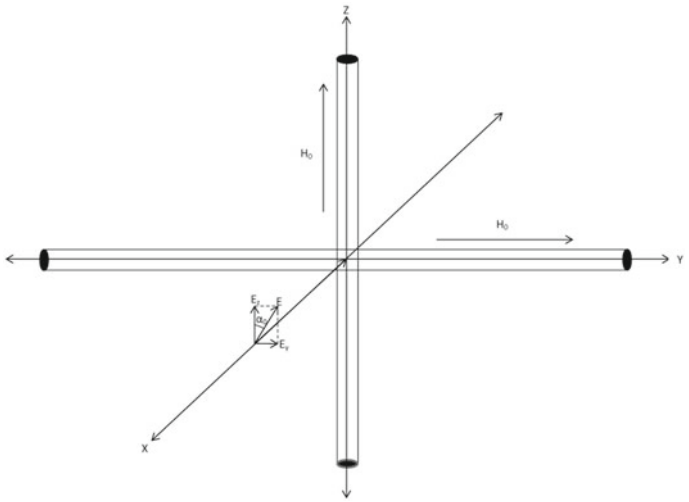


Fig. 1 Two crossed ferromagnetic microwires

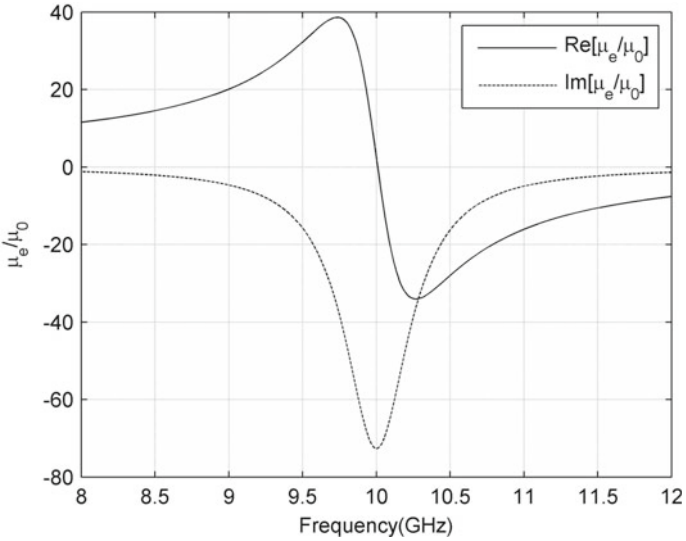


Fig. 2 $Re[\mu_e]$ and $Im[\mu_e]$ for the ferrite medium (Iñgio et al.) [8]

$$\chi_s = \frac{\omega_0\omega_m\alpha\left[\omega_0^2 + \omega^2\left(1+\alpha^2\right)\right]}{\left[\omega_0^2 - \omega^2\left(1+\alpha^2\right)\right]^2 + 4\omega_0^2\omega^2\alpha^2}, \tag{4}$$

$$K_p = \frac{\omega_0\omega_m\alpha\left[\omega_0^2 - \omega^2\left(1+\alpha^2\right)\right]}{\left[\omega_0^2 - \omega^2\left(1+\alpha^2\right)\right]^2 + 4\omega_0^2\omega^2\alpha^2}, \tag{5}$$

$$K_s = \frac{2\omega_0\omega_m\omega^2\alpha}{\left[\omega_0^2 - \omega^2\left(1+\alpha^2\right)\right]^2 + 4\omega_0^2\omega^2\alpha^2}. \tag{6}$$

Effective permeability μ_e and effective propagation constant β_e are, respectively, given as [8]

$$\mu_e = \frac{\mu^2 - \kappa^2}{\mu}, \tag{7}$$

$$\beta_e^2 = \omega^2\mu_e\epsilon_c, \tag{8}$$

where ϵ_c is the complex permittivity given by

$$\epsilon_c = \epsilon_0 - j\frac{\sigma}{\omega}. \tag{9}$$

1 Calculation of The Unknown Expansion Coefficients

Ferromagnetic microwires of an infinite length with radius "a" are aligned along the y- and z-axes, respectively as shown in Fig. 1. H_0 is the applied internal axial magnetization along their axis. The solution for microwire aligned along the z-axis is already available in the literature and we will use the solution as given in [8]. Here, only the microwire aligned along the y-axis is considered for the analysis and an EM wave is incident normally with angle of polarization α_0. Here the time dependence is considered to be $e^{j\omega t}$ by default. The scattered and incident field components along the y-axis are given by

$$E_y^{inc} = E_0 \sin\alpha_0 \sum_{n=-\infty}^{+\infty} j^n J_n\left(\beta_0 \rho_y\right) e^{-jn\phi_y}, \tag{10}$$

$$E_y^{s} = E_0 \sum_{n=-\infty}^{+\infty} b_n^s H_n^{(2)}(\beta_0 \rho_y) e^{-jn\phi_y}. \tag{11}$$

The ϕ-components can be easily obtained with the help of Maxwell's equations as

$$H_{\phi_y}^{inc} = -\frac{jE_0}{\eta_0} \sin\alpha_0 \sum_{n=-\infty}^{+\infty} j^n J_n'(\beta_0 \rho_y) e^{-jn\phi_y}, \tag{12}$$

$$H_{\phi_y}^{s} = -\frac{jE_0}{\eta_0} \sum_{n=-\infty}^{+\infty} b_n^s H_n^{(2)'}(\beta_0 \rho_y) e^{-jn\phi}. \tag{13}$$

Here b_n^s is the unknown scattering field coefficient, ρ_y and ϕ_y are the radial and azimuthal coordinates w.r.t to microwire along the y-axis, and β_0 and η_0 are propagation constant and characteristic impedance of the free space, respectively. The superscripts "inc" and "s" represent the incident and the scattered fields, respectively, J_n is the Bessel's function of the first kind and nth order, and $H_n^{(2)}$ is the Hankel's function of the second kind and nth order. The ' denotes derivative of first order with respect to the argument.

y and ϕ_y components of internal fields of (i.e., E_y^d and $H_{\phi_y}^d$) for the reference microwire are given in [10–12].

The boundary conditions at the surface (i.e., at $\rho = a$) of microwire positioned along the y-axis give

$$E_y^{inc} + E_y^{s} = E_y^{d}, \tag{14}$$

$$H_{\phi_y}^{inc} + H_{\phi_y}^{s} = H_{\phi_y}^{d}. \tag{15}$$

After substituting (10) to (13) into (14) and (15) and solving further yields the unknown scattering field coefficient

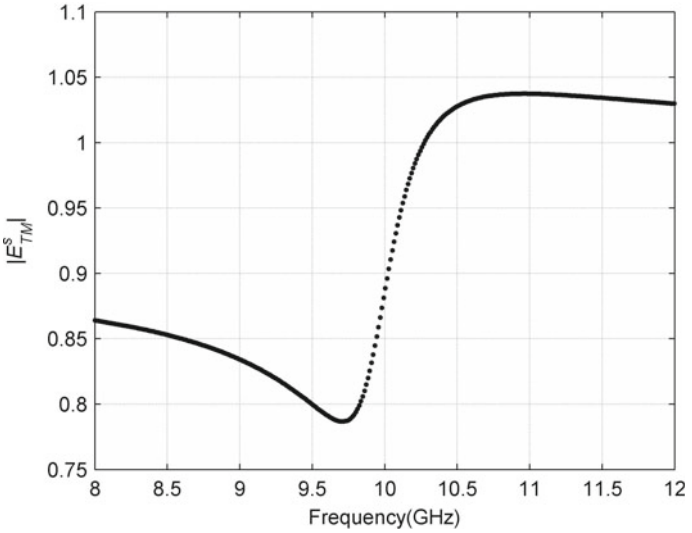


Fig. 3 Magnitude of the scattered field for TM_z polarization

$$b_n^s = j^n \sin\alpha_0 \left[\frac{D_n J_n(\beta_0 a) - J_n'(\beta_0 a) J_n(\beta_e a)}{H_n^{'(2)}(\beta_0 a) J_n(\beta_e a) - D_n H_n^{(2)}(\beta_0 a)} \right], \tag{16}$$

and the scattering field coefficient for the microwire placed along the z-axis is given in [8] as

$$a_n^s = j^n \cos\alpha_0 \left[\frac{D_n J_n(\beta_0 a) - J_n'(\beta_0 a) J_n(\beta_e a)}{H_n^{'(2)}(\beta_0 a) J_n(\beta_e a) - D_n H_n^{(2)}(\beta_0 a)} \right], \tag{17}$$

where

$$D_n = \frac{\beta_e \mu_0}{\beta_0 \mu_e} \left[J_n'(\beta_e a) - \frac{n\kappa}{a\mu\beta_e} J_n(\beta_e a) \right]. \tag{18}$$

The total scattered field can be obtained with the help of scattering field coefficients as

$$E^s_{(\text{total})} = \sqrt{[E_z^s]^2 + [E_y^s]^2}. \tag{19}$$

2 Discussion on Results

Numerical results are obtained for the microwires with following specifications as given in [8]: conductivity, $\sigma = 6.5 \times 10^4$ S/m, gyromagnetic ratio $\gamma = 1.33 \times 10^{11}\ \text{T}^{-1}\text{s}^{-1}$, saturation magnetization $\mu_0 M_s = 0.22$ T, loss factor $\alpha = 0.02$, and internal magnetization $H_0 = 213$ kA/m for X band (8–12 GHz). Figures 3 and 4

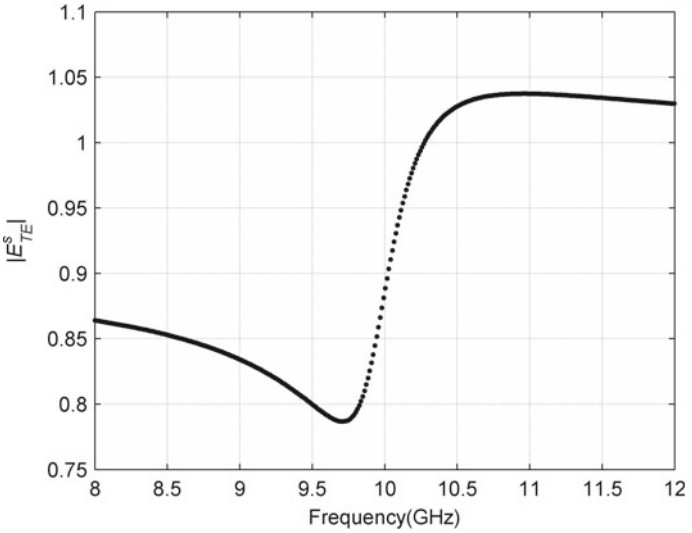


Fig. 4 Magnitude of the scattered field for TE_z polarization

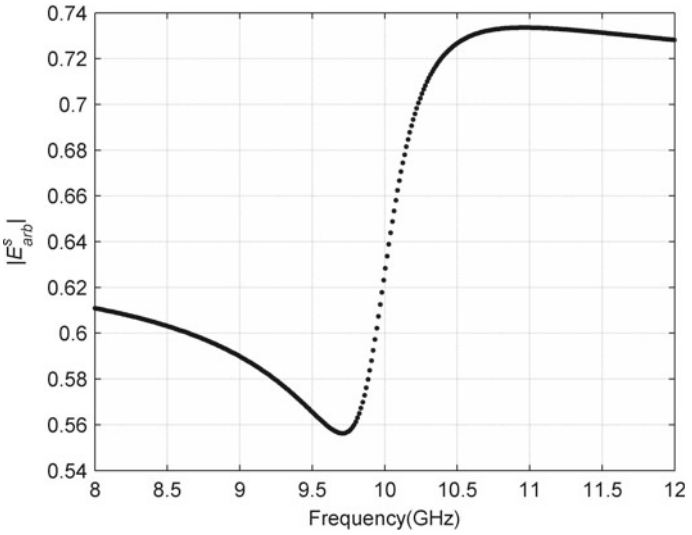


Fig. 5 Magnitude of the scattered field for $\alpha_0 = 45°$

show the magnitude of the scattered field plotted against the frequency for TM_z and TE_z polarization. The scattering behavior of the single microwire already is explained in [8] for frequencies lower and higher than the FMR frequency (i.e., 10 GHz) where $Re[\mu_e] > 0$ and $Re[\mu_e] < 0$, respectively (see in Fig. 2). For frequencies lower than the FMR frequency, $Re[\mu_e] > 0$ the ferrite medium acts like a lossy dielectric which results in weak scattering. However, for the values of frequencies above FMR, $Re[\mu_e] < 0$ due to which the imaginary part of propagation constant becomes negative. As a result, the medium essentially acts like plasma region and supports only evanescent wave which increases scattering. In both figures, it can be noticed that the scattered field is exactly same for both cases which is in contrast to the case of single microwire discussed in [8]. Figure 5 shows the magnitude of scattered field component plotted against the frequency for $\alpha_0 = 45°$. The shape of plot is similar to that shown in Figs. 3 and 4 although the magnitude is decreased by a small fraction in this case.

3 Conclusion

The numerical results presented in this paper verify that the crossed microwire structure acts in a similar manner for TM_z and TE_z polarized uniform plane wave. In the literature, scattering from a single microwire is discussed which shows different scattering behaviors for TM_z and TE_z polarization due to which the medium formed with ferromagnetic microwires produces different responses for TM_z and TE_z polarization. The proposed structure has proved the similar scattering irrespective of the polarizations. Hence, the proposed structure can be used as a unit cell in designing of a double negative metamaterial. The analysis proposed in this paper can be utilized in the designing of metamaterials.

References

1. J. Carbonell, M.H. García, D.J. Sánchez, Double negative metamaterials based on ferromagnetic microwires. Phys. Rev. B **81**, 024401-1–024401-6 (2010)
2. L. Carignan, A. Yelon, D. Ménard, Ferromagnetic nanowire metamaterials: theory and applications. IEEE Trans. Microwav. Theory Techn. **59**(10), 2568–2586 (2011)
3. H. Yongxue, H. Peng, G.H. Vincent, V. Carmine, Role of ferrites in negative index metamaterials. IEEE Trans. Magn. **42**(10), 2852–2854 (2006)
4. J.B. Pendry, A. Holden, W. Stewart, I. Youngs, Extremely low frequency plasmons in metallic mesostructures. Phys. Rev. Lett. **76**(25), 4773–4776 (1996)
5. D.R. Smith, W.J. Padilla, N.S.C. Nemat, S. Schultz, Composite medium with simultaneously negative permeability and permittivity. Phys. Rev. Lett. **84**, 4184–4187 (2000)
6. R.A. Shelby, D.R. Smith, S. Schultz, Experimental verification of a negative index of refraction. Science **292**, 77–79 (2001)
7. V.G. Veselago, The electrodynamics of substances with simultaneously negative values of ϵ and μ. Sov. Phys. Uspekhi **10**(4), 509–514 (1968)

8. I. Liberal, I. Ederra, C. Gómez-Polo, A. Labrador, J.I. Pérez Landazábal, R. Gonzalo, Theoretical modeling and experimental verification of the scattering from a ferromagnetic microwire. IEEE Trans. Microwav. Theory Techn. **59**(3), 517–526 (2011)
9. T. Kumar, N. Kalyansundaram, A novel DNG medium formed by ferromagnetic microwire grid. Prog. Electromagn. Res. B **74**, 155–171 (2017)
10. W.H. Eggimann, Scattering of a plane wave on a ferrite cylinder at normal incidence. IEEE Trans. Microwav. Theory Techn. **8**(4), 440–445 (1960)
11. A.M. Attiya, M.A. Alkanhal, Generalized formulation for the scattering from a ferromagnetic microwire. ACES J. **27**(5) (2012)
12. R.A. Waldron, Electromagnetic wave propagation in cylindrical waveguides containing gyromagnetic media. J. Br. Inst. Radio Eng. **18**(10), 597–612 (1958)

Performance Analysis of Wearable Textile Antenna Under Different Conditions for WLAN and C-Band Applications



Ashok Yadav, Vinod Kumar Singh and Himanshu Mohan

1 Introduction

The use of wearable fabric materials for the advancement of microstrip antenna section has been quick because of the ongoing scaling down of remote gadgets [1, 2]. The antenna is comprised of copper, printed at the best and the base of a dielectric substrate. In today's world, there are various types of wearable antennas that play a main task for a lot of wireless applications [3–6]. These antennas are usually printed antennas that are able to bend or be forced out of form. The antenna is formed from copper, printed at upper and therefore, the bottom of the substrate material. The copper print on the upper piece of the dielectric substrate performs the radiating patch, and the print on the lower portion of the substrate performs the ground plane of the antenna [7–9]. The near field propagation plays an important task, at whatever point the low power transmit and get devices are appended with body or attire. The far-field qualities of the antenna are additionally an essential capacity, when an application communication is set up between body-worn sensors and bigger units like PC, Laptop, Mobile Phones, PDA, etc. [10–13].

If we need the great impedance bandwidth over ultra wideband range, then this design is greatly recommended and it provides high data rate and low power utilization as well. Wearable textile antenna execution is diminished because of various qualities of the human body. The human body performs as resistive component

A. Yadav
Krishna Engineering College, Ghaziabad, Uttar Pradesh, India
e-mail: ashok.biet@gmail.com

V. K. Singh (✉)
S.R. Group of Institution, Jhansi, Uttar Pradesh, India
e-mail: singhvinod34@gmail.com

H. Mohan
College of Technology and Engineering, Udaipur, Rajasthan, India
e-mail: himanshumohan.dtu@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_49

and gives high transmission degradation between body and antenna [14–16]. In any case, in the event that the human body is incorporated with a viable remote systems administration framework, it offers great transmission between antennas. For the short distance remote inclusion, the body-centric communication system is especially helpful which can be connected with human bodies or garments. Improved comforts and services are applicable with the help of networking. On the other hand, when electromagnetic radiation is utilized, the well-being of a person is the primary concern [17].

In this article, a unique design of a wearable fabric antenna with microstrip feed line technique is exhibited to upgrade the antenna attributes and different wireless applications.

2 Design of Proposed Antenna

In this paper, a rectangular slot loaded patch with microstrip feed line is fabricated on a fabric substrate of jeans with a dimension of 40 × 40 mm^2. Adhesive tape of copper is used for development of radiating patch as well as partial ground plane. The copper tape comprises a foil of copper which is self-adhesive. The tape has double conductivity. This will allow the current to move through the two sides. The partial ground plane of 40 × 20 mm^2 presents a favorable task of impedance matching circuit. This enables the antenna to function for wideband applications. Calculated by Eq. 1, the radius of patch comes out to be 11 mm, while microstrip feed width and length are taken from basic equations of microstrip antenna. Practical utilization of fabric antennas is most advantageous when fabric antennas are fused into apparel. To excite the textile antenna, SMA connector is directly connected with microstrip line feed technique. The simulation work is completed in HFSS antenna design software [18]. The design of proposed antenna is shown in Fig. 1 and.the relative permittivity of jeans substrate material is 1.7 whereas height of jeans substrate is 1 mm, respectively. The partial ground plane is also designed with copper tape of

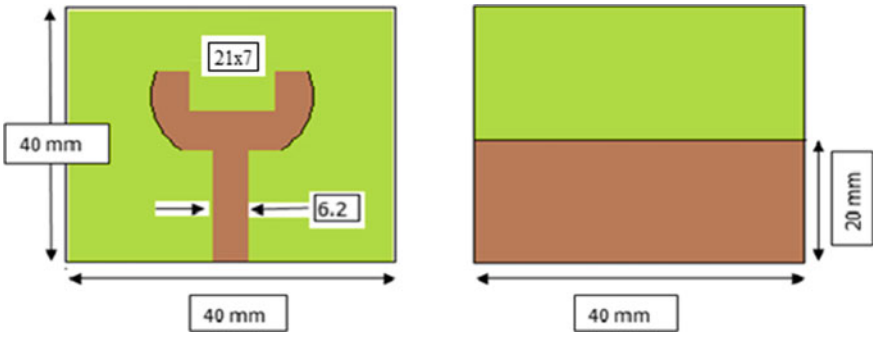


Fig. 1 Conducting patch and ground plane of the proposed antenna

Table 1 Proposed antenna parameters

Parameter	Dimension
Substrate thickness	1 mm
Dielectric constant	1.7
Length of square	40 mm
Width of square	40 mm
Partial ground	40 mm × 20 mm
Slot dimension	21 mm × 7 mm
Feed width	6.2 mm
Loss tangent	0.025

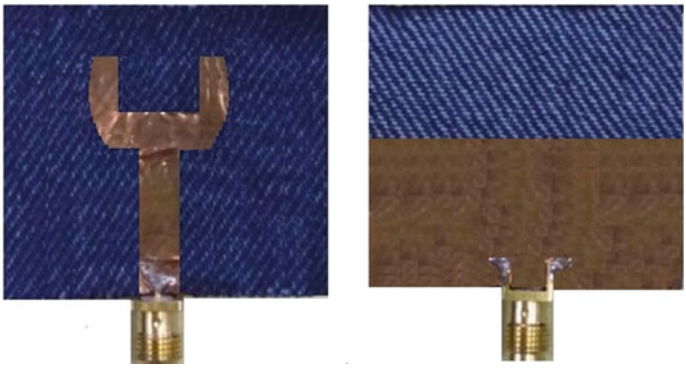


Fig. 2 Fabricated conducting patch and ground plane of the proposed antenna

0.03 mm thickness. All the antenna parameters are given in Table 1 and snapshot of fabricated proposed antenna is shown in Fig. 2.

$$r = \frac{87.94}{\sqrt[f_r]{\varepsilon_r}} \tag{1}$$

where

R Radius of circular patch in mm
f_r Resonating frequency in GHz
ε_r Relative permittivity

3 Effect on Scattering Coefficient Under Variation of Parameters

Figure 3 explains the parameter variation study of the proposed antenna. In this figure, two different parameters are varied and presented in this graph. By varying substrate height at 1 mm and microstrip feed width line at 7 mm, 8 mm, and 9 mm, respectively, antenna characteristic is analyzed. Similarly, in Figs. 4 and 5, effect of variation on different substrate height at 0.8, 1.2 mm and microstrip line feed at 7mm, 8 mm, 9 mm, respectively presents that the return loss is changed during variation in both situations. This change in return loss is applicable for the good impedance matching. It is seen in Fig. 4 that by the variation of substrate height at 0.8 mm, the resonating frequency is changed at 5.5 GHz on feed width of 9 mm but the other variation feed width 7, 8 mm there is no change in resonating frequency. In Fig. 5, slightly variation in resonating frequency has occurred on feed width of 7 mm, 8 mm, and 9 m, respectively.

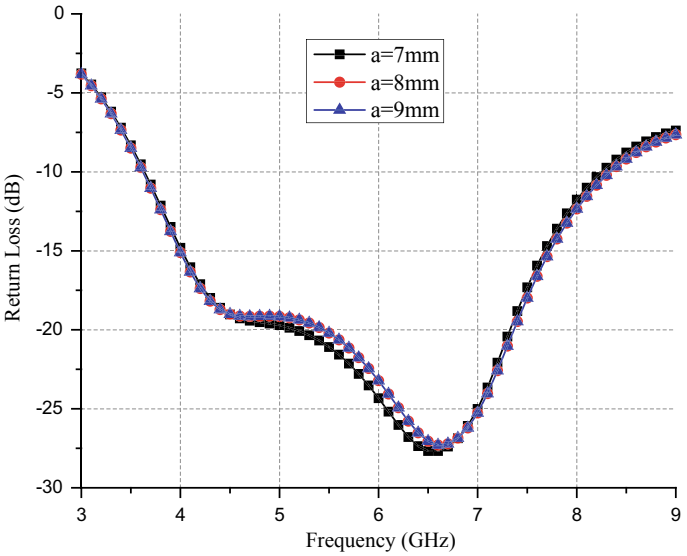


Fig. 3 Effect of variation on substrate height at 1 mm and different feed width variation

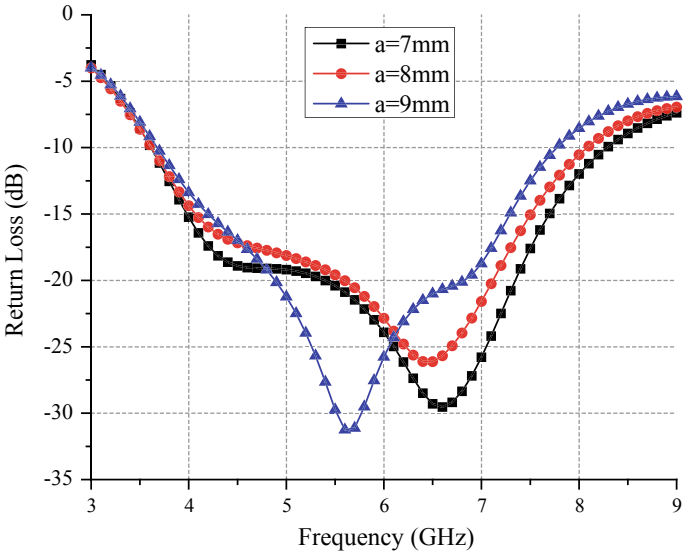


Fig. 4 Effect of variation on substrate height at 0.8 mm and different feed width variation

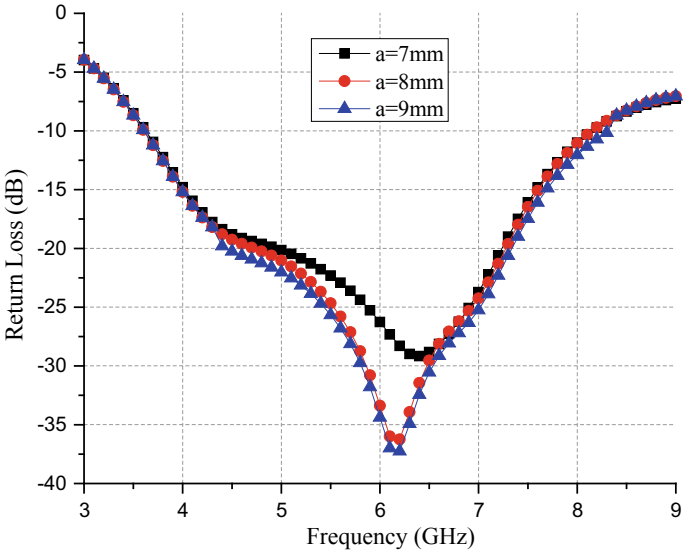


Fig. 5 Effect of variation on substrate height at 1.2 mm and different feed width variation

4 Results and Discussion

4.1 S_{11} Parameter of Proposed Antenna

Figure 6 explains about simulated and measured return loss versus frequency plot. From the graph, simulated impedance bandwidth of 77% is achieved between frequency 3.7–8.3 GHz and measured impedance bandwidth is 83.76% between 3.4–8.3 GHz, respectively under magnitude of S_{11} <–10 dB.

4.2 Performance of Proposed Antenna Under Various Conditions

The performance of the proposed antenna is analyzed on vector network analyzer under three different conditions. Figure 7 illustrates the antenna return loss characteristics under normal condition, wet condition, and bent condition, respectively. From Fig. 6, it is observed that the variations on resonance frequencies are obtained and impedance bandwidth is almost same (83.76% between 3.4–8.3 GHz).

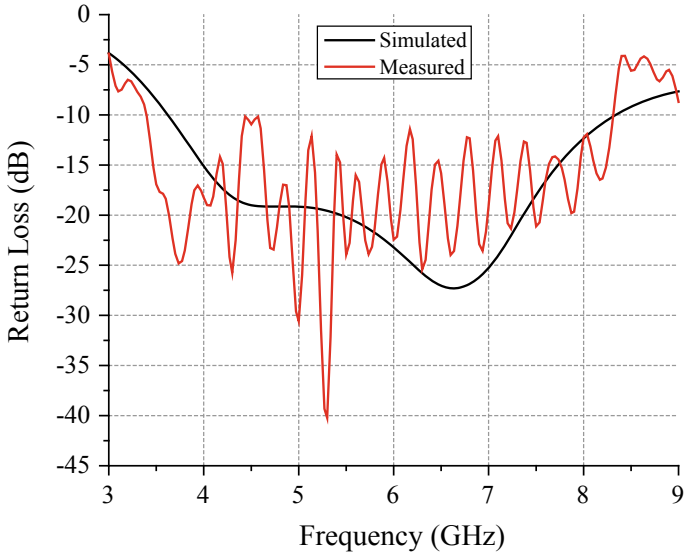


Fig. 6 Simulated and measured return loss versus frequency plot

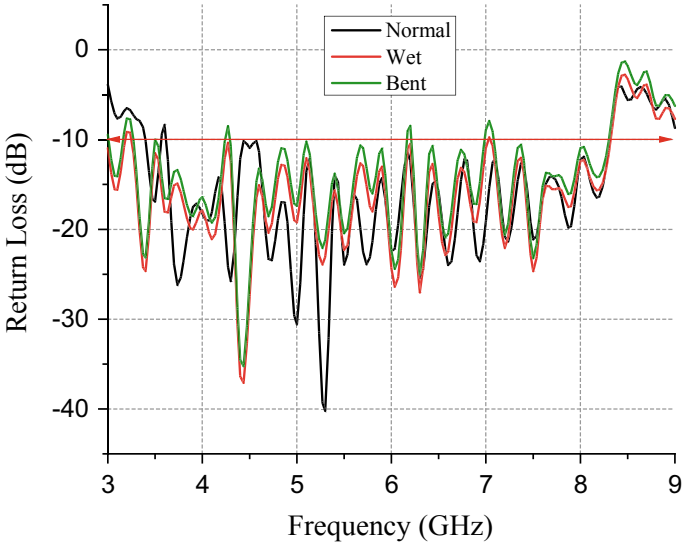


Fig. 7 Measured return loss versus frequency plot under various conditions

4.3 Comparison of Gain and Radiation Characteristics of Fabric Antenna

Simulated and measured gain versus frequency plot is presented in Fig. 8. It is found that the peak realized gain of 5.25dBi is obtained at 8.2 GHz. Figure 9 presents the radiation pattern at 6.13 GHz in x–z plane. The simulated and measured radiation patterns are almost closer and spread in all directions around 360° i.e., omni direction pattern. Figure 10 explains about the radiation pattern at 6.13 GHz in y–z plane. Simulated and measured patterns are almost closer.

5 Conclusion

The proposed textile antenna is designed and tested through a vector network analyzer. The results obtained from measurement and simulation are very much closer and established good agreement. The designed textile antenna is applicable for WLAN (5.15–5.35 GHz) HIPER LAN/2 (5.15–5.350 GHz, 5.470–5.725 GHz) and C-Band (4–8 GHz) applications. Notwithstanding, the simple reduced size of the textile antenna further affirms its propriety for compact wireless devices. The antenna also proves its good applicability for networking applications as it showed good performance in WLAN frequency band. Good performance in C-Band increases its applicability in satellite applications also.

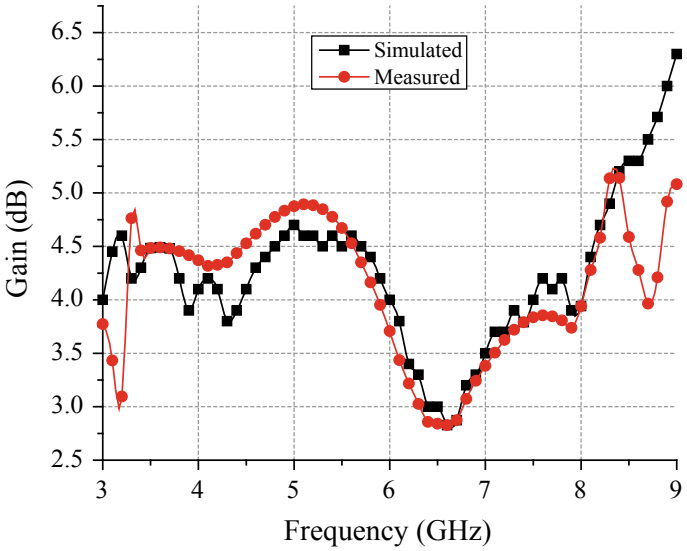


Fig. 8 Simulated and measured gain versus frequency plot

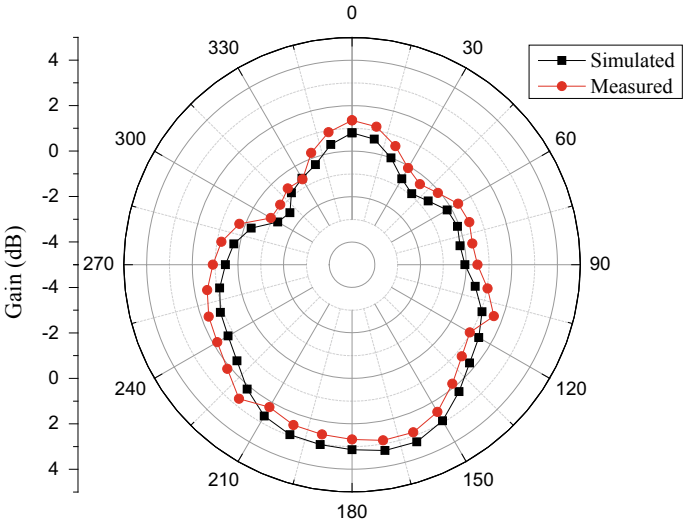


Fig. 9 Simulated and measured radiation pattern in x–z plane at 6.13 GHz

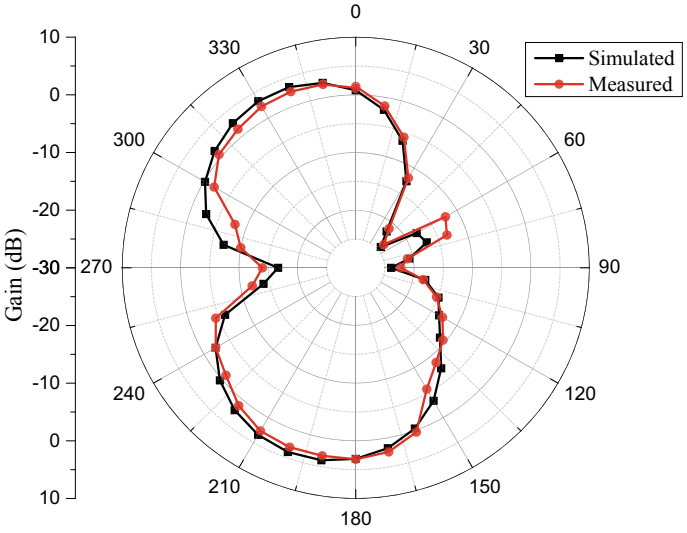


Fig. 10 Simulated and measured radiation pattern in y–z plane at 6.13 GHz

References

1. S. Hall, Y. Hao, Antennas and propagation for body-centric wireless communications. ARTECH HOUSE (2006)
2. Guo-Ping Gao, Hu Bin, Xiao-Long Tian et al. Experimental study of a wearable aperture-coupled patch antenna for wireless body area network. Microw. Opt. Technol. Lett. **59**(4), 761–766 (2017)
3. B. Hu, G.P. Gao, L.L. He, X.D. Cong, J.N. Zhao, Bending and on-arm effects on a wearable antenna for 2.45 GHz body area network. Antennas Wirel. Propag. IEEE Lett. **15,** 378–381 (2016)
4. V.K. Singh, S. Dhupkariya, N. Bangari, Wearable ultra wide dual band flexible textile antenna for WiMax/WLAN application. Int. J. Wirel. Personal Commun. **95**(2), 1075–1086 (2017)
5. Z.H. Jiang, D.E. Brocker, P.E. Sieber, D.H. Werner, A compact, low-profile metasurface enabled antenna for wearable medical body-area network devices. IEEE Trans. Antennas Propag. **62**(8), 4020–4030 (2014)
6. S. Sankaralingam, B. Gupta, Development of textile antennas for body wearable applications and investigations on their performance under bent conditions. Progress Electromagn. Res. B **22**, 53–71 (2010)
7. M.A.R. Osman, M.K.A. Rahim, M. Azfar, N.A. Samsuri, F. Zubir, K. Kamardin, Design, implementation and performance of ultra-wideband textile antenna. Progress Electromagn. Res. B **27**, 307–325 (2011)
8. M.A. Osman, M.K.A. Rahim, N.A. Samsuri, M.K. Elbasheer, M.E. Ali, Textile UWB antenna bending and wet performances. Int. J. Antennas Propag. 1–12 (2012)
9. A. Mersani, L. Osman, Design of dual-band textile antenna for 2.45/5.8 GHz wireless applications. IEEE Int. Conf. Multimed. Comput. Syst. (ICMCS) (2016), pp. 397–399
10. S. Sankaralingam, B. Gupta, Experimental results on HiperLAN/2 antennas for wearable applications. Progress Electromagn. Res. C **25**, 27–40 (2012)
11. M. Grilo, F.S. Correra, Rectangular patch antenna on textile substrate fed by proximity coupling. J. Microw. Optoelectron. Electromagn. Appl. **14**, 103–112 (2015)

12. W. Chaihongsa, C. Phongcharoenpanich, Performance of textile antenna using two layers of strip line and round-off circular patch, in *IEEE Conference on Antenna Measurements and Applications (CAMA)* (2015), pp. 1–3
13. N.A. Elias, N.A. Samsuri, Effects of human body and antenna orientation on dipole textile antenna performance and SAR, in *IEEE Asia-Pacific Conference on Applied Electromagnetics (APACE 2012)*, Melaka, Malaysia (2012)
14. S. Yan, P. Jack Soh, Wearable dual-band magneto-electric dipole antenna for WBAN/WLAN applications. IEEE Trans. Antennas Propag. https://doi.org/10.1109/tap.2015.2443863
15. N. Singh, A.K. Singh, V.K. Singh, Design and performance of wearable ultra wide band textile antenna for medical applications. Microw. Opt. Technol. Lett. **57**(7), 1553–1557 (2015)
16. S. Sankaralingam, B. Gupta, Use of electro-textiles for development of wibro antennas. Progress Electromagn. Res. **16**, 183–193 (2010)
17. Y. Ouyang, W.J. Chappell, High frequency properties of electro-textiles for wearable Antenna applications. IEEE Trans. Antennas Propag. **56**(2), 381–389 (2008)
18. HFSS Simulation software, version-12, Ansys, USA

Chip Resistor Loaded Fractal Frequency Selective Surface Based Miniaturized Broadband Microwave Absorber from 2 to 18 GHz



Tanveer Kaur Suchu, Arpit Sahu, Ravi Panwar and Rajesh Khanna

1 Introduction

Over contemporary decades, frequency selective surfaces (FSSs) are getting significant interest of researchers in the field of electromagnetics [1]. FSS are aperture/patch-shaped periodic structures used for filtering operations; depending on frequency selective characteristics, they may be high, low, band-pass or band-stop filters. FSS based advanced EM structures exhibit many interesting properties, giving rise to various practical applications, such as antennas, radomes, and microwave absorbers [1, 2]. FSS impacted and sandwiched MASs have been widely reported for their electromagnetic interference (EMI), electromagnetic compatibility (EMC), stealth technology, and anechoic chamber [3, 4].

The nature-inspired fractal FSS based EM structures are getting significant interest of researchers in order to obtain a multiband frequency response [1]. Fractal-based compact broadband polarization insensitive metamaterial absorber with a bandwidth of 3.5 GHz at the dielectric thickness of 2.0 mm was proposed by Munaga et al. [5]. Based on the antenna-filter-antenna (AFA) concept, the sub-wavelength fractal elements with multilayer FSS were proposed by Li. et al. [6]. It was considered that the self-similarity property was the most significant feature of fractal geometry. Briefly, the self-similarity can be depicted structure's replication of the geometry at an alternate scale inside a similar structure. In this manner, this property of fractal structure

T. K. Suchu · R. Khanna
Department of Electronics and Communication Engineering, Thapar Institute of Engineering and Technology, Patiala, Punjab, India

T. K. Suchu
Department of Electrical Engineering, San Jose State University, San Jose, CA, USA

A. Sahu · R. Panwar (✉)
Indian Institute of Information Technology, Design and Manufacturing, Jabalpur 482005, Madhya Pradesh, India
e-mail: rpanwar.iitr@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_50

brings about a multiband or wideband behavior [6, 7]. Liu et al. described a design technique for FSS incorporated with resistive components. It was demonstrated that the FSS resonating frequency is principally controlled by the estimations of the lumped elements and to a lesser degree by the separation between these elements [8]. Alexander et al. presented the use of a nonlinear dielectric sheet for designing and implementing RF limiter metamaterial It has been described that how the inductive–capacitive resonators behave like RF limiter cells, when loaded with p-i-n diodes [9]. Nan Wang reported the novel quadruple band metamaterial absorber design and conducted simulations at microwave frequencies [10]. The design consists of dielectric substrate separating the metallic plate and periodic patterned structures resulting in good absorption at four resonance frequencies by manipulating in the periodic structure.

It is well known that the FSS shape, size, inter-element spacing, and substrate EM properties greatly affect the frequency performance of MASs [1]. A critical analysis of such design variables is very much required for an efficient MAS design. Therefore, an effort has been made to optimize and analyze design variables in the range of 2–18 GHz. The higher iterated Minkowski loop FFSS geometry has been explored in order to have a multiband frequency response of the MAS. Further, chip resistors have been incorporated to enhance the bandwidth of the fractal-based MAS. The frequency-dependent RC characteristics have been studied within correlation to the substrate dielectric constant, substrate thickness, and spacer layer thickness.

2 Design of FFSS Based MASs

In this study, Minkowski loop FFSS geometry has been explored for microwave absorbing applications. The high iterated FFSS geometry has been modeled by constituting eight rectangular splits in each arm of the conducting layer on the top metallic layer as presented in Fig. 1a, b. The present work deals with two kinds of Minkowski loop FFSS configurations: (1) Minkowski loop FFSS without chip resistors and (2) Minkowski loop FFSS loaded with the chip resistors. The rectangular splits have been designed to insert chip resistors with variable values. The numerical analysis of such structures has been carried out using CST Microwave Studio in the frequency range of 2–18 GHz. Figure 1 (c) shows FFSS based MAS without chip resistors. The MAS comprising of a three-layer configuration backed with a ground plane, perfect electric conductor (PEC) with a thickness H. A space layer has been incorporated between the FSS loaded dielectric and PEC. The proposed design is symmetrical in both x- and y-axis. The line width and resistance of the outer loop are W_{out} and R_{out}, respectively, and same for the inner loop, i.e., W_{in} and R_{in}. There are six design parameters to be considered, which include an element spacing (b), W_{out}, R_{out},W_{in}, R_{in}, and the separation between two loops (L_s) for a second-order FFSS. Each unit cell of FFSS used in the structure is having dimensions of 10 × 10 mm with 0.5 mm boundary margin. These structures are impressed periodically on a dielectric substrate with relative permittivity ε_{r1}, loss tangent $\tan_{\delta 1}$, and thickness h1. A dielectric

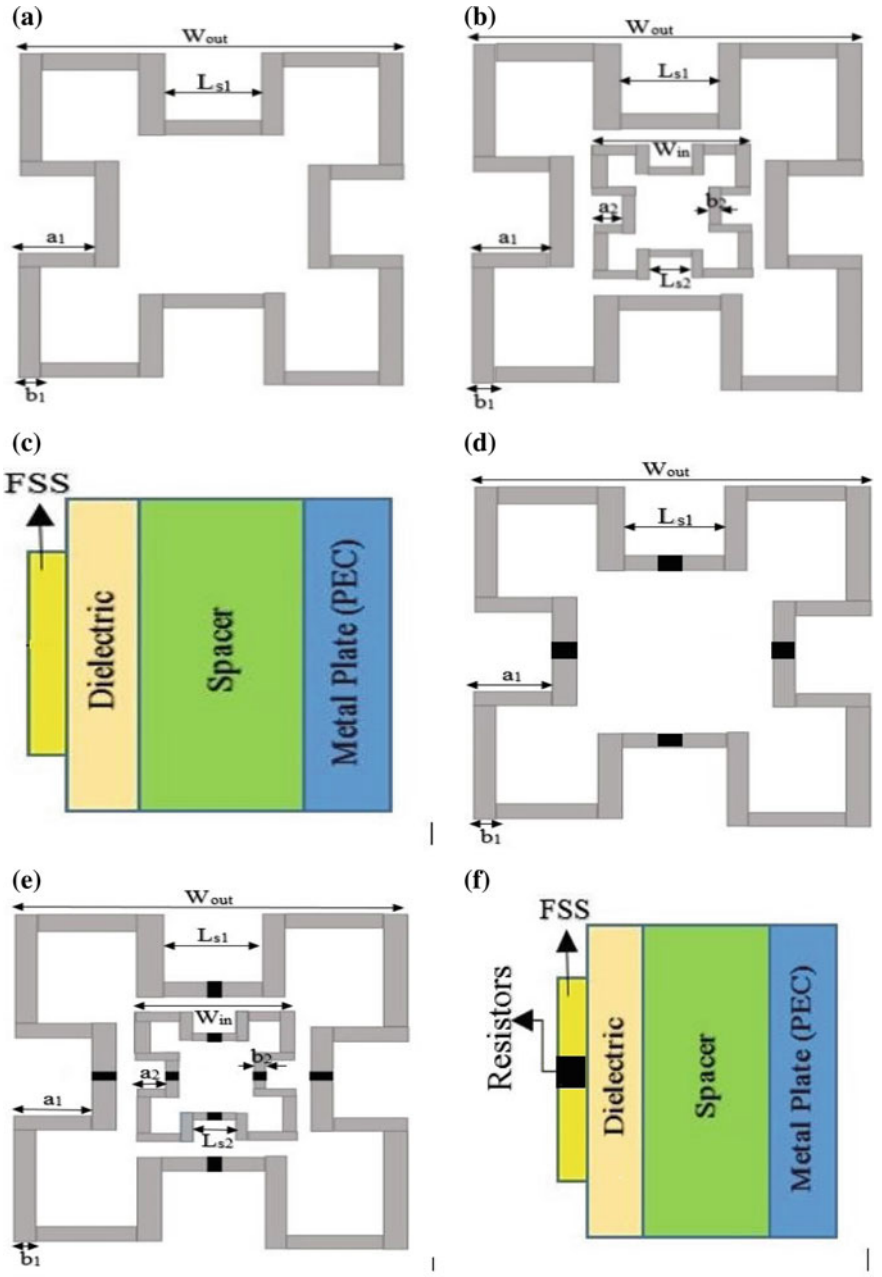


Fig. 1 Distinct MAS configurations **a** first-order FFSS, **b** second-order FFSS, **c** side view of FFSS based absorber without chip resistors, **d** chip resistor loaded first-order FFSS, **e** chip resistor loaded second-order FFSS, and **f** side view of chip resistor loaded absorber

Table 1 Optimized dimensions of chip resistor loaded FFSS based MASs

Parameters	Optimized values (mm)	Parameters	Optimized values
W_{out}	9.0	ε_{r1}	2.25
W_{in}	4.2	$\tan_{\delta 1}$	0.001
L_{S1}	3.0	$\tan_{\delta 2}$	0.0017
L_{S2}	1.4	H	2.0 mm
a1	1.5	h1	0.8 mm
a2	0.7	h2	3.8 mm
b1	0.5	R_{in}	150 Ω
b2	0.2	R_{out}	250 Ω
		ε_{r2}	1.05

spacer is placed in between the FFSS and the ground plane with permittivity ε_{r2}, loss tangent $\tan_{\delta 2}$, and thickness h2.

Similarly, the unit cell geometry of first- and second-order Minkowski loop FFSSs loaded with chip resistors are illustrated in Fig. 1d, e. Figure 1f shows chip resistor loaded Minkowski loop FFSS based MAS. It consists of a top, metallic structure (i.e., FFSS) impinged on a thin dielectric substrate with relative permittivity, i.e., ε_{r1}, loss tangent $\tan_{\delta 1}$, and thickness h1. A dielectric spacer was placed in between the FFSS and the ground plane with permittivity, i.e., ε_{r2}, loss tangent $\tan_{\delta 2}$, and thickness h2. The top layer was built of first-and second-order FFSS having split in the center of every arm of the loop. Further, the lumped resistors were inserted at the split positions of every loop. Rogers ($\varepsilon_{r1} = 2.25$) has been utilized as the dielectric substrate having dimensions mentioned previously. The optimized dimensions for the chip resistor loaded FFSS based MASs are presented in Table 1.

3 Results and Discussion

Figure 2 shows the frequency-dependent RC characteristics of first- and second-order FFSS based MASs without resistive elements. It has been observed that first-order FFSS achieves its peak RC value of –0.1 dB at the frequency of 4.6 GHz. On the other hand, when second-order FFSS is considered, it possess a resonant frequency of 10.4 GHz with the peak RC value of –1.2 dB. It is well known that the absorption bandwidth corresponding to the RC $\leq$ –10 dB is the foremost requirement to achieve better absorption characteristics. But, the require ment of wide bandwidth is not fulfilled by the FFSS loaded structures. However, the incorporation of the spacer layer can improve the bandwidth of MASs [3]. Therefore, an effort has been made to examine the effect of the spacer thickness over the RC characteristics of the proposed structures. An intact structure has been developed by adding the spacer layer ($\varepsilon' = 1.05$ and $\tan_{\delta} = 0.0017$) in between the ground plane and the dielectric

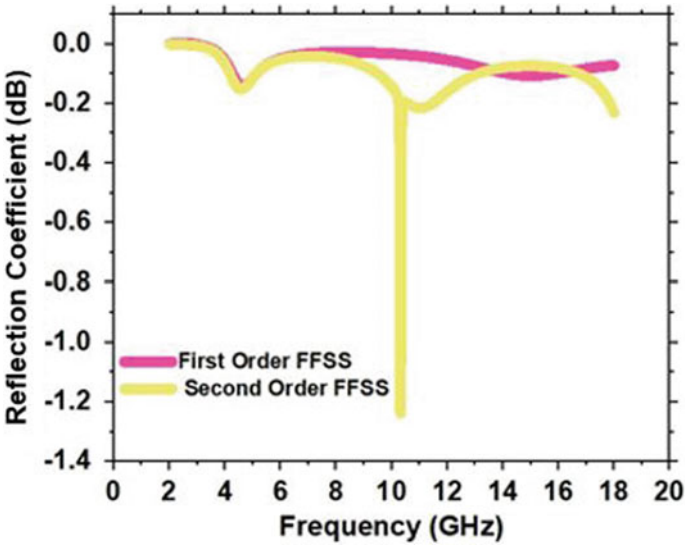


Fig. 2 Frequency-dependent RC characteristics of first- and second-order FFSS based MASs without the chip resistor in the range of 2–18 GHz

substrate. Figure 3a depicts the RC-frequency spectra of a spacer incorporated MAS. The result shows that RC and resonance frequency can be changed by introducing the spacer. Most importantly, a tremendous enhancement has been observed in the –10 dB bandwidth, as clear from Fig. 3a. MAS with 3.8 mm thick spacer layer has been found to provide a –10 dB bandwidth of 11.1 GHz (5.5–16.6 GHz). The peak RC value of –23.2 dB has been noticed for such configuration. Moreover, a shift in the resonant frequency has been noticed from higher to lower frequency regimes.

Figure 3b shows the frequency-dependent RC characteristics of the chip resistor loaded first- and second-order FFSS based MASs. The better RC characteristics are noticed for a higher iterated FFSS (i.e., second-order FFSS) as compared to first-order FFSS. A second-order FFSS structure is found to accomplish more extensive peak RC value of –23.1 dB at a frequency of 12.2 GHz with a wide bandwidth of 11.1 GHz (5.5–16.6 GHz). Now, the next step is to carry out a critical analysis of second-order FFSS in terms of substrate dielectric constant, substrate thickness, spacer layer thickness, and chip resistor resistance values over RC characteristics of MAS. The frequency-dependent RC characteristics can be easily controlled by changing the resistance values of the chip resistors [1]. In order to examine the effect of chip resistor insertion, two resistors, i.e., R_{in} and R_{out} have been considered, where R_{out} is embedded in the outermost loop, and R_{in} are incorporated in the inner loop as mentioned earlier in the design section. The RC characteristics of the MAS with variable resistance values of R_{out} and R_{in} are depicted in Fig. 4a, b, respectively. It has been observed that, as the value of the resistance varies, the RC characteristics are significantly affected. Moreover, the absorption bandwidth is broadened to the

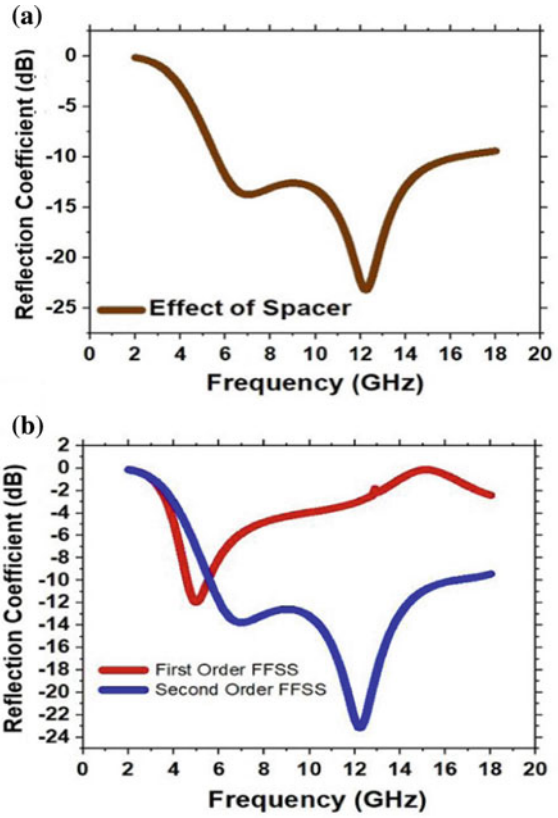


Fig. 3 Effect of space layer incorporation
a Frequency-dependent RC characteristics of MASs
b first- and second-order FFSS geometries loaded with chip resistors

greatest extent with the –10 dB bandwidth of 11.1 GHz. Figure 4a shows the RC characteristics of chip resistor loaded FFSS based MAS, where the resistance value of R_{out} is varied in the range of 150–450 Ω keeping R_{in} as constant. As the value of R_{out} varies, it influences the RC characteristics of the structure. At $R_{out} = 250$ Ω, the −10 dB bandwidth of MASs noticeably increased by the value of 11.1 GHz. Furthermore, the value of R_{in} is varied from 150 to 450 Ω, keeping R_{out} constant as shown in Fig. 4b, The improved RC characteristics have been achieved at an optimal resistance value of $R_{in} = 150$ Ω with a peak RC value of –23.1 dB at 12.2 GHz. From the abovementioned analysis, it can be concluded that enhanced RC characteristics and bandwidth can be achieved by the chip resistors incorporation with an optimal resistance value.

To examine the effect of dielectric constant over RC characteristics, the dielectric constant values have been varied in the range of 2–13. The materials with $\varepsilon' = 2.25$ (Rogers), 5.9 (ferro), 10 (Arlon 1000), and 12.9 (Gallium Arsenide) have been chosen as a material for the substrate. The frequency-dependent RC characteristics of MAS with distinct substrates are shown in Fig. 5a. It is clearly observed that the –10 dB bandwidth of the MAS decreases with increasing ε′. MAS with ε′ value of

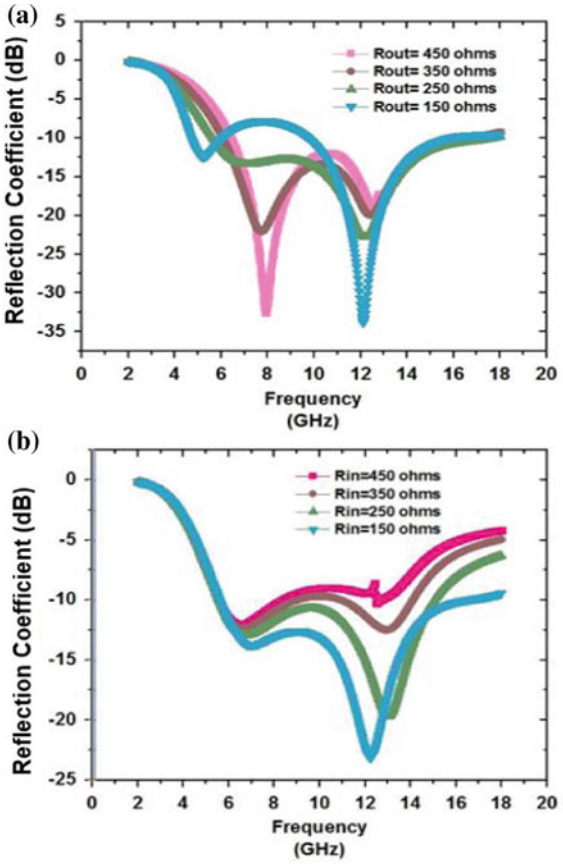


Fig. 4 Frequency-dependent RC characteristics of MASs with varied resistance **a** outer loop (R_{out}) and **b** inner loop (R_{in}) in the range of 2–18 GHz

2.25 is found to provide broadband absorption characteristics as compared to the high dielectric constant substrate.

The analyzed results of MAS for different values of substrate thickness (t) ranging from 0.8 to 3.2 mm are presented in Fig. 5b. It is seen that when t increases, absorption results get deteriorated. A peak RC value of –23.1 dB at a frequency of 12.2 GHz with a bandwidth of 11.1 GHz (5.5–16.6 GHz) has been observed in an optimal sample. A similar effort has been made to study the effect of space layer thickness and the corresponding results are depicted in Fig. 5c. When the spacer thickness increases from 1.8 to 3.8 mm, the resonant frequency shifts toward the higher frequency. It was clearly observed that bandwidth increased up to 11.1 GHz with a peak RC value of –23.1 dB at 12.2 GHz. Therefore, one can say that RC characteristics strongly depends on the design parameters of the MAS in a particular frequency regime.

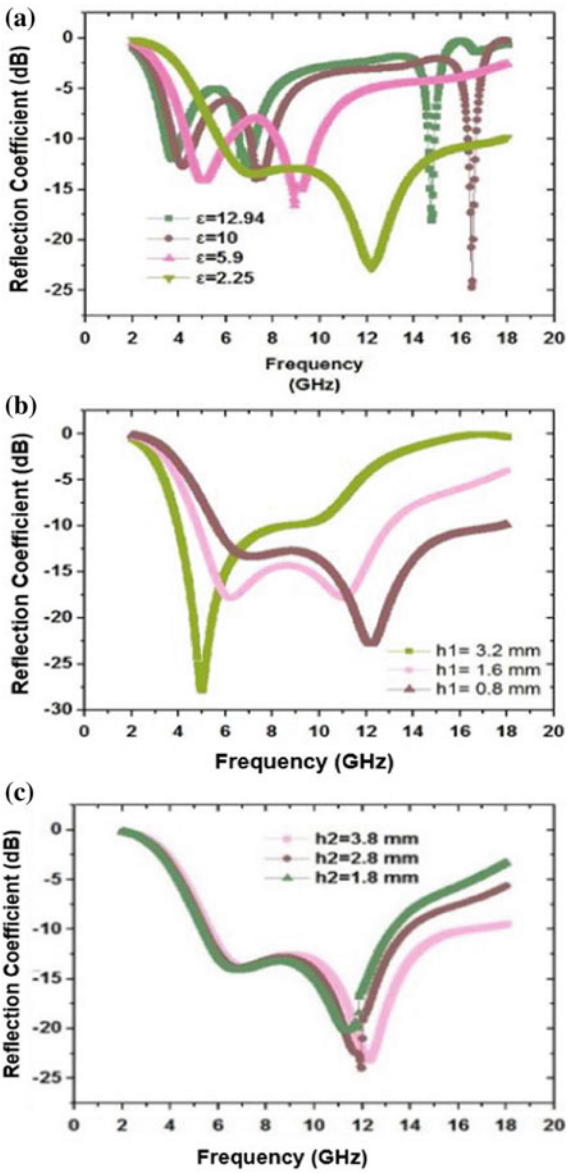


Fig. 5 RC characteristics of second-order FFSS based with varied **a** ε′, **b** dielectric substrate thickness, and **c** spacer thickness in the range of 2–18 GHz

4 Conclusion

The miniaturized and broadband MMSs were successfully developed by utilizing chip resistor and nature-inspired EM structures in the range of 2–18 GHz. A critical analysis of the design parameters such as fractal FSS iteration, substrate dielectric constant, substrate thickness, spacer thickness, and chip resistor resistance values has been carried out. An optimal MAS is found to accomplish more extensive peak RC value of –23.1 dB at a frequency of 12.2 GHz with a wide bandwidth of 11.1 GHz (5.5–16.6 GHz). In the present case, only the effect of chip resistor has been considered over the microwave absorption; in the future, the effect of diodes such as PIN, IMPATT, and RF switches. may be explored. In addition, a nondestructive microwave measurement setup may be developed for evaluating the performance of such EM structures.

Acknowledgements This research was supported by the Science and Engineering Research Board (SERB), Department of Science and Technology (DST), Government of India, under the Early Career Research Grant (ECR/2017/000676).

References

1. R. Panwar, J.R. Lee, Progress in frequency selective surfaces based smart electromagnetic structures: a critical review. Aerosp. Sci. Technol. **66**, 216–234 (2017)
2. R. Panwar, J.R. Lee, Performance and non-destructive evaluation methods of airborne radome and stealth structures. Meas. Sci. Technol. **29**(6), 062001 (2018)
3. M. Bakir, K. Delihacioglu, M. Karaaslan, F. Dincer, C. Sabah, U-shaped frequency selective surfaces for single- and dual-band applications together with absorber and sensor configurations. IET Microw. Antennas Propag. **10**(3), 293–300, 219 (2016)
4. S. Unaldi, S. Cimen, G. Cakir, U.E. Ayten, A novel dual-band ultrathin FSS with closely settled frequency response. IEEE Antennas Wirel. Propag. Lett. **16**(c), 1381–1384 (2017)
5. P. Munaga, S. Ghosh, S. Bhattacharyya, K.V. Srivastava, A fractal-based compact broadband polarization insensitive metamaterial absorber using lumped resistors. Microw. Opt. Technol. **58**(2016), 343–347 (2016)
6. Y. Li, L. Li, Y. Zhang, C. Zhao, Design and synthesis of multilayer frequency selective surface based on antenna-filter-antenna using minkowski fractal structures. IEEE Trans. Antennas Propag. **63**(1), 133–141 (2015)
7. W. Cao, B. Zhang, T. Yu, H. Li, A single-feed broadband circular polarized rectangular microstrip antenna with chip-resistor loading. IEEE Antennas Wirel. Propag. Lett. **9**, 1065–1068 (2010)
8. H.L. Liu, K.L. Ford, R.J. Langley, Design methodology for a miniaturized frequency selective surface using lumped reactive components. IEEE Trans. Antennas Propag. **57**(9), 2732–2738 (2009)
9. A.R. Katko, A.M. Hawkes, J.P. Barrett, S.A. Cummer, RF limiter metamaterial using p-i-n diodes. IEEE Antennas Wirel. Propag. Lett. **10**, 1571–1574 (2011)
10. N. Wang et al., Novel quadruple-band microwave metamaterial absorber. in IEEE Photonics J. **7**(1), 1–6 (2015). Art no. 5500506

Design and Analysis of Triple Split Ring Resonator-Based Polarization-Insensitive, Multiband Metamaterial Absorber



Arpit Sahu, Ravi Yadav, Trivesh Kumar and Ravi Panwar

1 Introduction

Metamaterials (MMs) are periodically arranged artificial structures composed of sub-wavelength constituent elements [1]. MMs are the center of the attraction for distinct commercialized applications due to their exotic electromagnetic properties. The key applications include stealth, radome, antenna, lens, cloaking, imaging, and shielding [2, 3]. MMs can be utilized for the realization of ultra-thin and wideband microwave absorbers [4]. The lower thickness, wide absorption bandwidth, multiband frequency response, polarization sensitivity, angular stability, mechanical strength, and environmental stability are important necessities for an efficient microwave absorber [5]. However, the simultaneous achievement of all these properties is a very difficult job for scientists working in this area.

MM-inspired microwave absorbers are well known due to the interesting EM properties of MM, which is not available in nature. MM makes the structure work as an effective homogenous medium with a negative effective permittivity (ε_{eff}) and effective permeability (μ_{eff}). Such a structural configuration results in a proper impedance matching, which is further responsible for the absorption of the incident EM wave. Starting from the Salisbury screen, which uses a continuous resistive sheet, placed a quarter wavelength above the metallic ground plane do not provide wide absorption bandwidth, however these are simple to manufacture [6]. The wide absorption bandwidth has been reported for the Jaumann absorber, but the fabrication complexity was the main issue behind their limited usage [7]. The other microwave absorber configurations include dielectromagnetic, nanocomposite, frequency selective surface (FSS) impinged and cascaded, high impedance surface (HIS), electromagnetic bandgap (EBG), active/smart, three-dimensional, nonplanar/curved absorbers [8, 9].

A. Sahu · R. Yadav · T. Kumar · R. Panwar (✉)
Discipline of Electronics and Communication Engineering, Indian Institute of Information Technology, Design and Manufacturing, Jabalpur, Jabalpur 482005, MP, India
e-mail: rpanwar.iitr@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_51

Each microwave absorbing configuration has its own merits and demerits depending upon the frequency regime utilized.

The microwave absorbing performance of an MMA depends on the periodicity, dielectric substrate characteristics, and thickness of the substrate as well as the structural geometry of the unit cell. Deep research interest in developing novel wide incident angle [10, 11], polarization insensitive [12], broadband [13, 14], reconfigurable [15], and narrowband [16] in microwave, terahertz, infrared, and optical regions is available in the open literature [17, 18, 19, 20 21]. But the optimal design of a polarization-sensitive, angularly stable MMA with a multiband frequency response is a complex task. Therefore, in this paper, an effort has been made to critically analyze the performance of a quadruple band polarization-insensitive MMA in the range of 7–18 GHz. This paper is structured as follows. Sect. 2 describes the absorber design and corresponding simulation steps. In Sect. 3, a critical analysis of the MMA design variables with corresponding physical significances is presented. Finally, the conclusion is discussed in Sect. 4.

2 Design of Absorber

TSRR-MMA unit cell has simulated and optimized using a full-wave EM simulation software CST microwave studio suite, as depicted in Fig. 1. Figure 1a depicts the unit cell of the absorbing structure, and a corresponding top view of the TSRR is given in Fig. 1b. The unit cell boundary conditions have been used in X-axis and Y-axis, while Z-axis set open to impinge TEM wave normal to the absorber surface. Considering the proposed MMA in an infinite array, this uses CST's full Floquet port implementation to simulate the unit cell. The TSRR etched on a 3.0 mm thick lossy FR-4 substrate (dielectric constant of 4.4 and loss tangent of 0.025). The top layer consists of an etched triple ring pattern, i.e., TSRR geometry. The bottom layer of

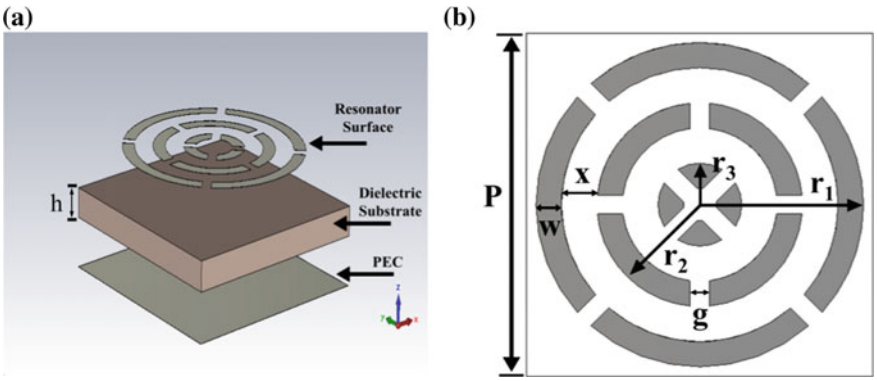


Fig. 1 Schematic of TSRR-MMA: **a** unit cell and **b** top view of the structure

the absorber is a perfect electric conductor (PEC), which is utilized in order to avoid any transmission through the sample.

3 Results and Discussion

3.1 Parametric Analysis for the Optimal Design of Absorber

It is well known that the performance of absorber is greatly influenced by its design variables. Therefore, in this section, an effort has been made to examine the effect of design variables, i.e., radius (r_1), strip width (w), substrate thickness (h), and periodicity (P) on the performance of the absorber. The corresponding simulated parametric sweep results of the proposed absorbing configuration are shown in Fig. 2a–d. Figure 2a presents radius (r_1) dependent RC spectra, where r_1 is varied from 2.9 mm to 3.5 mm with a step size of 0.2 mm. Interestingly, as the radius (r_1) of the unit

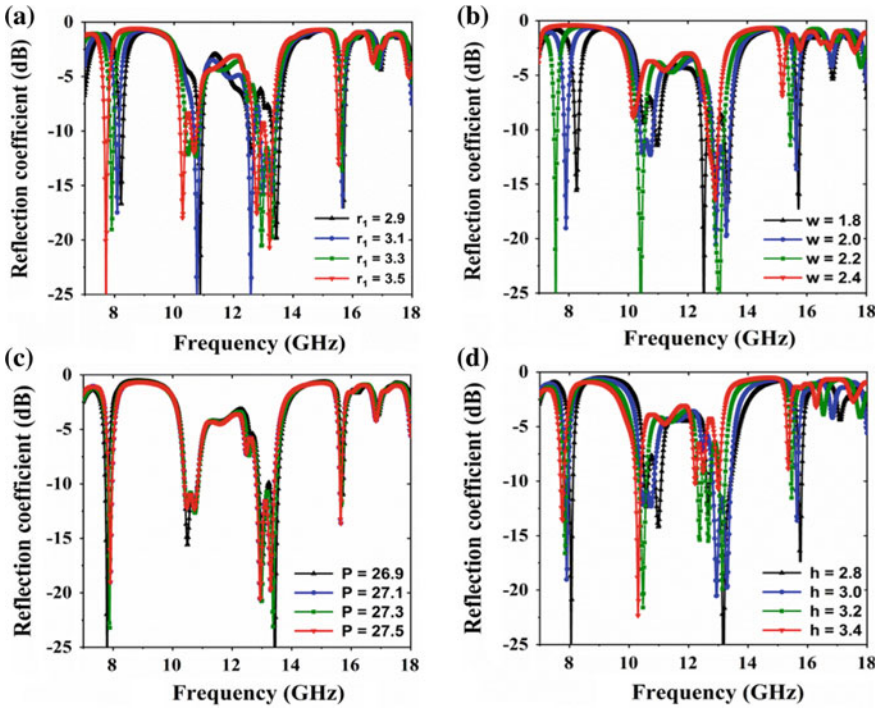


Fig. 2 RC spectra of TSRR-MMA: **a** radius *(r_1)* changed from 2.9 mm to 3.5 mm, **b** strip width *(w)* varied from 1.8 mm to 2.4 mm, **c** periodicity *(P)* varied from 26.9 mm to 27.5 mm, and **d** substrate thickness *(h)* changed from 2.8 mm to 3.4 mm

cell increases, the peak RC values shift toward lower frequency regime. In a similar manner, the RC properties are also found to be strongly influenced by "*w*", as shown in Fig. 2b, where "*w*" changed from 1.8 mm to 2.4 mm. Figure 2c shows the effect of "*P*" over RC characteristics of the structure. A minor effect in RC has been noticed with a corresponding variation in "*P*". Figure 2d shows the parametric analysis of "*h*" (ranging 2.8–3.4 mm) with a step width of 0.2 mm. As "*h*" increases, the minimum value of RC shifts from higher to lower frequencies. Hence, it may be concluded that the proposed absorber shows a significant deviation in the absorption with a corresponding variation in the structure design variables.

Polarization-independent and angularly stable absorbers have attracted immense research attention in recent years. Therefore, in order to demonstrate the effect of polarization angle variation (ϕ), the proposed absorber is designed at four different polarization angles at normal incidence. Figure 3a shows the RC frequency spectra of the absorber at distinct polarization angles of 0°, 10°, 20°, and 30°. The simulated results demonstrate the polarization-insensitive nature of the structure up to 90°. The symmetric geometry of the proposed structure might be a reason behind the

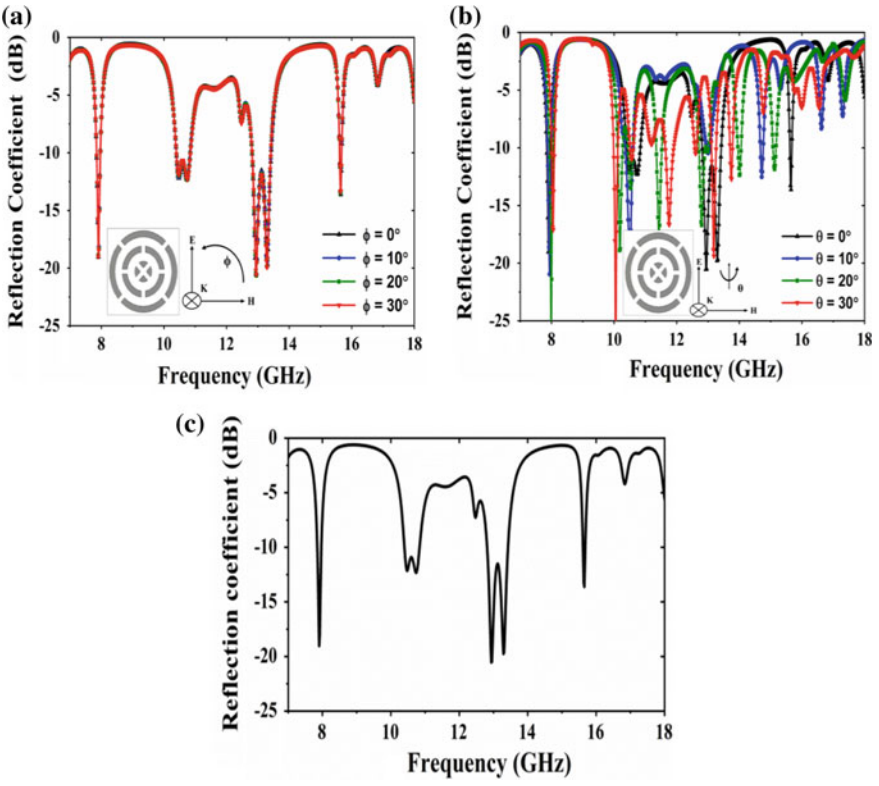


Fig. 3 Simulated response of **a** different polarization angles, i.e., 0°, 10°, 20°, and 30°; **b** different oblique angles, i.e., 0°, 10°, 20°, and 30°; and **c** proposed structure

Table 1 Optimal design variables of the proposed absorber

Parameters	Value (mm)	Parameter	Value (mm)
P	27.3	w	2.0
r_1	12.9	x	2.8
r_2	8.1	h	3.0
r_3	5.3	g	1.5

observed polarization-insensitive behavior. In a similar manner, an effort has been made to evaluate the angular stability performance of the absorber. Figure 3b shows the simulated RC spectra of the absorber at distinct oblique incidence angles (θ) of 0°, 10°, 20°, and 30° in the range of 7–18 GHz. The figure shows that, as the oblique angle increases, the RC of the absorber degrades. Finally, it may be concluded that the proposed TSRR-MMA shows excellent polarization neutral response, but an unstable oblique response. The optimized parametric values of the proposed unit cell are shown in Table 1. Figure 3c shows the RC of the proposed TSRR-MMA. The RC of the simulated results, which consists of four resonant frequencies at 7.90, 10.74, 12.93, and 15.65 GHz with the absorption coefficient of the 99%, 94.14%, 99%, and 95%, respectively. The total achieved frequency band below –10 dB in normal incidence is 1.182 GHz, which covers the X-band and Ku-band.

The optimized parametric values of the proposed unit cell are shown in Table I. Figure 3c shows the RC of the proposed TSRR-MMA. Further analyzing the RC of the simulated result which consists of four resonant frequencies at 7.90, 10.74, 12.93, and 15.65 GHz with the absorption coefficient of the 99%, 94.14%, 99%, and 95%, respectively. The total achieved frequency band below –10 dB in normal incidence is 1.182 GHz, which covers the X-band and Ku-band.

3.2 Constitutive Electromagnetic Parametric Retrieval and Physical Interpretation

To gain a better physical understanding of the absorption mechanism, the EM parameters, i.e., the normalized impedance (Z_{in}), ε_{eff}, and μ_{eff}, are retrieved, and the corresponding plots are shown in Fig. 4a–c. The retrieved constitutive EM parameters of the absorber at the resonating frequencies are also summarized in Table 2. The parametric retrieval algorithm is used to extract the parameters that are provided in [22–24]. It can be interpreted that the imaginary part of the Z_{in} includes negative values at the resonant frequencies. The Z_{in} at the four resonant frequencies, 7.90, 10.74, 12.93, and 15.65 GHz, reads 0.88–j0.42, 0.82–j0.51, 1.19–j1.17, and 0.25–j0.69, respectively. A perfect impedance matching is observed at the resonant frequencies, which may be considered as a main cause of the observed absorption. By comparing the real part of the ε_{eff} and μ_{eff}, the first resonance is due to electric resonance with ε_{eff} of 6.14 + j10.76 around 7.90 GHz. Magnetic resonance is responsible for the second resonance, with the μ_{eff} of 6.51 + j9.12 around 10.74 GHz. Moreover, the

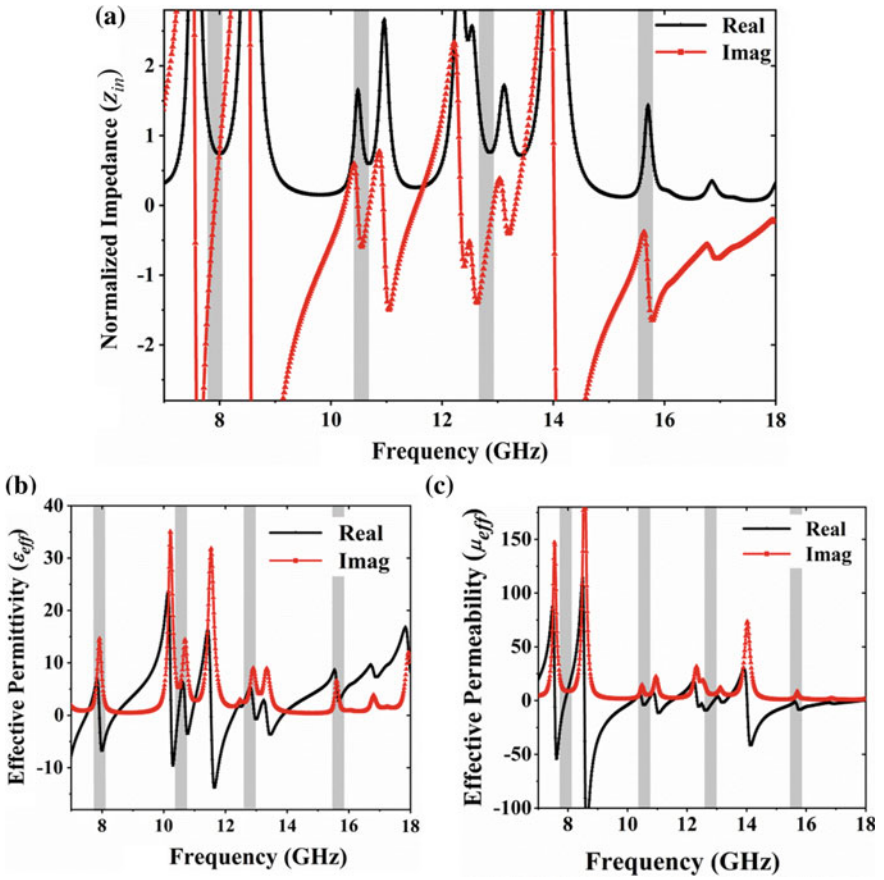


Fig. 4 Frequency-dependent plots of the extracted EM parameters: **a** Z, **b** ε_{eff}, and **c** μ_{eff}

Table 2 Extracted EM parameters of the absorber

Parameter	Frequency (GHz)			
	7.90	10.74	12.93	15.65
Real (ε_{eff})	6.14	−2.51	5.27	6.57
Imaginary (ε_{eff})	10.76	5.82	6.18	5.04
Real (μ_{eff})	−3.96	6.51	−2.91	−3.11
Imaginary (μ_{eff})	7.38	9.12	5.66	1.49

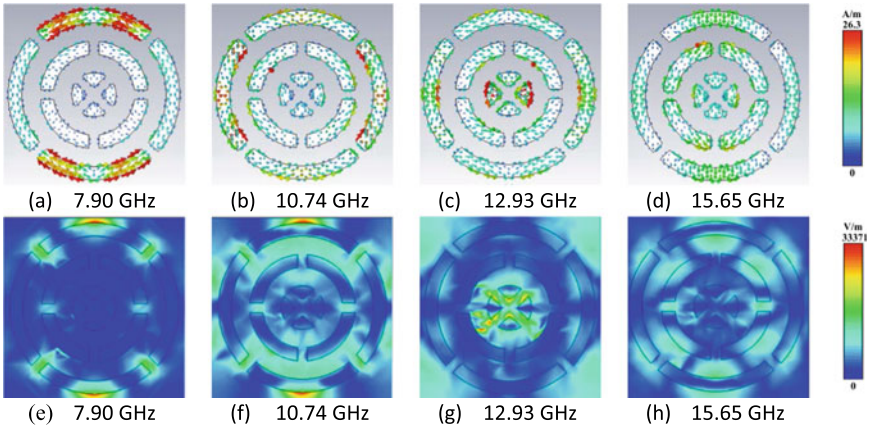


Fig. 5 Full-wave simulation results for the surface current distribution at the resonant frequencies of: **a** 7.90 GHz, **b** 10.74 GHz, **c** 12.93 GHz, and **d** 15.65 GHz; and the electric field at the resonance frequency of: **a** 7.90 GHz, **b** 10.74 GHz, **c** 12.93 GHz, and **d** 15.65 GHz

third (12.93 GHz) and fourth resonances (15.65 GHz) are due to electric resonance with $5.27 + j6.18$ and $8.57 + j2.74$, respectively [24]. After completely analyzing the retrieved EM parameter of the proposed structure, three resonance peaks due to an electric field and a single resonance due to the magnetic field are observed. The plots of ε_{eff} and μ_{eff} show the negative peak in the frequency range of 10.39–10.81 GHz that verifies the MM property of the TSRR-MMA.

A detailed analysis of the surface current and the electric field distribution is carried out at the resonant frequencies, i.e., 7.90, 10.74, 12.93, and 15.65 GHz in order to understand the physics behind the absorption. Figure 5 shows the surface current distribution, where the arrow in the figure shows the direction of the surface current. The color of the arrow shows the intensity of the developing current on the top surface of the resonator. In order to achieve minimum RC, the strong coupling should exist in between the bottom and top surface of the absorber. Moreover, on the top and bottom surfaces of the unit cell, the current direction should be antiparallel to each other.

From the plotted surface current distribution, the outer/first ring in the upper and lower parts of the TSRR causes the first resonance. However, at second resonance, the uniformly distributed surface current is associated with the first ring and at the edges of the second ring. Furthermore, the third resonant is achieved primarily due to surface current distributed at the inner ring of the resonator. Moreover, considering the fourth resonance, distribution of surface current significantly not confined to any particular resonator ring. Similarly, it may be seen that the electric field distribution of the first resonant mode mainly appears on the first ring of the TSRR. Strong coupling is observed between the split area and the outer ring of the contiguous unit cell. A high coupling between the first ring and the second ring of the TSRR is developed at the second resonance. Highly concentrated electric field coupling is observed in the

third ring in the third resonance frequency. Moreover, the fourth resonance mode is developed due to moderate coupling due to the first and second rings.

4 Conclusion

A multiband single layer TSRR-MMA has been proposed with a critical analysis of design variables. All the optimized geometrical parameters of the proposed absorber are showing minimum RC at the X- and Ku-bands. The physical interpretation of the observed absorption is explained based on the surface current and electric field distribution at the resonant frequencies. The structure is noticed to be polarization insensitive under normal incidence. Using the retrieved EM parameters, MM characteristics of the proposed TSRR absorber have been explained. The result proves the effectiveness of the absorber for stealth and radome applications in the specified frequency regime.

Acknowledgements This research was supported by the Science and Engineering Research Board (SERB), Department of Science and Technology (DST), Government of India, under the Early Career Research Grant (ECR/2017/000676).

References

1. N. Engheta, R.W. Ziolkowski, Metamaterials: Physics and Engineering Explorations (Wiley-IEEE Press, 2006)
2. R. Panwar, J.R. Lee, Performance and non-destructive evaluation methods of airborne radome and stealth structures. Meas. Sci. Technol. IOP Sci. **29**, 062001 (2018)
3. F. Capolino, Applications of Metamaterials (Taylor and Francis Press, 2009)
4. S. Ghosh, S. Bhattacharyya, K.V. Srivastava, Design, characterisation and fabrication of a broadband polarisation-insensitive multi-layer circuit analoue absorber. IET Microw. Antennas Propag. **10**(8), 850–855 (2016)
5. S. Bhattacharyya, K.V. Srivastava, Triple band polariazation-independent ultra-thin metamaterial absorber using electric field-driven LC resonator. J. Appl. Phys. **115**, 064508 (2014)
6. B.A. Munk, *Frequency Selective Surfaces: Theory and Design* (Wiley, New York, 2000)
7. R.C. Parida, D. Singh, N.K. Agarwal, Implementation of multilayer ferrite radar absorbing coating with genetic algorithm for radar cross-section reduction at X-band. Ind. J. Rad. Space Phys. **36**(2), 145–152 (2007)
8. R. Panwar, J. Ryul Lee, Progress in frequency selective surface-based smart electromagnetic structures: a critical review. Aerosp. Sci. Technol. (Elsevier) **66**, 216–234 (2017)
9. L. Liu, L.B. Kong, G.Q. Lin et al., Microwave permeability of ferromagnetic microwires composites/metamaterials and potential applications. IEEE Trans. Magn. **44**, 3119–3122 (2008)
10. L. Zhang, H. Lu, P. Zhou et al., Oblique incidence performance of microwave absorbers based on magnetic polymer composites. IEEE Trans. Magn. **51**, 7100604 (2015)
11. G.I. Kiani, K.L. Ford, K.P. Esselle et al., Oblique incidence performance of a novel frequency selective surface absorber. IEEE Trans. Antennas Propag. **55**, 2931–2934 (2007)
12. S. Ghosh, K.V. Srivastava, Polarization-insensitive dual-band switchable absorber with independent switching. IEEE Antennas Wirel. Propag. Lett. **55**, 2931–2934 (2017)

13. R. Panwar, S. Puthucheri, V. Agarwala, D. Singh, Fractal frequency selective surface embedded thin broadband microwave absorber coatings using heterogeneous composites. IEEE Trans. Microw. Theory Tech. **63**, 2438–2448 (2015)
14. A. Fardoost, F.G. Vanani, A. Amirhosseini et al., Design of a multilayer graphene-based ultra-wideband terahertz absorber. IEEE Trans. Nanotechnol. **16**, 68–74 (2017)
15. J. Li, J. Jiang, Y. He et al., Design of a tunable low-frequency and broadband radar absorber based on active frequency selective surface. IEEE Antennas Wirel. Propag. Letter **15**, 774–777 (2015)
16. A. Akhavan, S. Abdolhosseini, H. Ghafoorifard et al., Narrow band total absorber at near-infrared wavelengths using monolayer graphene and sub-wavelength grating based on critical coupling. J. Lightwave Technol. **36**, 5593–5599 (2018)
17. F. Costa, S. Genovesi, A. Monorchio et al., Low-cost metamaterial absorbers for sub-GHz wireless systems. IEEE Antennas Wirel. Propag. Lett. **13**, 27–30 (2013)
18. R. Mishra, R. Panwar, D. Singh, Equivalent circuit model for design of frequency selective, Terahertz-band, graphene based metamaterial absorber. IEEE Magn. Lett. **9**, 3707205 (2018)
19. Y. Dong, P. Liu, D. Yu et al., A tunable ultrabroadband ultrathin terahertz absorber using graphene stacks. IEEE Antennas Wirel. Propag. Lett. **16**, 1115–1118 (2016)
20. J.A. Mason, G. Allen, V.A. Podolskiy et al., Strong coupling of molecular and mid-infrared perfect absorber resonances. IEEE Photonics Technol. Lett. **24**, 31–33 (2012)
21. H. Zhu, F. Yi, E. Cubukcu, Plasmonic metamaterial absorber for broadband manipulation of mechanical resonances. Nat. Photonics Lett. **10**, 709–714 (2016)
22. X. Chen, T.M. Grzegorczyk, B.-I. Wu, J. Pacheco, J.A. Kong, Robust method to retrieve the constitutive effective parameters of metamaterials. Phys. Rev. E **70**, 016608 (2004)
23. D.R. Smith, D.C. Vier, T. Koschny, C.M. Soukoulis, Electromagnetic parameter retrieval from inhomogeneous metamaterials. Phys. Rev. E **71**, 036617 (2005)
24. C. Zhu, J.J. Ma, L. Li et al., Multiresonant metamaterial based on asymmetric triangular electromagnetic resonators. IEEE Antennas Wirel. Propag. Lett. **9**, 99–102 (2010)

Design and Analysis of Slot Loaded C-Shaped Trapezoidal Microstrip Antenna



Rishabh Kumar Baudh, Mahendra Kumar, Ravi Kant Prasad and Sonal Sahu

1 Introduction

From the last two decades, we are experiencing the changes in our lives because of wireless communication technology. The cordless phones took the place of short leashed handset cords in thousands of homes and offices. With the use of cell phones, we can communicate with each other throughout the world. The use of wireless local area network (WLAN) technology makes our lives free from management of yards of unsightly and expensive cable for Internet access. Antennas are a pillar element of communication. As per the explanation, the RF signal is transformed by antenna travel to a guide, into an EM wave in free space. Antenna significant feature is associated with reciprocity [1], which implies that an antenna characteristic remains in both conditions, i.e., transmitting and receiving. Mostly antennas work over on it, while a signal stuffed into an antenna, it will radiate EM wave toward space in a specific direction. The radiation pattern is a graphical description of the relativistic delivery of the emitted wave space.

The dielectric substrate of ground is of fixed size and at that point the radiation of a MPA determined through the impedance of specifically transmitted range of the patch and the diffracted field of surface waves of the edge within the dielectric substrate and air, the qualities like radiation example of a rectangular MPA for different dimensions of a square grounded dielectric material including different sub-

R. K. Baudh (✉) · M. Kumar · R. K. Prasad · S. Sahu
Bundelkhand Institute of Engineering and Technology, Jhansi, India
e-mail: rbaudha@yahoo.in

M. Kumar
e-mail: mahendra1812@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_52

strate thicknesses, recreation is finished by HFSS [2]. By notching a half-U-shape on edges of patch, a broadband rectangular microstrip antenna geometry is proposed. For development of transmission capacity, two resonant slots antenna is favored over a conventional geometry [3]. The geometry of wideband single layer coaxially encouraged conventional MPA with a twofold U-shape with legitimate measurements for the second U shape, data transfer capacity as high as 49.2% among 4.03 GHz to 6.66G Hz is attainable [4]. An antenna has been received by comprising a truncated corner square with a few round shape, L and U shape on patch. For the total C-band (4–8 GHz) inclusion, a nearness coupled nourishing technique is utilized by which a wide impedance data transmission is gotten conventional patch [5]. By expanding the substrate thickness as well as utilizing a low dielectric substrate, various impedance matching and sustaining systems, it can be improved reception apparatus transmission capacity [6]. The principle motivation behind this paper is to make a wideband microstrip antenna having microstrip line feed, which is utilized to improve the impedance bandwidth and minimize the return loss.

1.1 Antenna Material and Design

This antenna is designed by using glass epoxy (FR4) materials as substrate, radiating copper patch and finite ground plane. The prerequisite of designing an antenna is an appropriate dielectric material (substrate) of optimum thickness (t), dielectric constant, and loss tangent. The performance of microstrip antenna depends on the design (shape) of patch, feeding technique, thickness, and dielectric constant of the dielectric material. For designing, the dielectric constant of this proposed antenna is 4.4, height of the dielectric substrate (h) is 1.6 mm, and loss tangent (tan δ) is 0.012. The designing frequency (fr) is 2.4 GHz.

The designed antenna particularly consists of a ground plane with dimensions of 39.04 mm × 47.678 mm. The trapezoidal-shaped slot etched on a rectangular patch of size 29.24 mm × 37.878 mm is illustrated in Fig. 1. The most important parameters which are to be kept in mind while designing an antenna are its length and width, which are to be calculated very precisely for designing. The results of patch antenna may be affected by the dimensions of the patch. In order to design a four-sided MPA, the dimensions are computed [7]. The width of the four-sided patch may be evaluated as

$$W = \frac{C}{2fr\sqrt{\frac{2}{\varepsilon r + 1}}} \tag{1}$$

where

c speed of light (c) is 3×10^8m/s,
f_r design frequency, and
ε_r dielectric constant.

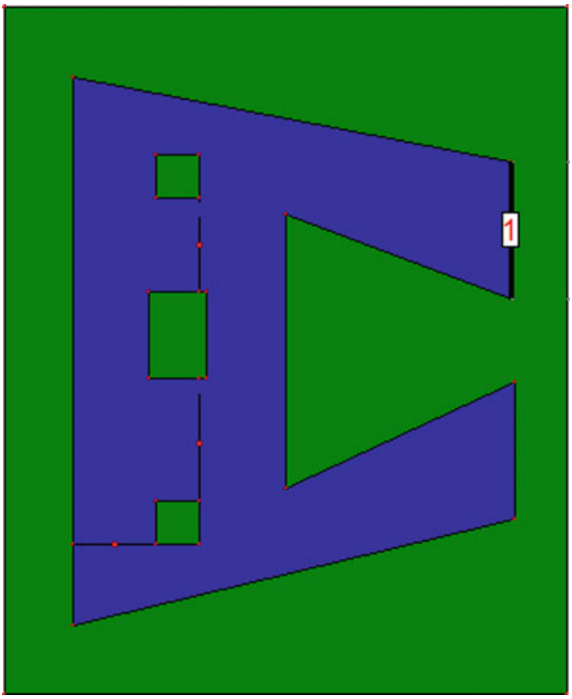


Fig. 1 Geometry of proposed antenna

Formula of effective dielectric constant is given as

$$\varepsilon_{ref} = \frac{\varepsilon r + 1}{2} + \frac{\varepsilon r - 1}{2}\sqrt{1 + \frac{12h}{w}} \tag{2}$$

where h = substrate thickness.

The extension length ΔL is given as

$$\frac{\Delta L}{h} = 0.412\frac{(\varepsilon_{ref} + 0.3)\left(\frac{w}{h} + 0.264\right)}{(\varepsilon_{ref} - 0.258)\left(\frac{w}{h} + 0.8\right)} \tag{3}$$

The length of patch antenna is given as

$L = \frac{C}{2fr\sqrt{\varepsilon ref}} - 2\Delta L$

The ground dimensions recognized by these formulas

$$Lg = L + 6h \tag{5}$$

$$Wg = W + 6h \tag{6}$$

A. *Antenna Geometry*

Geometry I: Conventional patch antenna excited with microstrip line feed gives a dual frequency band. The conventional antenna design is shown in Fig. 2a with its return loss displayed in Fig. 2b. The maximum return loss obtained is –16.5 dB. Two peak frequencies are observed at 1.75 GHz and 3.4 GHz. The bandwidth percentage of large frequency band is 17%.

Geometry II: In this geometry, X-offset and Y-offset are given to left edge; hence, it is inclined toward the center which is displayed in Fig. 3a. By notching at the edges, small bandwidth range is decreased as shown in Fig. 3b.

Geometry III: This is a previous step of final design of proposed antenna, which is obtained by cutting triangular slot around the corner of patch and one more trapezium slot at center, which has been shown in Fig. 4a. The dimensions of this design are shown in Table 1. The effect of this slot brings out the best result because it increases the bandwidth. The frequency range having return loss beyond –10 dB is from 1.92 GHz to 3.28 GHz having bandwidth 52.3%, and return loss is also improved to –24.2 dB at 2.4 GHz, which is displayed in Fig. 4b.

Geometry IV: This is a final design of proposed antenna, which is obtained by loading two squares and one rectangular slot on C-shaped trapezium patch, which has been shown in Fig. 5a. The dimension of square and rectangular slots is shown

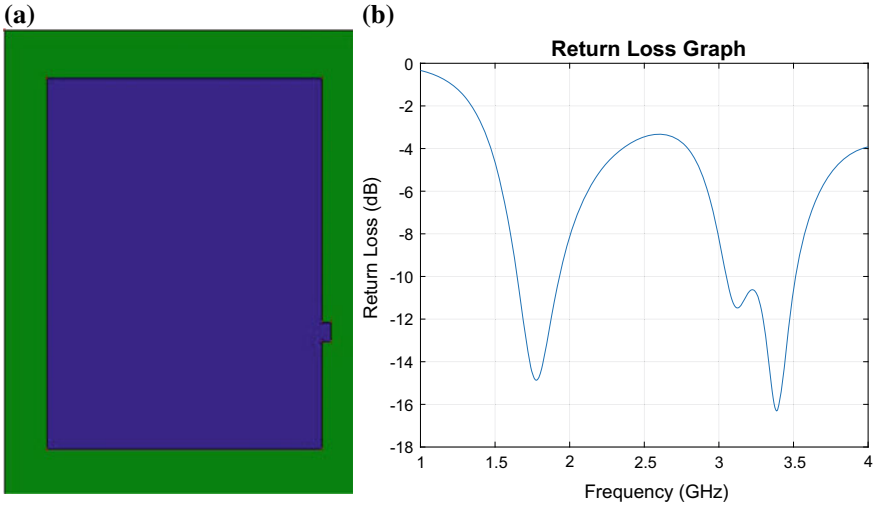


Fig. 2 **a** Rectangular patch antenna. **b** Return loss versus frequency plot of patch antenna

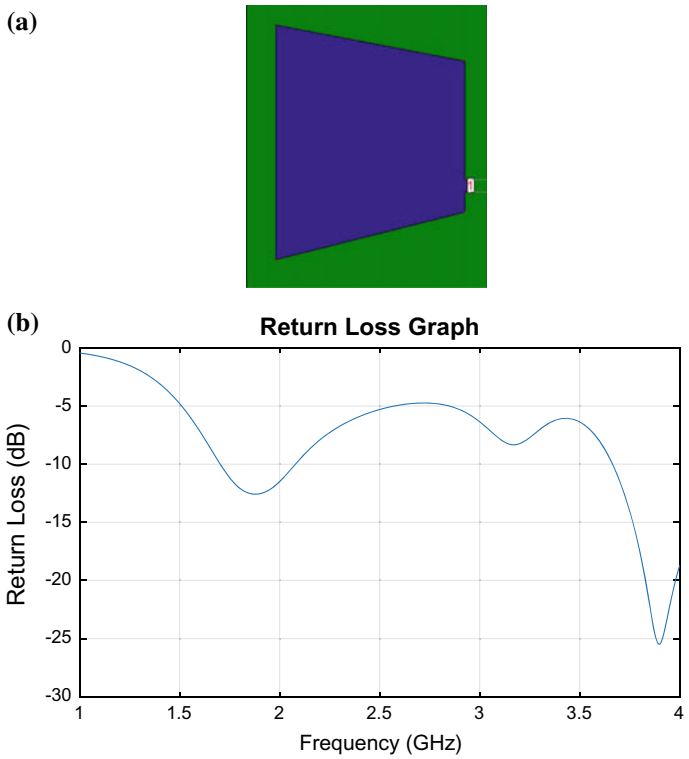


Fig. 3 **a** Dual slotted trapezium shape patch antenna. **b** Return loss versus frequency plot of dual slot trapezium shape patch antenna

in Table 2. The effect of this slot brings out the best result because it increases the bandwidth. The frequency range having return loss beyond –10 dB is from 1.86 GHz to 3.42 GHz having bandwidth 58.8%, and return loss is –18.9 dB at 2.65 GHz, which is displayed in Fig. 5b.

Table 3 shows a comparison of various parameters of antenna like bandwidth, return loss, and nature of frequency bands while Fig. 6 shows a comparison of S_{11} parameter or return loss (dB) against frequency (GHz).

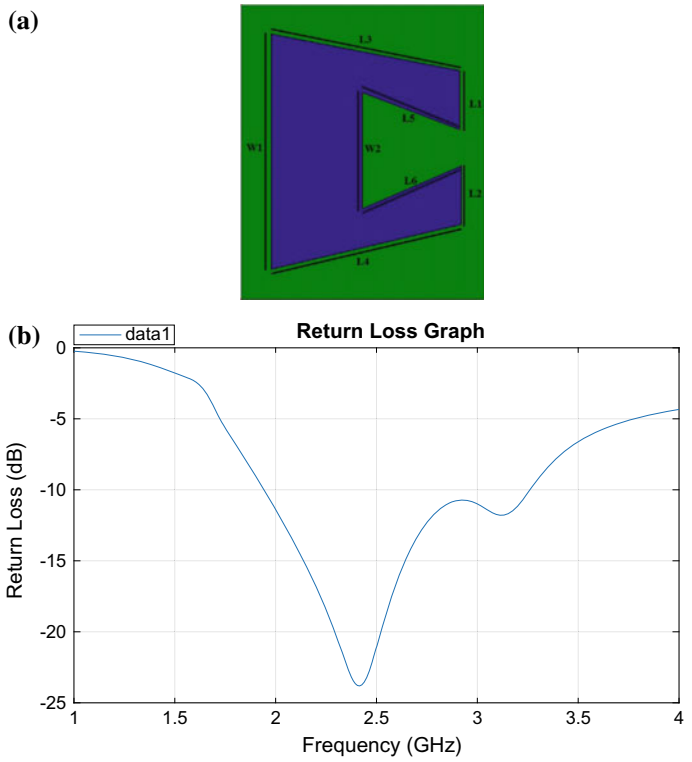


Fig. 4 **a** Previous step of final geometry of proposed antenna. **b** Return loss versus frequency plot of previous step of final proposed antenna

Table 1 Dimensions of previous step of proposed antenna

Parameter	Dimension (mm)
L_g	39.040
W_g	47.678
L_p	29.240
W_p	37.878
L_1, L_2	09.500
L_3, L_4	30.375
L_5, L_6	15.875
W_1	38.050
W_2	18.665

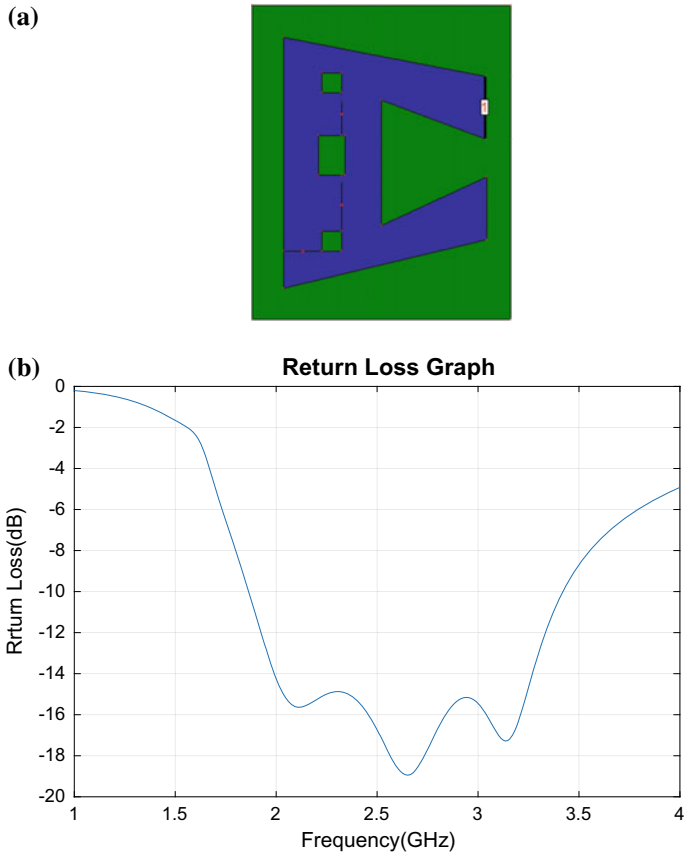


Fig. 5 **a**. Final geometry of proposed antenna. **b** Return loss versus frequency plot of final antenna

Table 2 Dimensions of proposed antenna

Parameter	Dimension (mm)
Square 1	3 × 3
Rectangle	4 × 6
Square 2	3 × 3

Table 3 Comparison between design steps with different parameters

Design	Bandwidth (%)	Maximum return loss	Nature of frequency band
Design I	17	−16.5 dB	Dual
Design II	10	−25.2 dB	Dual
Design III	52.3	−24.2 dB	Single
Final shape	58.8	−18.9 dB	Single

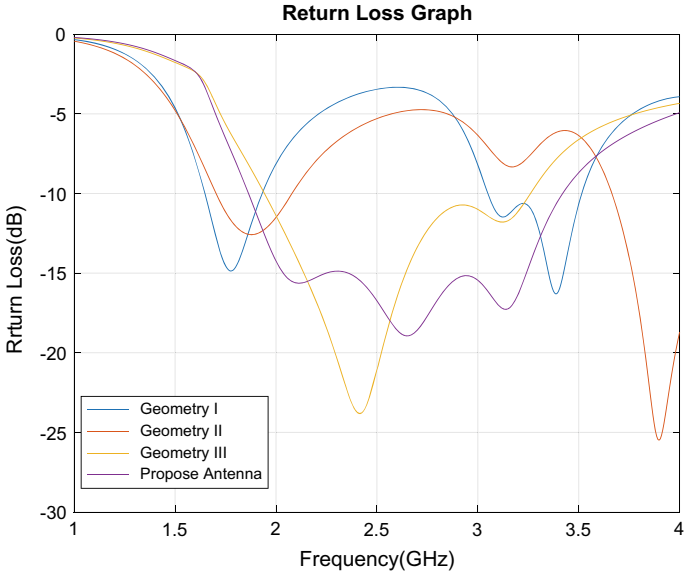


Fig. 6 Comparison of S11 parameter of different geometries

2 Results

The proposed slotted antenna is simulated through Zeeland IE3D simulator and having large bandwidth. The hardware of this antenna is shown in Fig. 7. The proposed antenna covers frequency range between 1.86 and 3.42 GHz with return loss beyond –10 dB. The simulation result shows the maximum return loss of −18.9 dB at

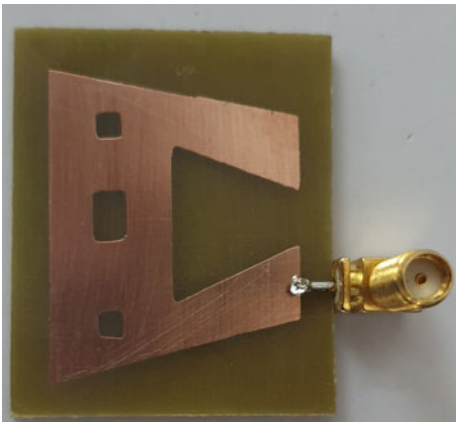


Fig. 7 Hardware of proposed antenna

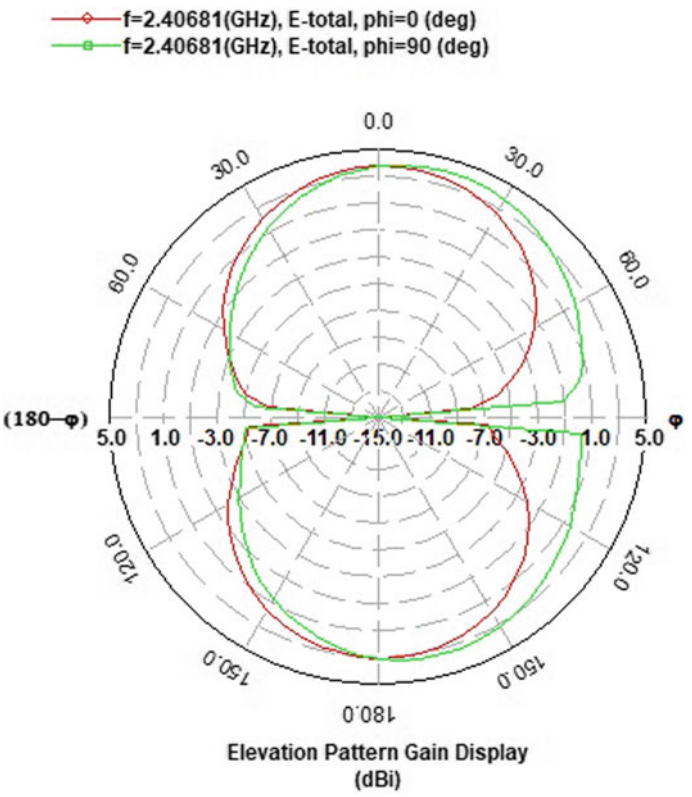


Fig. 8 Simulated radiation pattern of the proposed antenna

2.65 GHz. The return loss simulation result is shown in Fig. 5b. According to analysis of return loss graph, upper cutoff frequency is 3.42 GHz, and lower cutoff frequency is 1.86 GHz. Bandwidth percentage is calculated as 58.8% with resonating peak frequency at 2.65 GHz. The normalized pattern of radiation is simulated to obtain other parameters like gain and half power beamwidth. The stimulated radiation pattern is displayed in Fig. 8. The radiation pattern is measured at phi 0º and 90º.

The proposed antenna has been simulated by IE3D electromagnetic simulator. The simulated result has gain of 4.3 dBi and directivity of 4.8 dBi, which are displayed in Figs. 9 and 10, respectively. The simulated results by IE3D software of proposed antenna also have very high antenna efficiency, i.e., 99.98%, which is illustrated in Fig. 11. Large bandwidth has been obtained with respect to conventional rectangular patch antenna by slotting and selecting proper line feed technique.

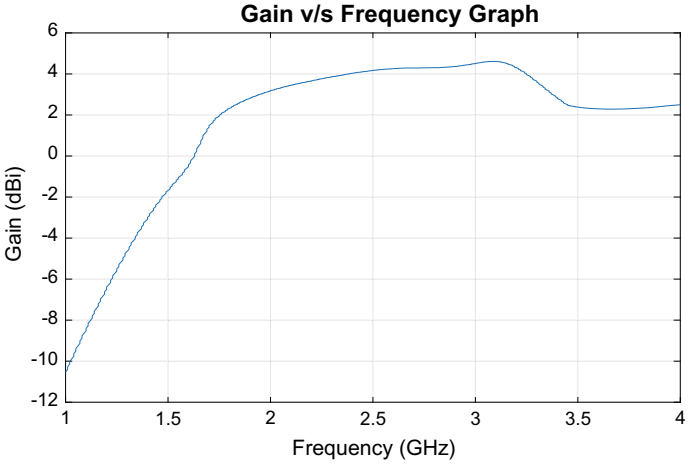


Fig. 9 Simulated gain of proposed antenna

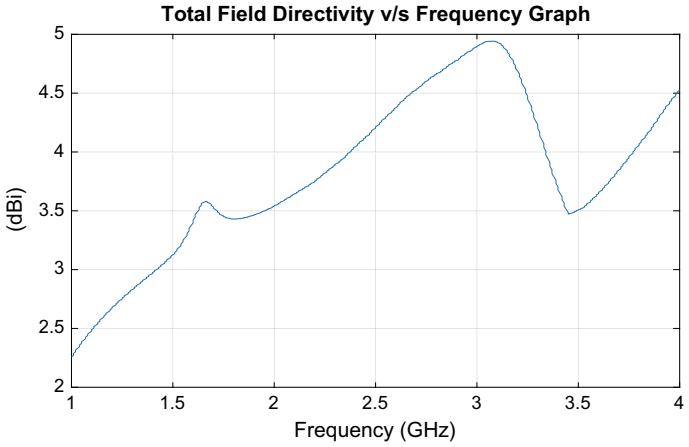


Fig. 10 Simulated directivity of proposed antenna

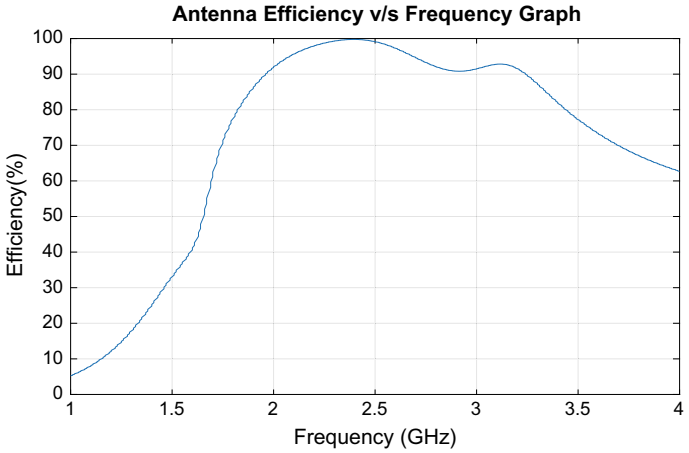


Fig. 11 Simulated antenna efficiency of proposed antenna

3 Conclusion

The proposed antenna covers the frequency range between 1.86 GHz and 3.42 GHz, which includes ultrawide band that is legally permissible in a frequency range of 3.1–10.6 GHz having bandwidth of 58.8%. This proposed antenna has very high antenna efficiency and radiation efficiency as calculated by IE3D software, i.e., 98.7% and 100%, respectively. The resonating peak frequency is obtained at 2.65 GHz with high return loss of –18.9 dB, and gain and directivity at this frequency are 4.86dBi and 4.96dBi, respectively. Due to ultrawide bandwidth range, its speed of transmitting data increases but the range decreases; therefore, high amount of data can be transmitted at low power, and it finds application in wireless communication, WiMAX, and Wi-Fi application.

References

1. M. Klemm, G. Troester, Textile UWB antennas for wireless body area networks. IEEE Trans. Antennas Propag. **54**(11), 3192–3197 (2006)
2. S. Shahid, M. Rizwan, M.A.B. Abbasi, H. Zahra, S.M. Abbas, M.A. Tarar, Textile antenna for body centric WiMAX and WLAN applications, in *International Conference on Emerging Technologies* (Islamabad, Pakistan, 2012)
3. P.B. Samal, P.J. Soh, G.A.E. Vandenbosch, UWB all-textile antenna with full ground plane for off-body WBAN communications. IEEE Trans. Antennas Propag. **62**(1), 102–108 (2014)
4. P. Salonen, Y. Rahmat-samii, K. Kamardin, P.S. Hall, Wearable antennas: advances in the design characteristics and application, in *Antenna and Propagation for Body-Centric Wireless Communication*, 2nd edn. (2012), pp. 209–270
5. Z. Wang, L. Zhang, D. Psychoudakis, J.L. Volakis, Flexible textile antennas for body-worn communication, in *IEEE International Workshop on Antenna Technology (iWAT)* (Tucson,

USA, 2012), pp. 205–208
6. T. Kaija, J. Lilja, P. Salonen, Exposing textile antennas for harsh environment, in *Milcom 2010 Military Communications Conference* (San Jose, USA, 2010), pp. 737–742
7. B. Roy, A.K. Bhaterchya, S.K. Chaudhary, Characterization of textile substrate to design a textile antenna, in *International Conference on Microwave and Photonics (ICMAP)* (Dhanbad, India, 2013), pp. 1–5
8. H.A. Rahim, M.F.A. Malek, I. Adam, K.M. Juni, M.I.M. Saleh, Basic characteristics of a textile monopole antenna for body-centric wireless communications, in *IEEE Symposium on Wireless Technology and Applications (ISWTA)* (Bandung, Indonesia, 2012), pp. 272–275
9. N. Amaro, C. Mendas, P. Pinho, Bending effects on a textile microstrip antenna, in *IEEE International Symposium on Antennas and Propagation (APSURSI)* (Spokane, USA, 2011), pp. 282–285
10. I. Locher, M. Klamm, T. Kirstein, G. Trster, Design and characterization of purely textile patch antennas. IEEE Trans. Adv. Packag. **29**(4), 777–788 (2006)
11. M.E. Jalil, M.K.A. Rahim, M.A. Abdullah, O. Ayop, Compact CPW-fed ultra-wideband (UWB) antenna using denim textile material, in *International Symposium on Antenna and Propagation* (Nagoys, Japan, 2012), pp. 30–33
12. S. Sankaralingam, B. Gupta, Determination of dielectric constant of fabric materials and their use as substrates for design and development of antennas for wearable applications. IEEE Trans. Instrum. Meas. **59**(12), 3122–3130 (2010)
13. R.K. Prasad, D.K. Srivastava, J.P. Saini, Gain and bandwidth enhancement of rectangular microstrip antenna by loading slot, in *2016 International Conference on Innovation and Challenges in Cyber Security (ICICCS-INBUSH)* (Noida, 2016), pp. 304–307

Calculation of SAR on Human Brain Model and Analysis of the Effect of Various Dielectric Shielding Materials to Reduce SAR



Pravesh Chaudhary and Ravika Vijay

1 Introduction

The rate at which radiation is absorbed by the human body is measured by the specific absorption rate (SAR). As per the international safety guidelines [1–3], the SAR must be below the limits. The high SAR value of mobile phone handset and mobile tower radiation cause adverse health effects like sleep disorder, headache, irritability, concentration problems, memory loss, depression, hearing loss, joint problems, etc. Cell tower radiation also affects birds, animals, plants, and environment [4–6]. In recent research [7], the author measured the radiation effect for GSM and CDMA, and he found that mobile should be kept away from the body during the conversation. Many countries in the world who have studied extensively the health hazards of cell tower radiation adopt strict radiation norms. The maximum permissible limit of SAR value in ICNIRP guidelines is not exceeding 2 W/kg, as averaged over any 10 g of tissue for any 6-min period while in IEEE standard, USA FCC guidelines, and TRAI (Telecom Regulatory Authority of India), the SAR value is 1.6 W/kg, as averaged over 1 g tissue for any 6-min period.

Some work on the similar topic has done recently in [8–10]. In [11], the author analyzed the radiation absorption and SAR through simulation and used three different dielectric materials for analysis. All the related researches are performed at lower frequency bands mostly 900 MHz to 2100 MHz, and they were not addressed the effect of dielectric layers on antenna gain. When we place a thin layer of dielectric material on the antenna to prevent radiation, its parameters like radiation pattern and gain are also changed. Here in this paper, we are calculated SAR along with the gain of the antenna.

For simulation, we are using a wideband and high-frequency antenna. It operates at 5G frequency bands. It covers 8.6 to 15 GHz frequency band.

P. Chaudhary (✉) · R. Vijay
Poornima University, Jaipur, India
e-mail: praveshphysics@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_53

1.1 SAR Definition

SAR is defined as

$$SAR = \frac{1}{2\rho} E^2$$

where "ρ" is the sample electrical conductivity, and "E" is the RMS electric field.

2 Human Head Model and SAR Calculation

In the first experiment, we used a monopole antenna, placed it near the human head model, and calculated the SAR. The antenna was placed 1 mm apart from the head model and then shifted away from the human head. Then, we analyzed the reduction in SAR. Similarly, we also use the different dielectric material layers for shielding and analyze the SAR. The simulation work was done in HFSS simulator.

2.1 Monopole Antenna and Human Head Model Characteristics

To design the printed monopole antenna, we have used RT-Duroid 5880 as substrate material having dielectric constant 2.2 and substrate thickness 1.6 mm. The substrate dimension is 26.6 × 21.8 mm. The length and width of the monopole antenna are 10 mm and 2.4 mm, respectively. To make the antenna wideband, we use the defective ground, and dimension of the ground is 26.6 mm × 4 mm. The antenna model is shown in Fig. 1.

The return loss and radiation pattern of the antenna are shown in Figs. 2 and 3, respectively. The antenna is resonant from 8.6 GHz to 15 GHz. It covers a wide band of frequencies.

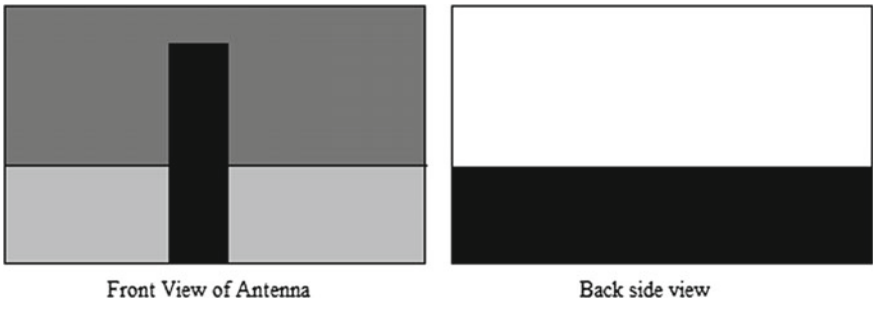


Fig. 1 Monopole antenna

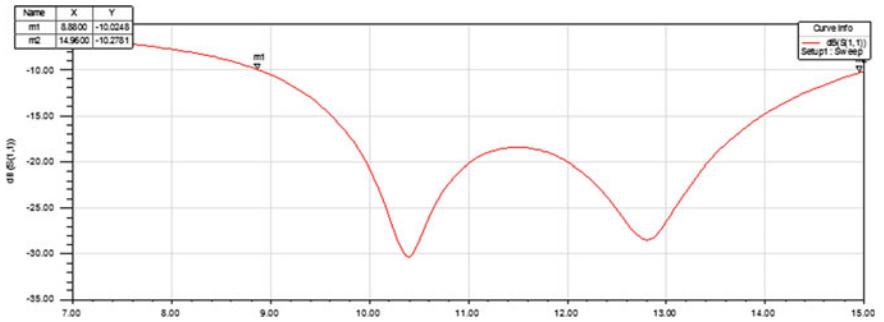


Fig. 2 Return loss diagram of the monopole antenna

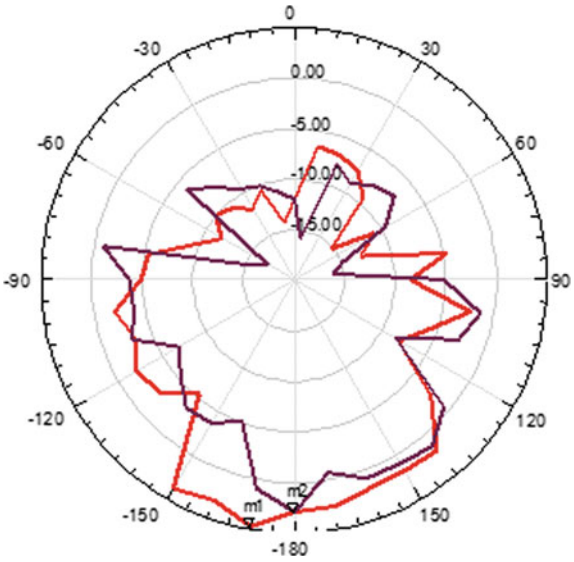


Fig. 3 The radiation pattern of the antenna

The antenna has a gain of 2.9 dB. Similarly for the human head model, first we design the outer shell and then the human brain liquid. The dimension of the inner radius of the head is 106.5 ± 5 mm, and the thickness is 5 ± 0.5 mm. The opening is 170 mm and $\varepsilon r = 4.6$. The specifications of the outer shell and brain liquid are shown in Table 1.

Table 1 Specification of the outer shell and brain liquid

Parameters	Outer shell	Brain liquid
Relative permittivity	4.6	41.5
Relative permeability	1	1
Dielectric loss tangent	0.01	0.9
Bulk conductivity	0	0.9 s/m
Density	N.A.	$\rho = 1$ g/cm^3

2.2 Calculate the SAR for the Human Brain with Variable Distance of the Antenna

In this experiment, we place the antenna at 1 mm distance from the human brain model and calculate the average SAR on the human brain. The experimental setup is shown in Fig. 4. We have calculated the average SAR and electric field linked with the human brain. The value of the linked electric field is 3.2e + 002 V/meter, and the average SAR is 2.06 e + 003 W/kg. The recommended value of SAR is 1.6 W/kg as decided by TRAI in India, and our result shows 2.06 e + 003 W/kg so it is very high from the recommended value. We perform this experiment with the variable distance. We place the antenna from 1 mm to 10 cm to see the effect of distance on SAR.

Table 2 shows the effect of mobile radiation on the human brain when the mobile is placed on various distance from the human brain. By analyzing Table 2, we can conclude that up to 15 mm distance from human brain, the mobile antenna produces much more SAR but when this antenna is placed at 10 cm distance from the brain, it gives the SAR value of 0.68 w/kg, which is under control and less value as recommended by TRAI. So the one recommended solution to reduce the SAR of the

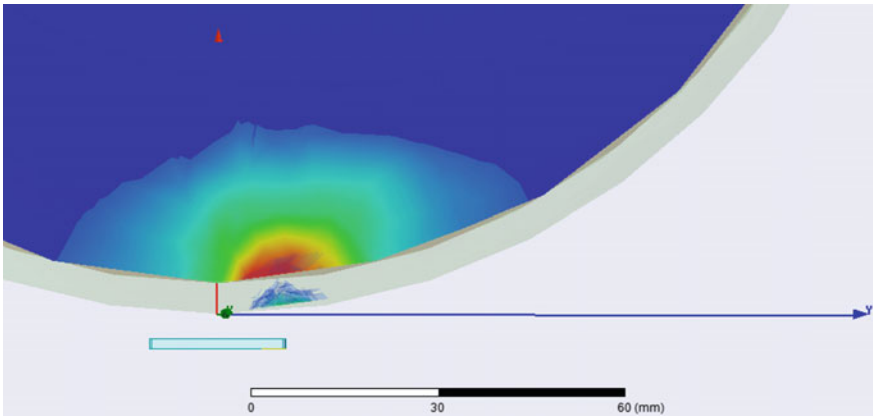


Fig. 4 Antenna radiation near the human brain and heating effect

Table 2 Effect of the variable distance of antenna on SAR

S. no	The distance of antenna from brain	Avg (W/Kgm)	E field V/m
1	1 mm	2.06 e + 003	3.2 e + 002
2	2 mm	2.03e + 003	2.69 e + 002
3	3 mm	1.7e + 003	2.35e + 002
4	6 mm	2.2 e + 003	5.1e + 001
5	10 mm	3.8e + 001	7.7 e + 001
6	15 mm	2.1 e + 001	5.8 e + 001
7	10 cm	6.85 e–001	4.16 e + 000

antenna is that we can place the antenna at the bottom of the mobile phone, and due to this, heating in the brain can be reduced.

3 Calculate the SAR for Human Brain When Antenna is Shielded by Various Dielectric Materials

In this experiment, we have placed a thin dielectric layer over the antenna of different dielectric and see the effect on SAR value. The thickness of the dielectric layer is 0.1 mm, and it is placed above 1 mm high from the antenna. Here we get a different value of the SAR but radiation pattern of the antenna and gain is changed. Here we have done 10 different experiments, and the experimental results are shown in Table 3.

Here the initial SAR value of antenna is 2.06e + 003w/kg, and the gain is 2.9 dB. By Table 3 if we put the conductor sheet of 0.1 mm over the antenna, we can reduce the SAR significantly. Bakelite and ferrite materials can be used as a shielding material

Table 3 Effect of the different dielectric layers (1 mm apart) on SAR

S. no	Material	Gain in dB	Avg. SAR
1	Mica	1.0	4.8e + 003
2	Silicon	0.9	4.6e + 003
3	Gallium arsenide	2.7	4.4e + 003
4	Glass	6	4.2e + 003
5	Graphite	8.5	1.89e + 002
6	Arlon	6.2	1.2e + 002
7	Bakelite	3.2	1.16e + 002
8	Duroid	−2.02	1.16e + 002
9	Ferrite	5.05	3.15e + 001
10	Copper	−15	2.28e + 001

because they reduce the SAR without much affecting the antenna gain. If we use a copper shield, it reduces the maximum SAR but at the cost of the gain of the antenna.

4 Conclusion

The monopole antenna has good radiation characteristics. After designing the antenna, we put it near the human brain model (1 mm apart) and checked the SAR of the antenna, which was very high (2.06e + 003 W/kg) from the recommended limit, that is, 1.6 W/kg. We performed two different experiments to reduce SAR and found that if antenna is placed at the bottom of mobile phone, we can reduce the SAR, and in a properly designed antenna if we shield the antenna with a dielectric layer of ferrite, we can reduce the SAR of the antenna, which is 3.15e + 001 W/kg.

References

1. IEEE standard for safety levels with respect to human exposure to radio frequency electromagnetic fields, 3 kHz to 300 GHz, IEEE C95.1-1991 (Institute of Electrical and Electronics Engineers, Inc., New York, 1992)
2. Guidelines on limits of exposure to radio frequency electromagnetic fields in the frequency range from 100 kHz to 300 GHz, Health Phys. **54**(1), 115–123, International Non-Ionizing Radiation Committee of the International Radiation Protection Association, (1988)
3. ICNIRP, Guidelines for limiting exposure to time-varying electric, magnetic and electromagnetic fields (up to 300 GHz). Health Phys. **74**(4), 494–522 (1998)
4. G. Kumar, Newsletter-cell phone/tower radiation hazards and solutions (1) (2013)
5. G. Kumar, Newsletter-cell phone/tower radiation hazards and solutions (2) (2013)
6. G. Kumar, Newsletter-cell phone/tower radiation hazards and solutions (3) (2014)
7. J. Isabona, V.M. Srivastava, Cellular mobile phone—a technical assessment on electromagnetic radiation intensity on human safety, in *2017 IEEE 3rd International Conference on Electro-Technology for National Development (NIGERCON)* (Owerri, 2017), pp. 271–274
8. P. Dhanesh, A review on SAR reduction methods used for mobile application. IOSR J. Electron. Commun. Eng. (IOSR-JECE) **10**(5), 25–27 (2015)
9. IEEE C95.1-1991: Safety Levels with Respect to Human Exposure to Radio Frequency Electromagnetic Fields, 3 kHz to 300 GHz (IEEE, Piscataway, NJ, 1992)
10. K. Divyashre, Low-cost solution for reducing the effect of mobile radiations. in *International Conference on Circuits, Controls, Communications and Computing, IEEE Conferences* (2016)
11. A. Imran, M.R. Islam, M.N. Hassan, N.H. Shiblu, S. Ahmad, M.T. Ali, Absorption rate analysis of cellular phone radiation on the human head, in *2016 5th International Conference on Informatics, Electronics and Vision (ICIEV)* (Dhaka, 2016), pp. 496–501
12. S. Dey, SAR reduction of bio-tissue in cellular communication, in *3rd International Conference on Electrical Information and Communication Technology (EICT)* (Khulna, Bangladesh, 2017)

Equivalent Circuit Model Analysis of Single Square Loop FSS for Transmission Mechanism



Rahul Shukla and Garima Tiwari

1 Introduction

Frequency selective surfaces (FSS) are the thin layer with repetitive behavior of the structure which can absorb, transmit, and reflect the electromagnetic frequency of the field. In this sense, FSS is the periodic structure. There are a number of structures available but in this article, we consider a structure of single square loop FSS. In this paper, we observe the transmission mechanism of FSS structure through software simulation by the parametric analysis [1]. In the case of the parametric analysis, every parameter of FSS design is different and tries to obtain the most optimized result. This convenient structure basically consists of the two layers as the dielectric substrate of FR4 layer and perfect conducting material layer over the surface of the FR4 form a layer of perfectly conducting material and design of single square loop. This single square loop is basically made up of conducting material. In this structure of FSS on the upper part of substrate FSS is designed and the lower side of substrate no PEC is available so that this structure is suitable for the observation of the transmission coefficient S21 [2].

In this paper, we also use the equivalent circuit modeling method which is efficient and simple technique to validate the software simulation results. FSSs can be modeled in the equivalent circuit as the energy storing component as inductive and capacitive, which depends on the design and structure of the frequency selective elements. In equivalent circuit model has been proposed or derive the efficient and reliable equations which depend on the parameter of FSS structure provide the appropriate result which helps to understand the transmission mechanism of single square loop

R. Shukla (✉) · G. Tiwari
EC Department, Jabalpur Engineering College, Jabalpur, Madhya Pradesh, India
e-mail: rahul.shukla.25m@gmail.com

G. Tiwari
e-mail: garima.tiwari29@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_54

FSS. In this case, with the help of the equation, we calculate the inductance (L) and capacitance (C) of the equivalent circuit in which capacitance and inductance are connected in series or FSS follow the periodicity so that this combination of capacitance and inductance is connected in parallel, so we calculate its transmission coefficient of single square loop FSS. As per the observation for the high transmission coefficient, there must be appropriate matching of with the free space. By observing input impedance of structure for the high transmission mechanism, the real part of the input impedance should be very close to the value of the free space impedance (377 Ω) [3].

2 Design of FSS

The frequency selective surface is the periodic array of structure. There are various designs available either it is radiation or non-radiating or slot which act as the band-pass or band-stop filters for the respective electromagnetic waves [4]. There are a number of possible designs available for designing the FSS array but we consider single square loop. Now we study the behavior of the single square loop structure; it presents a band-reject frequency response which depends on its dimensions. The square loop is a microstructure which is designed using HFSS as the full-wave simulation tool in the configuration of the square loops as shown in Fig. 1a, which consists of the FSS on the one side of the substrate, and there is no PEC (perfect electric conductor) on the other side. This square loop FSS is designed using master–slave boundary condition which is defined by the periodicity of the unit cell of single square loop. This structure is designed over the thin layer of the substrate FR4 (loss tangent 0.0002 and dielectric constant 4.4) with thickness of 1.6 mm in frequency range of 2–18 GHz.

As the transmission property of the square loop FSS is evaluated through the simple and accurate equivalent circuit model as shown in Fig. 1b. ECM method is the best alternative way for the full-wave simulation analysis for the FSS design. The FSS can be equivalent circuits modeled as the resonant circuits consist of the parallel combination of the two series inductive (L) and capacitive (C) resonator. The value

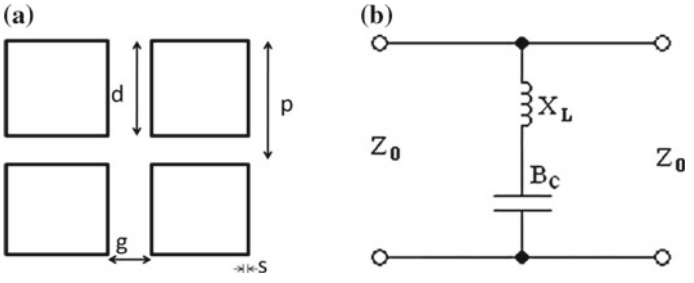


Fig. 1 Single square loop: **a** HFSS designs structure, and **b** equivalent circuit design

of L and C is basically affected by the dimension of the conducting strips and the spacing between them. If the transverse electric (TE) wave incident then vertical line act as the L (impedance) and the horizontal grating act as the C (impedance) [5, 6].

In the case of the EC modeling, the value of the L and C can be obtained by Eqs. (1) and (2), in which the d, g, p, and s (thickness of conducting strip) are the dimension of the single square loop as shown in Fig. 1a, and θ is the incidence angle which is related to normal incidence. The parameters "p" and "W" as shown in Eqs. (3) to (6) should be modified on the basis of Eqs. (1) and (2) [7].

$$\frac{X_L}{Z_0} = wL = \frac{d}{p}\cos\theta F(p, 2s, \lambda, \theta) \tag{1}$$

$$\frac{B_C}{Y_0} = wC = 4\frac{d}{p}\sec\theta F(p, g, \lambda, \theta)\xi eff \tag{2}$$

where

$$F(p, \omega, \lambda, \theta) = \frac{p}{\lambda}\left[\ln\left(csc\frac{\pi\omega}{2p}\right) + G(p, \omega, \lambda, \theta)\right] \tag{3}$$

$$G(p, \omega, \lambda, \theta) = \frac{1}{2} \times \frac{(1-\beta^2)^2\left[\left(1-\frac{\beta^2}{4}\right)(A_+ + A_-) + 4\beta^2 A_+ A_-\right]}{\left(1-\frac{\beta^2}{4}\right) + \beta^2\left(1+\frac{\beta^2}{4}-\frac{\beta^4}{8}\right)(A_+ + A_-) + 2\beta^6 A_+ A_-} \tag{4}$$

with

$$A_\pm = \frac{1}{\sqrt{\left[1 \pm \frac{2p\sin\theta}{\lambda} - \left(\frac{p\cos\theta}{\lambda}\right)^2\right]}} - 1 \tag{5}$$

$$\beta = \sin\left(\frac{\pi\omega}{2p}\right) \tag{6}$$

The factor Ɛeff in Eq. (2) is introduced by Munk with value of Ɛeff, and it uses for square loop FSS [8].

$$\varepsilon\text{eff} = 0.5(\varepsilon\text{r} + 1) \tag{7}$$

3 Result and Discussion

Parametric simulation is basically used to optimize the simulation result. In this perform parametric simulation of electromagnetic (EM), simulation software HFSS and equivalent equation which yield the result for the single square loop FSS. A wide set

of parametric simulations are performed over the different parameters of the square loop FSS, such as d, g, s, and p. The first critical parametric simulation is performed over the dimension d; significant changes appear at the resonance frequency as we increase or vary the dimension of d from 6 mm to 22 mm then resonance frequency shifts toward lower frequency, as shown in Fig. 2. If the perform parametric simulation only on the dimension s, then observe the effect mostly in the bandwidth of the FSS, as increase the value of the s from 0.1 mm to 3.5 mm then observed bandwidth is increased and resonance frequency peak shifted toward higher frequency as shown in Fig. 3. On the other part as perform parametric simulation on the dimension g of the FSS which make the small shift in the resonance frequency peak in the right side as increase the value of g from 1 mm to 5 mm but there is very small effect show on bandwidth of the FSS design as shown in Fig. 4. The most optimized result as shown in Fig. 7 is obtained after the parametric simulation on the square loop FSS structure dimension that is s = 2.00 mm, d = 10.00 mm, and g = 1.00 mm, and the "p" is the length of unit cell which is the combination of conducting strip length d and distance between two unit cell g with substrate thickness of t = 1.6 mm of FR4 with permittivity of 4.4 and loss tangent of 0.002.

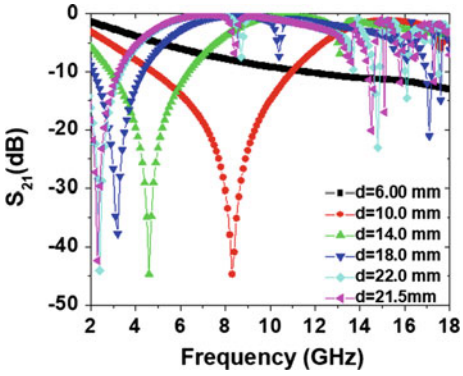


Fig. 2 Parametric simulation of dimension d for square loop FSS with frequency response

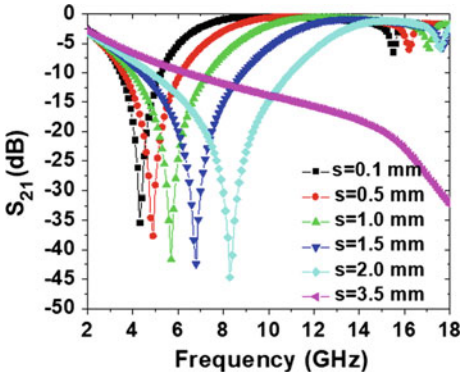


Fig. 3 Parametric simulation of dimension s for square loop FSS with frequency response

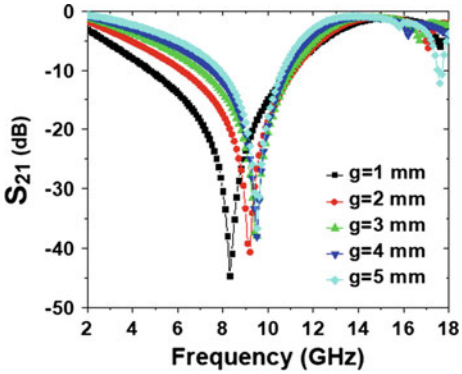


Fig. 4 Parametric simulation of dimension g for square loop FSS with frequency response

The input impedance of the FSS can be computed using Eqs. (1) and (2) but there is a term present in the relation which can be varied to the oblique incidence of the transverse magnetic (TM) and transverse electric (TE) field polarization of incident wave. The incoming wave from the source whose incident angle is represented it as θ basically consider it as a normal incidence of the incoming wave but on the other hand if influences the value of θ, observe that resonance frequency of the SL_FSS is less sensible to incoming wave at different oblique angles which is shown in Fig. 5 [9].

In the case of EC modeling, susceptance and reactance are reduced by the factor d and p (d/p) from the corresponding array strip length. The equivalent circuit inductive reactance and capacitive susceptance with respect to different dimensions of the FSSare obtained from Eqs. (1) and (2) as observed from Figs. 6 and 7. Graphical relationship of equivalent circuit component and FSS structure dimension capacitance basically depends on the gap between two unit cells; if gap between two unit cells increases, then we observe a decrease in capacitance and inductance depends on the width and the length of the strip of SL_ FSS. As per Fig. 7 which represent the strip length and inductance as increases the value of the d, inductance also increases. For the square loop FSS in Fig. 1a, the capacitance and inductance were calculated by C

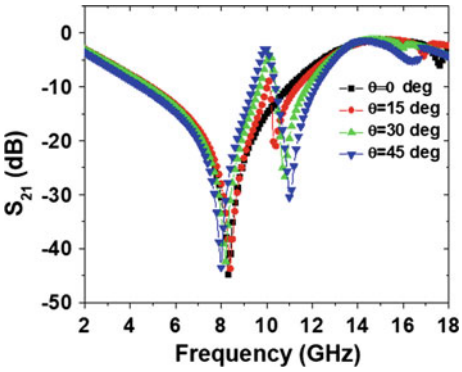


Fig. 5 Influence of incident angle of incoming wave on a square loop FSS

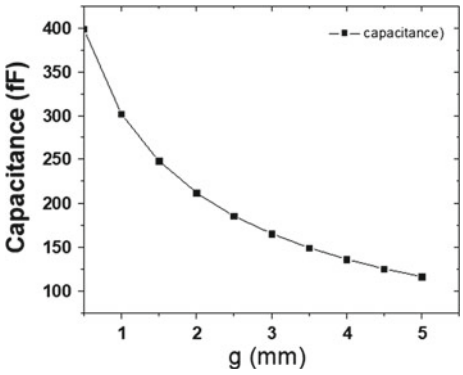


Fig. 6 Analysis of capacitance as variation in g dimension of FSS

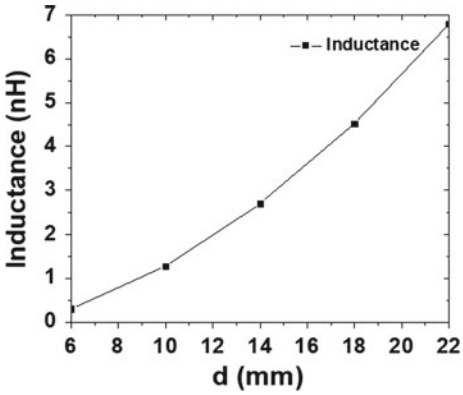


Fig. 7 Analysis of inductance as variation in d dimension of FSS

= 303.39 fF and L = 1.2784 nH with the help of these values of components C and L resonance frequency of the SL_FSS was calculated by the $f = 1/[2\Pi\text{sqrt}(CL)] =$ 8.08 GHz, which is similar to the simulation of most optimized results as shown in Fig. 8.

The most optimized result is getting after the parametric simulation on the parameter of d, g, s. This optimized result is compared with the ECM results for the validation of the band-reject FSS curve. In both of the cases, HFSS simulation and the ECM results are in agreement with each other as shown in Fig. 8.

4 Conclusion

In this article, we observe the simulation results for the single loop square FSS transmission mechanism, and the same result is also observed for the equivalent circuit model (ECM). Also, we found that the resultant output from both of the method will show the necessary frequency refinement that will provide a good agreement

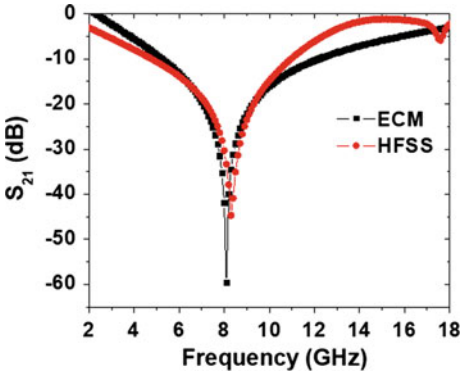


Fig. 8 Optimized square loop frequency response

between the HFSS simulation and the ECM results, which are very close to each other. This result demonstrates and suitable for the army, 3 generation, 4 generation, and 5 generation application.

References

1. R. Panwar, J.R. Lee, Progress infrequency selective surface-based smart electromagnetic Structures: article review. Aerosp. Sci. Technol. **66**, 216–234 (2017)
2. M. Yang, A.K. Brown, A hybrid model for radio wave propagation through frequency selective structures (FSS). IEEE Trans. Antennas Propag. **58**(9), 2961–2968 (2010)
3. D. Ferreira, F.S. Caldeirinha, I. Cuinas, Telmo R. Fernandes, Square loop FSS and slot FSS study for equivalent circuit model optimization. IEEE Trans. Antennas Propag. **63**(9), 3947–3955 (2015)
4. N. Qasem, R. Seager, Indoor band pass frequency selective wall paper equivalent circuit and ways to enhance wireless signal, in *Proceedings of Loughborough Antennas Propagation Conference (LAPC)* (Loughborough, UK, 2011), pp. 1–4
5. R.J. Langley, E.A. Parker, Equivalent circuit model for arrays of square loops. Electron. Lett. **18**(7), 294–296 (1982)
6. C.K. Lee, R.J. Langley, Equivalent-circuit models for frequencyselective surfaces at oblique angles of incidence. IEE Proc. **132**(6), 395–399 (1985)
7. F. Costa, A. Monorchio, G. Manara, Efficient analysis of frequency-selective surface by a simple equivalent-circuit model. IEEE Antennas Propag. Mag. **54**, 35–48 (2012)
8. B.A. Munk, *Frequency Selective Surfaces: Theory and Design* (Wiley, Hoboken, USA, 2000). ISBN 0-471-37047-9
9. G.H.H. Sung et al., A frequency selective wall for interference reduction in wireless indoor environments. IEEE Antennas Propag. Mag. **48**(5), 29–37 (2006)

Design of a Dual Annular Ring CPW Fed UWB Antenna for Wireless Applications



Hitanshu Katiyar, Alaknanda Ashok, Abhinav Sharma and Tarun Kumar

1 Introduction

After the allocation of frequency band from 3.1 to 10.6 GHz (7.5 GHz bandwidth) by Federal Communications Commission (FCC) in 2002, there has been massive growth in the wireless communication system. 3G, 4G, and 5G technology have revolutionized the wireless technology, and this leads to the development of broadband, multiband, and reconfigurable antennas for ultra-wideband (UWB) [1] applications. Presently researchers are emphasizing on the design and analysis of small size, high gain UWB antenna which can suit the demand of laptops, smartphones, modems, and other high-end electronic gadgets. This range of frequency is very useful for many short-range applications like worldwide interoperability for microwave access systems (Wi-MAX) for IEEE 802.16 (3.3–3.6 GHz), wireless local area network (WLAN) for IEEE 802.11a (5.15–5.35 and 5.725–5.825 GHz), C-band (3.7–4.2 GHz) for satellite and radar communication signals, radio frequency identification (RFID) (6.7–6.9 GHz) communication signals, and international telecommunication union (ITU) (8.025–8.4 GHz). Ultra-wideband is a short-range communication technology that has gained great attraction due to its features such as simple structure, compact size, relatively low spectral power density, high data interaction, minimum

H. Katiyar · A. Ashok
Department of Electronics and Communication Engineering, Women Institute of Technology, UTU, Dehradun, Uttarakhand, India
e-mail: hitanshukatiyar@gmail.com

A. Sharma (✉) · T. Kumar
Department of Electrical and Electronics Engineering, University of Petroleum and Energy Studies, Dehradun, Uttarakhand, India
e-mail: abhinav.sharma@ddn.upes.ac.in

T. Kumar
e-mail: tkumar@ddn.upes.ac.in

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_55

interference, excellent data transmission rate, stable omnidirectional radiation pattern, and constant gain over the entire operating band.

In [2], the author has designed a compact UWB antenna with defective ground plane structure (DGS) structure. In [3], the author proposed a miniature low profile UWB antenna with defective ground plane. In [4], the author incorporated CPW feeding mechanism for designing UWB antenna which significantly reduces the size of the antenna maintaining a wide operating bandwidth. In [5], the author proposed a circularly polarized CPW fed antenna with radiating patch composed of hexagonal ring linked by two annular rings over two corners. In [6], the author designed a leaf-shaped CPW fed UWB antenna and observed a good agreement between simulated and measured results. CPW fed antennas are attracting the interest of researchers as it possesses key features such as wide bandwidth, low cross-polarization, and radiation loss and easy integration with active devices [7].

Considering all these silent features, in this paper, a novel annular ring-shaped CPW fed compact UWB microstrip antenna is proposed. The proposed antenna achieved a wide bandwidth ranging from 3 to 14 GHz (150%). The paper is organized in the following manner. Section 2 describes the structure and design of antenna, Sect. 3 discusses the result, and Sect. 4 provides the conclusive remarks to summarize the paper.

2 Antenna Structure and Design

The geometrical configuration of the proposed dual annular ring CPW fed UWB antenna is shown in Fig. 1. The proposed antenna is designed on FR-4 substrate with a relative permittivity ε_r of 4.4 and loss tangent tan δ of 0.02 having the overall dimensions of 25 mm × 25 mm × 1.6 mm. The anticipated antenna has two annular rings, out of which bigger annular ring has radii r_1 and r_2 and smaller annular ring has radii r_3 and r_4. Further, the two annular rings have two rectangles which are 90° apart. The thickness of the feed line is 2.2 mm and achieve characteristic impedance of 50 Ω at the design frequency. Table 1 shows all the design parameters.

3 Results and Discussion

The simulation results were obtained using HFSS software, by Ansoft Corporation which is based on finite element method (FEM) technique. It is clear from Fig. 2 that the return loss is less than 10 dB throughout the ultra-wideband. The lowest dip obtained is –37.39 dB at 11.89 GHz. The second dip is at 7.68 GHz and comes out to be approximately –20.39 dB. Thus, the result is in accordance with the application of the design. It can be observed that the proposed antenna exhibits an ultra-wideband performance from 3.1 GHz to 10.7 GHz with a fractional bandwidth of 173.84%. The VSWR is as depicted in Fig. 3, which varies between 1 and 2 throughout the

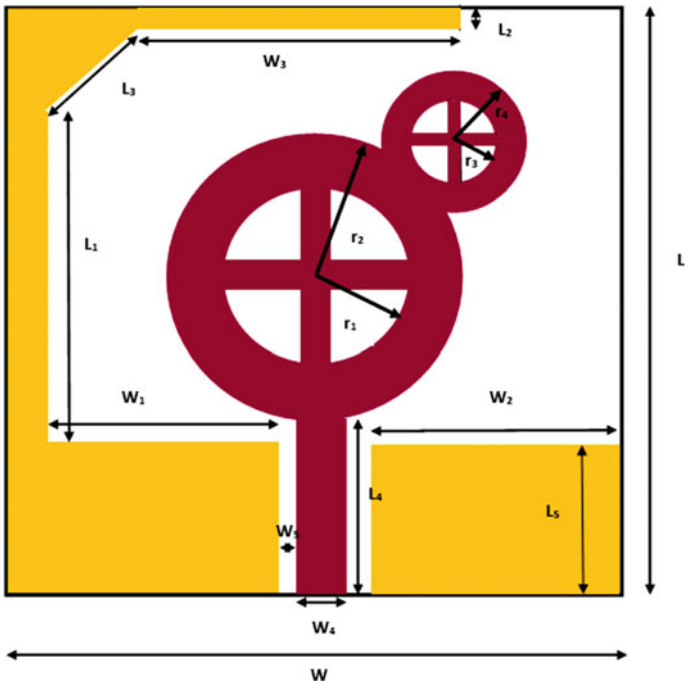


Fig. 1 Geometric presentation of the proposed dual annular ring CPW fed UWB antenna

Table 1 Design parameters of proposed UWB antenna

Parameters	L	W	L_1	L_2	L_3	L_4	L_5	W_1
Unit (mm)	25	25	11.5	1	8.13	8.6	7.5	9.8
Parameters	W_2	W_3	W_4	W_5	r_1	r_2	r_3	r_4
Unit (mm)	10.8	14.5	2.2	0.6	4.5	6.5	2	3

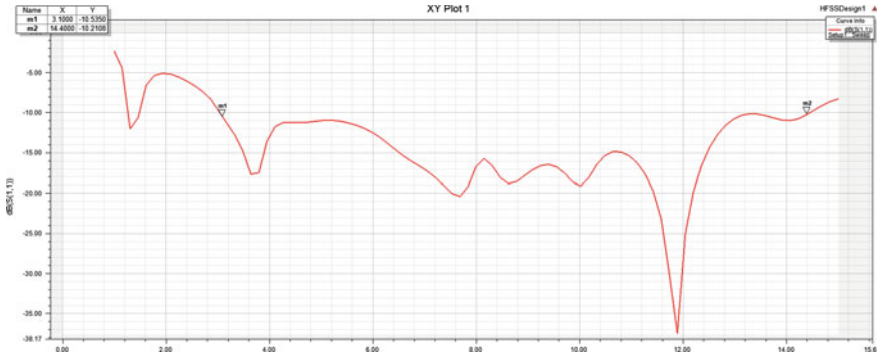


Fig. 2 Simulated results of proposed dual annular ring CPW fed UWB antenna

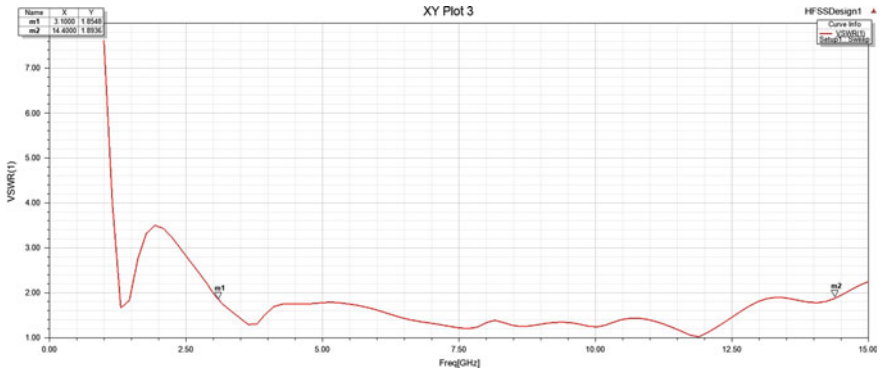


Fig. 3 Simulated results of VSWR of proposed dual annular ring CPW fed UWB antenna

operating bandwidth. VSWR is a matching coefficient whose value indicates how perfectly an antenna is matched with the transmission line. Therefore, at 11.89 GHz and 7.68 GHz, VSWR is approximately 1 as perfect matching is achieved at these two resonant frequencies.

Figure 4 shows the gain versus frequency 2D plot of the proposed antenna. The simulation result shows that average gain is 3.24 dB and the peak gain is 5.24 at 14.37 GHz. Figure 5 shows the 3D polar plot which depicts that maximum gain is 3.91 dB.

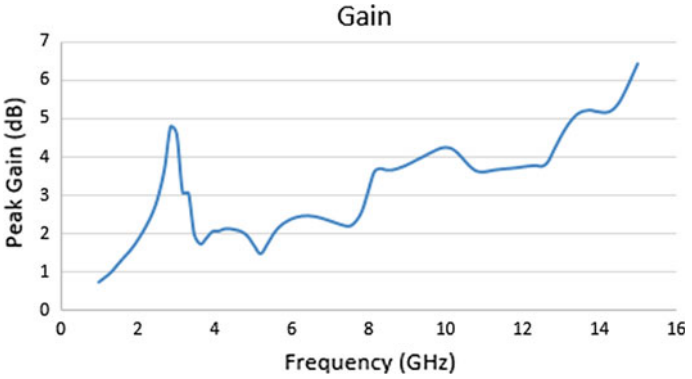


Fig. 4 2D simulated gain of proposed dual annular ring CPW fed UWB antenna

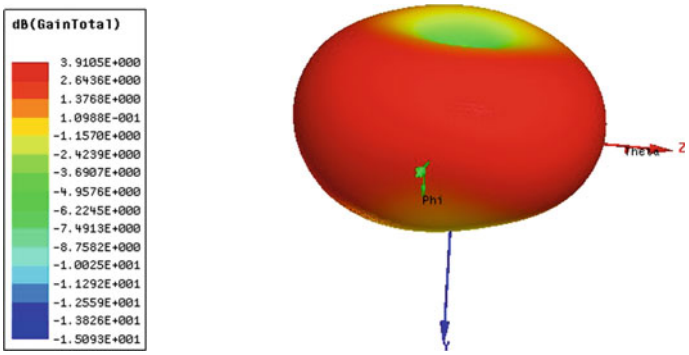


Fig. 5 Simulated gain of proposed dual annular ring CPW fed UWB antenna

4 Conclusion

A novel dual annular ring CPW fed UWB antenna is being built and simulated. The prototype has dimensions of 25 × 25 × 1.6 mm^3 and is fabricated on an inexpensive FR-4 substrate. The simulation results show that throughout the UWB the S_{11} parameter (reflection coefficient) is less than 10 dB and voltage standing wave ratio is less than two. The future scope of this work is to reduce the size of the antenna and to incorporate band notching for certain range of frequencies. In UWB antenna, Wi-Max (3.3–3.7 GHz), C-Band (3.7–4.2 GHz) and wireless local area network (5.15–5.825 GHz) causes undesirable interferences which can be eliminated using band notching techniques.

References

1. H.G. Schantz, A brief history of UWB antennas. IEEE Aerosp. Electron. Syst. Mag. **19**(4), 22–26 (2004)
2. V. Pant, et al., A novel compact UWB antenna for wireless applications, in *Third International Conference on. Intelligent Communication, Control and Devices* (2018)
3. M.A. Trimukhe, B.G. Hogade, Miniature low profile UWB antenna with defective ground plane. in *Second International Conference on Inventive Communication and Computational Technologies (ICICCT)*, (2018), pp. 111–114
4. A.K. Gautam, R. Chandel, B.K. Kanaujia, A CPW-fed hexagonal-shape monopole-like UWB antenna. Microw. Opt. Technol. Lett. **55**(11), 2582–2587 (2013)
5. R. Dhara, M. Midya, M. Mitra, S.K. Jana, CPW-fed tetra band circular polarized antenna for wireless communication applications. in *Applied Electromagnetics Conference (AEMC)* (2017), pp. 1–2
6. S. Kundu, S.K. Jana, A new compact leaf shaped CPW fed UWB antenna, in *International Conference on Antenna Innovations and Modern Technologies for Ground, Aircraft and Satellite Applications* (2017)
7. S. Chilukuri, K. Dahal, A CPW fed antenna with semicircular patch on triangle for wideband wireless applications, in *11th International Conference on Software, Knowledge, Information Management and Applications (SKIMA)* (2017)

Path Loss Calculation at 900 MHz and 2.4 GHz in WBAN



Purnima K. Sharma, T. Vijay Sai and Dinesh Sharma

1 Introduction

A multiple body parameters-computing device which can be worn on or kept inside human bodies is called the wireless body area network (WBAN). A mobile hub which acquires the data from the user and transmits it to a remote database is required. A smartphone in a pocket or bag serves the purpose. A body area network consists of a mobile or any other electronic device in a pocket or pouch which is attached to the waist. This device acts like a data hub receiving the body parameter signals from the patient and transmitting them to a remote database [1]. Communication is the transmission of information from one place to another without the help of wires, cables or any such conductors. The distance of transmission can be anywhere between a few metres to thousands of kilometres [2]. Path loss can be defined as the attenuation suffered by a signal travelling along a line between a transmitter and a receiver [3]. The causes of path loss include physical phenomena such as reflection of signal by striking any obstacle, refraction through a medium and diffraction at the edges of a body. The general expansion of the radio frequency signal front in free space causes the propagation losses which the path loss generally constitutes of. A tabular form including different cases along with their descriptions and frequency bands which are used in IEEE 802.15.6 device's operation is given in Table 1.

P. K. Sharma · T. Vijay Sai (✉)
ECE Department, SVEC, Tadepalligudem, AP, India
e-mail: tvijaysai1@gmail.com

P. K. Sharma
e-mail: purnima.kadali@gmail.com

D. Sharma
ECE Department, CCET, Chandigarh, India
e-mail: sharma82dinesh@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_56

Table 1 Different channel models in WBAN

Case	Description	Frequency band	Channel models
Case S1	In between implant and implant	401–405 MHz	CM1
Case S2	In between implant and body surface	401–405 MHz	CM2
Case S3	In between implant and external	401–405 MHz	CM2
Case S4	In between body surface and body surface (LOS)	13.5, 50, 400, 600, 900 MHz 2.4, 3.1–10.6 GHz	CM3
Case S5	In between body surface and body surface (NLOS)	13.5, 50, 400, 600, 900 MHz 2.4, 3.1–10.6 GHz	CM3
Case S6	In between body surface and external (LOS)	900 MHz 2.4, 3.1–10.6 GHz	CM4
Case S7	In between body surface and external (NLOS)	900 MHz 2.4, 3.1–10.6 GHz	CM4

2 Literature Survey

The natural expansion of the radio wave front in free space causes the propagation losses. The path loss constitutes of these propagation losses [4]. Body area network has its applications mainly in the medical field. These applications include electronic sensors which monitor patients for different health conditions. A system that can act as a data book in the field of diagnosis and treatment of ailments is also proposed [1]. This technology can also blood pressure (BP) and other patient vital signs. Research on wireless body area networks can be gone back to 1961 with work from Mackay [5] on radio telemetry which is kept inside the body. T G Zimmerman initially gave the wireless body area networks in the editorial from 1996 [2]. A homogeneous loss medium having an insulated dipole antenna is tested for path loss models in paper [6]. And the effect of permittivity and conductivity on the path loss is studied. Path loss models become widely acceptable and applicable if extraction of the propagation channel properties is possible for body-centric systems [7]. The possible thickness of the tissues in an adult and a child at various regions of the body is studied in [8]. Some sensors are designed such that they can be comfortably implanted inside the body [9]. A portal receives the data from the sensors which sends the data through the mobile system to the medics [10]. In [11], a way to calculate the conducting medium antenna gain is developed and by using this calculated gain, a path loss model is proposed. Several path loss models have been proposed for implant and implant-surface propagation by [12, 13]. Antennas for implant and surface of body propagation have been designed and implemented [3] at 2.45 and 5.8 GHz, and a way to find the gain of these antennas can be used in developing PL models. This layered model consists of the skin (2 mm), subcutaneous adipose tissue (SAT) (23 mm),

breast tissue (BT) (30 mm) and muscle tissue (6 mm), and is terminated in the liver (infinite) [12].

3 Experimental Setup and Measurements

The Yagi and dipole antennas are used for calculating the loss between transmitter and receiver in WBAN. N9320B Spectrum Analyzer and N9310A RF Signal Generator manufactured by Agilent Technologies are used, and the analysis is carried out by different antennas like dipole antenna and Yagi antenna. The RF signal generator is used to generate a signal with desired frequency and power. The spectrum analyzer is used to display the signal with the specifications like power in dB and frequency in the range of 100 Hz to 3 GHz. Three different positions on the body like chest to pocket, pulse to pocket and spine to pocket are checked which are pictorially represented in Figs. 1, 2 and 3, respectively.

Initially, two antennas of same kind are taken. One is connected to the RF Signal generator using a connector. Another one is connected to the spectrum analyzer with the connector. Both the antennas are adjusted in such a way that they are in line of sight (LOS). It is observed that when the distance between the antennas is increased, the loss is gradually increased. Figure 4 represents the measurement of path loss between the Yagi antennas placed in LOS position with some distance.

Figure 6 shows the experimental setup of spectrum analyzer and RF signal generator.

Figure 5 shows the waveform on the spectrum analyzer with a frequency of 900 MHz. Similarly, dipole antenna and Yagi antenna are tested using the analyzers.

Now the position of the antenna is changed to non-line of sight (NLOS). The transmitting antenna is placed in the shirt pocket, the receiving antenna is placed

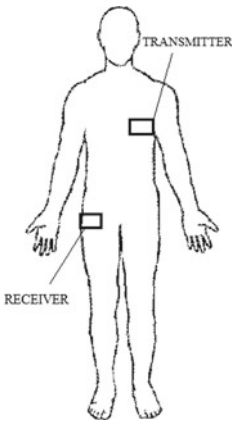


Fig. 1 Chest (heart) to pocket

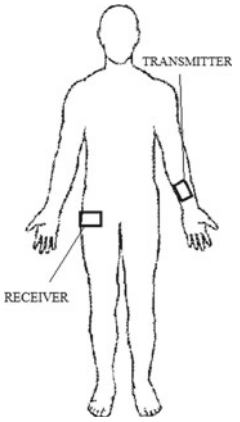


Fig. 2 Pulse to pocket

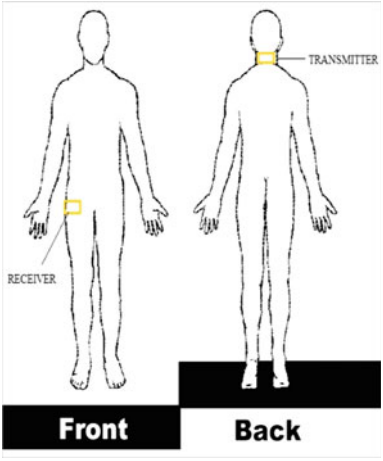


Fig. 3 Spine to pocket

in the pant pocket as shown in Fig. 7 and path loss is calculated. The transmitting antenna is changed to another position, like spine as shown in Fig. 8 and wrist as shown in Fig. 9 and the path loss is calculated.

Table 2 shows the transmitted and received power by using Yagi antenna at 2.4 GHz frequency.

Table 3 shows the transmitted and received power by using Yagi antenna at 900 MHz frequency.

Table 4 shows the transmitted and received powers by using Dipole antenna at 900 MHz frequency.



Fig. 4 Antennas line of sight (LOS) setup using Yagi antenna



Fig. 5 Experimental setup

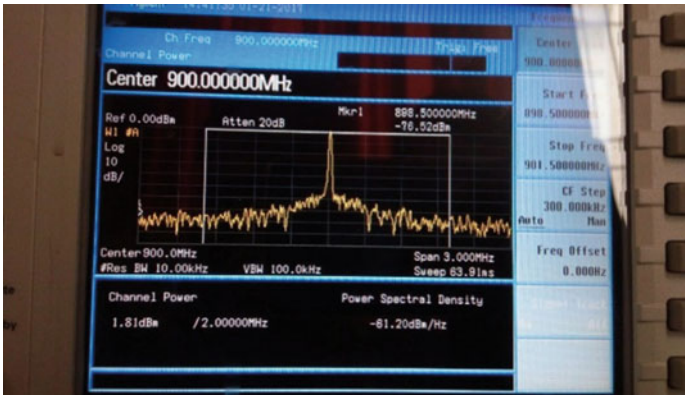


Fig. 6 Waveform in spectrum analyser



Fig. 7 Chest (heart) to pocket



Fig. 8 Spine to pocket



Fig. 9 Wrist (pulse) to pocket

Table 2 Measurement by Yagi antenna at 2.4 GHz

Measurement by Yagi antenna at 2.4GHZ			
Frequency (GHz)	Transmitted power	Received power	Position
2.4	10	−5.97	LOS
2.4	−10	−24.67	
2.4	−27	−52	
2.4	10	−54.25	Spine to pocket
2.4	−10	−57.25	

Table 3 Measurement by Yagi antenna at 900 MHz

Measurement by Yagi antenna at 900 MHz			
Frequency (MHz)	Transmitted power	Received power	Position
900	10	−52.88	Heart to pocket
900	−10	−55	
900	−27	−61.20	
900	10	−57.88	Spine to pocket
900	−10	−58.88	

Table 4 Measurement by dipole antenna at 900 MHz

Measurement by dipole antenna at 900 MHz			
Frequency (MHz)	Transmitted power	Received power	Position
900	10	3.96	LOS
900	−10	−16.88	
900	−27	−37.67	

4 Conclusion

Wireless body area network is a budding technology which offers numerous advantages and ample of applications to many patients and to the society as well in one way or the other. Applications of WBAN include medical health care, sports and fitness monitoring, and military and security. In the whole process of finding the path loss, dipole antenna gave a good result compared to Yagi antenna, patch antenna and other antennas. Among the propagation models such as Okumura model, Hata model, free space path loss model and other models, it is found that a log distance propagation model proved to be the model with the lowest path loss. Analysing different antennas for selecting the best antenna is achieved using experimentation. From our analysis, dipole antenna is found to be the best antenna for the field of WBAN.

Acknowledgements We gratefully acknowledge Sri Vasavi Engineering College, Tadepalligudem, Andhra Pradesh management, staff and students for their cooperation during our work in their Wireless Communication Lab.

References

1. S. Varghese, R.M. Nair, A wireless physiological parameter monitoring system. IJARCET **4**(6) (2015)
2. D. Kurup, W. Joseph, G. Vermeeren, L. Martens, Path loss model for in-body communication in homogeneous human muscle tissue. IET Electron. Lett. **45**, 453–454 (2009)
3. K. SaiSanath Kumar, K. Naveen Kumar Reddy, V. Pushpavathy, P. Raja Reddy, D. Sharma, P.K. Sharma, B. Vamsi Krishna, Calculation of path losses at CM3 for wireless body area networks (WBAN) by using different types of antennas. Int. J. Appl. Eng. Res. (IJAER) **11**(7), 5210–5217 (2016). (Print ISSN 0973-4562, Online ISSN 1087–1090)
4. M.A. Alim, M.M. Rahman, M.M. Hossain, A. Al-Nahid, Analysis of large-scale propagation models for mobile communications in urban area. IJCSIS **7**(1) (2010)
5. R. Stuart Mackay, Radio telemetering from within the body. Science 135, 1410 and 1961. https://doi.org/10.1126/science.134.3488.1410-a
6. P.S. Hall, Y. Hao, Antennas, and Propagation for Body-Centric Wireless Communications
7. A. Christ, A. Klingenbock, T. Samaras, C. Goiceanu, N. Kuster, The dependence of electromagnetic far-field absorption on body tissue composition in the frequency range from 300 MHz to 6 GHz. IEEE Trans. Microwave. Theory Tech. **54**, 2188–2195 (2006)
8. IEEE. IEEE 802. 15 TG-6 Body Area Networks (2011)
9. J. Xing, Y. Zhu, A survey on body area network, in *Fifth International Conference on Wireless Communication, Networking and Mobile Computing* (2009), pp. 1–4
10. D. Kurup, W. Joseph, E. Tanghe, G. Vermeeren, L. Martens, Extraction of antenna gain from path loss model for in-body communication. IET Electron. Lett. **47**(23),1262–1263 (2011)
11. K. Ito, W. Xia, M. Takahashi, K. Saito, An implanted cavity slot antenna for medical communication systems, in *Proceedings of the 3rd European Conference on Antennas and Propagation (EuCAP)* (Berlin, Germany, 2009), pp. 718–721
12. P.K. Sharma, D. Sharma, Mobile wireless technology is boon or curse? Indian J. Sci. Technol. **9**(40), 0974–5645 (2016). (ISSN (Print): 0974-6846 ISSN (Online) https://doi.org/10.17485/ijst/2016/v9i40/84293
13. Wireless BANs for healthcare: a survey. Int. J. Adhoc, Sensor Ubiquitous Comput. (3) (2012)

14. A. Alomainy, Y. Hao, X. Hu, C.G. Parini, P.S. Hall, UWB on-body radio propagation and system modelling for wireless body-centric networks. in *IEE Proceedings Communications* (IET: Stevenage, UK, 2006), pp. 107–114
15. J. Tak, K. Kwon, S. Kim, J. Choi, Dual-band on-body repeater antenna for in-on-on WBAN applications. Int. J. Antennas Propag. **2013**, 1993–1997 (2003)
16. P.K. Sharma, D. Sharma, R.K. Singh, Evolution of mobile wireless communication networks (0G-8G). Int. J. Appl. Eng. Res. (IJAER) **10**(6), 14765–14778 (2015). (Print ISSN 0973-4562, Online ISSN 1087–1090)

Ladder Mat Shape Microstrip Patch Antenna for X Band



Nivedita Dash and Sunil Kumar Singh

1 Introduction

An antenna is a radiator which is used for transmitting or collecting electromagnetic waves. Microstrip antenna fabricates on printed circuit board using microstrip technique [1]. Microstrip antenna is an internal antenna and mostly used at microwave frequency. Microstrip antenna is well known as a patch antenna. Shapes of transmitting patch can be design in form of triangle, square, rectangle, dipole, circle, etc. Thin strip, circle, square, and rectangle shapes are mostly used because these shapes are easily examined and manufacture [2, 3].

For wireless application, microstrip patch antenna is used because of its lightweight and low planer configuration [4]. Figure 1 shows the side view of microstrip patch antenna, which is fed by microstrip transmission line. High conductive material (typically copper) is used for construct ground, patch antenna, and microstrip transmission line. L, W, and h show the patch's length, patch's width, and height of substrate, respectively. Typically, substrate height h is very smaller but not much smaller than 0.05 of wavelengths. Aperture coupling, microstrip, coaxial, and proximity coupling are some common techniques used to provide feeding. In microstrip line feeding, the coupling feeding exists between the patch and the microstrip line could be in the form of butt/edge [2].

In this case, a square size with ladder shape microstrip antenna with different rectangular and square slots designed and studied, and also studied return loss, radiation pattern, gain, and VSWR [5].

N. Dash (✉) · S. K. Singh
EC Department, Jabalpur Engineering College, Jabalpur, Madhya Pradesh, India
e-mail: nivedita.333@gmail.com

S. K. Singh
e-mail: sspjec2007@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_57

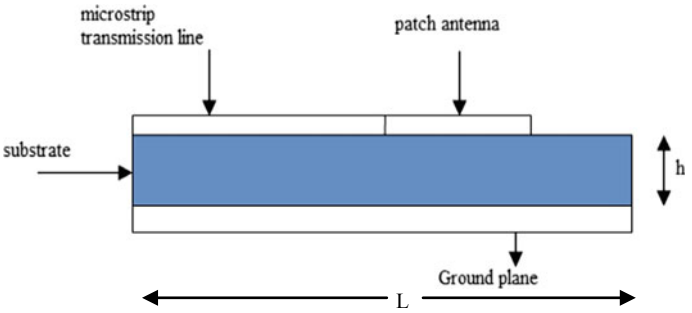


Fig. 1 Side view of microstrip patch antenna

2 Antenna Design Configuration

To design microstrip patch antenna, first we have to find out the practical width of patch, which provides good radiation [6].

$$W = \frac{1}{2f_r\sqrt{\mu_\circ \epsilon_\circ}}\sqrt{\frac{2}{1+\epsilon_{\text{reff}}}} = \frac{C}{2f_r}\sqrt{\frac{2}{1+\epsilon_{\text{reff}}}}$$

where

ϵ_{reff} = substrate's dielectric constant
C = velocity of light

After determining width of patch, the effective dielectric ϵ_{reff} can be represented by

$$\epsilon_{\text{reff}} = \frac{\epsilon_r - 1}{2} + \frac{\epsilon_r + 1}{2}\sqrt{\left[12\frac{h}{W} + 1\right]}$$

Patch's width and effective dielectric constant found, now the extension of the length ΔL can be expressed as

$$\Delta L = 0.412h\frac{(\epsilon_{\text{reff}} + 0.3)\left(\frac{W}{h} + 0.214\right)}{(\epsilon_{\text{reff}} - 0.258)\left(\frac{W}{h} + 0.8\right)}$$

Patch's length can be determined by solving for L [6]

$$L = \frac{C}{2f_r\sqrt{\epsilon_{\text{reff}}}} - 2\Delta_L$$

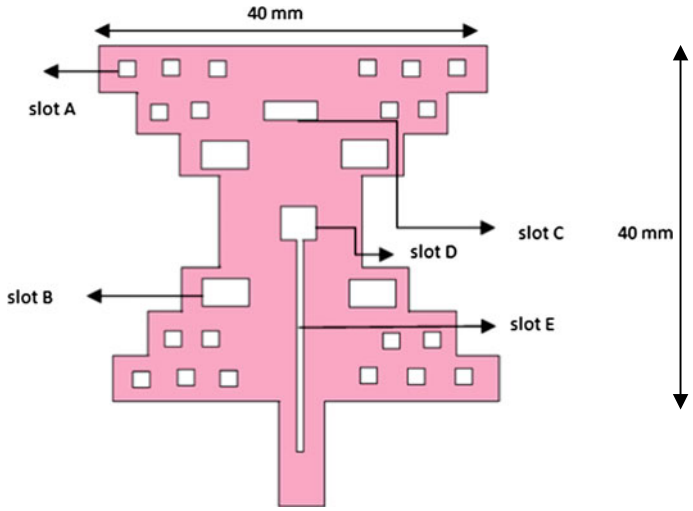


Fig. 2 Design of ladder mat shape patch antenna

In this paper, the proposed microstrip patch antenna resonant frequency is 8 GHz, and it is consisting of ground, substrate, and patch. For substrate, FR4 epoxy material is used whose relative permittivity is $\epsilon_r = 4.4$. Patch is in shape of ladder mat. Square patch has some different sizes of rectangular and square slots. These slots are named as sot A, B, C, D, and E. Slot A consists of 20 slots with same dimensions, and it is placed left and right, and up and down side of patch. Slot A length and width is 2 mm. Slot B is consist of four slots with same dimensions, and it is placed between slot A. Slot B length is 4 mm and width is 6 mm. Slot C length is 2 mm and width is 4 mm. Slot D length and width is 4 mm. Slot E length is 24 mm and width is 0.3 mm. Length and width of conducting ground square is 60 mm. Length and width of patch is 40 mm. The 50 Ω microstrip feed line is applied. In this antenna, patch and ground are made finite conductor with the help of boundary condition of perfect E. Boundary condition of radiation is also provided to radiation box. To get radiation pattern, a radiation box is also made in infinite free space which will minimize the reflection and maximize the absorb from the antenna. Figure 2 shows the proposed antenna structure.

3 Result and Discussion

3.1 Return Loss S_{11}

In this paper, the performance of the proposed antenna is measured through high-frequency structural simulator (HFSS 13). The simulation of return loss S11 is shown

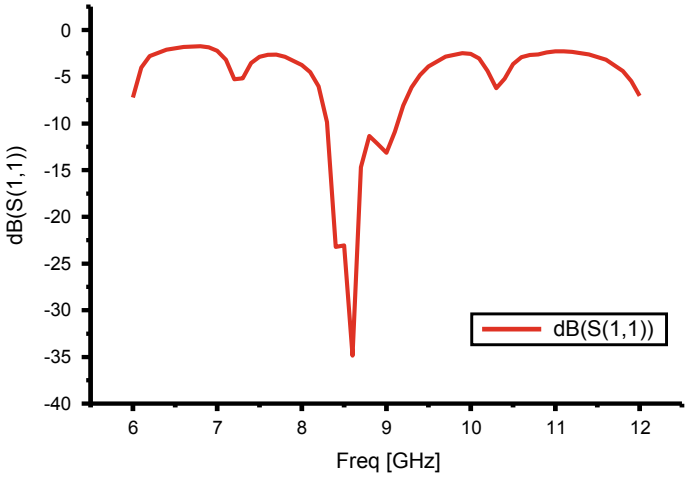


Fig. 3 Simulated result of return loss S_{11}

in Fig. 3. Return loss defines the ratio from power reflected to the launched power from the antenna. For better performance of the antenna, the return loss should be less than –10 dB. In this paper, return loss of –34.87 dB is obtained at frequency 8.6 GHz, which is simulated frequency.

3.2 VSWR

The simulation of the VSWR is shown in Fig. 4. VSWR defines mismatch between transmitting or receiving element to transmission cable. Ideally VSWR should be less than 2, and in this proposed antenna, it is obtained 0.3135.

3.3 Radiation Pattern

Radiation pattern shows radiated energy by the transmitting or receiving radiator. Figure 5 shows the achieved radiation pattern of this proposed antenna. Achieved radiation pattern is close to omnidirection.

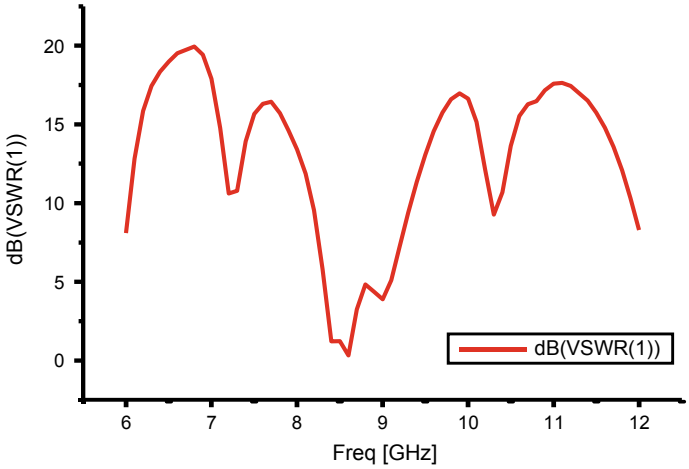
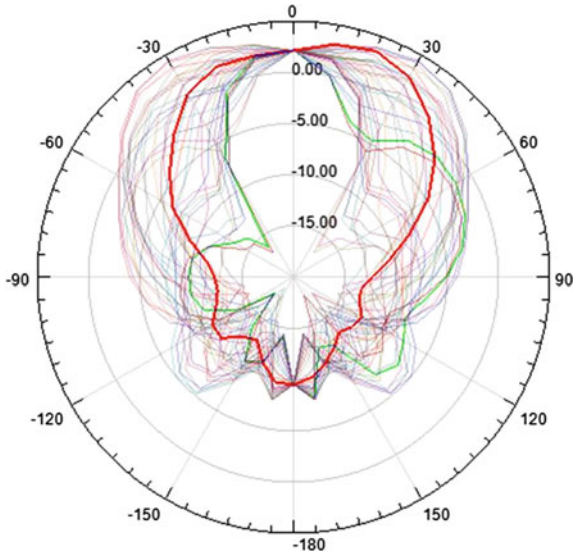


Fig. 4 VSWR of the antenna

Fig. 5 Radiation of antenna



3.4 Directivity

An antenna's directivity is the ratio of the radiation intensity in a given direction from the antenna to the radiation intensity averaged in all directions. The average radiation intensity is equal to the total power radiated by the antenna divided by 4π. If the direction is not specified, the direction of maximum radiation intensity is

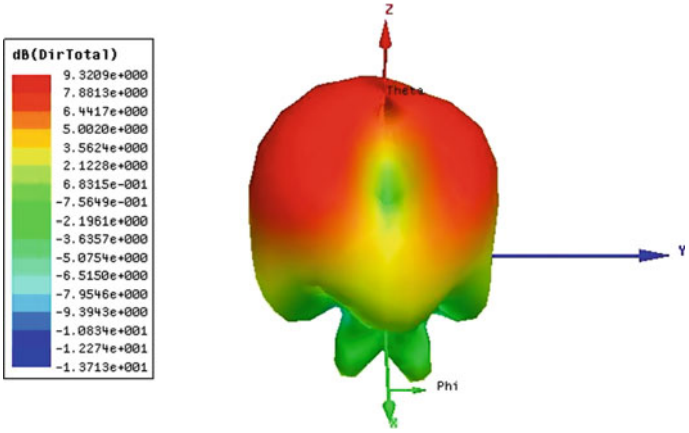


Fig. 6 Directivity of antenna

implied. In this article, antenna's directivity achieved 9.3209 dB. Figure 6 shows the proposed antenna's directivity.

3.5 Gain

The antenna gain defines how antenna sends or receives signal in a specified direction. Basically, it states the quantitative relation between greatest amount of radiated power in a specified direction to the overall input power. Gain of this proposed antenna is 4.420 dB (Fig. 7).

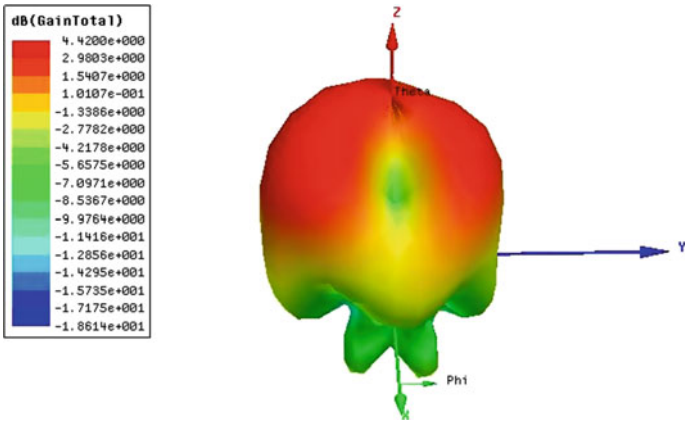


Fig. 7 Gain of antenna

$$\text{Gain} = 4\Pi \frac{\text{radiation intensity}}{\text{total input(accepted) power}}$$

3.6 Bandwidth

An antenna's bandwidth defines, a particular range of frequencies in which antenna can work properly. For this proposed research article bandwidth is obtained 0.9.

4 Conclusion

In this proposed paper a square size with ladder mat shape microstrip patch antenna is designed. It has both square and rectangular sots. The proposed antenna can reduce the cost of antenna, because slots reduced the copper. We have done the same simulation in CST software for better interpretation.

Figure 8 shows the compared simulation of return loss S_{11} between the HFSS and CST. The designed frequency of this antenna is 8 GHz, and the simulated result is near to 8 GHz. With the help of HFSS and CST software, we can conclude that the simulated result is approximately same. In this band range proposed antenna can be used for radar and satellite application. In this paper, the proposed antenna has high gain, low return loss, and small size. The designed radiator is suitable for "X" band application.

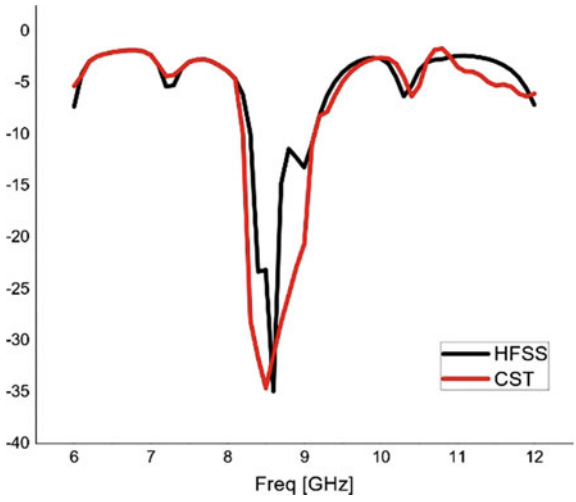


Fig. 8 Compared simulated return loss between HFSS and CST

References

1. H. Saini, A. Kaur, A. Thakur, R. Kumar, A parametric analysis of ground slotted patch antenna for X-band applications, in *2016 3rd International Conference on Signal Processing and Integrated Networks (SPIN)*
2. C.A. Balanis, *Antenna Theory, Analysis and Design*, 2nd edn. (Wiley, 1997)
3. V. Singh, B. Mishra, R. Singh, A compact and wide band microstrip patch antenna for X-band applications, in *2015 Second International Conference on Advances in Computing and Communication Engineering*. 978-1-4799-1734-1/15 $31.00 © (IEEE, 2015). https://doi.org/10.1109/icacce.2015.135
4. M.M. Islam, M.T. Islam, M.R.I. Faruque, W. Hueyshin, Design of an X-band microstrip patch antenna with enhanced bandwidth, in *Proceedings of 2013 2nd International Conference on Advances in Electrical Engineering (ICAEE 2013)* 19–21 December 2013, Dhaka, Bangladesh
5. S.S. Sran, J.S. Sivia, Quad staircase shaped microstrip patch antenna for S, C and X band applications, in *International Conference on Computational Modeling and Security (CMS 2016)*. Procedia Computer Science vol. 85 (2016), pp. 443 – 450
6. A.K. Sidhu, J.S. Sivia, Microstrip rectangular patch antenna for S and X band applications, in *IEEE WiSPNET 2016 conference*. 978-1-4673-9338- 6/16/$31.00 (IEEE, c2016)

Frequency Agile Slotted Diagonally Sliced Elliptically Polarized Square Patch Antenna



Shipra Bhatia and M. V. Deepak Nair

1 Introduction

Since the past few years, reconfigurable patch antennas have been widely investigated, designed, and used to meet the stringent requirements of modern communication standards like 4G and proposed 5G standards [1–5]. Due to the possible reconfiguration properties of the antenna parameters like frequency, radiation pattern, and polarization, reconfigurable antennas are getting importance among many researchers and are also been widely adopted in the military where abovementioned functions can be fulfilled by a single antenna [2]. Reconfigurable properties of the antenna include frequency reconfiguration (i.e., a shift in frequency w.r.t. change in diode bias) [3], radiation pattern reconfiguration (done by controlling the direction of current flow in the antenna) [4], and polarization reconfiguration (switching between linear, elliptical, and circular polarization) [5]. These reconfiguration properties can be achieved by electronic means, optical means (using photoconductive materials), physical means (by changing the structure of the antenna), or by using ferrites or liquid crystals as antenna material [6, 7]. Reconfiguration through electronic means is achieved by using electronic switching elements like PIN diode, Varactor diode, FETs, and Micro-Electro-Mechanical Systems (MEMS) switches. By the use of switching elements in the patch, length of the antenna can be tuned to some extent causing it to resonate at different frequencies [8, 9]. As compared to other solid-state switches such as FET, MEMS, and Varactor diodes, PIN diode offers better linearity and higher isolation (ranging 20–40 dB). It also provides very high

S. Bhatia (✉) · M. V. Deepak Nair
Department of ECE, The LNM Institute of Information Technology, Jaipur, Rajasthan, India
e-mail: 17mec004@lnmiit.ac.in

M. V. Deepak Nair
e-mail: deepak.nair@lnmiit.ac.in

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_58

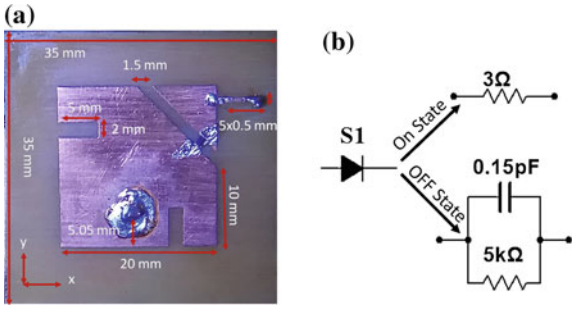


Fig. 1 **a** Antenna fabricated. **b** Equivalent circuit diagram of PIN diode

switching speed, in nanoseconds, and larger bandwidth [10, 11]. Several reconfigurable antenna designs based on different reconfiguration properties and techniques have been proposed in [6–11]. In this work, frequency agile microstrip patch antenna has been proposed by diagonally slicing one corner of the square patch antenna. The proposed antenna is a single-layer, single-feed design with two slits incorporated in it and the dimensions of the substrate is $35 \times 35 \times 1.58$ mm^3. The switching between two different frequencies is achieved by inserting a PIN diode at an appropriate position between the two parts of the patch.

2 Antenna Design and Geometry

The geometry of the proposed diagonally trimmed frequency agile antenna is shown in Fig. 1a. Many rectangular reconfigurable antennas with a larger size have been proposed earlier but this square-shaped elliptically polarized antenna is small and reconfigurable [4, 5]. The antenna is fabricated using FR4 with a dielectric constant of 4.4 and loss tangent of 0.02. The square patch of size 20×20 mm^2 is co-axially fed at the center of the x-axis and 5.05 mm from the edge along y-axis which is found to be best suited using Ansys HFSS [12]. Two slits of dimensions 5×2 mm^2 are introduced in the antenna patch to reduce the size of the antenna and also to shift the polarization from linear to elliptical. The dimension of slits has been selected after parametric variation using optimetrics tool in ANSYS HFSS. Along with the PIN diode, other biasing components are also used in the design which is discussed in the following section.

3 Biasing Circuit and Effect of PIN Diode's Position

For simulations and verification's, ANSYS HFSS is used where PIN diode is modeled as a rectangular metal strip of impedance 3 Ω with the same length as that of diagonal slot whereas, for experimental validations, RF PIN diode BAR50-02V from Infineon

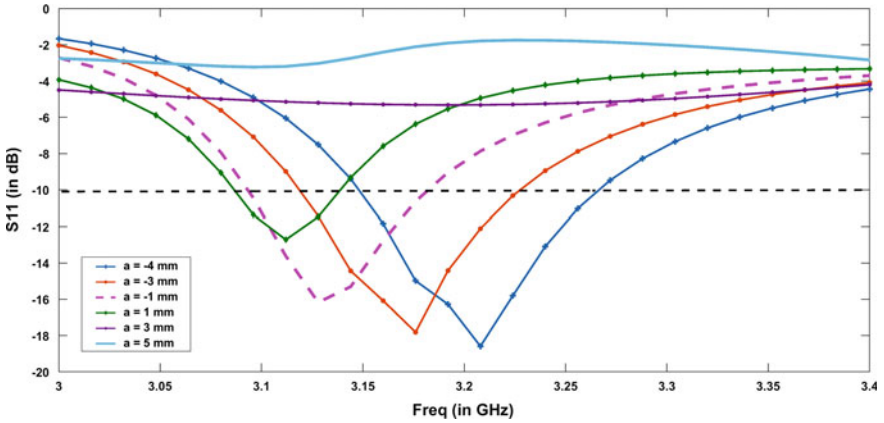


Fig. 2 Effect on S11 with respect to PIN diode's position

Technologies is placed in the diagonal slot which has a typical forward current of 100 mA and forward voltage of 0.95 V. Figure 1b shows the equivalent circuit of BAR50-02V PIN diode as depicted in the datasheet provided by the Infineon Technologies. Bias Tee CBTEE 01-50-6000 has been connected to the coaxial feed of the patch antenna to provide both DC biased voltage and RF incident power to the antenna. The maximum RF power that can be applied to Bias Tee is 2 W at a maximum DC port voltage of 15 V. To block the RF power and allow the DC current to flow seamlessly to ground via the sliced part of practical antenna, RF surface mount inductor of 51 nH is also used in the antenna design during the fabrication process which has a current rating of 250 mA and typical maximum DC resistance of 1.2 Ω.

Figure 2 shows the effect of changing the position of PIN diode from right to left corner of the diagonal slot. The diode position is labeled as parameter "a" and is made to sweep across frequency range using parametric variation in the optimetrics tool of Ansys HFSS. It is shown that PIN diode at rightmost position of slot, i.e., at –4 mm gives highest return loss value with a larger bandwidth. As the diode is shifted toward the left side of the slot, the return loss and bandwidth starts to decrease as compared to –4 mm position of the diode. After a specific position of the PIN diode, i.e., at leftmost corner of the diagonal slot, return loss is above –10 dB value.

4 Results and Observations

For experimental validations of the proposed design, Keysight Fieldfox Microwave Vector Network Analyzer N9925A was used. All the above and the following experimental results are obtained from the VNA mentioned before. The results and observations of the proposed antenna design are shown in the subsections below.

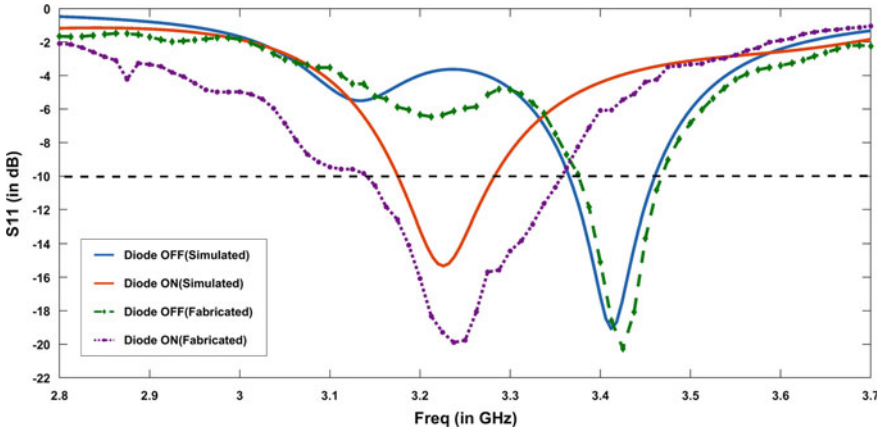


Fig. 3 Reflection coefficient curve for proposed design

4.1 Reflection Coefficient

Figure 3 depicts the simulated and fabricated return loss curve of the proposed reconfigurable antenna when PIN diode is in ON and OFF states, respectively. When the PIN diode is in OFF state, the main part of the patch resonates at 3.40 GHz providing a bandwidth of 102 MHz and –19 dB return loss for VSWR <2. PIN diode switches to ON condition when it gets DC voltage above the threshold voltage and at that condition, both parts of patch gets connected and the electrical length (size) of patch increases which shifts the frequency to 3.21 GHz with a bandwidth of 210 MHz and return loss of –15 dB for VSWR <2. Good agreement has been found between simulated and experimental results of the antenna design. The choice of the frequencies is made such that the proposed antenna can be used for applications like radio navigation, satellite communications, and mobile communications.

4.2 Gain and Axial Ratio (AR)

It is observed from the simulations conducted in ANSYS HFSS that the proposed reconfigurable antenna provides a gain of 2.58 dBi when the diode is in OFF state and when the diode is in ON state, the antenna gain reduces to 1.6 dBi. The axial ratio is the ratio of the magnitudes of the major and minor axis defined by the electric field vector. For elliptical polarization, AR value is between 1 and infinity [2]. An axial ratio value of 10.21 dB is achieved when the diode is OFF. It reduces to 7.33 dB as PIN diode switches ON. Elliptical polarization is achieved in both the switching cases which can also be used in FM stations to improve reception of the signal.

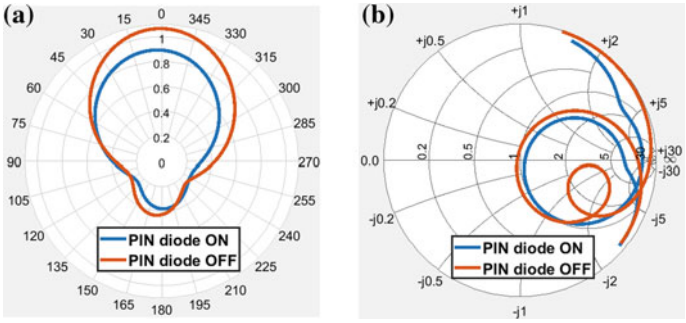


Fig. 4 **a** E plane radiation pattern. **b** Smith chart for proposed antenna

4.3 Radiation Pattern and Smith Chart

Figure 4a depicts the E plane radiation pattern for both the switching cases. It is observed that antenna has maximum forward radiation with small minor lobes as well. Figure 4b shows radiation patterns in Smith chart when the diode is in ON and OFF state. It is also visible from the provided charts that the antenna is perfectly matched to 50 Ω value as the resonant frequency of the antenna lies very close to the center of the plot, i.e., value 1 in the Smith chart.

4.4 E Field Patterns

Figure 5a shows the magnitude of E field when the diode is in OFF state and corresponding to that no current flows to the trimmed part of the patch. As the diode gets ON, as shown in Fig. 5b, we can see some parts of electric field travels to trimmed patch via PIN diode.

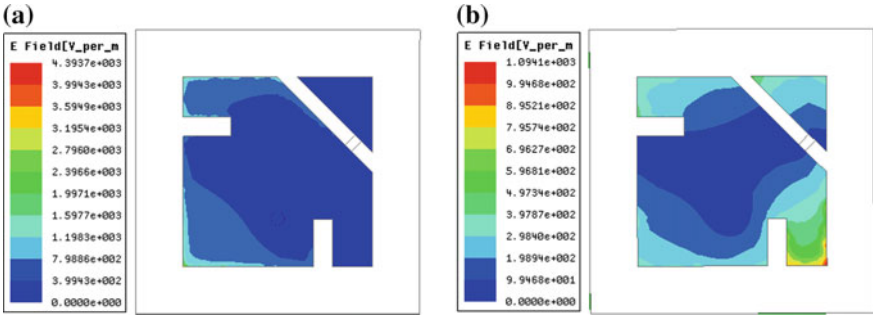


Fig. 5 **a** E field pattern when: **a** Diode is OFF. **b** Diode is ON

5 Conclusions

Proposed diagonally trimmed frequency reconfigurable square patch antenna with two slits incorporated in it, resonates at two different frequencies (3.40 and 3.21 GHz). The switching between two different frequencies is achieved by using a PIN diode inserted between two parts of the patch. This antenna provides elliptical polarization in both the switching cases of the diode. The proposed antenna experimentally provides a maximum gain of 2.58 dBi and highest bandwidth of 210 MHz which is around 80 MHz higher than the simulated design. This is achieved using a simple structure of the microstrip antenna with single-layer, single-feed diagonally trimmed patch antenna. It is evident that this design is very much suitable for applications like radio navigation, satellite communication where currently two separate antennas are used to achieve such properties.

References

1. J. Costantine, Y. Tawk, S.E. Barbin, C.G. Christodoulou, Reconfigurable antennas: design and applications. Proc. IEEE **103**(3), 424–437 (2015)
2. Y. Shi, Y. Cai, X.-F. Zhang, K. Kang, A simple tri-polarization reconfigurable magneto-electric dipole antenna. IEEE Antennas Wirel. Propag. Lett. **17**(2), 291–294 (2018)
3. S. Soltani, P. Lotfi, R.D. Murch, A port and frequency reconfigurable MIMO slot antenna for WLAN applications. IEEE Trans. Antennas Propag. **64**(4), 1209–1217 (2016)
4. N. Nguyen-Trong, L. Hall, C. Fumeaux, A frequency- and pattern- reconfigurable center-shorted microstrip antenna. IEEE Antennas Wirel. Propag. Lett. **15**, 1955–1958 (2016)
5. A. Khidre, K.-F. Lee, F. Yang, A.Z. Elsherbeni, Circular Polarization Reconfigurable Wideband E-Shaped Patch Antenna for Wireless Applications, IEEE Trans. Antennas Propag. **61**(2), 960–964 (2013)
6. S. Foo, Liquid-crystal-tunable metasurface antennas, p. 5 (2017)
7. Y. Tawk, J. Costantine, S. Hemmady, G. Balakrishnan, K. Avery, C.G. Christodoulou, Demonstration of a cognitive radio front end using an optically pumped reconfigurable antenna system (OPRAS). IEEE Trans. Antennas Propag. **60**(2), 1075–1083 (2012)
8. Y.P. Selvam, L. Elumalai, M.G.N. Alsath, M. Kanagasabai, S. Subbaraj, S. Kingsly, Novel frequency- and pattern-reconfigurable rhombic patch antenna with switchable polarization, IEEE Antennas Wirel. Propag. Lett. **16**, 1639–1642 (2017)
9. R.K. Singh, A. Basu, S.K. Koul, A novel reconfigurable microstrip patch antenna with polarization agility in two switchable frequency bands. IEEE Trans. Antennas Propag. **66**(10), 5608–5613 (2018)
10. W. Lin, H. Wong, Wideband circular-polarization reconfigurable antenna with L-shaped feeding probes. IEEE Antennas Wirel. Propag. Lett. **16**, 2114–2117 (2017)
11. Z.-N.C. Nasimuddin, X. Qing, Circularly polarized slotted/slit-microstrip patch antennas, in *Microstrip Antennas*, ed. by N. Nasimuddin, InTech (2011)
12. S.E. Jasim, M.A. Jusoh, M.H. Mazwir, S.N.S. Mahmud, Finding the best feeding point location of patch antenna using HFSS, **10**(23), 6 (2015)

Drone Shape Slot Array Microstrip Patch Antenna for X-Band



Palak Jain and Sunil Kumar Singh

1 Introduction

There are many applications like vehicles radar system, sensing monitoring system, wireless mobile communication, bioimaging system, and broadcasting have been the requirements of wideband technology. The remarkable proposal of proper transmission and reception of wideband frequencies is to design microstrip patch antenna with various shapes and sizes. Microstrip antenna operates printed patch on a ground microwave substrate of distinct relative permittivities with many interesting appearance like low radiation loss, compact in size, featherweight, and ease of fabrication [1]. An antenna is a transducer which changes over radio frequency into substituting currents. An antenna is a conductor that transmits and receives signals such as microwave, radio, or satellite signals. There are various sorts of antenna one of sort is microstrip patch antenna. The microstrip antenna is a microwave antenna and its a place of security antenna. Microstrip patch antenna is a metallic strip or patch mounted on a substrate with a dielectric material on a ground plane. A substrate is sandwiched between a patch and ground. There are different substrates that can be used for the structure of microstrip antennas, and their dielectric relentless are normally to the extent of $2.2 \leq \varepsilon r \leq 12$. Often, microstrip antennas are moreover insinuated as patch antennas [2]. Common patch antenna shapes are square, rectangle, circular, elliptical, and triangle in light of effortlessness of examination and creation. In recent years there is so much demand for high gain antenna which increased to use a high-frequency range [3].

P. Jain (✉) · S. K. Singh
EC Department, Jabalpur Engineering College, Jabalpur, Madhya Pradesh, India
e-mail: jainpranjal549@gmail.com

S. K. Singh
e-mail: sspjec2007@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_59

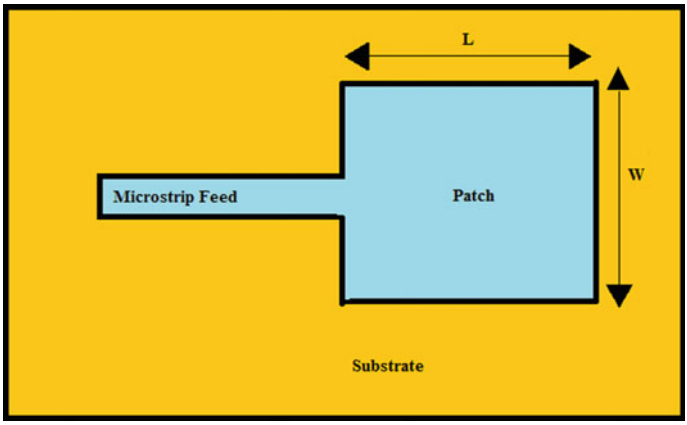


Fig. 1 Front view of microstrip patch antenna

In this paper, a rectangular shape of patch antenna is used. Figure 1 shows a microstrip patch antenna, which is feed by microstrip transmission line. The width of patch W, length land mounted on a substrate. There are many supporting strategies that can be used to reinforce microstrip antennas. There are four frameworks which are microstrip feedline, coaxial feedline, proximity-coupled feed, and opening coupled feed [4].

- Low creation cost.
- Less size and low cost [5].
- Linear and circuitous polarization.

A rectangular patch antenna is arranged with the help of different rectangular openings, poly line spaces, square openings and considered return loss, VSWR, gain, and radiation pattern.

2 Antenna Design Configuration

The geometry of the proposed antenna at frequency 9 GHz, as shown in Fig. 2. An antenna arrangement includes three segments ground plane, substrate, and patch. Rogger duroid material is used as a substrate of relative permittivity ($\varepsilon r = 2.2$), the segment of patch is 12×10 mm (L × W). An antenna depends on the Rogger duroid (tm) substrate of thickness 0.79 mm, rectangular patch has some exceptional size of rectangular and square slots. There are five openings, namely A, B, C, D, and E, and space A contains two rectangle shape spaces 1×0.5 mm. Opening B contains eight square openings with a component of size 0.5 mm. Slot C contains eight poly line openings having different points. Slot D has 0.2 mm length and 5 mm width and Slot E has 3.4 mm length and 0.1 mm width.

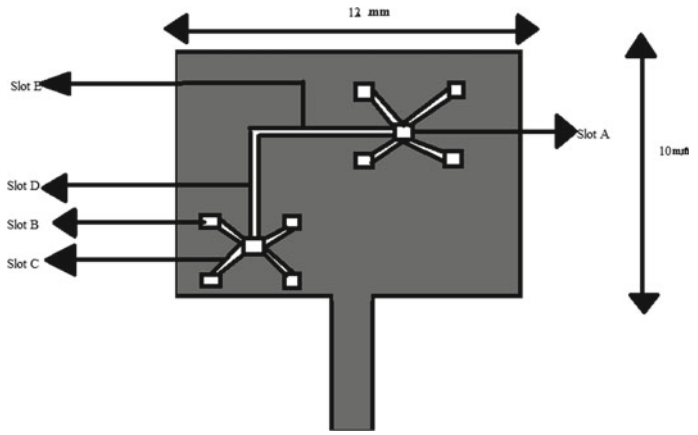


Fig. 2 Design of drone shape slot array patch antenna

A 50 Ω microstrip feed line is associated in a transmission line. In ground and patch flawless electric field limit condition is associated, in a radiation box limit condition of radiation is provided. To get radiation pattern, a radiation encase-like manner is made in terminable space, which will reduce the reflection and in addition, the digestion from the antenna. For designing a microstrip patch antenna, some mathematical calculation are there to find out width and length of the patch by considering the following Eqs. (1)–(6) [6]. The patch antenna width is calculated by the following equation:

$$W = \frac{C}{2fr}\sqrt{\frac{2}{\epsilon_{reff}}} + 1 \tag{1}$$

where

$$C = \frac{1}{\sqrt{\mu_0\ \epsilon_0}} \tag{2}$$

C free space velocity of light
ϵ_{reff} dielectric constant of substrate.

After finding out the width of patch, we determine the effective dielectric constant of patch antenna

$$\epsilon_{reff} = \frac{\epsilon_r + 1}{2} + \frac{\epsilon_r - 1}{2}\frac{1}{\sqrt{1 + 12\frac{h}{W}}} \tag{3}$$

Now calculate extension of the length ΔL, which can be stated as

$$\Delta L = \frac{0.412h(\in_{reff} + 0.3)\left(\frac{W}{h} + 0.214\right)}{(\in_{reff} - 0.258)\left(\frac{W}{h} + 0.8\right)} \tag{4}$$

The effective length of the patch is calculated as

$$L_{eff} = L + 2\Delta L \tag{5}$$

where

$$L = \frac{C}{2f_r\sqrt{\in_{reff}}} - 2\Delta L \tag{6}$$

L_{eff} effective length.

3 Results

The proposed antenna configuration is reenacted utilizing high-frequency structure simulator. The aftereffect of reproduction gives misfortune which is not exactly 10 dB. In this paper, return loss is shown in Fig. 3 Here, the return loss obtained is 31.16 dB which is near to resonating frequency 9 GHz. The voltage standing wave proportion is acquired if impedance is coordinated to a transmission line which is ≤ 2, and the proposed antenna in this paper got 0.4807, as shown in Fig. 4.

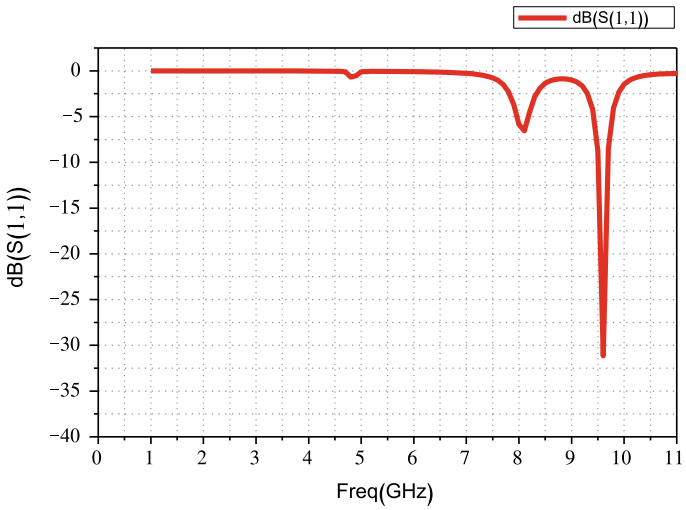


Fig. 3 Simulated return loss

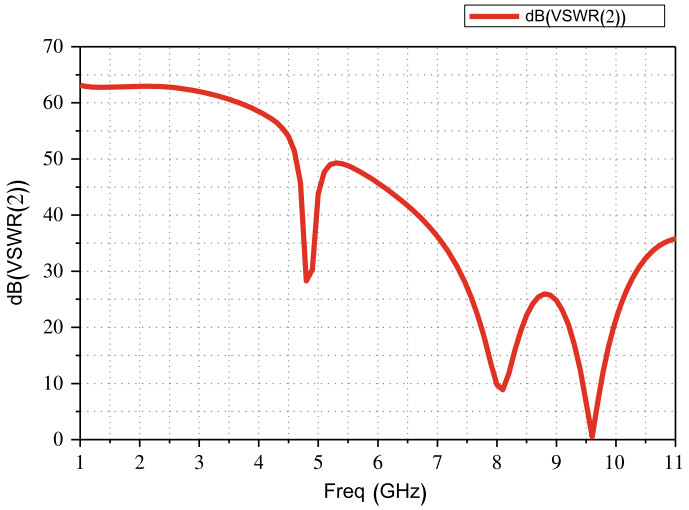


Fig. 4 Simulated VSWR

The radiation example can be characterized as a graphical representation of power flowing from transmitting antenna or inflowing to the receiving antenna as a function if direction angles from the antenna. Radiation pattern of proposed antenna is shown in Fig. 5. The gain is observed and is shown in Fig. 6. A gain is defined as the ratio of the intensity in a given direction, to the radiation intensity that would be obtained if

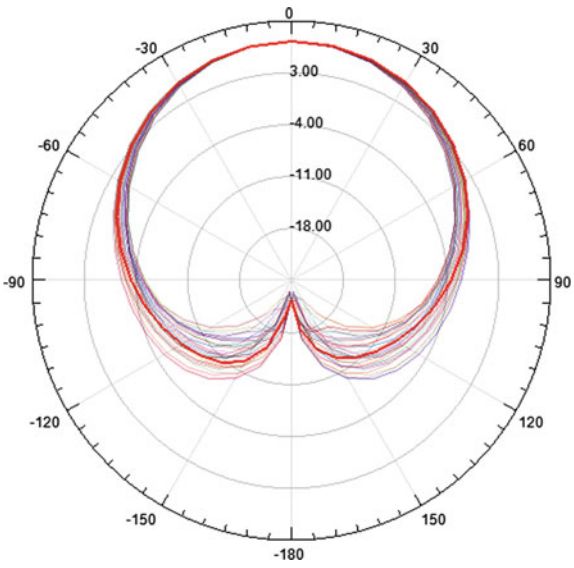


Fig. 5 Radiation pattern of an antenna

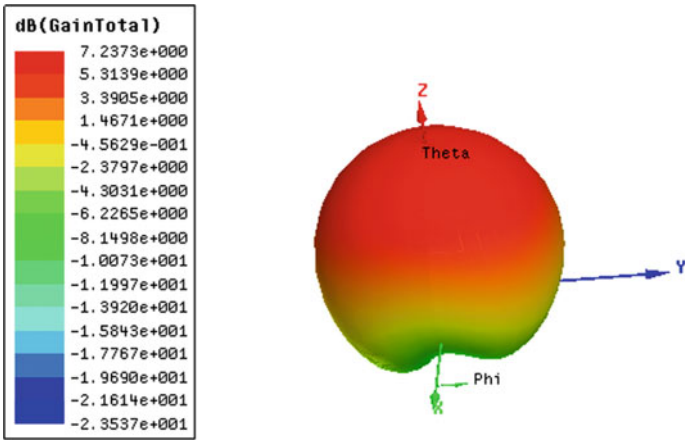


Fig. 6 Gain of an antenna

the power accepted by the antenna was radiated isotropically. The proposed antenna obtained gain is 7.2373 dB.

$$\text{Gain} = 4\pi \frac{\text{radiation intensity}}{\text{total input power}}$$

4 Conclusion

This paper proposed a drone shape slot array microstrip patch antenna for designed frequency 9 GHz and simulated frequency is near to 9 GHz, which is in range of X-band. The antenna is having a decreased size of design and it is reasonable to use in radar and satellite correspondence systems. The antenna structure for this paper uses a 0.79 mm thick substrate by using microstrip feed line, and for the excitation purpose, 50 Ω feed line is used and gives a low return loss, high gain, and minimal size.

References

1. V. Singh, B. Mishra, R. Singh, A compact and wide band microstrip patch antenna for X band applications, in *2015 Second International Conference on Advances in Computing and Communication Engineering* 978-1-4799-1734-1/15 $31.00 © (IEEE, 2015). https://doi.org/10.1109/icacce.2015.135
2. C.A. Balanis, *Antenna Theory, Analysis and Design*, 2nd ed. Wiley, Inc. (1997)

3. M. Aly Aboul-Dahab, H.H. Mahmoud Ghouz, A.Z. Ahmed Zaki, High gain compact microstrip patch antenna for X band applications. Int. J. Antennas (JANT) **2**(1) (2016)
4. V. Thaipal, P.K. Sharma, R.K. Thaipal, Design and analysis of microstrip patch antenna at 9 GHz. Int. J. Innov. Eng. Technol. (IJIET), **6**(3) Feb (2016)
5. A. Goel, M.R. Tripathy, S.A. Zaidi, Design and simulation of inverted T shaped antenna for X band applications. Int. J. Inf. Comput. Technol. **4**(6), 565–570 (2041) © International Research Publication House, ISSN 0974-2239. http://www.irphouse.com
6. M.M. Islam, M.T. Islam, M.R.I. Farque, W. Hueyshin, Design of an X band microstrip patch antenna with enhanced bandwidth, in *Proceedings of 2013 2nd International Conference on advances in Electrical Engineering (ICAEE 2013),* Dhaka, Bangladesh, 19–21 Dec 2013
7. M.F. Islam, Design and simulation of X band micro machined microstrip antenna, in *The 5th Student Conference on Research and Development-SCORed 2007*, Malaysia, 11–12 Dec 2003
8. M.K. Aghwariya, P.K. Singhal, Dual director microstrip planar yagi uda antenna for X band. Int. J. Electron., Comput. Commun. Technol. IJECCT **3**(2) (2013)
9. A.D. Farhood, M. Kamilnaji, S.A. Hasheem, A microstrip fed UWB antenna for WCDMA and X band applications. Int. J. Comput. Trends Technol. (IJCTT) **48**(1) (2017)
10. D.M. Pozar, *Microwave Engineering*, 4th ed. Wiley, Inc. (1998)

An Optimal Design of Split-Ring Resonator and Electronic Waste Composite-Based Cost-Effective Microwave Absorber for Low Observable Applications



Arpit Sahu and Ravi Panwar

1 Introduction

Microwave absorbers are the materials, which can attenuate the electromagnetic (EM) wave such that the reflection coefficient (RC) from the particular object reduced significantly [1]. Microwave absorbing material has various practical applications, which involves stealth, electromagnetic interference (EMI) shielding, radomes, and camouflage, etc. [2]. During the past few years, the demand for thin, cost-effective, lightweight, and broadband radar absorbers in military applications have been increased due to progressive development in the stealth technology. In spite of that, it is still very difficult to exploit a thin broadband absorber with enhanced absorption capability due to the existence of strong thickness–bandwidth trade-off. Microwave absorber based on Salisbury screen, Jaumann absorber, and frequency selective surfaces (FSS) have been the focus for the wide commercial and defense applications [3]. The use of Salisbury screen, which uses a continuous resistive sheet, was limited due to its narrow bandwidth and bulky structure. The Jaumann absorber has been greatly utilized to achieve wide absorption bandwidth by replacing the Salisbury screen. However, Jaumann absorber involves fabrication complexity due to multi-layering. Similarly, it is considerably easy to achieve thin wideband microwave absorber with the help of FSS [4]. The other microwave absorber includes nanocomposites, cascaded, electromagnetic bandgap (EBG), active/smart, metamaterial absorber (MMA), three-dimensional absorbers, etc. [5].

The requirement to provide cost-effective method/material, which involves wide absorption bandwidth with minimum thickness, is still under research. Therefore, to bridge this gap between cost-effectiveness and thin wideband absorber alternative

A. Sahu · R. Panwar (✉)
Discipline of Electronics and Communication Engineering, Indian Institute of Information Technology, Design and Manufacturing, Jabalpur, Jabalpur 482005, Madhya Pradesh, India
e-mail: rpanwar.iitr@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_60

approach has been involved by using E-waste composite. The production of electronic equipment has increased progressively, resulting in an enormous amount of E-waste [6, 7]. Great research interest has been seen toward the utilization of E-waste composite as microwave absorbing material. E-waste absorber is capable to achieve all these attributes together due to the presence of a precious dielectric, magnetic conducting elements, which may be helpful for impedance matching [8, 9]. Ease of availability, heterogeneous nature of composite and low cost is some of the few points, which proves the reliability of the proposed method [10]. Therefore, in this paper, an effort has been made to critically analyze the performance of various resonating geometries in the range of 8–12 GHz (X-band). This paper is structured as follows: Sect. 2 describes the absorber design and corresponding simulation steps. In Sect. 3, a critical analysis of results with a suitable discussion is presented. Finally, the conclusion is discussed in Sect. 4.

2 Design of Absorber

Figure 1 depicts the proposed unit cell of the perfect electric conductor (PEC) backed three layers absorbing structure comprising of a metamaterial, FR4 dielectric, and E-waste composite. The proposed absorber unit cell structure has been simulated using full-wave EM simulation software Computer Simulation Technology (CST) microwave studio suite. In order to have an infinite periodic structure, the unit-cell boundary condition has been applied to the cell. Furthermore, the unit cell is excited by the Floquet ports. In the proposed unit cell, SRR loaded on the multilayer substrate in which FR4 (dielectric constant of 4.4 and loss tangent of 0.025) is placed at the top of the E-waste composite layer. The frequency-dependent electromagnetic analysis of E-waste composite is provided in [10]. The ground plane in the lower layer prevents the wave to pass through the absorber.

In the extension of the proposed optimized SRR structure analysis, several other geometries have been studied in order to understand the reflection properties of the

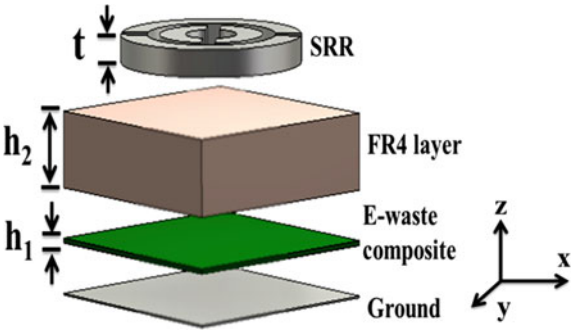


Fig. 1 Schematic of proposed SRR-MMA unit cell

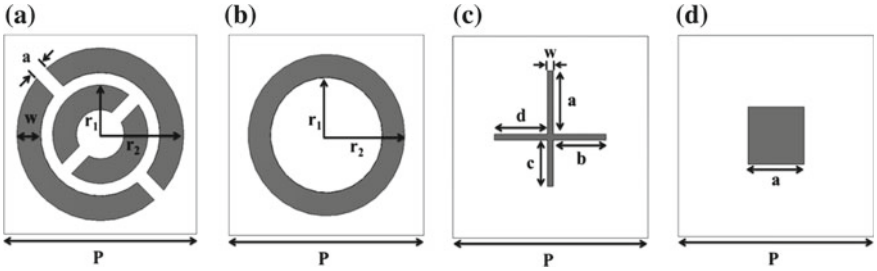


Fig. 2 Front view of the other designs **a** SRR, **b** circular, **c** plus, and **d** square patch

E-waste composite-based absorbers. Especially, deep research has been carried out to understand the effect of the proposed E-waste composite layer with nonresonant, single resonant, and multi resonating elements. In order to achieve the target, the four conducting geometries, i.e., metamaterial and FSSs including SRR, circular, plus, and square patch geometries have been considered in the present study. The front view of the proposed designs (i.e., SRR, circular, plus, and square patch) is shown in Fig. 2.

3 Results and Discussion

Initially, an effort has been made to visualize the effect of the E-waste composite layer thickness over the microwave absorption. For this purpose, the composite layer thickness is varied from 0.5 to 2.5 mm with a step size of 0.5 mm. Figure 3 shows the thickness-dependent RC spectras of SRR-MMA at X-band. It can be interpreted from the results that the RC of the absorber keeps increasing as the thickness of the E-waste layer is decreasing. It is observed that the absorber shows significantly improved RC with a 0.5 mm thick composite layer. By maintaining $h_1 = 0.5$ mm, absorber achieved 2.50 GHz of bandwidth below −10 dB threshold with the peak resonating frequency of −27.4 dB at 8.23 GHz.

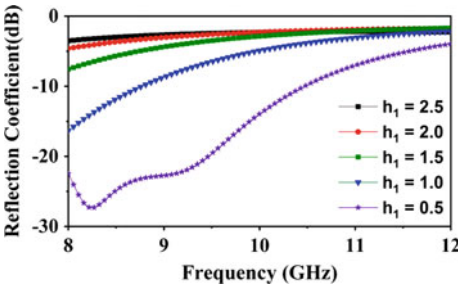


Fig. 3 RC spectras of SRR-MMA with E-waste layer thickness (h_1) changed from 2.5 to 0.5 mm

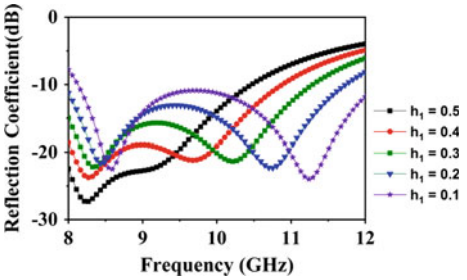


Fig. 4 RC spectras of SRR-MMA with E-waste layer thickness (h_1) changed from 0.5 to 0.1 mm

As the better absorption properties are observed for 0.5 mm thick composite, therefore, in the next step, the thickness is further reduced up to 0.1 mm with a step size of 0.1 mm. Figure 4 shows the RC of the absorber with the variation in the thickness of h_1. As the thickness of h_1 is reduced, interestingly, the bandwidth of the absorber is improved progressively by maintaining the minimum RC threshold of −10 dB. From the plot, it is clear that the bandwidth of 3.93 GHz is achieved with the maximum RC of −24.17 dB at 11.32 GHz for a 0.1 mm thick composite. Hence, it can be concluded that E-waste layer has the ability to provide wide absorption bandwidth by maintaining minimal thickness.

Furthermore, it is very crucial to understand the overall effect and performance of the absorber in the presence or by altering the positioning of the E-waste composite layer. Figure 5 demonstrates the effect of the E-waste composite layer over the microwave absorption capability of an absorber. Figure 5a shows the RC-frequency spectra of the absorber when the E-waste composite layer is incorporated and with the absence of E-waste composite layer. It is clearly proved from the plot that after incorporating the thin E-waste, composite layer bandwidth is further improved from 3.676 to 3.93 GHz with significant RC below −10 dB. Similarly, the effect of positioning of the E-waste composite layer has been studied independently as shown in Fig. 5b.

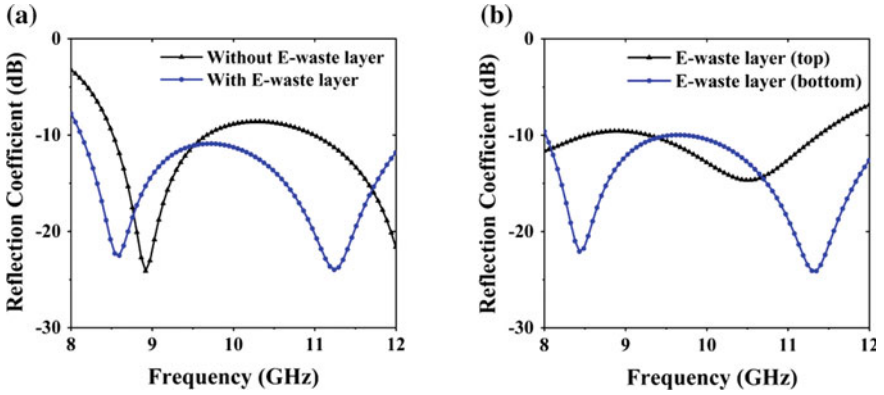


Fig. 5 Simulated response of SRR-MMA **a** with incorporating and without incorporating E-waste composite layer **b** with E-waste composite layer position either at the top or bottom layer

By positioning the E-waste layer at the top of the FR4 substrate, the absorption spectrum greatly weakened with the difference of 1.93 GHz. Likewise, the resonating RC shows the difference of −9.48 dB. Therefore, it can be concluded that an effective position to incorporate the E-waste composite layer will be at the bottom of the FR4 layer/impedance matching layer of the proposed absorbing structure.

Several other geometries like plus, circular, and square patch have also been used to examine their effect on the RC characteristics of the absorber. The primary concern is to investigate the behavior of the E-waste composite in the different types of resonating structures (non-resonating type, single resonating type, and multi-resonating type structures). Figure 6 depicts frequency-dependent RC characteristics of the absorbing structures loaded with circular, plus, square patch, and SRR geometries. The optimized dimensions of the proposed structures are shown in Table 1. A single resonating-type circular geometry shows deprived response after optimization with no available bandwidth below −10 dB in X-band. It is interesting to observe that a non-resonating structure as square patch shows very strong RC of −22.80 dB at 10.23 GHz with the fractional bandwidth of 40.98%. Similarly, a multi-resonating structure (plus geometry) is also been optimized and the fractional bandwidth of 38.71% with RC of −17.70 dB is observed at 9.78 GHz.

Furthermore, in order to achieve wide absorption bandwidth with a strong RC, the author took the liberty to use metamaterial structure in order to show the potential of E-waste composite layer. Hence, after optimization of SRR–MMA, very wide absorption bandwidth with a fractional bandwidth of 47.40% is achieved with resonating frequencies of 8.44 GHz with the RC of −22.12 dB. In addition to the above merits, the optimized MMA requires very less thickness of 1.62 mm (0.0611 λ_o). Table 2 depicts the comparison of the proposed geometries, which is made on the basis of minimal RC, resonating frequency (f_o), and % fractional bandwidth with respect to the lower cutoff frequency and the total thickness in terms of free-space wavelength. The main aim of comparison is to examine the feasibility of geometry for the real-time practical electromagnetic applications. It is noticed that the performance of SRR-loaded absorber is best in terms of its lower thickness and wide absorption bandwidth.

An extensive comparison of the proposed absorber is carried out to better understand the performance with respect to the previously reported works in Table 3. The proposed absorber achieves wide bandwidth at low frequency having a fractional

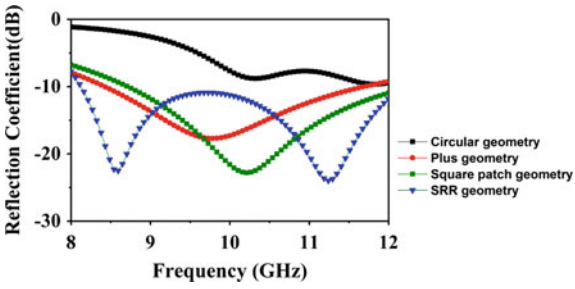


Fig. 6 Simulated RC of different optimized structures as circular, plus, square patch, and SRR geometries

Table 1 Optimal design variables of the proposed absorbers

S. no.	Structure	Design variables (in mm)
1.	Circular	$r_1 = 5$, $r_2 = 7$, $h_1 = 1$, $h_2 = 2$, $t = 1$, and $P = 17.5$
2.	Plus	$a = 1.1$, $b = 0.95$, $c = 0.8$, $d = 0.95$, $h_1 = 1$, $h_2 = 2$, $t = 2$, $w = 0.1$ and $P = 3.5$
3.	Square patch	$a = 0.857$, $h_1 = 2$, $h_2 = 1$, $t = 0.1$, and $P = 3$
4.	SRR	$r_1 = 1.8$, $r_2 = 3.15$, $w = 0.9$, $a = 0.5$, $h_1 = 0.1$, $h_2 = 1.52$, $t = 0.55$, and $P = 7.3$

Table 2 Comparison between the different unit-cell structures

S. no.	Structure type	RC (dB)	Resonating frequency (f_o) (GHz)	% Fractional bandwidth	Total thickness
1	Circular FSS	−9.63	11.82	–	3 mm (0.0871 λ_o)
2	Plus FSS	−17.70	9.78	38.71%	3 mm (0.0978 λ_o)
3	Square patch FSS	−22.80	10.23	40.98%	3 mm (0.1023 λ_o)
4	SRR-MMA	−22.12	8.44	47.40%	1.62 mm (0.0611 λ_o)

Table 3 Comparison between the performance of the proposed SRR-MMA with the other reported works

Ref.	RC (dB)	Resonating frequency (f_o) (GHz)	% Fractional bandwidth (%)	Total thickness
[11]	−24	10.8	6.85	0.8 mm (0.0288 λ_o)
[12]	−23	10	45.6	5.6 mm (0.1866 λ_o)
[13]	−18	11	51.16	3 mm (0.11 λ_o)
This work	−22.12	8.44	47.40	1.62 mm (0.0611 λ_o)

bandwidth of 47.40%. The advantage of the proposed structure compared to all the previously proposed structures is the wide frequency bandwidth with a minimum substrate thickness of the structure. The cost of the absorber with respect to the other available absorbers is also less.

4 Conclusion

A cost-effective disassembled electronic waste composite material is successfully studied for their improved microwave absorption capability. Further, the multilayered structure is utilized for analyzing various resonating geometries to improve the absorption behavior of E-waste composite. Results show that a double layer absorber with split-ring resonator (SRR) possess a strong RC of −22.12 dB at 8.44 GHz for the thickness of 1.62 mm (0.0611 λ_o). The % fractional bandwidth of 47.40% (3.93 GHz) in the range of 8–12 GHz, which is interesting. Hence, E-waste composite having low cost shows enormous potential for thin, lightweight, and broadband absorber for a range of microwave-absorbing application in a cost-effective manner.

Acknowledgements This research was supported by the Science and Engineering Research Board (SERB), Department of Science and Technology (DST), Government of India, under the Early Career Research Grant (ECR/2017/000676).

References

1. K.J. Vinoy, R.M. Jha, *Radar Absorbing Materials: From Theory to Design and Characterization, Boston* (1996)
2. R. Panwar, J.R. Lee, Performance and non-destructive evaluation methods of airborne radome and stealth structures. Meas. Sci. Technol. IOP Sci. **29**, 062001 (2018)
3. B.A. Munk, *Frequency Selective Surfaces: Theory and Design* (Wiley, New York, NY, USA, 2000)
4. R. Panwar, J.R. Lee, Progress in frequency selective surface-based smart electromagnetic structures: a critical review. Aerosp. Sci. Technol. (Elsevier) **66**, 216–234 (2017)
5. N. Engheta, R.W. Ziolkowski, *Metamaterials: Physics and Engineering Explorations* (Wiley-IEEE Press, 2006)
6. S. Sthiannopkao, M.H. Wong, Handling electronic-waste in developed and developing countries: initiatives, practices, and consequences. Sci. Total Environ, 463–464 (2013)
7. K. Huang, J. Guo, Z. Xu, Recycling of waste printed circuit boards: a review of current technologies and treatment status in China. J. Hazard. Mater. **164**, 399–408 (2009)
8. M.K. Amirhosseini, Use of dielectric inhomogeneous planar layers as optimum microwave absorbers. IET Microwaves Antennas Propag. **4**(12), 2228–2233 (2010)
9. K.Y. Park, S.E. Lee, C.G. Kim, J.H. Han, Fabrication and electromagnetic characteristics of electromagnetic wave absorbing sandwich structures. Compos. Sci. Technol. **66**(3–4), 576–584 (2006)
10. R. Panwar, V. Agarwala, D. Singh, A cost effective solution for development of broadband radar absorbing material using electronic waste. Ceram. Int. **41**, 2923–2930 (2015)
11. H. Yuan, B.O. Zhua, Y. Feng, A frequency and bandwidth tunable metamaterial absorber in x-band. J. Appl. Phys. **117**, 173103 (2015)
12. M. Yoo, S. Lim, Polarization-independent and ultrawideband metamaterial absorber using a hexagonal artificial impedance surface and a resistor-capacitor layer. IEEE Trans. Antennas Propag. **62**(5), (2014)
13. T.T. Nguyen, S. Lim, Design of metamaterial absorber using eight-resistive-arm cell for simultaneous broadband and wide-incidence-angle absorption. Sci. Rep. **7**, 14814 (2017)

Design of Rectangular Microstrip Patch Antenna and Array for Broadband Applications



Vivek Kumar, D. K. Parsediya and Anamika Gupta

1 Introduction

One of the indispensable parts for microwave communication is an antenna as it aids the transmission–reception process of the information [1]. It is a tool made, especially to radiate and receive the electromagnetic waves efficiently. It behaves as a transducer converting the current and voltage on a transmission line into an electromagnetic field in a space, which consists of electric field and magnetic field travelling at right angles to each other [2, 3]. The rapid progress in mobile communication technology today needs essential improvements and enhancements of the antennas. Generally for simplification in scrutiny and performance prediction, the patch is square, rectangular, circular, triangular and elliptical or some other common type in shape [4, 5].

The rectangular-shaped patch is most widely encountered for microstrip antenna in practical uses [6].

Even after various useful properties such as small size, lightweight, ease of installation and low cost of fabrication, the microstrip antenna suffers from the narrow bandwidth property. In this paper, a rectangular array patch antenna is designed at a resonant or central frequency of 1.8 GHz [7].

The imperative antenna characteristics like return loss, antenna radiation pattern, Voltage Standing Wave Ratio (VSWR), gain and directivity are taken into consideration in this work.

V. Kumar (✉) · D. K. Parsediya · A. Gupta
Department of Electronics, MITS Gwalior, Gwalior, India
e-mail: vks7384@gmail.com

D. K. Parsediya
e-mail: parsediyadeep@mitsGwalior.in

A. Gupta
e-mail: anamika061995@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_61

Global System for Mobile (GSM) is second-generation cellular standard that was evolved to create voice and data delivery using digital modulation. GSM is based on a wireless network technology known as Time Division Multiple Access (TDMA). The operating frequency ranges are 900, 1800 and 1900 MHz [8]. GSM is mainly used with the mobile devices and is a worldwide accepted standard for digital communication.

In Sect. 2, the details of design process for patch antenna are given. It exhibits the geometry of antenna design with various dimensional parameters. The simulation and designing of the proposed patch antenna is mainly done using software Computer Simulation Technology (CST) version 2018.

In Sect. 3, the comparison of all simulated results for single antenna designs and antenna array is discussed. The comparison is mainly based on gain, return loss, directivity and VSWR.

Section 4 concludes and discusses the potential future work related to microstrip patch antenna for GSM application.

2 Antenna Designs

Basically, the microstrip patch antenna consists of two metal plates separated by the substrate. The top plate is known as patch and bottom plate as ground plane. The antenna is excited through feed line. Generally for good transmission and reception, patch size is small [9].

Figure 1 illustrates the rectangular microstrip patch antenna with line feed and inset feed mechanism. All the design parameters for MPA using line feed and inset feed are shown in Fig. 1a and b, respectively. In both of the designs, the material used for the ground plane is Perfect Electric Conductor (PEC) with dimensions of 60×60 mm and the thickness of 0.035 mm. FR-4 with permittivity 4.3 and height 1.6 mm is used as a substrate. The dimensions for both the antenna designs are given

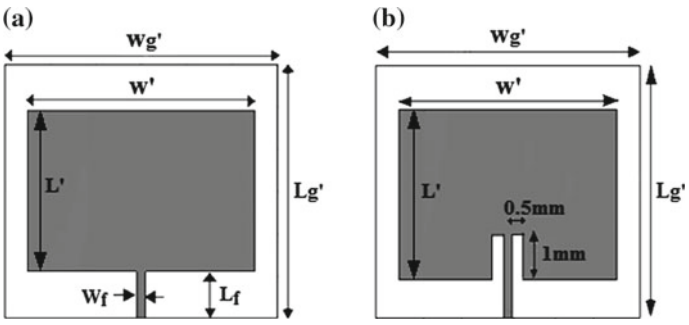


Fig. 1 **a** Rectangular microstrip patch antenna design. **b** Rectangular microstrip patch antenna design using inset feeding

Table 1 Dimension of rectangular MPA design with line and inset feed

Parameter	Dimension (in mm)
Wg'	60
Lg'	60
W'	50
L'	38
wf	2
lf	11

in Table 1. In the insect feed, the two slots are cut near the feed line. The dimensions of patch, feed and ground are exactly the same except insect feed cut.

These designs are simulated for the frequency range from 1 to 2.6 GHz with central frequency of 1.8 GHz.

The array of MPA using line feed is shown in Fig. 2. The size of array is 120 × 70 × 1.670 mm. This array encompasses two patches of the same size with a simple line feed. The patch size is 50 × 38 mm. The fabricated array of MPA using line feed is shown in Fig. 3. Table 2 shows the dimensions of array antenna.

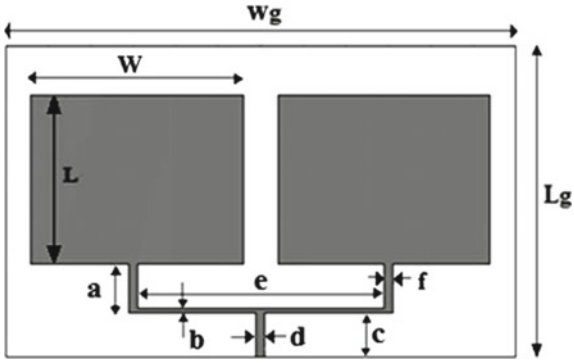


Fig. 2 Array antenna using rectangular MPA with line feed

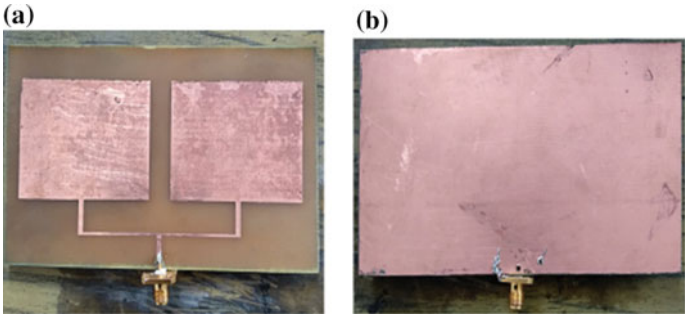


Fig. 3 Fabricated Array antenna using rectangular MPA with line feed. **a** Front view. **b** Back view

Table 2 Dimensions of array antenna using first proposed design

Parameter	Dimension (in mm)	Parameter	Dimension (in mm)
Wg	120	b	1
Lg	70	c	10
W	50	d	2.5
L	38	e	58
a	11	f	2

3 Results and Comparison

The simulation of rectangular Microstrip Patch antenna design with line feed and insect feed are carried out by simulation software CST. The return loss/S11 of microstrip patch antenna with line feed is shown in Fig. 4a which is at frequency 1.8192 GHz. Figures 4b and 5 illustrate the VSWR plot and polar radiation pattern for microstrip patch antenna with line feed, respectively.

The return loss/S11 plot for MPA with insect feeding is shown in Fig. 6a which is at frequency 1.8208 GHz. Figures 6b and 7 show the VSWR and radiation pattern of MPA design with inset feeding, respectively.

The return loss/S11 of array using rectangular MPA with line feed is shown in Fig. 8a which is at frequency 1.8 GHz. VSWR and radiation pattern of the array

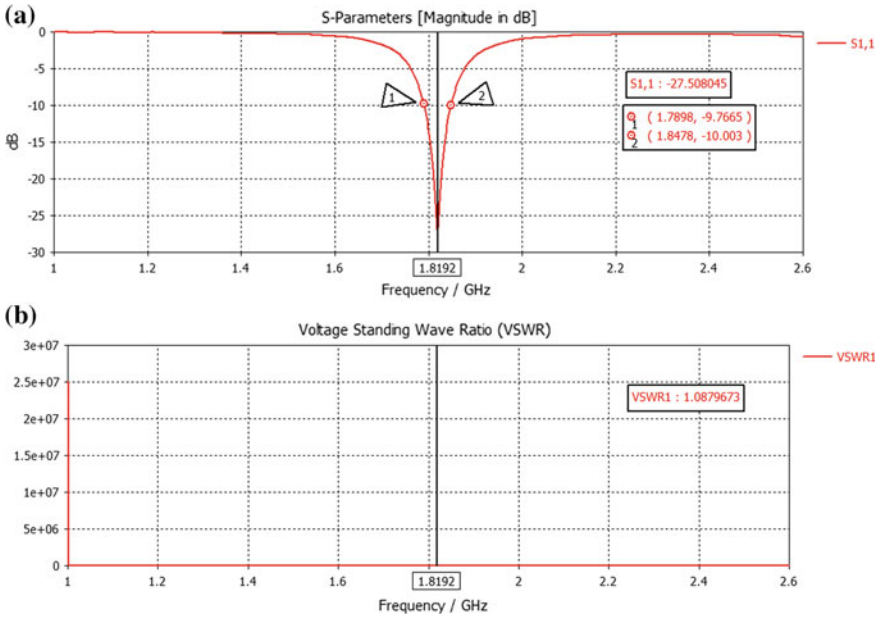


Fig. 4 **a** S11/Return loss and. **b** VSWR plot for rectangular MPA with line feed

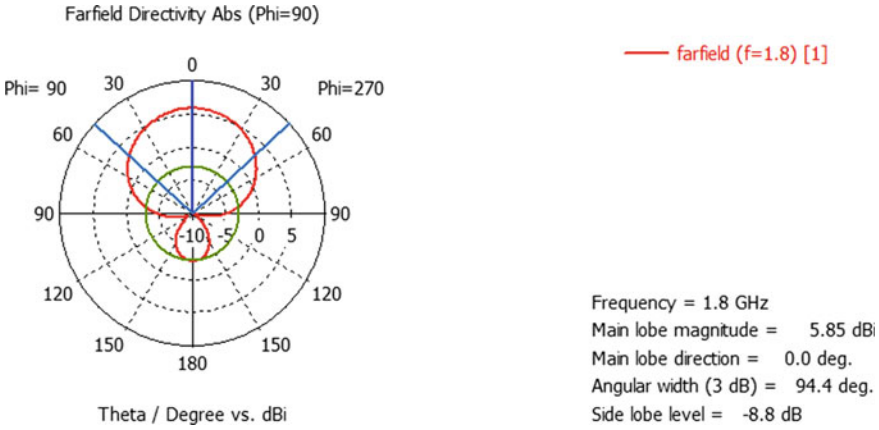


Fig. 5 Polar radiation pattern of rectangular MPA with line feed design

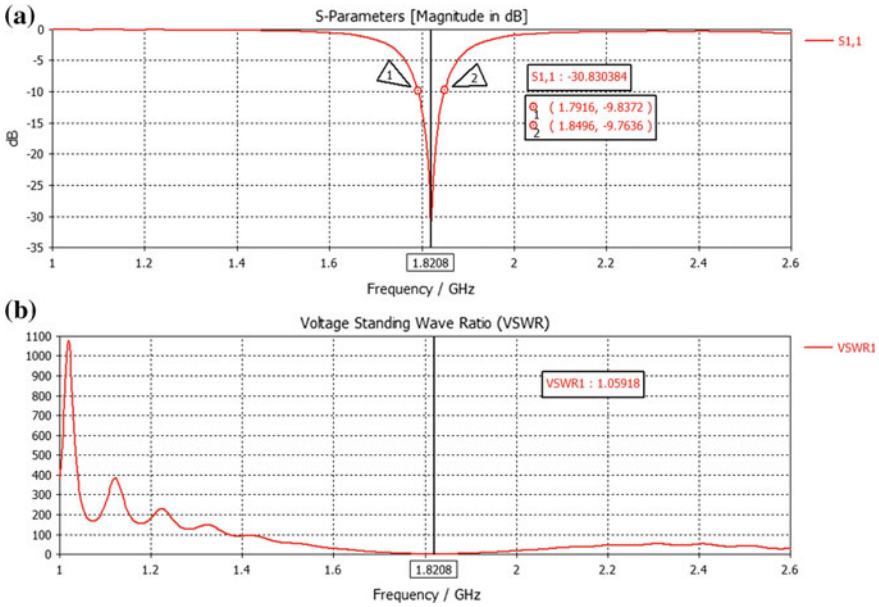


Fig. 6 **a** Return loss/S11 and **b** VSWR plot for MPA design with inset feeding

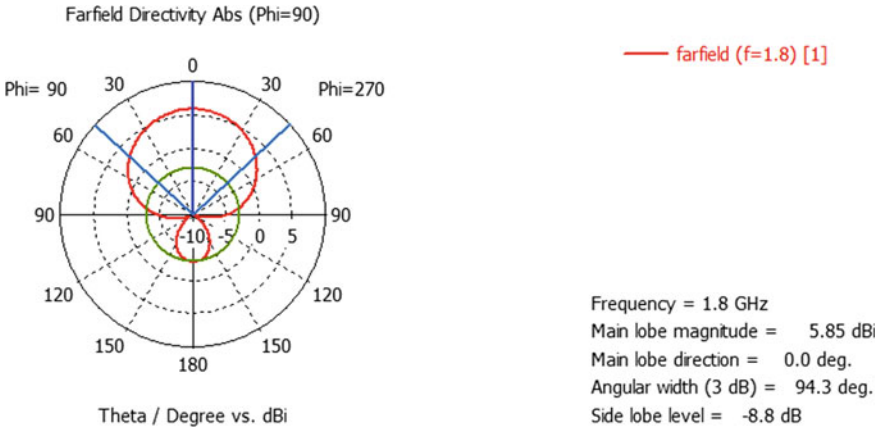


Fig. 7 Polar radiation pattern of MPA design with inset feeding

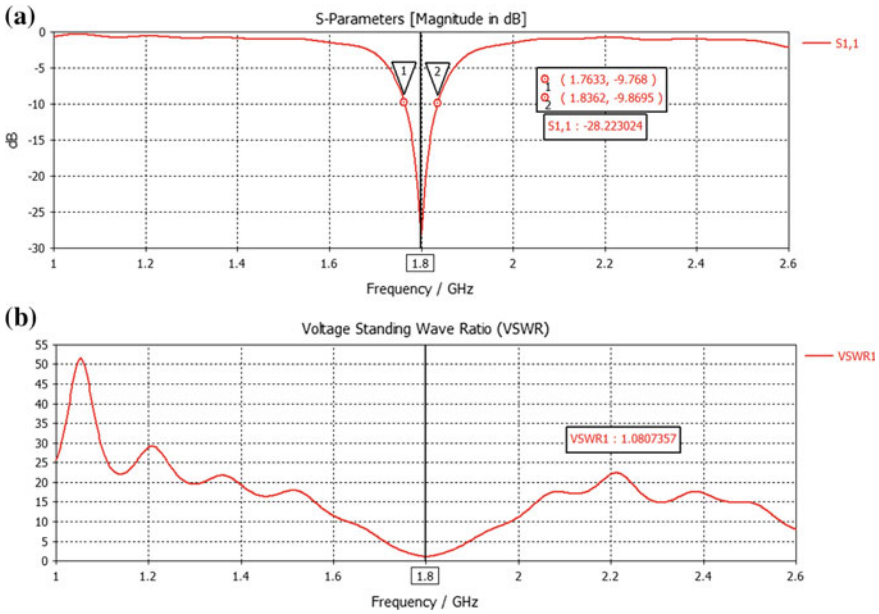


Fig. 8 **a** Return loss/S11 and **b** VSWR plot for array antenna

antenna are shown in Figs. 8b and 9, respectively. Figure 10 depicts the simulated and measured results of array antenna.

The comparison of rectangular MPA with line feed, with inset feed and antenna array is given in Table 3. It is observed that the Gain, directivity and bandwidth are improved in array antenna. The values of gain, directivity and bandwidth for array antenna are 4.26 dB, 7.73 dBi and 72.9 MHz, respectively.

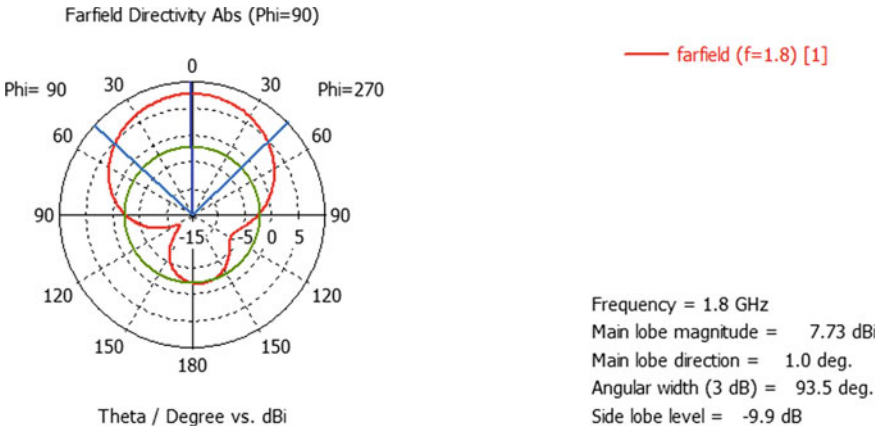


Fig. 9 Polar radiation pattern of array antenna

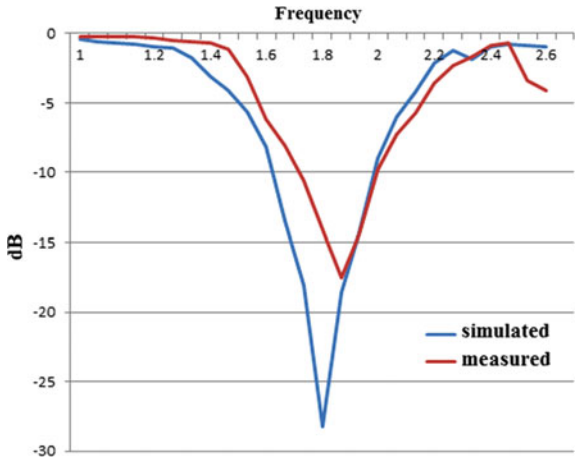


Fig. 10 Simulated and measured return loss of array antenna

Table 3 Comparison of parameters

Antenna parameter	Rectangular microstrip patch antenna with line feed	Rectangular microstrip patch antenna with inset feed	Array of rectangular microstrip patch antenna with line feed
S_{11} (dB)	−27.51	−30.83	−28.22
VSWR	1.09	1.06	1.08
Gain (dB)	2.22	2.23	4.26
Directivity (dBi)	5.85	5.85	7.73
Bandwidth (MHz)	58	58	72.9
Frequency (GHz)	1.8192	1.8208	1.8

4 Conclusion

In this paper, the rectangular microstrip patch antenna with line feed and inset feed mechanism are presented which are mainly useful in GSM application. The array of same antenna design with simple line feed is also analysed. The proposed array is simple structure, low cost and operated at single frequency band. The simulation results show the central frequency at 1.8 GHz with bandwidth 72.9 MHz ranging from 1.7633 to 1.8362 GHz. The S11/return loss of proposed array antenna is – 28.22 dB. The Gain and Directivity are 4.26 dB and 7.73 dBi, respectively. In future, the size of antenna may reduce which will operate in same frequency band.

References

1. B.S. Sandeep, S.S. Kashyap, Design and simulation of microstrip patch array antenna for wireless communications at 2.4 GHZ. Int. J. Sci. Eng. Res. **3** (2013)
2. H. Werfelli, K. Tayari, M. Chaoui, M. Lahiani, H. Ghariani, Design of rectangular microstrip patch antenna, in *2016 2nd International Conference on Advanced Technologies for Signal and Image Processing (ATSIP)*, (Monastir, 2016), pp. 798–803
3. J.W. Uduwawala, Simple design of multiband microstrip patch antenna robust to fabricate tolerances for GSM, UMTS, LTE and BLUTOOTH application. Prog. Electromagn. Res. **27** (2012)
4. A. Afridi, M.A. Tarar, Design of dual band microstrip antenna. J. Microw. Power Electromagn. Energy **7** (2013)
5. N. Alpaslan, S. Can, A.E. Yılmaz, 1.8 GHz conformal rectangular microstrip patch antenna design, in *2016 24th Signal Processing and Communication Application Conference (SIU)*, (Zonguldak, 2016), pp. 505–508
6. A.R.O. Mumin, J. Abdullah, R. Alias, R.A. Abdulhasan, J. Ali, S.H. Dahlan, Square ring microstrip patch triple band antenna for GSM/ WiMAX/ WLAN systems, in *2017 International Conference on Control, Electronics, Renewable Energy and Communications (ICCREC)*, (Yogyakarta, 2017), pp. 70–74
7. L. Peng, C. Ruan, X. Wu, Design and operation of dual/triple-band asymmetric M-shaped microstrip patch antennas. IEEE Antennas Wirel. Propag. Lett. **9**, 1069–1072 (2010)

8. C.-M. Wu, C.-N. Chiu, C.-K. Hsu, A new nonuniform meandered and fork-type grounded antenna for triple-band WLAN applications. IEEE Antennas Wirel. Propag. Lett. **5**, 346–348 (2006)
9. S. Maci, G. BiffiGentili, P. Piazzesi, C. Salvador, Dual-band slot-loaded patch antenna, in *IEE Proceedings—Microwaves, Antennas and Propagation*, vol. 142, no. 3, pp. 225–232 (1995)

Solar Rectenna to Power Wireless Sensors and Implanted Electronic Applications



B. Naresh, Vinod Kumar Singh and V. K. Sharma

1 Introduction

The energy harvested being is used to power the integrated devices into clothing like sensors, LED displays and to charge the supercapacitors. Mostly the energy can be harvested from solar, piezoelectric, microwave (Radio Frequency) energy sources to power low, ultra-low-power electronics, sensors or to recharge thin supercapacitors. The energy harvesting technology will reduce weight burden on dismounted soldiers in terms of batteries to carry, no maintenance cost, mostly, all power can regenerate from environment which extends the charging cycle. Researchers are inventing new materials which has less weight and flexible to wear. Flexible solar cell-based energy harvesting have been reported using polyamide thin-film photovoltaic smart bracelet for health care application, measuring the heartbeats of a patient through photoplethysmography (PPG) [1, 2].

Inductively, power transfer circuit is designed on plastic with flexible solar cell as a power source [3], an intelligent hardware-based charging controller circuit is presented for wireless sensor application [4]. The conversion of ambient microwave power into DC power has been done by a circuit called rectenna.

Usually, rectenna consists of a radio frequency (RF) receiving antenna and a rectifying circuit. The concept of wearable or textile antenna is having a possible way to power the incorporated electronic devices with maximum efficiency.

In hybrid energy, harvesting system energy can be harvested from multiple sources, such as a low-cost compact solar/electromagnetic harvester designed on flexible polyester [5], dual-band rectenna was implemented on flexible polyethylene

B. Naresh · V. K. Singh (✉) · V. K. Sharma
Department of Electrical, Bhagwant University, Ajmer, Rajasthan, India
e-mail: singhvinod34@gmail.com

B. Naresh
e-mail: nareshbangari@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_62

terephthalate PET substrate [6], multi-input multi-output three-source solar, vibration and thermal harvester in [7], and vibration-based micro-electrometrical energy harvester [8, 9].

In this paper, wearable dual-band RF energy harvester is proposed to power up the wearable electronic gadgets. The proposed dual-band rectenna can harvest the microwave energy at 1.85 and 2.45 GHz bands with power level starts from –20 dBm. A thin film amorphous solar cell encapsulation was used as substrate for the dual-band hollow rectangle antenna, without affecting the solar cell performance.

2 Design of Wearable Antenna

The planar microstrip antenna is a low weight, easy to design and fabrication, and wideband and multiband operation are obtained at low cost. There are two ways of design methods in RF solar energy harvesting. The first one is autonomous integration of antenna with solar cell, in which solar cell is a separate power source and require additional surface area to implant. The second one is full integration of antenna with solar cell, in which antenna is designed on the solar cell without affecting the solar efficiency.

In this paper, to power the wearable electronics full integration technique was used; which means thin-film solar cell encapsulation layers which involves ethylene-vinyl acetate (EVA) and nylon fabric are used as a substrate for the microwave antenna. The thin-film solar cell is placed between stacked protective transparent layer on top and a dielectric layers at the bottom, and the aim of the encapsulation is protecting the solar cell from environmental degradation and ensures its operation and reliability. The microstrip antenna designed on solar cell has a rectangle ring shape. Without affecting the electrical efficiency of solar cell, radiating element lays on boundaries of the solar cell, at which stainless steel foil is not present as shown in Fig. 1a.

Nowadays, the efficiency of the thin-film solar cell are almost same as the rigid silicon panels and their flexible characteristic, lightweight, easier installation are the main motives for flexible solar system [10–14]. The following layers are detected in encapsulation, airside nylon layer with thickness 300 μm, EVA middle layer thickness of 300 μm, cell side nylon with 300 μm, and cell side EVA with 300 μm; these all layers are shown in Fig. 1b. Microstrip antenna mainly required radiating element and dielectric substrate, encapsulation of the solar cell have the dielectric material and the radiating rectangular ring is designed by a copper foil tape. Geometry of a rectangular loop microstrip antenna is shown in Fig. 1c. The solar cell electrical parameters and dimension are presented in Table 1.

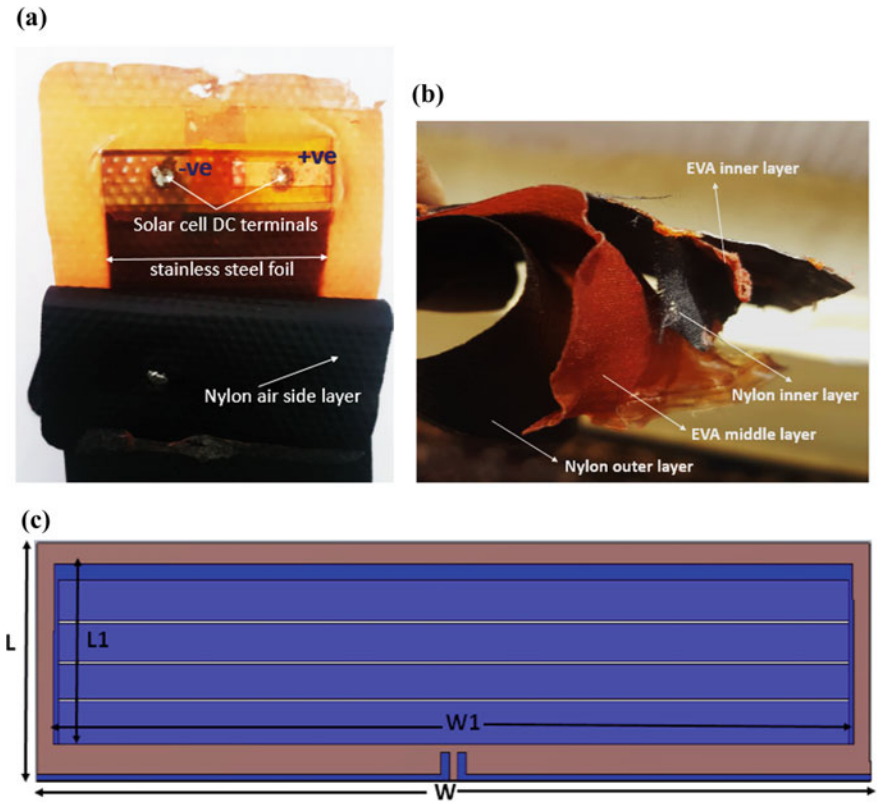


Fig. 1 **a** Solar cell back view, **b** encapsulation layers, **c** geometry of a rectangular loop microstrip antenna

Table 1 Antenna parameters

S. No.	Parameter	Value
1	W [mm]	196
2	$W1$ [mm]	194
3	L [mm]	58
4	$L1$ [mm]	50
5	Feed line [mm]	8 × 2
6	Ground [mm]	24 × 189

3 Microwave Rectification

3.1 Analysis of Voltage Doubler

There are several types of topologies in the rectification of RF to DC such as single series, single shunt, full-wave diode rectifiers [15], and voltage double rectifiers. The efficiency of the rectenna depends on frequency, RF diode type, input power level, matching between antenna and rectifier circuit; power losses if rectifier offers minimum losses electrical efficiency is high. Therefore, standard voltage doubler is the noble option over half-wave rectifier at ultra-low-power levels.

Microwave signal arrived at the input of the rectifier $v_{\text{in}} = V_{\text{AC}} \sin \omega t$ where V_{AC} is the amplitude and ω is the frequency of the input signal. Figure 2 shows the diode D1 comes into forward bias to negative half-wave and capacitor $C1$ charges, the diode D_2 rectifies the positive half-wave and capacitor $C2$ is charged. Finding the input impedance of the RF (voltage doubler) diode for design of matching circuit is the vital point since diode has nonlinear characteristics. The output DC voltage (V_{DC}) at the load is given by Eq. (1) and input impedance of the voltage doubler is calculated by using Eq. (2) [16], where I_{S} is the diode saturation current, m is the ideality factor, and V_{T} is the thermal voltage.

$$V_{\text{DC}} = 2V_{\text{AC}} \sin \omega t - 2V_{\text{F}} \tag{1}$$

$$Z_{\text{D}} = \frac{V_{\text{AC}} \sin \omega t - 0.5V_{\text{DC}}}{I_{\text{S}}\left[B_0\left(\frac{V_{\text{AC}}}{mV_{\text{T}}}\right) \exp\left(\frac{-0.5V_{\text{DC}}}{mV_{\text{T}}}\right) - 1\right]} \tag{2}$$

The voltage doubler circuit is designed in ADS software and the Simulink circuit is shown in Fig. 3a. The Large signal S-parameter simulation (LSSP) tool is used to

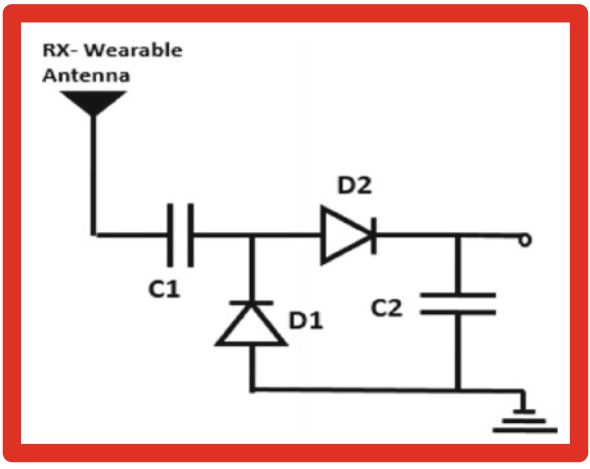


Fig. 2 Schematic of the voltage doubler circuit

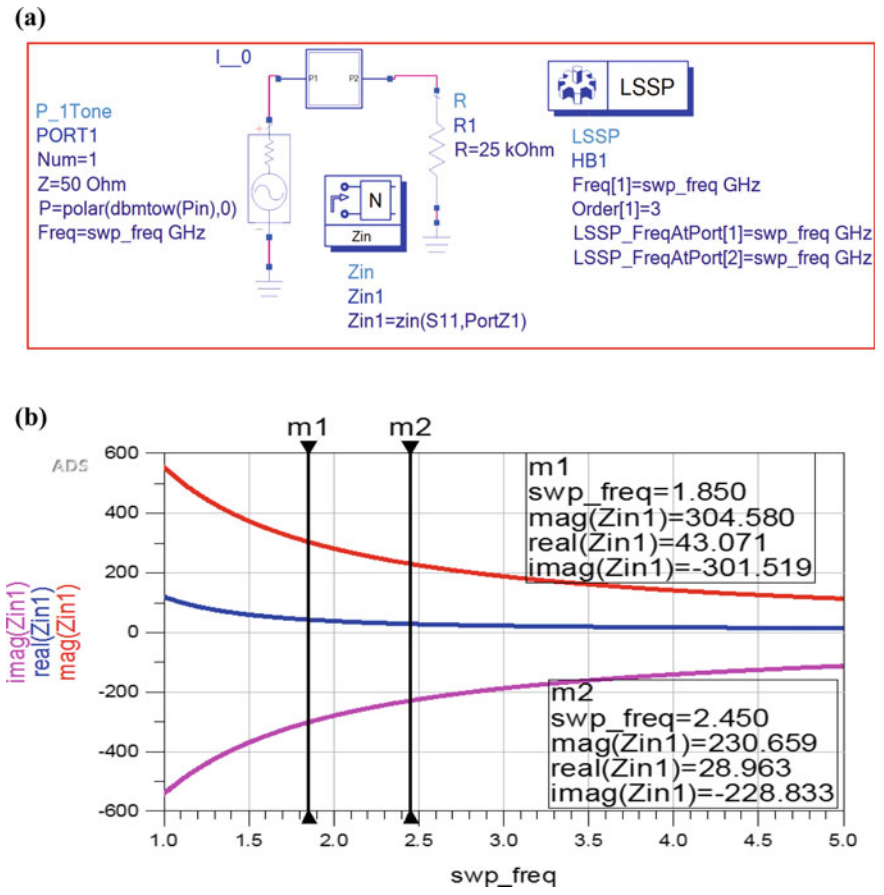


Fig. 3 **a** Voltage doubler block designed in ADS software, **b** simulated input impedance of the voltage doubler

find the input impedance of the voltage doubler that has been designed with Schottky diode SMS7630 [17]. Voltage doubler input impedance (Z_{VD}) obtained from ADS simulation is depicted in Fig. 3b, at 1.85 GHz, the impedance is $Z_{VD} = 43.071 - j301.519$ and 2.45 GHz it is $Z_{VD} = 28.963 - j228.833$ with −30 dBm as input power for a load resistance of 25 K Ω. It is clear that real part of the voltage doubler is affected by the frequency. The simulated input impedance of the voltage doubler is varied from 43.071 ohm to 28.963 ohm; the imaginary part effected more it rapidly changes from 301.519 to 228.833 Ω.

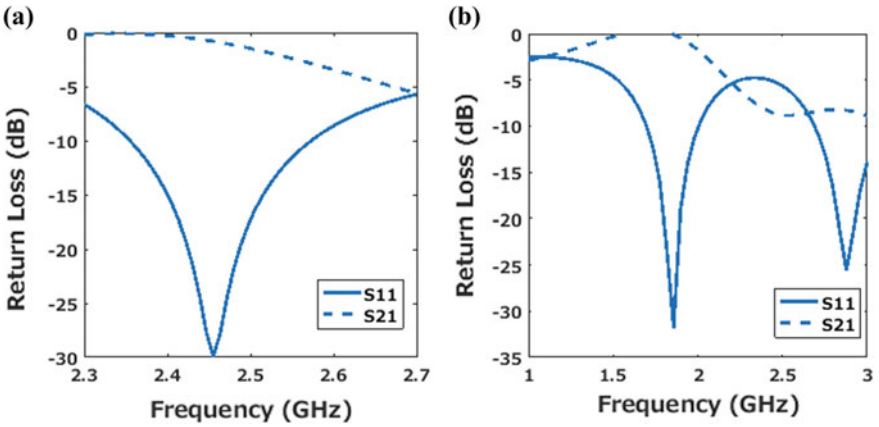


Fig. 4 Simulated S11 plot of matching circuit at 2.45 and 1.86 GHz

3.2 Dual-Band Impedance Matching

The primary goal of the matching circuit is to minimize the reflection from the voltage doubler, maximize the received microwave power, and matching is necessary for power transfer from one stage to another [18–21]. In this paper, matching circuit is implemented with microstrip transmission lines (Distributed elements). ADS Smith chat utility is used to design the filters by considering the source impedance as 50 Ω and load impedance is the diode impedance at respective frequencies. Smith chart utility is used for selecting the filter elements, later, they are modified into open and short stubs in ADS software. Two individual filters are designed for each frequency, the upper filter is for 1.85 GHz and the lower one is for 2.45 GHz. The simulated return loss plot of the matching circuits is shown in Fig. 4a, b. Both the matching circuit has the reflection coefficient (S11) −30 dB at resonance frequencies, the S21 is near the 0 dB which means the signal noise generated by the nearby components will be rejected hence losses will be reduced.

4 Experimental Setup and Results

The antenna measurement is done with the help of a key sight microwave analyzer to find the return loss (S11) in dB of the fabricated antenna. The testing setup of the flexible antenna is shown in Fig. 5a. The return loss measured in the experiment is shown in Fig. 5b. The energy harvesting frequencies are 2.45 and 1.86 GHz and bandwidths are 1.70–1.90 GHz (200 MHz) and 2.35–2.48 GHz (130 MHz). The primary band harvesting frequency is 1.85 GHz with return loss magnitude of − 27 dB. The second band frequency is 2.45 GHz with the return loss magnitude of − 28 dB.

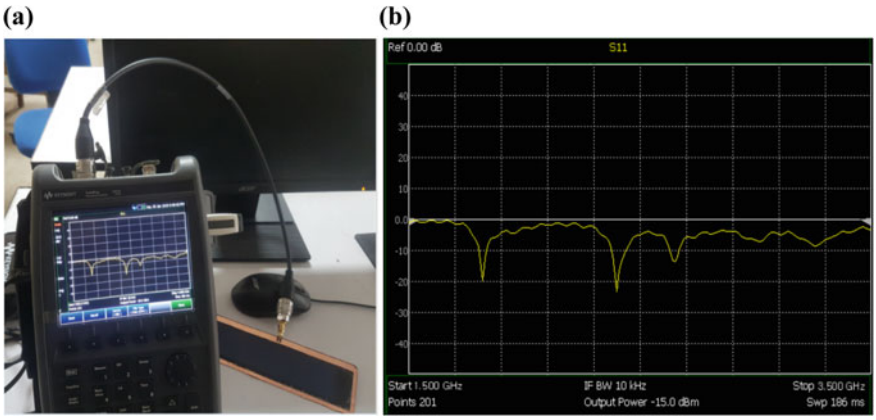


Fig. 5 **a** Return loss testing of designed flexible antenna, **b** screenshot of measured S11

The fabricated rectenna hardware is shown in Fig. 6. Voltage doubler long with matching circuit was built on PCB with 1.6 mm thickness and its conversion efficiency performance at resonance frequencies was studied when the input RF power varied. Two zero-bias Schottky diode (SMS7630) was chosen and the bypass capacitors (C1, C2) were chosen to be 100 pF and storage capacitors (C3) 100 μF, both from Panasonic Electronics. The conversion efficiency, η, of the voltage doubler for RF power levels varying from −30 to 5 dBm is shown in Fig. 7a for different load values and maximum efficiency was obtained for 7.5 K Ω is observed. The voltage doubler conversion efficiency at 1.85 GHz and 2.45 GHz are 58% and 43%, respectively, for a load of 7.5 K Ω as shown in Fig. 7b.

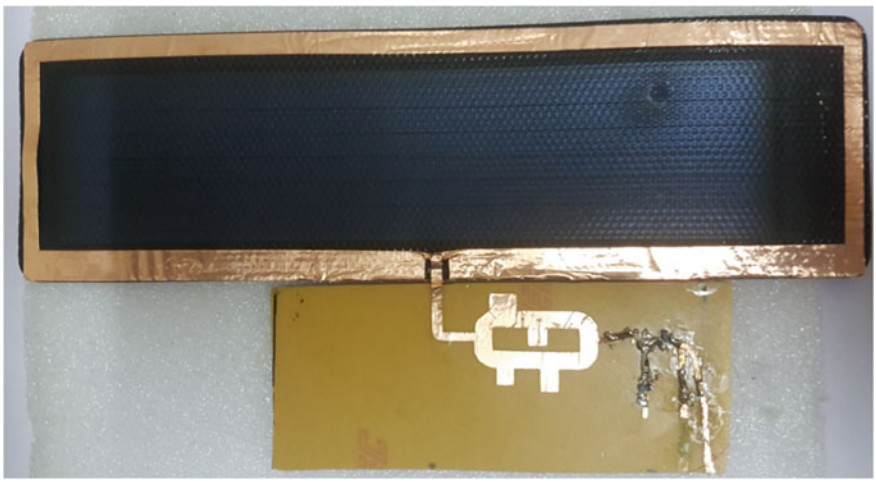


Fig. 6 Wearable rectenna on thin-film solar cell

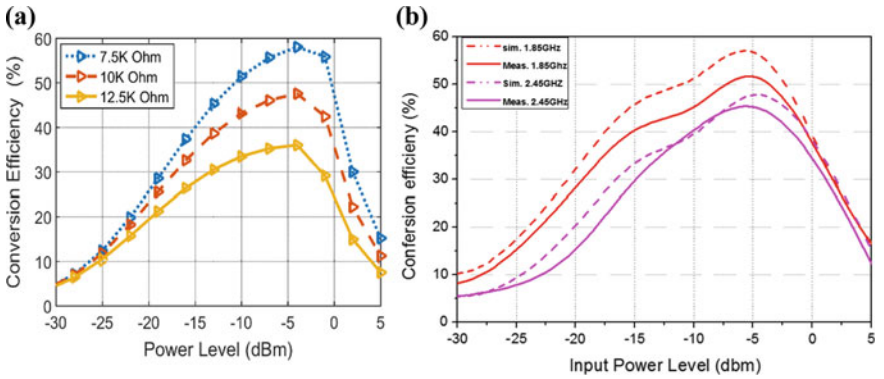


Fig. 7 **a** Simulated conversion efficiency for different load resistances, **b** measured and simulated rectifier conversion efficiency at dual frequencies for 7.5 K Ω

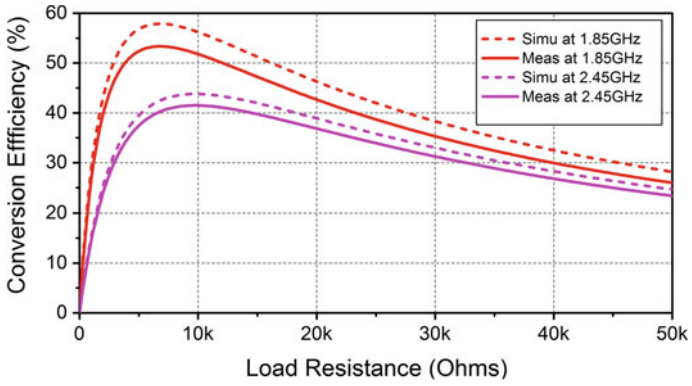


Fig. 8 Conversion efficiency with respect to load resistance in K Ohms

The maximum power curves with, respectively, load resistance at center frequencies are shown in Fig. 8 for an input power level of −5 dBm. The graph explains that maximum efficiency measured at first frequency is 42.35% and at the second frequency is 56.48% for a load resistance of 7.5 K Ω.

5 Conclusion

In this work, the design of flexible antenna on a thin-film solar cell (as a substrate) without affecting the solar efficiency is reported. The rectangular loop antenna is fabricated on solar cell encapsulation materials are EVA, nylon, and top layer is ETFE as substrate. Thus, it is a multilayer microstrip antenna with the thickness of 0.8 mm and antenna has a dual-band operation. The operation frequencies are 1.85

and 2.45 GHz with measured S11 below 20 dB and above 50 MHz band width at resonance frequencies. Designed rectenna is having good matching characteristic at dual band with maximum efficiency of 58% at −5 dBm power level. The output voltage measured at is maximum power level is 420 mV.

References

1. P. Jokic, M. Magno, Powering smart wearable systems with flexible solar energy harvesting, in *2017 IEEE International Symposium on Circuits and Systems (ISCAS),* Baltimore, MD (2017)
2. T. Wu, M.S. Arefin, J. Redouté, M.R. Yuce, Flexible wearable sensor nodes with solar energy harvesting, in *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Seogwipo (2017), pp. 3273–3276
3. Y. Hu, Flexible solar-energy harvesting system on plastic with thin-film LC oscillators operating above ft for inductively-coupled power delivery, in *Proceedings of the IEEE 2012 Custom Integrated Circuits Conference,* San Jose, CA (2012), pp. 1–4
4. Y. Li, R. AU-Shi, An intelligent solar energy-harvesting system for wireless sensor networks. EURASIP J. Wirel. Commun. Netw. DA **1**, 179 (2015)
5. G. Andia Vera, A. Georgiadis, A. Collado, S. Via, Design of a 2.45 GHz rectenna for electromagnetic (EM) energy scavenging, in *IEEE Radio and Wireless Symposium, RWW 2010—Paper Digest* (2010), pp. 61–64
6. A. Collado, A. Georgiadis, Conformal hybrid solar and electromagnetic (EM) energy harvesting rectenna. IEEE Trans. Circuits Syst. I Regul. Pap. **60**(8), 2225–2234 (2013)
7. S. Bandyopadhyayand, A.P. Chandrakasan, Platform architecture for solar, thermal, and vibration energy combining with MPPT and single inductor. IEEE J. Solid-State Circuits **47**(9), 2199–2215 (2012)
8. X. Chen, Y. Cao, J. Xie, Piezoelectric and electromagnetic hybrid energy harvester for powering wireless sensor nodes in smart grid. J. Mech. Sci. Technol. **29**, 4313
9. M. Salim, H.S.S. Aljibori, D. Salim, M.H.M. Khir, A.S. Kherbeet, A review of vibration-based MEMS hybrid energy harvesters. J. Mech. Sci. Technol. **29** (2015)
10. K.L. Chopra, P.D. Paulson, V. Dutta, Thin-film solar cells: an overview. Prog. Photovoltaics **12**, 69–92 (2004). https://doi.org/10.1002/pip.541
11. D.M. Chapin, C.S. Fuller, G.L. Pearson, A new silicon p–n junction photo cell for converting solar radiation into electrical power. J. Appl. Phys. **25**, 676 (1954)
12. K.L. Chopra, S.R. Das, *Thin-Film Solar Cells* (Plenum, New York, 1983)
13. R.W. Birkmire, E. Eser, Polycrystalline thin-film solar cells: present status and future potential. Annu. Rev. Mater. Sci. **27**, 625–653 (1997)
14. A.L. Fahrenbruch, R.H. Bube, *Fundamentals of Solar Cells* (Academic Press, New York, 1983)
15. V. Marian, B. Allard, C. Vollaire, J. Verdier, Strategy for microwave energy harvesting from ambient field or a feeding source. IEEE Trans. Power Electron. **27**(11), 4481–4491 (2012)
16. C. Song, Y. Huang, P. Carter, J. Zhou, S. Yuan, Q. Xu, M. Kod, A novel six-band dual CP rectenna using improved impedance matching technique for ambient RF energy harvesting. IEEE Trans. Antennas Propag. **64** (2016). https://doi.org/10.1109/tap.2016.2565697
17. *Surface Mount Mixer and Detector Schottky Diodes*, *Data Sheet* (Skyworks Solutions, Inc., Woburn, MA, USA, 2013)
18. D. Pavone, A. Buonanno, M.D'. Urso, F. Corte, Design considerations for radio frequency energy harvesting devices. Prog. Electromagn. Res. B, **31**, 19–35 (2012)
19. G. Chaudhary, P. Kim, Y. Jeong, J.H. Yoon, Design of high efficiency RF-DC conversion circuit using novel termination networks for RF energy harvesting system. Microwave Opt. Technol. Lett. **54**(10), 2330–2335 (2012)

20. Zohaib Hameed, Kambiz Moez, Design of impedance matching circuits for RF energy harvesting systems. Microelectron. J. **62**, 49–56 (2017)
21. X. Wang, L. Zhang, Y. Xu, Y.-F. Bai, C. Liu, X.-W. Shi, A tri-band impedance transformer using stubbed coupling line. Prog. Electromagn. Res. **141**, 33–45 (2013)

Design and Performance Enhancement of Wearable Textile Antenna



Monika Parmar, Mohit Gaharwar and Jaydeep Dewangan

1 Introduction

Another period for the attire industry is the reconciliation of gadgets into textile messengers. The article of clothing of things to come won't just ensure the human body against the limits of nature, yet, additionally give data about the wearer's condition of well-being and condition [1] is shown in Fig. 1. This work not just goes for growing such wearable textile systems mostly for expert firemen and crisis calamities workforce, yet, in addition to regular citizen casualties of common and different catastrophes. The fireman's inward and external article of clothing is being outfitted with an assortment of sensors [1] whose information are handled in a wearable electronic box and transmitted to a base station. This off-body correspondence requires the improvement of appropriate antenna that consolidates adaptability with heartiness and unwavering quality. An assortment of antennas for body-driven correspondence has been presented as of late. Due to the shape, low profile, and conservativeness, textile antennas have turned out to be most appropriate for coordination into articles of clothing [2, 3].

The fast improvement of remote power transmission has expanded the interest for material minimized reception apparatus with high addition and broadband attributes. Material reception apparatuses are favorable in view of their successful cost and straightforward acknowledgment process. There are a great deal of strategies to build

M. Parmar (✉) · M. Gaharwar
Department of ECE, Sachdeva Institute of Technology, Farah, Mathura, India
e-mail: mohitgaharwarcoe@gmail.com

M. Gaharwar
e-mail: parmarmonika249@gmail.com

J. Dewangan
Department of EEE, Bhilai Institute of Technology, Raipur, India
e-mail: jaideepdewangan11@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_63

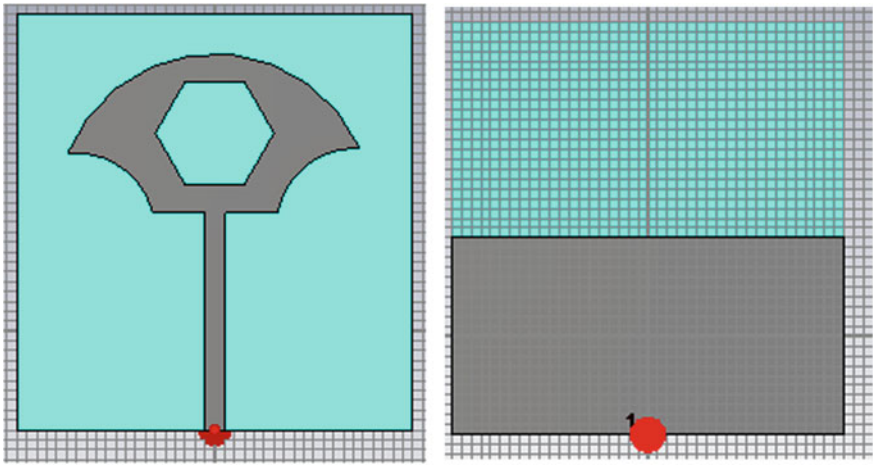


Fig. 1 Configuration of the antenna

the data transfer capacity of flexible antenna, by expanding the tallness of substrate, utilization of low dielectric substrate, utilizing various encouraging procedures and by taking incomplete ground [4–7]. The proposed reception apparatus configuration utilized jeans whose ε_r is around 1.7. Due to low dielectric constant, the losses and attenuation are very less and it also enhances the bandwidth of textile antenna. Decision of substrate selection of the most appropriate substrate for an antenna involves prime significance. This is on the grounds that numerous restrictions of the microstrip wearable antenna, for example, exceptional reflection coefficient, low gain, and low efficiency can be overwhelmed by choosing a proper substrate for creation of the textile antenna. The substrate properties, for example, its dielectric constant and loss tangent pronouncedly affect the reception antenna attributes. A portion of the basic properties that are to be dealt with while choosing a dielectric are homogeneity and dampness retention. Consequently, we pick an uncommon sort of fabric (jeans) with a thickness of 1 mm substrate. Low permittivity of the reception apparatus substrate permits the structure of textile antenna with a vast gain and a high effectiveness [8–14].

2 Process of Antenna Design

The textile antenna substrate is made using a 1 mm thick jeans textile. Jeans is chosen in this work based on its cost-effectiveness, ease of accessibility, and ease of fabrication in comparison to other textile materials. The copper tape is used to form the ground plane and radiating elements of the anticipated antenna is shown in Fig. 1. The final resulting structure is fabricated fully using textiles, except for the SMA connector which is integrated for evaluation purpose. The antenna design and

Table 1 Design parameters

S. No.	Design parameters	Values
1	Substrate dimension (mm)	40 × 52
2	Ground dimension	40 × 10
3	Patch radius (mm)	16
4	Hexagonal slots radius (mm)	6
5	Dimension of microstrip line (mm)	22 × 2

simulations were carried out in CST microwave. The length and width of ground plane are 10 mm and 40 mm. The general permittivity of jeans substrate is 1.7 with tangent loss of 0.025. The radius of patch is 16 mm and hexagonal slots with measurement 6 mm and the width of microstrip feed line is 2 mm which have been given in Table 1.

3 Simulated Results

The recreation results were acquired by CST software and antenna highlights were studied. Figure 2 indicates return misfortune versus recurrence chart having the triple band at full frequencies 3.8499, 7.8399, and 10.09 GHz individually. The 3D characteristics of proposed antenna are shown in Fig. 3, which shows directivity of 3.528, 3.989, and 4.782dBi. Figure 4 shows the polar plot of the anticipated antenna at resonant frequencies 3.8499 GHz, 7.8399 GHz, and 10.09 GHz, respectively.

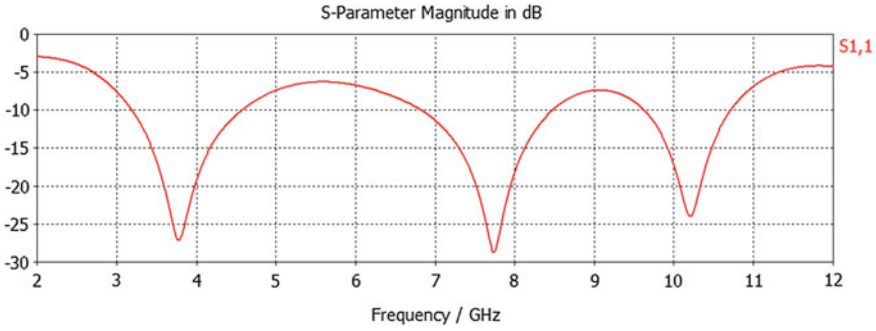


Fig. 2 Return loss versus frequency simulated graph

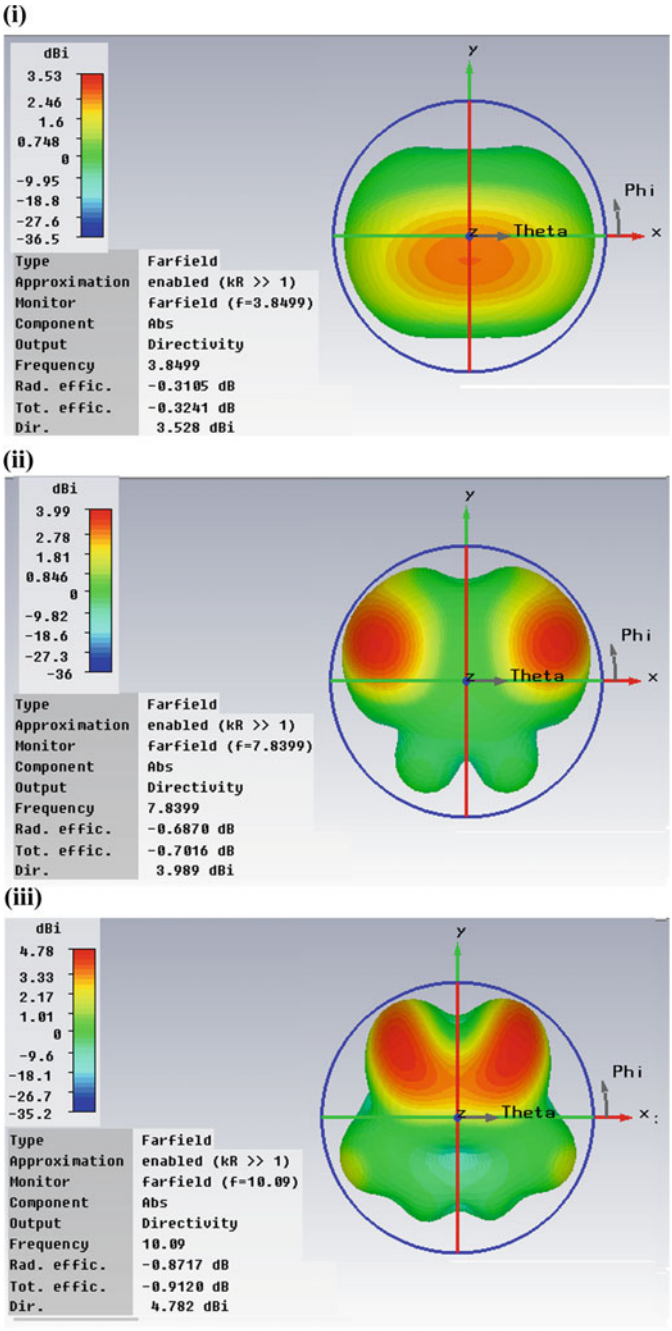


Fig. 3 3D radiation characteristics at (i) 3.8499 GHz, (ii) 7.8399 GHz, and (iii) 10.09 GHz

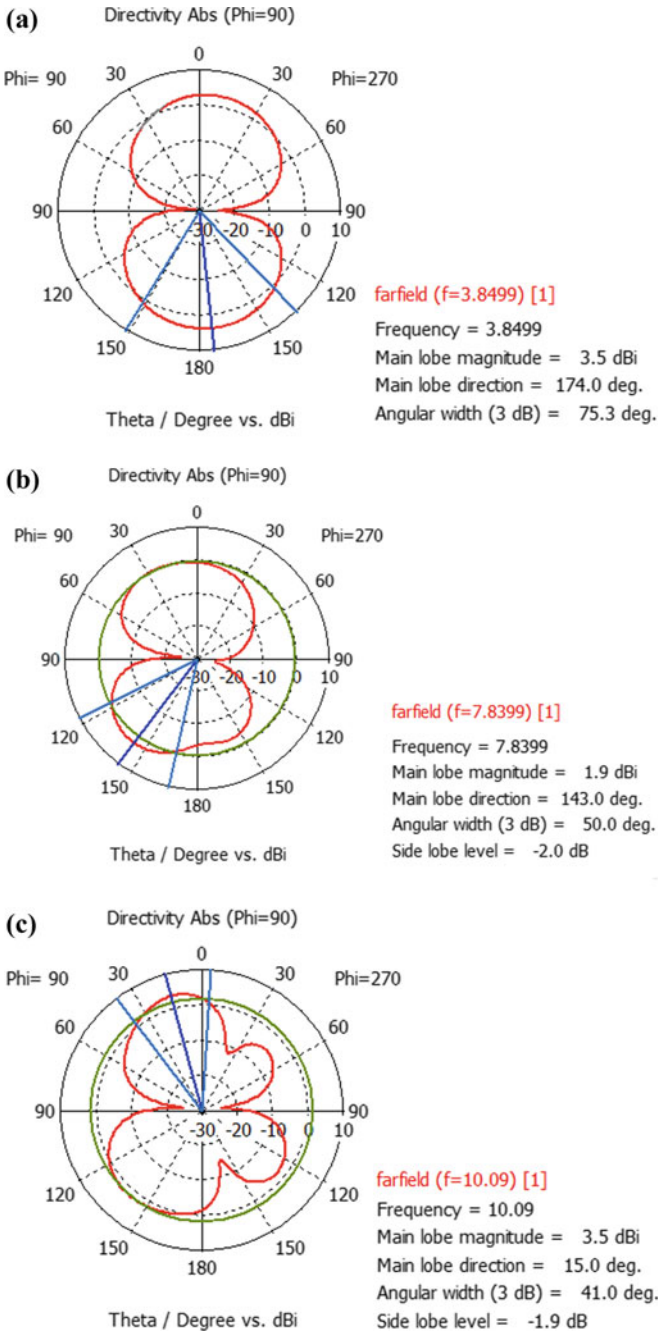


Fig. 4 Polar plot at resonant frequencies 3.8499, 7.8399, and 10.09 GHz

4 Conclusions

The material reception apparatus utilizing textile as a substrate with triple-band attributes is displayed and examined. An expected reception apparatus gives three diverse wide groups with an impedance data transmission of 35.56, 21.58, and 10.71% with the addition of 3.528, 3.989, and 4.782 dBi separately. The radio wire is best appropriate for various remote correspondence framework such WiMAX application.

References

1. S. Sankaralingam, B. Gupta, Development of textile antennas for body wearable applications and investigations on their performance under bent conditions. Prog. Electromagn. Res. B **22**, 53–71 (2010)
2. M.A.R. Osman, M.K.A. Rahim, M. Azfar, N.A. Samsuri, F. Zubir, K. Kamardin, Design, implementation and performance of ultra-wideband textile antenna. Prog. Electromagn. Res. B **27**, 307–325 (2011)
3. R. Singh, V.K. Singh, N.K. Singh, Wide band and miniaturized partial ground plane microstrip antenna for X & Ku band applications. Int. J. Control Theory Appl. **10**(8), 477–486 (2017)
4. V.K. Singh, Z. Ali, S. Ayub, A.K. Singh, A wide band compact microstrip antenna for GPS/DCS/PCS/WLAN applications, intelligent computing, networking, and informatics, vol. 243 (Springer, Berlin, 2014), pp. 1107–1113. (ISBN: 978-81-322-1664-3)
5. R. Srivastava, V.K. Singh, S. Ayub, Comparative analysis and bandwidth enhancement with direct coupled C slotted microstrip antenna for dual wide band applications, in *Frontiers of Intelligent Computing: Theory and Applications*, vol. 328 (Springer, Berlin, 2015), pp. 449–455. (ISBN: 978-3-319-12011-9)
6. S. Srivastava, V.K. Singh, Z. Ali, A.K. Singh, Duo triangle shaped microstrip patch antenna analysis for WiMAX lower band application, in *International Conference on computational Intelligence: Modeling Techniques and Applications (CIMTA-2013), Procedia Technology*, vol. 10, (Elsevier, 2013), pp. 554–563
7. N.M. Din, C.K. Chakrabarty, A. Bin Ismail, K.K.A. Devi, W.Y. Chen, Design of RF energy harvesting system for energizing low power devices. Prog. Electromagn. Res. **32**, 49–69 (2012)
8. R. Srivastava, S. Ayub, V.K. Singh, Dual band rectangular and circular slot loaded microstrip antenna for WLAN/GPS/WiMax applications, in *IEEE Conference on Communication Systems and Network Technologies (CSNT-2014)*, pp. 45–48 (2014)
9. J. Loni, V.K. Singh, Development of bandwidth enhanced microstrip patch antenna for UMTS application. J. Microw. Eng. Technol. **2**(1), pp. 01–07 (2015); V.K. Singh, Z. Ali, S. Ayub, A.K. Singh, Bandwidth optimization of compact microstrip antenna for PCS/DCS/Bluetooth application. Cent. Eur. J. Eng. Springer, **4**(3), pp. 281–286 (2014)
10. N. Singh, A.K. Singh, V.K. Singh, Design and performance of wearable ultra wide band textile antenna for medical applications. Microw. Opt. Technol. Lett. **57**(7), 1553–1557 (2015)
11. N.K. Singh, N. Sharma, Z. Ali, V.K. Singh, A.K. Bhoi, Inset fed circular microstrip antenna with defected ground, in *Advances in Electronics, Communication and Computing*, ed. by A. Kalam, S. Das, K. Sharma. Lecture Notes in Electrical Engineering, vol. 443 (Springer, Berlin, 2018), pp. 605–611
12. R. Kushwaha, V.K. Singh, N.K. Singh, A. Saxena, D. Sharma, A compact pentagonal textile microstrip antenna for wide band application, in *Advances in Electronics, Communication and Computing*, ed. by A. Kalam, S. Das, K. Sharma. Lecture Notes in Electrical Engineering, vol. 443 (Springer, Berlin, 2018), pp. 793–799

13. V.K. Singh, S. Khan, A. Verma, A. Vishnoi, R. Tiwari, Textile antenna for wireless health monitoring system applications. Int. J. Control Theory Appl. **10**(8), 487–493 (2017)
14. Z.H. Jiang, D.E. Brocker, P.E. Sieber, D.H. Werner, A compact, low-profile meta surface enabled antenna for wearable medical body-area network devices. IEEE Trans. Antennas Propag. **62**(8), 4020–4030 (2014)

Wideband E-Shaped Planar Antenna for Cellular, GPS, and Wireless Applications



Shobit Agarwal, Umair Rafique and Vasu Jain

1 Introduction

Nowadays, most of the communication devices preferred printed antennas for different applications, because they are small in size, lightweight and able to provide omnidirectional radiation characteristics. Furthermore, with the introduction of different communication standards, there is a need to design a compact wideband or multiband antennas, which can operate at multiple frequency bands and can easily be integrated into one unit.

To achieve these requirements, several antennas with wideband and multiband characteristics have been reported, such as planar inverted F-antennas (PIFAs) [1, 2] and inverted F- antennas (IFAs) [3, 4]. Both, PIFA and IFA are compact in size and provide a multiband and wideband response, which is suitable for many communication devices. But these kinds of antennas exhibit additional height in the system, which increases design complexity. If the height between the radiating element and ground plane is reduced, then the bandwidth of PIFAs and IFAs will decrease to some extent. On the other hand, planar antennas are one of the preferred candidates for compact communication devices, because they can be manufactured on the same circuit board of any communication-related device.

In [5], a multiband planar monopole antenna was designed and presented for mobile wireless communication applications. A crescent-shaped radiator with defected ground structure (DGS) was used to achieve an impedance bandwidth

S. Agarwal (✉)
Department of Electrical Engineering, Indian Institute of Technology Ropar, Rupnagar, India
e-mail: 2018eez0015@iitrpr.ac.in; shobitagarwal@ieee.org

U. Rafique
Department of Electrical Engineering, Capital University of Science and Technology, Islamabad, Pakistan
e-mail: umair-rafique@hotmail.com

V. Jain
Allengers Medical Systems Limited, Chandigarh, India
e-mail: vasujn92@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_64

of 58.3% over the frequency range of 1.7–3.1 GHz. Ahmed et al. [6] presented a circular-shaped printed antenna for cellular phone applications. Through this structure, the authors achieved the bandwidth of 960 MHz from 1790 to 2750 MHz. But for practical applications, their design was not suitable due to poor return loss. In [7], a butterfly-shaped wideband patch antenna was realized for wireless communication. The presented design was able to provide a bandwidth of 40%, but due to its three-dimensional (3D) structure, it cannot be used in compact devices. A compact co-planar waveguide (CPW)-fed wideband planar antenna was presented in Ref. [8] for wireless applications. An axe-like structure with modified CPW feed line was utilized to achieve a frequency response from 2.7 to 5.7 GHz.

Bharti et al. [9] presented a small-size printed monopole antenna design for slim mobile handsets. The presented antenna design was able to cover operating frequencies of GSM, GPS, Wi-Fi/WLAN, and WiMAX bands. But their presented design offered a negative gain in the bands of interest. In [10], the authors presented a triple wideband patch antenna design with DGS for WLAN/WiMAX applications. They utilized two inverted L-shaped stubs and a rectangular stub to achieve frequency response in three different ranges, such as 2.085–2.72 GHz, 2.89–4.35 GHz, and 4.59–7.055 GHz, respectively.

In this paper, the planar antenna design is presented for wideband cellular, GPS, and WLAN applications. An E-shaped radiating element with a partial ground plane is used to achieve a response from 0.9–2.63 GHz. Rest of the paper is organized in this way. Section 2 briefly explains the design of the proposed wideband antenna. Section 3 provides a parametric study. Section 4 describes the results and finally, the conclusion of the paper is given in Sect. 5.

2 Wideband Antenna Design

Figure 1 depicts the design of the proposed wideband antenna, while Fig. 2 shows the prototype of the fabricated antenna. The overall optimized design parameters of the proposed antenna are given in Table 1. A low-cost FR-4 substrate of thickness, $h = 1.6$ mm and relative permittivity, $\varepsilon_r = 4.4$ is chosen for the design of proposed antenna. An E-shaped patch radiator is used to achieve a wideband response, which is connected with a microstrip feed line of 50 Ω having width and length, W_f and L_f, respectively. The feed line width is calculated by using the equation given below [11]:

$$W_f = \left[\frac{8e^A}{e^{2A} - 1}\right] \times h \tag{1}$$

where

$$A = \frac{Z_{\text{in}}}{60}\left(\frac{\varepsilon_r + 1}{2}\right)^{1/2} + \frac{\varepsilon_r - 1}{\varepsilon_r + 1}\left(0.23 + \frac{0.11}{\varepsilon_r}\right) \tag{2}$$

where Z_{in} is the input impedance.

A partial ground plane is designed on the other side of the substrate for impedance matching. The length L_1 and width W_1 of the outer arms of E-shaped patch radiator are playing an important role to achieve a wideband response. The value of parameter L_1 is approximately equal to $\lambda_g/4$, where λ_g is the guided wavelength and can be calculated as

$$\lambda_g = \frac{c}{f_r \sqrt{\varepsilon_{\text{reff}}}} \tag{3}$$

where

$$\varepsilon_{\text{reff}} = \frac{\varepsilon_r + 1}{2} \tag{4}$$

where c is the speed of light, f_r represents resonant frequency, which is 900 MHz, and $\varepsilon_{\text{reff}}$ denotes effective permittivity of the substrate.

The parameters, such as L_2, W_2, L_3, W_3, and the gap between the ground plane and main radiator are also playing an important role in impedance matching.

3 Parametric Study

In this section, a parametric study is provided for different antenna parameters.

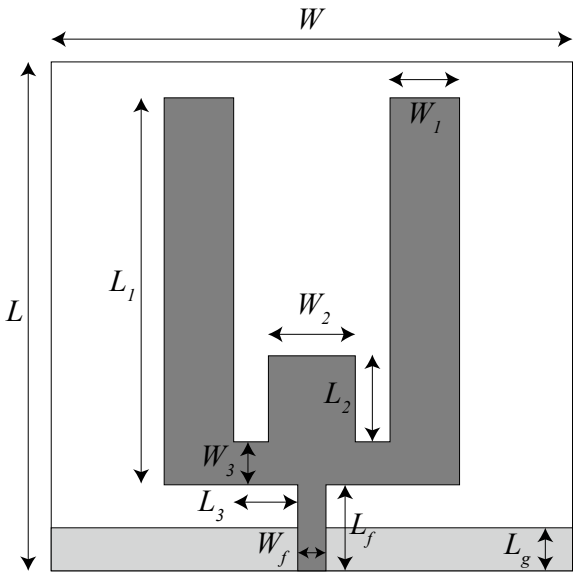


Fig. 1 Design of the proposed wideband E-shaped planar antenna

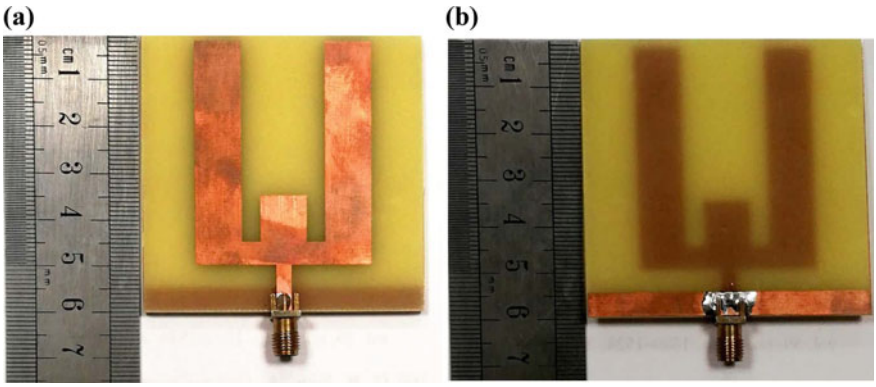


Fig. 2 Prototype of the fabricated antenna. **a** Front view. **b** Back view

Table 1 Design parameters of the proposed wideband planar antenna (in mm)

L	W	L_1	W_1	L_2	W_2
60	60	48	10	10	10
L_3	W_3	L_f	W_f	L_g	–
7.375	5	10	3.25	5	–

The effect of the parameter W_1 on the antenna's performance is shown in Fig. 3a. It is observed from the figure that for lower values of W_1, i.e., for 2 and 4 mm, the antenna is providing dual-band response. For values greater than 4 mm, the antenna offers a wideband response in the frequency range of 0.9–2.6 GHz. Figure 3b shows the effect of parameter L_1 on antenna's performance. From the figure, it is observed that for L_1 = 23 and 38 mm, antenna is resonating only at one frequency band. After that the bandwidth shift toward lower frequency bands when parameter L_1 is increased up to 48 mm.

The effect of design parameters W_2 and L_2 on the antenna's performance are illustrated in Fig. 4a, b. From Fig. 4a, it is observed that by changing W_2 from 2 to 10 mm, there is a minor change in return loss value is observed at higher frequencies. The same behavior is observed for L_2 values as shown in Fig. 4b.

Figure 5a, b illustrates the effect of W_3 and L_3 on the performance of antenna. It is observed from Fig. 5a that except W_3 = 2.5 mm, there is no significant effect is observed on antenna's performance. For L_3 = 2.125 and 4.75 mm, shown in Fig. 5b, antenna is not able to provide any resonance in the band of interest. For values greater than 4.75 mm, the antenna is able to provide a wideband response from 0.9 to 2.63 GHz.

From the above-presented discussion, it can be noted that the design parameters W_1, L_1, W_3, and L_3 are playing an important role in achieving wide impedance bandwidth.

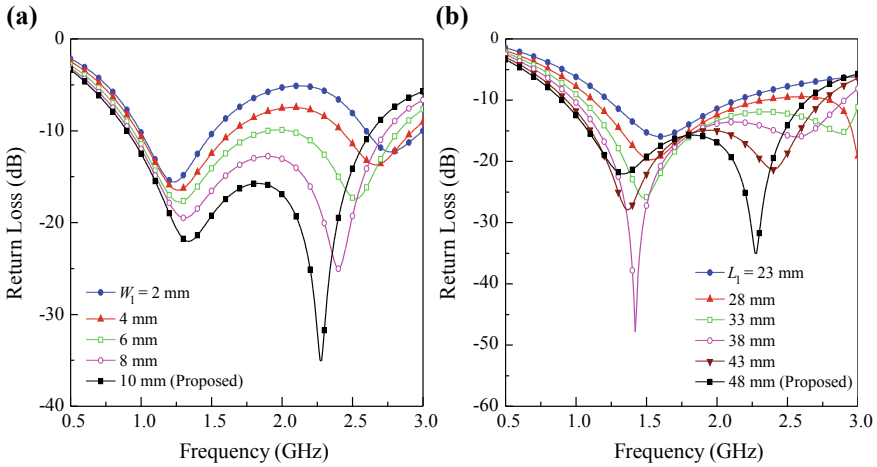


Fig. 3 Effect of W_1 and L_1 on antenna's performance

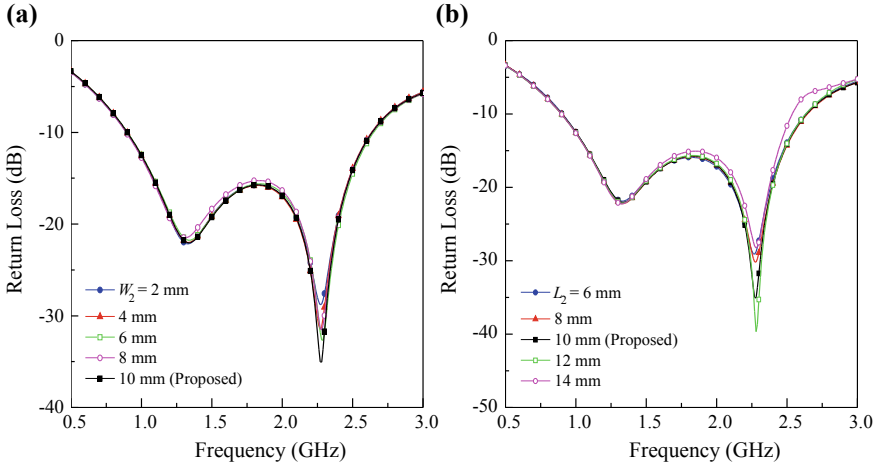


Fig. 4 Effect of W_2 and L_2 on antenna's performance

4 Results and Discussion

The proposed antenna is designed and simulated in commercially available electromagnetic software ANSYS HFSS. From Fig. 6, it is observed that the proposed antenna provides resonance in the frequency range of 0.9–2.63 GHz. For the verification of simulation results, the fabricated antenna prototype, shown in Fig. 2, is measured using Agilent Technologies Vector Network Analyzer (VNA) N5232A in the frequency range of 0.5–3 GHz. The measured return loss characteristics are also illustrated in Fig. 6. It is observed from the figure that the simulated and measured

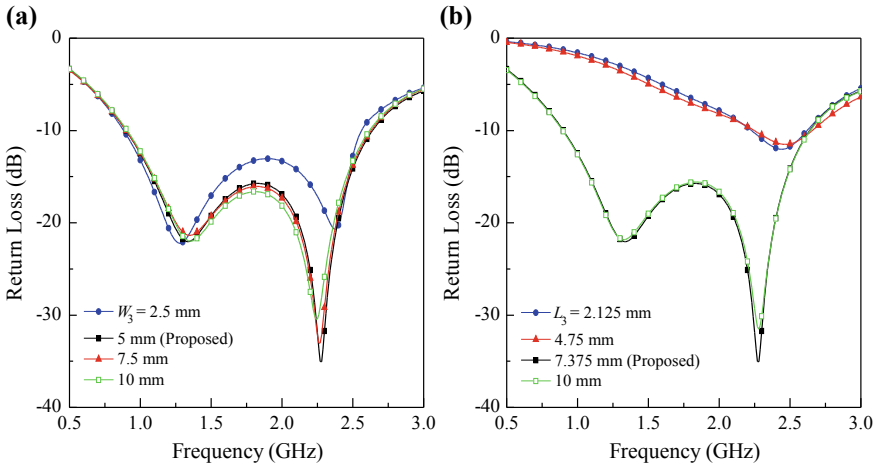


Fig. 5 Effect of W_3 and L_3 on antenna's performance

results are in good agreement. The measured impedance bandwidth of the proposed antenna is 1.48 GHz from 0.99 to 2.47 GHz. The deficiency between the results is due to imperfect soldering of the SMA connector and fabrication tolerances. From the result of Fig. 6, one can observe that the proposed wideband antenna is suitable for GSM (950/1800 MHz), GPS (1227/1575 MHz), WCDMA/UMTS (2100 MHz), and WLAN (2.45 GHz) applications.

The radiation characteristics of the proposed antenna for resonant frequencies 1227 MHz, 1575 MHz, 1800 MHz, 2100 MHz, and 2.45 GHz are illustrated in Fig. 7a–e. It is observed that for E-plane, antenna offers bidirectional radiation characteristics. For H-plane, omnidirectional radiation properties are observed.

Figure 8 depicts the simulated gain result of the proposed wideband antenna. From the figure, it is observed that for lower frequencies (0.9–1.25 GHz), the gain value is less than 1 dBi. For the rest of the band, a linear increase in gain value is observed with the increase in resonant frequency. The minimum value of gain is noted to be 0.8 dBi at 0.9 GHz and the maximum value is 3.34 dBi at 2.7 GHz. The radiation efficiency of the proposed wideband antenna is also depicted in Fig. 8. For lower resonant frequencies, the average radiation efficiency is noted to be 33.3%. For resonant frequencies greater than 1.5 GHz, the average radiation efficiency is 80%.

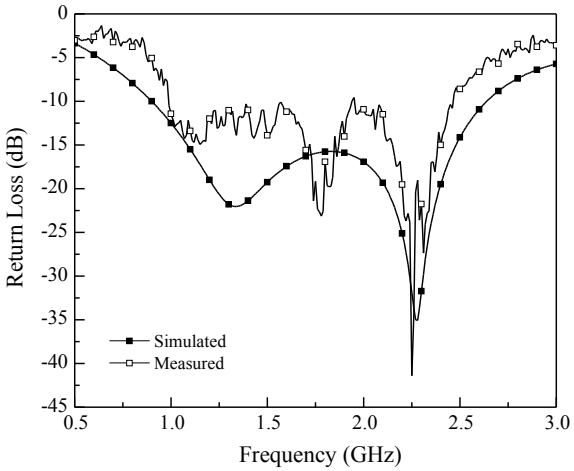


Fig. 6 Simulated and measured return loss characteristics of the proposed wideband planar antenna

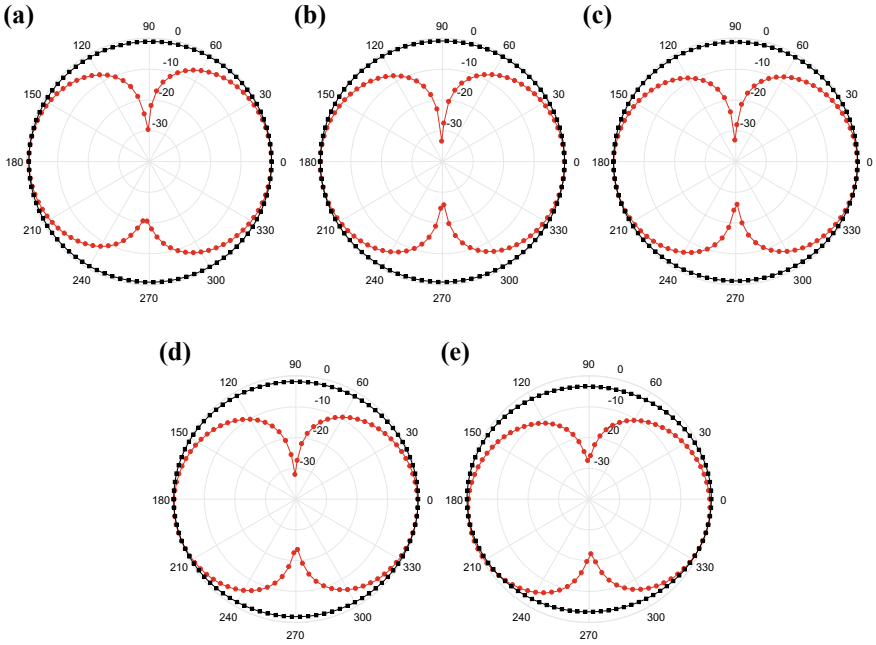


Fig. 7 Radiation characteristics of the proposed wideband planar antenna for **a** 1227 MHz **b** 1575 MHz **c** 1800 MHz **d** 2.1 GHz **e** 2.45 GHz (solid circle (●) E-plane, solid square (■) H-plane)

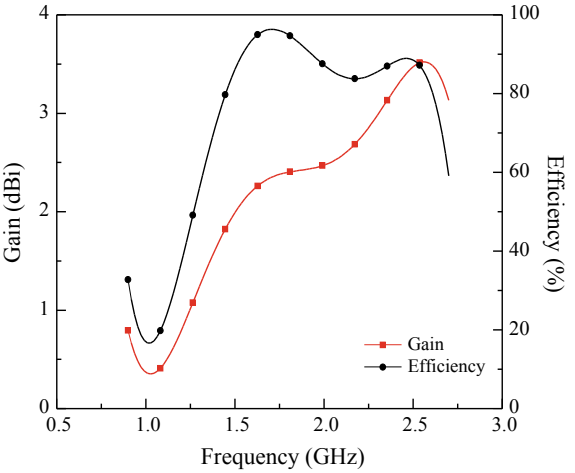


Fig. 8 Gain and radiation efficiency of the proposed wideband planar antenna

5 Conclusion

An E-shaped printed antenna is designed for wideband applications. It is observed from simulation results that the designed antenna is providing a bandwidth of 1.73 GHz from 0.9 to 2.63 GHz, while the measured impedance bandwidth is 1.48 GHz in the frequency range of 0.99–2.47 GHz. A brief parametric study is also provided to evaluate the antenna's performance. Furthermore, the proposed wideband antenna exhibits good radiation characteristics with acceptable gain. From the presented study, it can be noted that the proposed antenna design is suitable for GSM, GPS, WCDMA/UMTS, and WLAN communications.

References

1. M.K. Karkkainen, Meandered multiband PIFA with coplanar parasitic patches. IEEE Microw. Wirel. Compon. Lett. **15**(10), 630–632 (2005)
2. H.-S. Yoon, S.-O. Park, A new compact hexaband internal antenna of the planar inverted F-type for mobile handsets. IEEE Antennas Wirel. Propag. Lett. **6**, 336–339 (2007)
3. C.-W. Chiu, Y.-J. Chi, Planar hexa-band inverted-F antenna for portable device applications. IEEE Antennas Wirel. Propag. Lett. **8**, 1099–1102 (2009)
4. J. Lee, Y. Sung, Heptaband inverted-F antenna with independent resonance control for mobile handset applications. IEEE Antennas Wirel. Propag. Lett. **13**, 1267–1270 (2014)
5. C.H. See, R.A. Abd-Alhameed, D. Zhou, T.H. Lee, P.S. Excell, A crescent-shaped multiband planar monopole antenna for mobile wireless applications. IEEE Antennas Wirel. Propag. Lett. **9**, 152–155 (2010)
6. M.B. Ahmed, M. Bouhorma, F. El Ouaai, S. Bri, A. Mamouni, A.H. Boudhir, A compact circular-shaped antenna for the multi 3G cellular phones applications, in *International Conference on Multimedia Computing and Systems* **2012**, pp. 222–225 (2012)
7. L. Sun, M. He, J. Hu, Y. Zhu, H. Chen, A butterfly-shaped wideband microstrip patch antenna for wireless communication. Int. J. Antennas Propag. **2015**, (2015)

8. M. Nishamol, V. Sarin, S.M. Nair, M. Jayakrishnan, P. Mohanan, K. Vasudevan, CPW fed wideband printed antenna for wireless communications, in *Fifth International Conference on Advances in Computing and Communications* **2015**, pp. 241–244 (2015)
9. P.K. Bharti, H.S. Singh, G.K. Pandey, M.K. Meshram, Thin profile wideband printed monopole antenna for slim mobile handsets applications. Prog. Electromagn. Res. C **57**, 149–158 (2015)
10. A.Z. Manouare, S. Ibnyaich, A.E. Idrissi, A. Ghammaz, Miniaturized triple wideband CPW-fed patch antenna with a defected ground structure for WLAN/WiMAX applications. J. Microw. Optoelectron. Electromagn. Appl. **15**(3), 157–169 (2016)
11. A.B. Constantine, *Antenna Theory: Analysis and Design*, 3rd edn. (wiley, 2005)

Radial Velocity Distortion Reduction for NLFM-Based Radar System Using a Notch Filter



Ankur Thakur, Salman Raju Talluri and Sandeep Kumar

1 Introduction

Radar systems are used to detect the targets and then track them [1]. This is done by transmitting a pulse signal into space and then observing the reflected signal from the target. Usually, small duty cycle pulses with high pulse repetitive interval are used in practical applications [1, 2]. For tracking a target, its position and velocity are required. Tracking is done by observing the current position and current velocity of the target over a period of time. Then, predicting the future position and velocity of the target. The time difference between the transmitted signal and the reflected signal is used to find the range of the target, which is equal to the round trip time. If there is a relative velocity between the radar system and target, the frequency of the transmitted pulse signal will be varied which is called as Doppler frequency. This shift in the original transmitted signal frequency is used to measure the radial velocity of the target and for a stationary target, the Doppler shift is zero [1].

A good radar system has to determine the range and velocity more accurately. At the same time, the radar system must also capable of resolving two nearby targets as distinct which are moving with relatively close velocities [2]. This means range and frequency resolution of the radar system must be good. To meet these requirements, a very narrow pulse is required which is good for improved range resolution and for resolving the velocities more accurately the pulse repetitive frequency has to

A. Thakur · S. R. Talluri
Department of Electronics and Communication Engineering, Jaypee University of Information Technology, Solan, Waknaghat 171001, India
e-mail: thakur9569@gmail.com

S. R. Talluri
e-mail: salmanraju.taluri@juit.ac.in

S. Kumar (✉)
Central Research Laboratory, Bharat Electronics Limited, Ghaziabad, India
e-mail: sann.kaushik@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_65

be high. If the pulse repetitive frequency is high, the maximum range that radar can detect will be small. For meeting these requirements, modification in the transmitting waveform is required and can be done with pulse compression (PC). The effective short duration pulse is produced by transmitting long duration pulse with an angle modulation is called PC [2, 3].

The reflected signal from the target is passed through the matched filter (MF), which maximizes the peak signal-to-noise ratio [2]. This filter response must be matched to the transmitted pulse [2]. If there is only a time delay between the transmitted pulse and reflected pulse (for a stationary target), the output of the MF will peak at the time instant that corresponds to the round trip time of the target. This peak amplitude of the output is used to determine the range of the target. For a very accurate range measurement, this peak must be very narrow and sharp [2]. Ideally, it must be similar to the impulse function with no spread in the time domain. If there is any difference between the MF response and the received signal from the target, the peak of the MF output shifts and magnitude will also be reduced. This causes distortion and hence limits the radar system performance to measure the range accurately [2].

1.1 Related Works

To overcome radial velocity distortion many techniques have been developed such as Linear frequency-modulated (LFM), Nonlinear frequency-modulated (NLFM), and polyphase codes. Hyperbolic frequency modulated signals are used for target tracking with minimizing the distortion at the MF output [4]. Vizitiu [5] has formed a technique to overwhelm the problems of LFM signal that is widening of the main lobe width which disturbs the range resolution by using nonlinear laws. Doppler ambiguity resolution techniques have been proposed to further utilize the signals in high-velocity regime applications [6]. In [7], a method is generated for a series of deceiving coherent false targets based on periodic 0-pi phase modulation for defying LFM radar. In [8], the trouble of more than one passive coherent location radar structures hooked up on shifting platforms for range resolution enhancement is addressed. It also described how these various shift impedes proper imaging. A technique of PC based on an upper order solution results in a left-handed nonlinear transmission line with collection nonlinear capacitances is developed in [9].

Conventionally, LFM signals are used for range and velocity measurements. But they are more prone to distortion in the output of the MF if the radial velocity of the target is large and hence NLFM signals are used when the velocity of the object is very high [10, 11]. In order to reduce the distortion, the NLFM signals are explored, which gives relatively smaller side lobes in the MF output [12, 13]. The main advantage of the NLFM signals over the LFM signals is that they are invariant to target velocity. Along with this, these NLFM signals give less side lobe levels. In [14], the NLFM signals are used in order to have relatively side lobe levels. This can be further improved by using the windowing (which incorporates the amplitude

tapering) [15]. In [16], the Doppler invariant waveforms using the hyperbolic NLFM signals are presented. Polyphase code-based waveforms are presented whose range resolution capability is good [17, 18]. In [19], a technique was projected to project the code sequences in order to maximize the radar signal-to-interference plus noise ratio. It can reduce the mutual interference between radar and communication, but it has a huge computational complexity for the high resolution. In [20–22], to achieve low-range side lobes, the authors propose a mismatched filter for the transmitted waveform with frequency nicks has similar range side lobes. These side lobes are the auto-correlation without frequency notches at the cost of a signal-to-noise ratio loss.

In [23–26] different classical orthogonal polynomials are used to design the waveforms that optimized the time–bandwidth product in PC. So, hyperbolic NLFM can be approximated with the polynomial-based nonlinearities for the low frequency transmitted signal. In this paper, analysis with polynomial-based NLFM signals is done to reduce the radial velocity distortion. This improvement has been suggested with the help of the notch filter. The main objective of this paper is to analyze the Doppler-invariant property of the polynomial-based NLFM signals with notch filtering. The notch filter used in the analysis is of the six-pole filter and since it is very easy to implement thereby it can be used in real-time practical applications.

1.2 Motivation

The challenges which persevere during PC decoded from the literature survey comprise:

- High distortion in the response of MF for LFM signals due to the radial velocity of the target.
- Less reduction in PSLs is achieved that creates the problem of false alarm detection.

The proposed method is developed to overcome all these issues during the PC.

1.3 Contribution

The author's involvement is summarized as follows:

- This article uses the NLFM signal of orders $n = 3$ to $n = 6$, for the observation of distortion in the MF response.
- The transmitting signal is passed through a six-pole Butterworth notch filter, and then a reduction in the side lobe is compared with the unfiltered signal using auto-correlation function (ACF).
- For approaching and retarding target, Doppler invariance properties are compared with and without filtering of the transmitting NLFM.

The remaining paper is ordered as follows: In Sect. 2, the models are built and the output from the notch filter is theoretically deduced. Section 3 addresses the simulation and the results of the method which is presented in this article. Finally, part 4 concludes the paper.

2 Problem Formulation

The transmitted signal $x(t)$ in a radar system experiences changes in its amplitude, phase, and frequency. This depends on the range, velocity of the target and the medium in which the target is present. Due to the velocity, the received reflected signal from the target can be expanded or can be compressed into the time domain, depending on the approaching target or receding from the radar system. For a receding target, the signal will be expanded while for an approaching target the signal will be compressed. The received signal can be expressed as [10]

$$x_r(t) = x(st - \varphi_0) \tag{1}$$

where $s = 1 \pm 2v/c$ (+sign is for approaching the target and—sign is for retarding target). Here, s is called as the scaling factor and φ_0 is the phase factor coming from the initial range R_0 and reference time t_0. It can be obtained as

$$\varphi_0 = \frac{2R_0}{c} + \frac{2v}{c} t_0 \tag{2}$$

This received signal $x_r(t)$ is applied to the MF whose impulse response has been taken as the communicated signal. Most radar systems operate in pulse mode. The transmitted pulse with NLFM can be represented in mathematical form as

$$x(t) = A(t) \cdot \exp^{j\phi(t)} \cdot \mathrm{Rect}\left(\frac{t - (\tau/2)}{\tau}\right) \tag{3}$$

where $A(t)$ represents the amplitude modulation while $\varphi(t)$ represents the phase modulation. Here $\mathrm{Rect}((t - (\tau/2))/\tau)$ represents the rectangular pulse of duration τ seconds. Amplitude modulation is not preferred in radar systems due to the noise sensitivity. However, the amplitude modulation improves the response of the MF output slightly [18, 25]. Generally, the phase function $\varphi(t)$ is a nonlinear function of time. This phase function represents the angle modulation of the carrier which in turn can be considered as phase modulation and frequency modulation. The relationship between instantaneous frequency and phase function are given as

$$f_{in} = \frac{1}{2\pi} \frac{d\phi(t)}{dt} \tag{4}$$

For the NLFM signals, this phase function must be monotonically increasing the function of time. This paper considers this monotonic variation with a polynomial-based nonlinear and hence the phase function have been expressed as

$$\phi(t) = 2\pi f_{\min} t + \alpha t^n \tag{5}$$

$f_{\min}$ is the minimum frequency of the signal and α has to be derived depending on the value of n, in order to keep the minimum and maximum frequencies of all the NLFM signals kept constant. The parameter α can be obtained using

$$\alpha = \frac{f_{max} - f_{\min}}{n\tau^{(n-1)}} \tag{6}$$

where τ is the overall duration of the pulse.

2.1 Notch Filter Modeling

To observe the MF output for a notch filter, LFM waveform analysis has been done mathematically. Here two assumptions are made for analysis: In first, the transmitted LFM signal is as follows:

$$x(t) = \mathrm{Rec}t\left(\frac{t}{\tau}\right) \cdot \exp(j\pi\beta^2), \quad \beta = B/\tau \tag{7}$$

$$X(\omega) = \mathrm{Rec}t\left(\frac{\omega}{2\pi B}\right) \cdot \exp\left(-j\frac{\omega^2}{4\pi\beta}\right) \tag{8}$$

where B is the bandwidth and $X(\omega)$ is the spectrum of $x(t)$. In second, interference in the echo signal is considered on the receiver side whose response is narrow with the carrier of ω_1 and B_1 bandwidth. These interferences should be less because they may liable for false target finding. The MF with a cut at ω_1 is designed for suppression purposes and its spectrum is represented in (9).

$$H(\omega) = X^*(\omega) - \mathrm{Rec}t\left(\frac{\omega - \omega_1}{2\pi B_1}\right) \cdot \exp\left(j\frac{\omega^2}{4\pi\beta}\right) \tag{9}$$

Usually, the received radar signal includes the echo signal and the interference signal [27, 28]. Considering only one high-frequency interference signal ($x_{\mathrm{int}}(t)$) and the overall spectrum of the received signal is

$$X_r(\omega) = X(\omega) \cdot \exp(-j\omega\tau) + X_{\mathrm{int}}(\omega) \tag{10}$$

$X_{\text{int}}(\omega)$ is the spectrum of the interfering signals. The received signal is processed with the MF and then overall output becomes

$$Y(\omega) = X_r(\omega) \cdot H(\omega) \tag{11}$$

$$\begin{aligned} Y(\omega) = &\left(\text{Rect}\left(\frac{\omega}{2\pi B}\right) - \text{Rect}\left(\frac{\omega-\omega_1}{2\pi B_1}\right)\right) \cdot \exp(-j\omega\tau) \\ &+ \left(\text{Rect}\left(\frac{\omega}{2\pi B}\right) - \text{Rect}\left(\frac{\omega-\omega_1}{2\pi B_1}\right)\right) \cdot X_{\text{int}}(\omega) \cdot \exp\left(j\frac{\omega^2}{4\pi\beta}\right) \end{aligned} \tag{12}$$

As per assumptions, $x_{\text{int}}(t)$ is a narrow band signal with carrier ω_1, and bandwidth B_1 thereby overall output spectrum is

$$Y(\omega) = \left(\text{Rect}\left(\frac{\omega}{2\pi B}\right) - \text{Rect}\left(\frac{\omega-\omega_1}{2\pi B_1}\right)\right) \cdot \exp(-j\omega\tau) \tag{13}$$

Its time domain equivalent is

$$y(t) = B \cdot x(B(t-\tau)) - B_1 \cdot x(B_1(t-\tau)) \exp(-j\omega(t+\tau)) \tag{14}$$

For multiple frequency interference, the MF output for N notches is given by

$$H(\omega) = X^*(\omega) - \sum_{\beta=1}^{N} \text{Rect}\left(\frac{\omega-\omega_\beta}{2\pi B_\beta}\right) \cdot \exp\left(j\frac{\omega^2}{4\pi\beta}\right) \tag{15}$$

Corresponding MF output spectrum is

$$Y(\omega) = \left(\text{Rect}\left(\frac{\omega}{2\pi B}\right) - \sum_{\beta=1}^{N} \text{Rect}\left(\frac{\omega-\omega_\beta}{2\pi B_\beta}\right)\right) \cdot \exp(-j\omega\tau) \tag{16}$$

It is observed from (14) that the output of the MF with a notch has two parts. One is the desired part and the other one is the undesired part, which is a high-frequency signal named as clutter and it is determined by ω_1 and B_1 parameters. In this article, these undesired signals for NLFM waveforms are suppressed by notch filtering process as observed in simulations. However, LFM is not preferred when the velocity of the target is high because it gives more distortion in the MF output [29]. In this article, the performance of the Butterworth notch filter on NLFM is analyzed. It is observed that the performance of the notch filter is influenced by the notches width and carrier value, but it little affects the resolution of system impulse response as shown in (14). The main objective of this analysis is to observe the effect of notch filtering on the NLFM signals before transmitting into the free space with the constraints of having the same pulse duration with transmitted signal frequency variations kept constant.

3 Simulations and Results

In the simulations, four nonlinear polynomial-based variations have been considered. For all these NLFM modulated signals, the initial frequency and frequency deviation have been taken as a constant.

The initial frequency is set to 50 Hz and the final frequency has been set to 100 Hz. The duration of the pulse has been taken as 2 sec. The values taken can be changed according to the practical scenario, but the conclusions will not be changed. These signals pass through a six-pole Butterworth notch filter for different fractional bandwidths. The center frequency of the notch filter has been selected as the frequency at which the maximum peak appears in the unfiltered signal spectrum. The spectrum of the NLFM signal without filtering is decreasing in a hyperbolic manner with oscillatory nature giving minimal and maximal. The differential bandwidth has been taken such as to suppress the first peak only, first two peaks only and first three peaks, and so on. These cutoff frequencies are selected such that the highest peaks in the spectrum are suppressed so that the normalized spectrum will appear in a linear manner rather than hyperbolic variation. Using Fast Fourier Transform, the spectrums of the NLFM signals without filtering and with filtering are obtained for $n = 3$ to $n = 6$. In this paper, graphs are shown only for $n = 3$ and $n = 6$. Figure 1 represents the spectrum of the NLFM signal with $n = 3$ with filtering and without filtering. Four

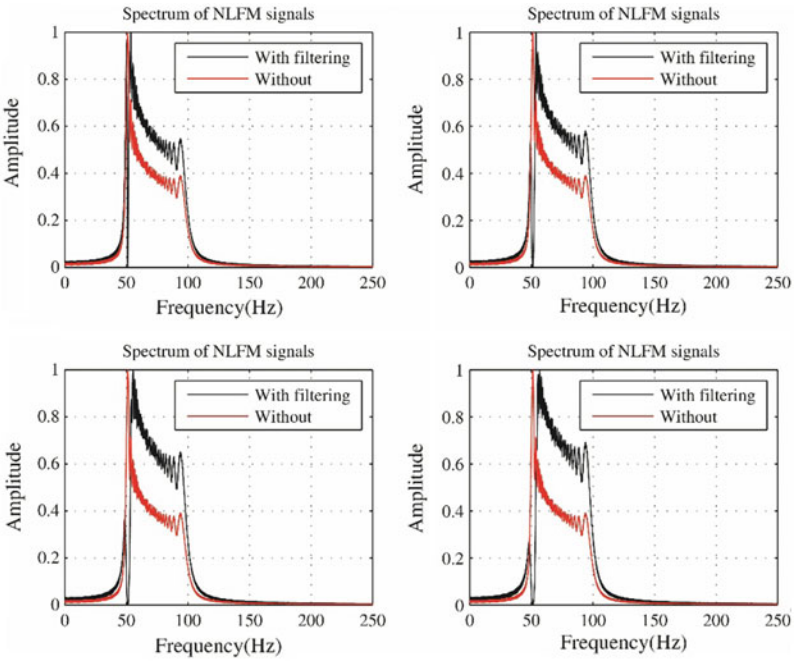


Fig. 1 Spectrum of the NLFM signals with and without filtering $n = 3$

different differential bandwidths have been used. The cutoff frequency for $n = 3$ of the notch filter has been taken as 51.2 Hz.

The bandwidths have been taken at 1.57 Hz (for the suppression of the first peak), 3.47 Hz (for the suppression of the first two peaks), 4.98 Hz (for the suppression of first three peaks), and 6.37 Hz (for the suppression of first four peaks) in the cyclic order. From this figure, it is observed that increasing the fractional bandwidth of the filter is improving the normalized spectrum close to unity. However, it has an adverse effect on the auto-correlation function since this increases the peak in the side lobe level. Figure 2 represents the auto-correlation function (MF output), it is observed that increasing the cutoff frequency has a positive effect on reducing the relative side lobe levels far from the main lobe but the peaks near the main lobe is increasing. Hence, from this figure, it is suggested that there is an optimal cutoff frequency of the notch filter.

Similar analysis has been carried out for the other orders for $n = 4$, $n = 5$, and $n = 6$. Figure 3 represents the spectra and Fig. 4 the auto-correlation functions for $n = 6$. From these figures, it is observed that, if the order of the nonlinearity is increasing, there is a great improvement in the reduction of side lobe levels with filtering. The main conclusion from this is that the notch filtering can be effectively used to reduce the relative side lobe levels in the output of the MF for better range resolutions.

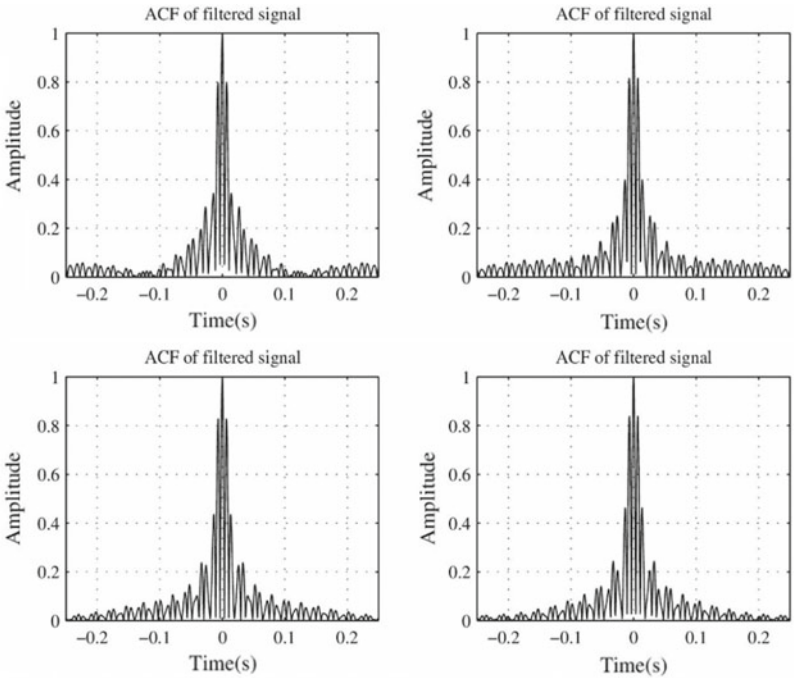


Fig. 2 ACF of filtered NLFM signals $n = 3$

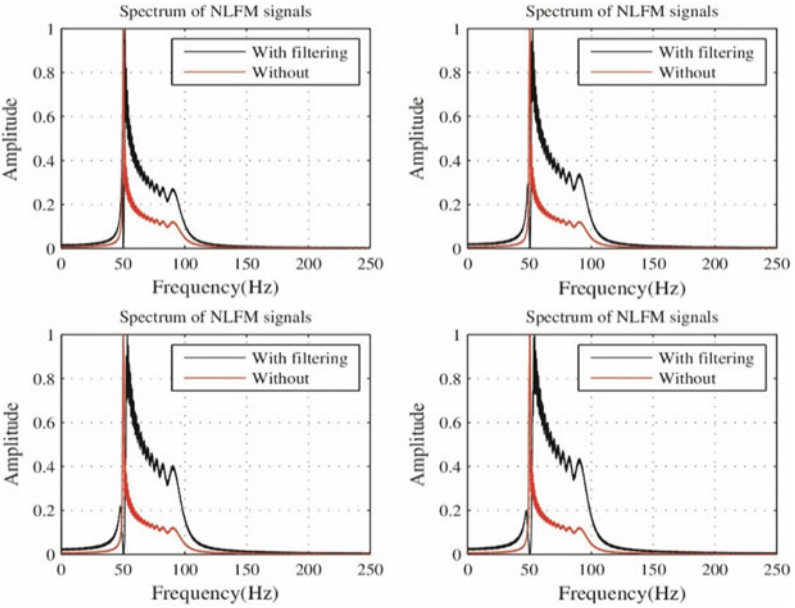


Fig. 3 Spectrum of the NLFM signals with and without filtering $n = 6$

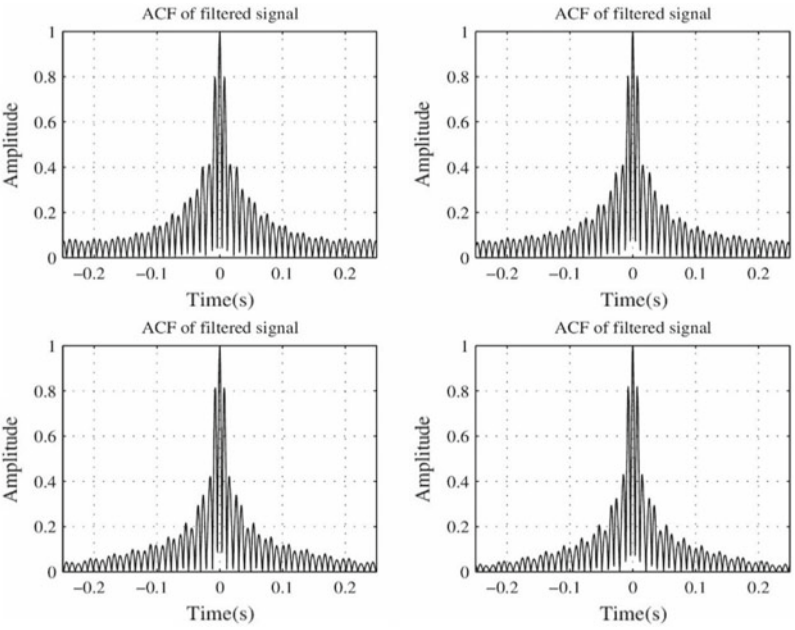


Fig. 4 ACF of filtered NLFM signals $n = 6$

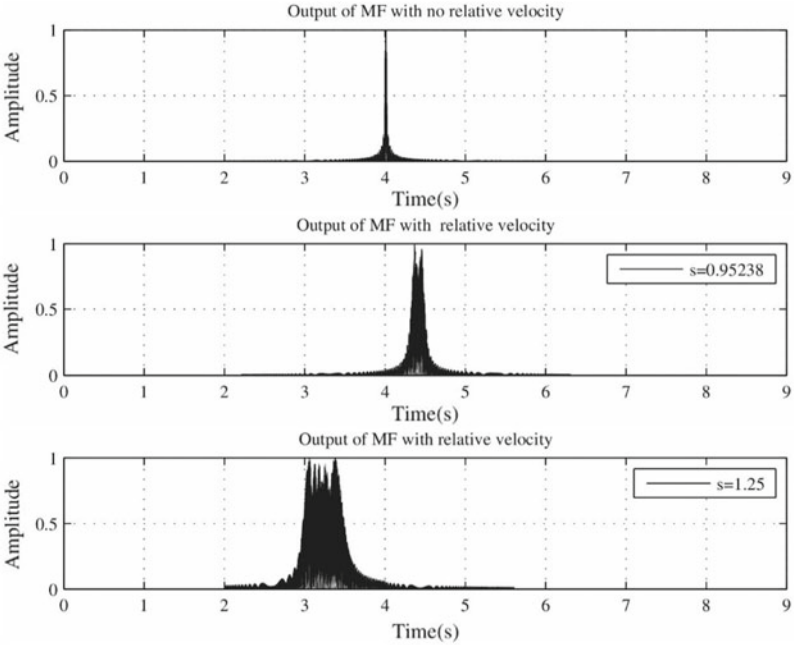


Fig. 5 Output of MF for different scaling factors of NLFM signals with and without filtering $n = 2$

The other important factor of the NLFM is the Doppler invariance property. To check this invariance property, the outputs of the MFs are observed for two scaling factors (s) for all the orders. One is for approaching the target ($s = 1.25$) and the other is for receding target ($s = 0.95238$) from the radar system. Figures 5, 6, 7, 8, and 9 represents the MF outputs for two scaling factors. From these figures, it is observed that the filtering has a positive effect on the decoupling of the range and Doppler effect which is a main important property from the analysis. Moreover, around -12 dB reduction in side lobe is achieved after filtering for an approaching target when $n = 3$ and more reduction is achieved for receding target as shown in Fig. 6.

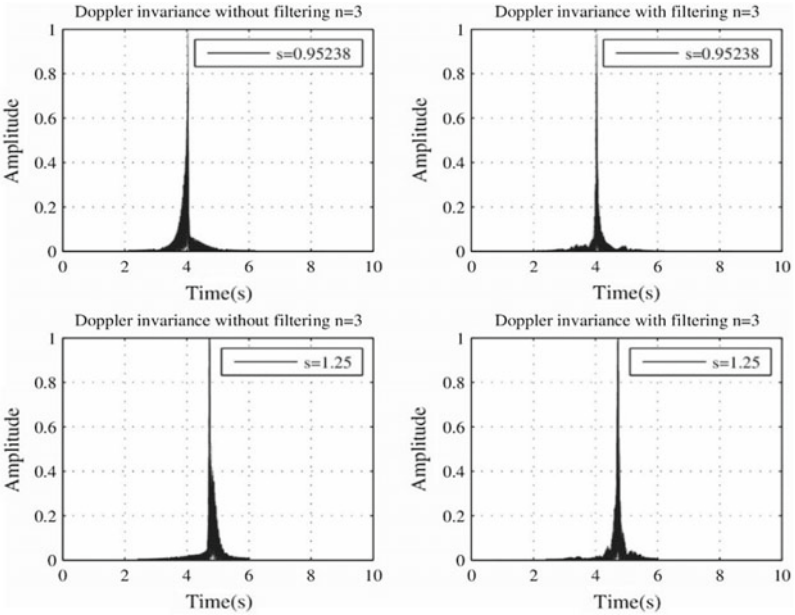


Fig. 6 Output of MF for different scaling factors of NLFM signals with and without filtering $n = 3$

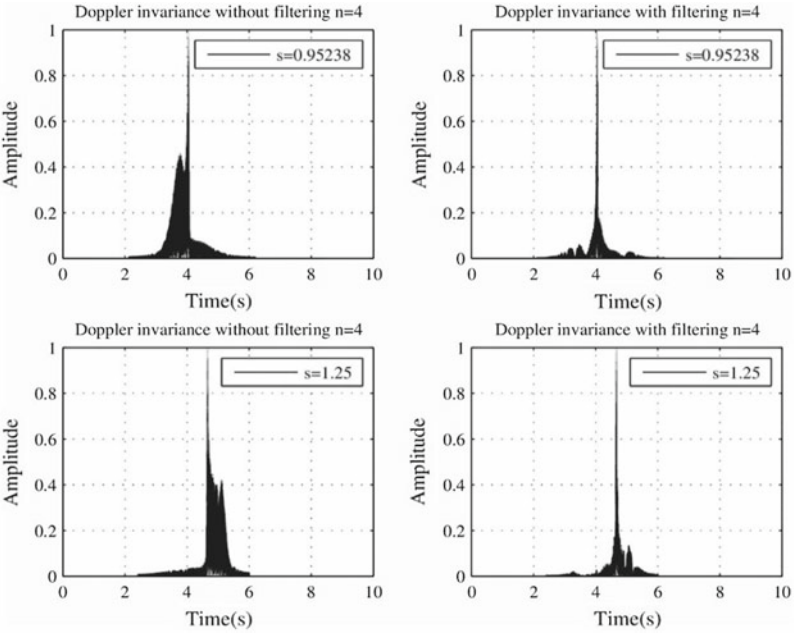


Fig. 7 Output of MF for different scaling factors of NLFM signals with and without filtering $n = 4$

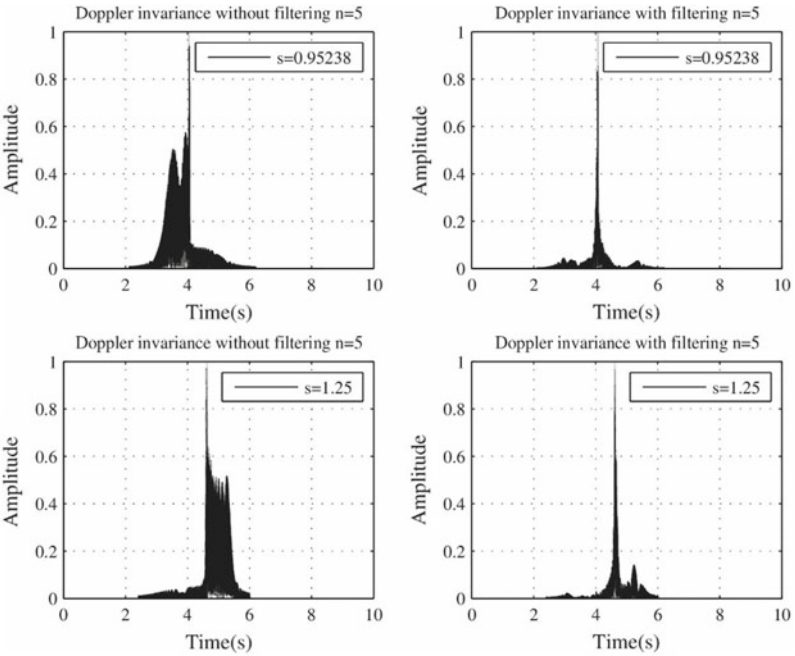


Fig. 8 Output of MF for different scaling factors of NLFM signals with and without filtering $n = 5$

In Fig. 7, receding target for $n = 4$ gives side lobe peak at -6.55 and -13.55 dB without filtering and with filtering respectively. Hence -7 dB side lobe reduction is achieved that improve the false alarm detection. So, the notch filtering process gives more reduction in the side lobes, while incorporating less distortion in the MF response. Similar observations are done for $n = 5$ and $n = 6$. It is also observed that after filtering the side lobes exists away from the main peak, which indicates better range resolution. By observing Figs. 5, 6, 7, 8 and 9, it indicates that the third-order NLFM signal is much more Doppler invariant in comparison with the other nonlinearities given the bandwidth and the starting and ending frequencies of the modulated signal are the same. Hence, it is better to use the third-order polynomial NLFM along with the notch filtering for better range and velocity resolutions.

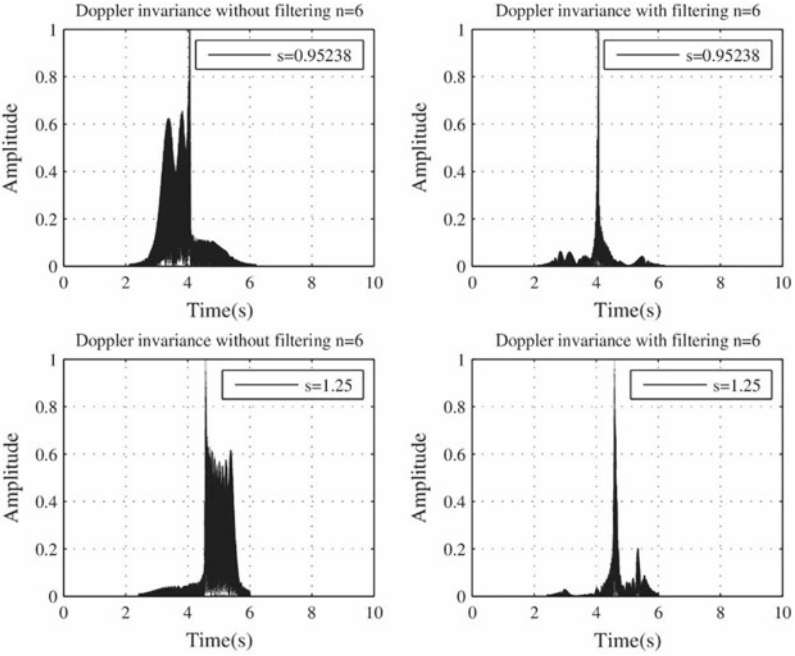


Fig. 9 Output of MF for different scaling factors of NLFM signals with and without filtering $n = 6$

4 Conclusion

Since the output of the MF is more prone to distortion due to the target velocity for the LFM signals, the reduction in the MF output for polynomial-based NLFM signals using the notch filter has been proposed. Doppler-invariant property has been compared for the LFM and four NLFM signals. From all the NLFM signals considered, third-order NLFM signals with the notch filtering are a better choice in the higher target velocity applications where distortion due to target velocity is more. Moreover, after filtering the MF output side lobes exist away from the main peak which indicates better range resolution. Further research is required to design NLFM-based signals which improve the response of the MF in a radar system.

References

1. P.Z. Peebles, *Radar Principles*, 3rd edn. (Wiley 2007)
2. M.I. Skolnik, *Introduction to Radar System*, 3rd edn. (McGraw-Hill, 2001)
3. V.C. Chen, F. Li, S.S. Ho, H. Wechsler, Micro-doppler effect in radar phenomenon model and simulation study. IEEE Trans. Aerosp. Electron. Syst. **42**(1), 2–21 (2006)

4. X. Song, P. Willett, S. Zhou, Range bias modelling for hyperbolic-frequency modulated waveforms in target tracking. IEEE J. Ocean. Eng. **37**(4), 670–679 (2012)
5. I.C. Vizitiu, Some aspects of sidelobe reduction in pulse compression radars using NLFM signal processing. Progress Electromagn. Res. **47**, 119–129 (2014)
6. A. Thakur, S.R. Talluri, R.K. Panigrahi, Sidelobe reduction in pulse compression having better range resolution. Comput. Electron. Eng. (2018)
7. S. Qingzhan, T. Ning, W. Chao, Y. Naichang, On deception jamming for countering LFM radar based on periodic 0-pi phase modulation. Int. J. Electron. Commun. **83**, 245–252 (2018)
8. W. Philipp, C. Diego, Optimal trajectories for range resolution improvement in multi-PCL SAR. Int. J. Electron. Commun. **73**, 173–182 (2017)
9. F. Gharakhili, M. Shahabadi, M. Hakkak, Higher-order soliton-effect pulse compression in a nonlinear left-handed transmission line. Int. J. Electron. Commun. **65**, 848–851 (2011)
10. B.R. Mahafza, *Radar Systems Analysis and Design Using MATLAB* (Chapman and Hall, 2013)
11. Y.K. Chan, M.Y. Chua, V.C. Koo, Side-lobes reduction using simple two and tri-stages nonlinear frequency modulation (NLFM). Progress Electromagn. Res. **65**, 33–52 (2009)
12. W.K. Lee, A pair of asymmetrical weighting receivers and polyphase codes for efficient aperiodic correlations. IEEE Commun. Lett. **10**(5), 387–389 (2006)
13. Y. Jie, K.S. Tapan, A novel Doppler-tolerant polyphase codes for pulse compression based on hyperbolic frequency modulation. Digit. Signal Process **17**, 1019–1029 (2007)
14. W. Zhou, C.M. Yeh, K. Jin, J. Yang, Y.B. Lu, ISAR imaging based on the wideband hyperbolic frequency-modulation waveform. Sensors **15**(9), 23188–23204 (2015)
15. M. Zhou, J.J. Zhang, A.P. Suppappola, Hyperbolic frequency modulation for multiple users in underwater acoustic communications. in *IEEE Symposium on Acoustics, Speech and Signal Processing* (2014) (pp. 3498–3502)
16. W. Yue, Y. Zhang, A novel nonlinear frequency modulation waveform design aimed at side-lobe reduction. in *IEEE Symposium on Signal Processing, Communications and Computing (ICSPCC)* (2014) (pp. 613–618)
17. A. Thakur, S.R. Talluri, A novel pulse compression technique for side-lobe reduction using woo filter concepts. in *IEEE International Conference on Communication and Signal Processing (ICCSP)* (2017) (pp. 1086–1090)
18. L. Xiao, K. Ling, C. Guolong, Y. Wei, A low complexity coherent integration method for maneuvering target detection. Digit. Signal Process **49**, 137–147 (2016)
19. X. Zhe, C. Baixiao, Y. Minglei, Transmitter polarization optimization with polarimetric MIMO radar for mainlobe interference suppression. Digit. Signal Process **65**, 19–26 (2017)
20. P.P. Ping, H. Li, Y.X. Zhang, W. Qi, Z.M. Deng, Range, radial velocity, and acceleration MLE using frequency modulation coded LFM pulse train. Digit. Signal Process **60**, 252–261 (2017)
21. M.A. Richards, *Fundamentals of Radar Signal Processing* (McGraw-Hill, 2014)
22. A. Aubry, V. Carotenuto, M.A. De, Optimization theory-based radar waveform design for spectrally dense environments. IEEE Trans. Aerosp. Electron. Syst. Mag. **31**(12), 14–25 (2016)
23. A. Thakur, S.R. Talluri, Comparative analysis on pulse compression with classical orthogonal polynomials for optimized time-bandwidth product. Ain Shams Eng. J. **9**(4), 1791–1797 (2018)
24. H.E. Najafabadi, M. Ataei, M.F. Sabahi, Chebyshev chaotic polynomials for MIMO radar waveform generation. IET Radar Sonar Navig. **11**(2), 330–340 (2017)
25. L. Aifei, S.L. Yee, C.T. Kah, G. Caicai, Mismatched filter for transmit waveform with frequency notches. IET Radar Sonar Navig. **12**(3), 332–340 (2018)
26. S.D. Blunt, E.L. Mokole, Overview of radar waveform diversity. IEEE Trans. Aerosp. Electron. Syst. Mag. **31**(11), 2–42 (2016)
27. O. Rabaste, L. Savy, Mismatched filter optimization for radar applications using quadratically constrained quadratic programs. IEEE Trans. Aerosp. Electron. Syst. **51**(4), 3107–3122 (2015)
28. T.R. Kishore, K.D. Rao, Automatic intra-pulse modulation classification of advanced low probability of intercepts radar waveforms. IEEE Trans. Aerosp. Electron. Syst. **53**(2), 901–914 (2017)
29. H. Griffiths, L. Cohen, S. Watts, Radar spectrum engineering and management: technical and regulatory issues. Proc. IEEE **103**(1), 85–102 (2015)

Design of Compact Two-Element MIMO F-Antenna



Anamika Gupta, Laxmi Shrivastav and Santosh Sharma

1 Introduction

In modern-day MIMO, technology is utilized in spacious amount in wireless communication system [1]. Limited frequency spectrum and input power level, it is necessary to provide furnished data. Nowadays, 4G long-term evolution (LTE) and 5G network are in trend, so multiple-input multiple-output (MIMO) antenna technology is one of the promising solutions for wireless communication system [2, 3]. No need for sacrificing the additional spectrum or power so that it can improve the channel capacity and communication speed [4]. Between the individual element pattern, MIMO antenna system should have high inter element isolation and low correlation while maintaining the low inter element spacing [5]. The frequency of 3.732 GHz is super high frequency. Its application is point-to-point communication, wireless communication (LAN), mobile, and digital TV broadcasting in kV band (DTH services—direct to home). It is compact antenna by which it required small size [6]. MIMO antennas have been required to fulfill various requirement bandwidth, gain, isolation, and correlation. MIMO antenna are used for implementing wideband structure using a different technique like common radiating element in microstrip-fed slots. Ring-shaped ground and V-shaped ground are used for implementing the wideband MIMO antenna [7]. In this F-monopole, two-element antenna are arranged in a symmetric order in which ground planes are interconnected to each other [7].

A. Gupta (✉) · L. Shrivastav · S. Sharma
Department of Electronics, MITS Gwalior, Gwalior, India
e-mail: anamikaa061995@gmail.com

L. Shrivastav
e-mail: lselex@mitsgwalior.in

S. Sharma
e-mail: santos_gwl@mitsGwalior.in

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_66

These MIMO structures brace a lower frequency resonance. MIMO antenna has a compact footprint. CST software is used for designed antenna.

2 F-Single Antenna Element

Figure 1a shows the front part of the proposed F-microstrip antenna, In F-microstrip antenna, the parameters are $d = 5.5$ mm, $e = 11$ mm, $f = 2$ mm, $g = 3$ mm, $h = 8$ mm, $i = 10.5$ mm, $j = 11$ mm, $A = 40$ mm, $B = 20$ mm, the substrate size is 20×20 mm which is represented by A and B. FR4 lossy substrate is used in this antenna ($\epsilon_r = 4.4$, $\tan\delta = 0.02$).

3 Result and Discussion

3.1 Single F-Antenna

This is a simple patch antenna with ground plane as 20×20 mm and its shape is F the same as F, hence, the name is F-antenna. The return loss of this F-antennas is -24.1 dB. The voltage standing wave ratio (VSWR) is 1.12. Gain of MIMO antenna is 1.8 dB and the bandwidth of MIMO antenna is 291.5 MHz, which is very useful and large. The directivity of this antenna is 2.429 dBi. The frequency of this antenna is dual, which means it can be used in different applications that is 3.668 GHz frequency which is used in wireless communication LAN and useful in mobile. And another frequency is 4.553 GHz whose application is (UWB) ultra wideband and the size of this antenna is very small.

In single antenna not finding the sufficient gain that why using same antenna in MIMO, for finding sufficient gain. The material of substrate used is FR-4lossy ($\varepsilon = 4.4$, $\tan\delta = 0.02$) but it doesn't give large amount of gain which is its drawback. It is used in low frequency and is easily available.

Inquiry on Equilibrium Interconnected Ground Plane MIMO Configuration

The following design criteria must be satisfied in MIMO system:

1. In entire operating band, the isolation must be greater than 10 dB.
2. Number of elements must be accommodated.
3. In MIMO configuration, ground plane are connected continuously to be useful in practical application.

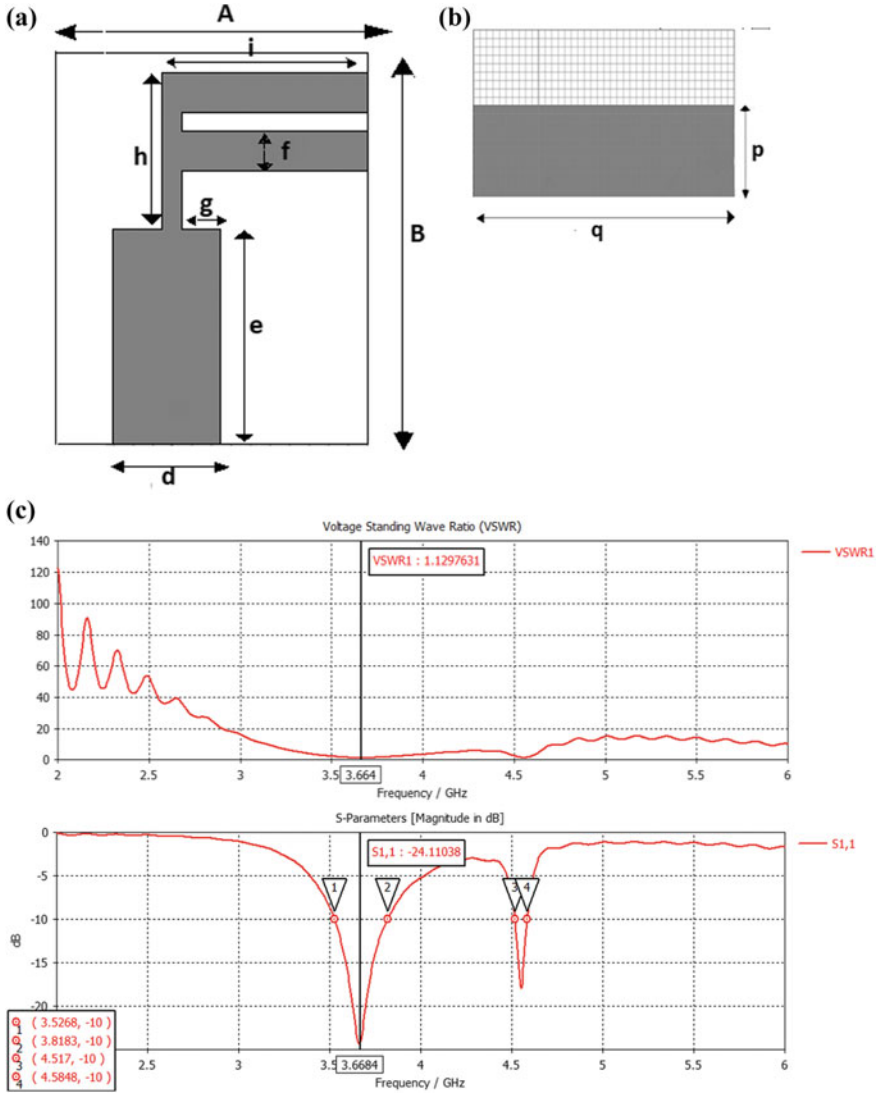


Fig. 1 **a** Single F-antenna, **b** back part of single F-antenna, **c** S11/return loss and VSWR for single F-antenna, **d** polar radiation pattern of single F-antenna

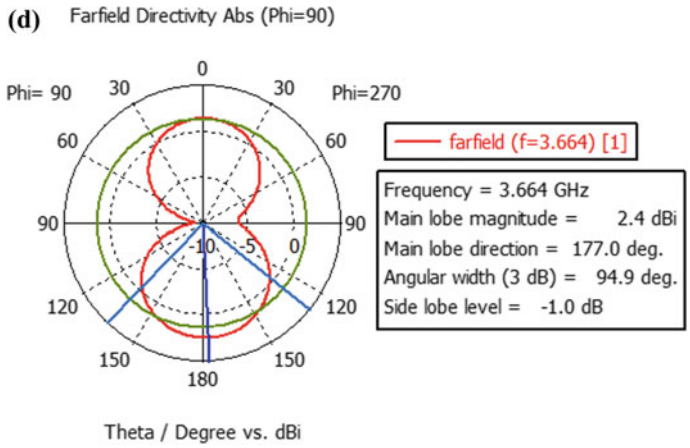


Fig. 1 (continued)

3.2 Design of Double Antenna Element

In the caption mentioned the designing parameter of two-element MIMO antenna system. There is two element of antenna that we called MIMO antenna and its shape is F so we called as F-antenna. In F-Antenna, we take two patches which have symmetric geometry and between these two F-Antenna, we design one parasitic element whose shape is U. Because of these U-shaped antennas, all the surface radiation reduces. Shape of ground is rectangular. The proposed antenna has overall size of 40×20 mm^2. The antenna is synthesized on commercially available 1.6 mm thick FR4 substrate ($\varepsilon_r = 4.2$ & $\tan\delta = 0.02$). In this MIMO antenna, we connected two antennas in F shape which we called F-shaped antenna. In that ground plane are interconnected. Comparing with the paper [1], its ground size is 20×20 and ground size is 40×20 in ground size 1 mm [1]. Its frequency is 4.41 GHz and this paper frequency is 3.732 and 4.5961 GHz. In this paper, we try to reduce frequency that is why it is used in some other applications and the antenna band is dual band and for that dual band, it is used in different application (like DTH services, digital TV broadcasting, point-to-point communication, and satellite communication), and 4.5961 is used in satellite communication, wireless line. So, these antenna frequency bands are better that antenna [1]. Range of MIMO antenna is 3.585–3.872 and other frequency range is 4.5638–4.6255, and it is useful for both of the ranges.

Figure 2a in MIMO F microstrip antenna the parameters are $a = 5$ mm, $b = 4.5$ mm, $c = -4$ mm, $d = 5.5$ mm, $e = 11$ mm, $f = 2$ mm, $g = 3$ mm, $h = 8$ mm, $i = 10.5$ mm, $j = 11$ mm, $A = 40$ mm, $B = 20$ mm.

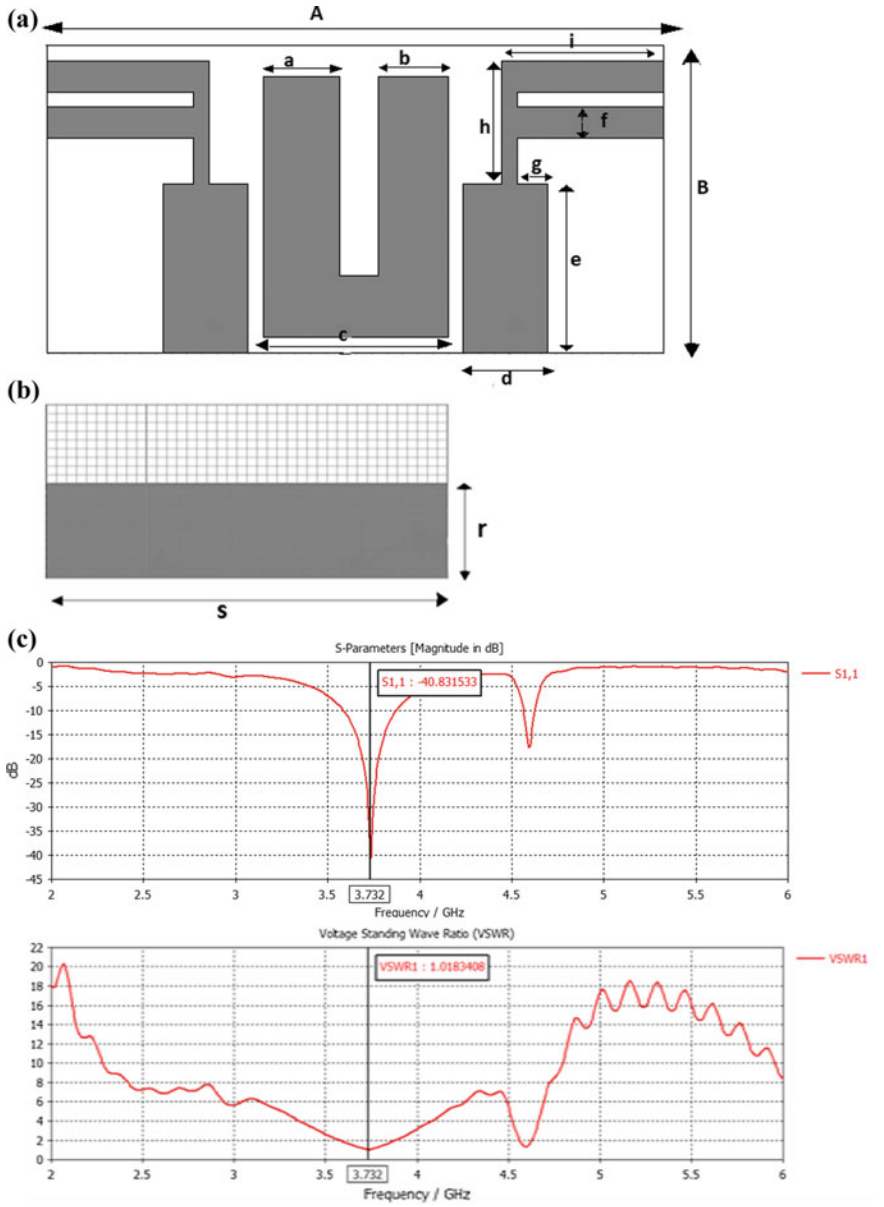


Fig. 2 **a** MIMO F-antenna front part, **b** back side of MIMO F-antenna **c** S11/return loss and VSWR for MIMO F-antenna, **d** coupling of MIMO F-antenna, **e** polar plot of MIMO F-antenna

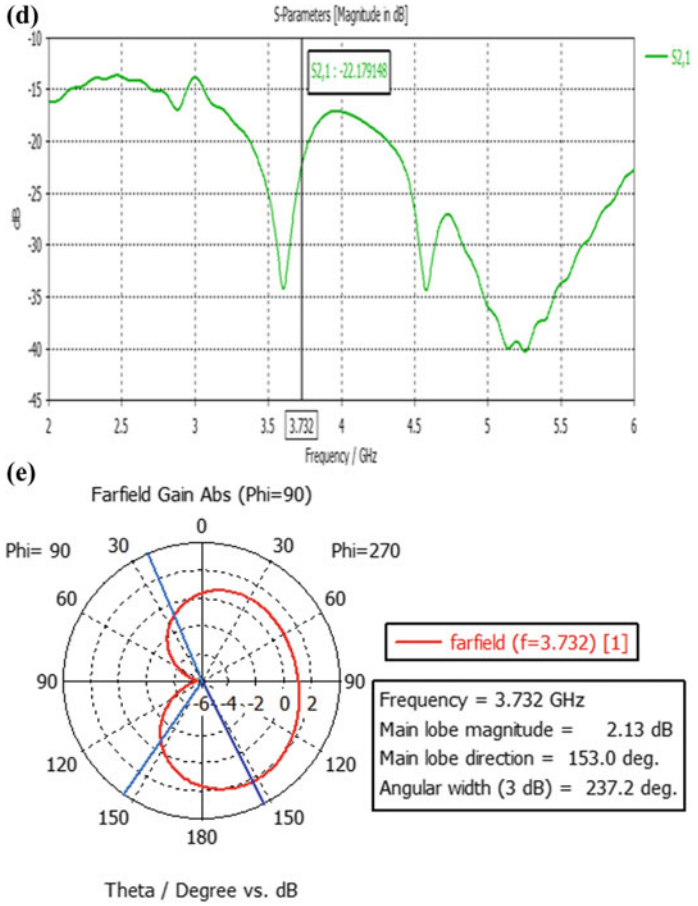


Fig. 2 (continued)

Figure 2b shows the S-parameter corresponding to the excitation of element in the MIMO system. Figure 2d shows that the mutual coupling between pair of element is $|S21| = -22$ dB. Thus, the overall inter element isolation performance's satisfactory for the proposed two-element MIMO antenna.

Figure 2e shows radiation pattern when element -1 is excited when the other ports at matched load termination. Simulated peak gain is 2.306.

Average peak gain is greater than 2 & radiation efficient is 57.97% is observed for the entire antenna operating band. The return loss of single antenna is much greater then MIMO antenna [1]. In Fig. 2c the VSWR of the antenna is 1 (Table 1).

Table 1 Comparison of parameter

Parameter	Single F-antenna	MIMO F-antenna
S_{11} (dB)	−24.1	−40.8
VSWR	1.12	1.01
Gain (dB)	1.8	2.3
Directivity (dBi)	2.429	2.886
Bandwidth (MHz)	291.5	284.6
Frequency (GHz)	3.668 and 4.5534	3.732 and 4.5993

4 Conclusion

It has a smaller footprint as compared to the antenna [2, 3] in MIMO system. In this antenna, the achieved dual band frequency is 3.732 GHz. 3.732 is super high frequency and 4.5961 is called extremely high frequency. The super high frequency (SHF) bands are frequently referred to as the microwave spectrum. In this antenna, the return loss is −40.831 and −17.7129 dB. Coupling between two antennas is −22.1390 dB. The future scope is enhancement in gain and directivity.

References

1. M.S. Sharawi, Printed multi-band MIMO antenna systems and their performance metrics. IEEE Antennas Propag. Mag. **55**, 218–232 (2013)
2. Q. Li, G. Li, W. Lee, M. Lee, D. Mazzarese, B. Clerckx, Z. Li, MIMO techniques in WiMAX and LTE: a feature overview. IEEE Commun. Mag. **48**(5), 86–92 (2010)
3. O. Bejarano, E.W. Knightly, M.Y. Park, IEEE 802.11ac: from channelization to multi-user MIMO. IEEE Commun. Mag. **51**(10), 84–90 (2013)
4. N. Herscovici, C. Christodoulou, E.R. Iglesias, O.Q. Teruel, M.S, Compact multimode patch antennas for MIMO applications. IEEE Antennas and Propag. Mag. **50**(2), 197–205 (2008)
5. A. Ramachandran, S.V. Pushpakaran, M. Pezholil, V. Kesavath, A four-port MIMO antenna using concentric square-ring patches loaded with CSRR for high isolation. IEEE Antennas Wirel. Propag. Lett. **15**, 1196–1199 (2016)
6. D. Sarkar, A. Singh, K. Saurav, K.V. Srivastava, A four element quad-band MIMO antenna employing split ring resonator and interdigital capacitor. IET Microw. Antennas Propag. **9**(13), 1453–1460 (2015)
7. M.U. Khan, M.S. Sharawi, A. Steffes, D.N. Aloi, A 4-element MIMO antenna system loaded with CSRRs and patch antenna elements. in *Proceedings of 7th European Conference on Antennas and Propagation (EuCAP)* (Gothenburg, 2013) (pp. 2016–2019)

Chebyshev Polynomials-Based Pulse Compression Waveform for Modern Microwave Radars



Ankur Thakur, Salman Raju Talluri and P. K. Verma

1 Introduction

Radar system is normally used for detection and to find the location of the object. For this purpose, long-duration waveforms are communicated toward object and then echo signal is captured. This echo waveform is managed at the radar receiver with signal processing techniques to find the features of the object. Transmitting waveform must be more energized to get powerful echo signal from the target, otherwise echo waveform may get weakened during transmission within the environment. Range resolution, $R = \frac{c\tau}{2}$ is the ability of the radar system to distinguish two or more objects in terms of range. Here, τ is the communicated pulse duration and short-period signals are used to acquire high-range resolution, which is required in radar system [1].

Basically, transmitting waveform duration cannot be reduced too much because short-duration waveforms require high peak power to get adequate energy for long-distance applications. Thereby, overall radar system becomes bulkier and costlier which is undesired. To overwhelm this issue, the waveforms having less peak power and lengthier duration are necessary at the conveying end for long-range detection [2]. With an angle modulation long-duration waveforms, bandwidth can be made larger and this larger bandwidth shows thin effective waveform duration that are

A. Thakur · S. R. Talluri
Department of Electronics and Communication Engineering, Jaypee University of Information Technology, Solan, Waknaghat 171001, India
e-mail: thakur9569@gmail.com

S. R. Talluri
e-mail: salmanraju.taluri@juit.ac.in

P. K. Verma (✉)
Electronics Engineering Department, Rajkiya Engineering College, Sonbhadra 231206, UP, India
e-mail: vermapk87@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_67

used for better range resolution. In radar system, this signal processing method is called as pulse compression [2].

The echoed waveform from the object is passed through the matched filter then receiver of the radar system can obtain both range and detection of the target in accurate manner. Matched filter is a filter that exploits the signal-to-noise ratio (SNR) at radar receiver which, in turn, maximizes the detectability of an object. Matched filters are normally used in radar receiver [2, 3] in which some reference waveform is communicated and echo is observed for mutual elements of the conveyed signal. Usually, matched filter output contains the main response as well as some side lobes. These side lobes are unwanted because they may contain the information related to the nearby target. So, side lobes should be small as much as possible in order to make the system more accurate. The AF in the radar system [3] is used to determine the range and velocity resolution through the delay–Doppler plane. It consists of time delay and Doppler shift as two measuring parameters through which range and velocity of the target can be determined, respectively.

There are numerous pulse compression methods that have been established in the literature. Phase-coded pulse compression waveform provides good range resolution than linear frequency modulated (LFM) [4, 5] because in LFM, to compensate the side lobes level, windowing is used but using windowing main lobe, width is increased which interrupts the range resolution [5]. In [6], the authors has produced a hyperchaotic coding scheme and equivalent optimal selection process to obtain the phase code signal, which shows great performance for side lobe reduction. In [7], the author has shaped a method to overawe the glitches of frequency-modulated waveform that is widening the main lobe width. Also, many techniques have been developed for the side lobe reduction in pulse compression whose results are good [8, 9]. Techniques based on binary sequence are better than polyphase pulse compression techniques because latter one is more complex and newly settled Woo filters [10, 11] are much better choices for better PSL, which uses two correlators to create a single filter in radar system.

This paper addresses the autocorrelation function for better PSL, TB product of newly derived signals is obtained from Chebyshev polynomials of different orders and a detailed analysis is carried out including AF of these polynomials through which range and velocity resolution can be determined. This paper has been prearranged as follows: The problem formulation is explained in Sect. 2, while Sect. 3 gives the description of simulations and results. Lastly, Sect. 4 concludes the work presented in this paper.

2 Problem Formulation

This pulse compression technique gets the benefits of both long-range detection as well as good range resolution capability for the pulse having small duration. The transmitted signal is represented as

$$x(t) = \alpha(t)\cos(2\pi f t + \varphi(t)) \tag{1}$$

Here, $\alpha(t)$ signifies the amplitude modulation whose amplitude varies with respect to time and $\varphi(t)$ represents an angle modulation. Angle modulation is preferred over amplitude modulation due to the advantage in noise immunity [12]. The received signal can be represented as

$$r(t) = \cos(2\pi f t + \Omega(t) - T_d) \tag{2}$$

Due to relative motion of target with radar system, the phase difference of $\Omega(t)$ exist, while T_d is the round trip time delay [3]. This reflected signal is delivered through matched filter to improve the SNR. The matched filter output gives PSL behavior of the received signal. Lower the PSL then better is the matched filter response.

Figure 1 shows the pulse compression-based radar system block diagram. Initially, a microwave source generates the continuous carrier signal which is modulated by a pulse modulator. The carrier signal is modulated in angle to get the pulse compression and then applied to transmitting antenna. The received signal has to be passed through the matched filter that gives the side lobe behavior and using signal processing or notch filtering techniques, characteristics of the target has to be observed on the radar indicator.

In order to examine signal suitability for range and velocity resolution, the AF has been observed and this is represented as

$$X(t_d, f_d) = \int_{t=-\infty}^{\infty} x(t)x^*(t - t_d)e^{j2\pi f_d t}dt \tag{3}$$

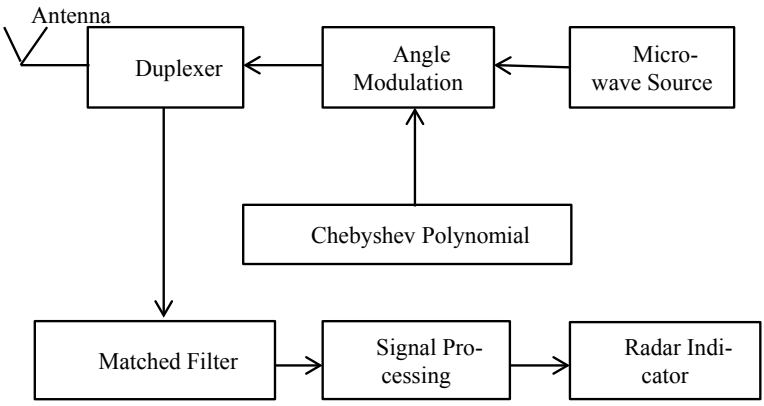


Fig. 1 Block diagram of pulse compression radar

$X(t_d, f_d)$ is an AF, t_d and f_d are the time delay and Doppler shift, respectively. AF behaves like a matched filter when there is no Doppler shifts exist and its response is at all times symmetrical around the origin [12–14]. In most cases, this function offers single central peak at the origin while the rest of energy can be spread uniformly within the delay–Doppler plane and this narrow peak which lies at the origin implies good resolution in each range and velocity [14]. At the receiving end, when filter is matched in Doppler to the reflected signal and is sampled at a delay similar to the precise range of goal, then the response of this filter will likely be maximum on the origin which gives good decision property. If this function is sampled at some distinctive delay, then the response of the filter will be deviated from the foremost and possibility of ambiguity in velocity and range may arise.

As PSL of matched filter are of major considerations in radar receiver, probably in two approaches, these side lobes can be lowered, in one windowing methods can be utilized but it increases the relative main lobe width that influences the range resolution and in another, transmit such signals which automatically give less PSL. This paper presents the second method in which Chebyshev polynomials of specific kinds are analyzed and in addition their AF are observed. So, these modified Chebyshev polynomials can be used for transmission purpose in real time applications in place of $\varphi(t)$ in Eq. 1.

3 Simulations and Results

Chebyshev polynomial introduces a trigonometric mapping method which produces a real-valued sequence having better autocorrelation property [15–17]. Here, two cases has been considered to check the PSL behavior of matched filter, in the first case, Chebyshev polynomial are analyzed, after that cycle of the polynomials are modified by multiplying the polynomials cycle with some $+1$ or -1 polarity which are generated by cycles zero crossing. Recursive equation for Chebyshev polynomial of order n can be calculated as

$$T_{n+1} = 2xT_n - T_{n-1} \tag{4}$$

Basically, two possibilities exist to change the cycle of polynomial, i.e., half-cycle and full complete cycle can be changed. For nth order, $\frac{n-1}{2}$ and $n-1$ cycles occur for full and half-cycle alteration, respectively. In this paper, both the modification has been done and their matched filter responses are analyzed. Subject to the polynomials order, different possibility exists for cycle modification. We considered that modification in cycles which gives improvement in PSL value out of all the possibilities. Subsequently, matched filter response is analyzed for these modified polynomials. To produce full-cycle alteration, odd number ordered polynomial has been considered. For even-order polynomials, some part of the waveform is shortened which modify the anticipated outcomes. While half-cycle modification are possible for both (even and odd) ordered polynomials.

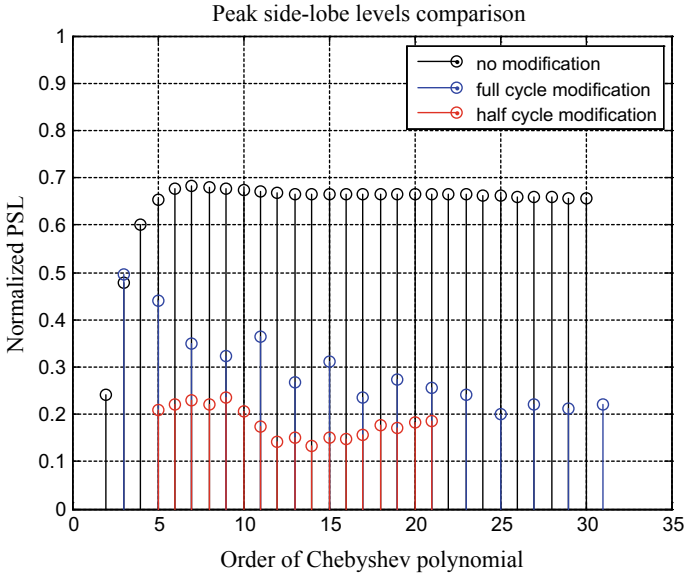


Fig. 2 Peak side lobe levels comparison

PSL behavior of ordinary Chebyshev polynomial and modified cycle's polynomials are characterized in Fig. 2. This figure signifies the matched filter response for the finest PSL level out of all the options of the different Chebyshev polynomials and it is perceived that half-cycle-altered polynomial provides better PSL than other two cases.

3.1 Performance Parameters Analysis

As PSL and main lobe width are the main parameters to observe the performance of the signal for pulse compression. PSL is calculated as

$$\mathrm{PSL(dB)} = 20\log_{10}\frac{\max i_{1\le i<N}|M(i)|}{|M(0)|} \tag{5}$$

where *M(i)* is the matched filter amplitude of the *i*th sample number for Chebyshev polynomials and *N* represents overall length of signal. These parameters are calculated for different order polynomials and isrepresented in Table 1.

In Fig. 3, first row represents the time domain representation and matched filter response for Chebyshev polynomial having order n = 15. It is observed that the PSL are not suppressed, which may result in the unwanted target detection. To suppress the PSL, adjustments in the cycle of Chebyshev polynomial is incorporated. In second

Table 1 The quantitative parameter calculation for different polynomials

Order (n)	No cycle modification		Full-cycle modification		Half-cycle modification	
	PSL (dB)	Main lobe width	PSL (dB)	Main lobe width	PSL (dB)	Main lobe width
11	−3.47	0.22	−8.77	0.25	−15.17	1.04
12	−3.51	0.20	–	–	−17.07	0.96
13	−3.53	0.18	−11.42	0.22	−16.43	0.90
14	−3.54	0.17	–	–	−17.57	0.62
15	−3.55	0.16	−10.13	0.2	−16.38	0.22

– Represents no full cycle modification exist for even order polynomials

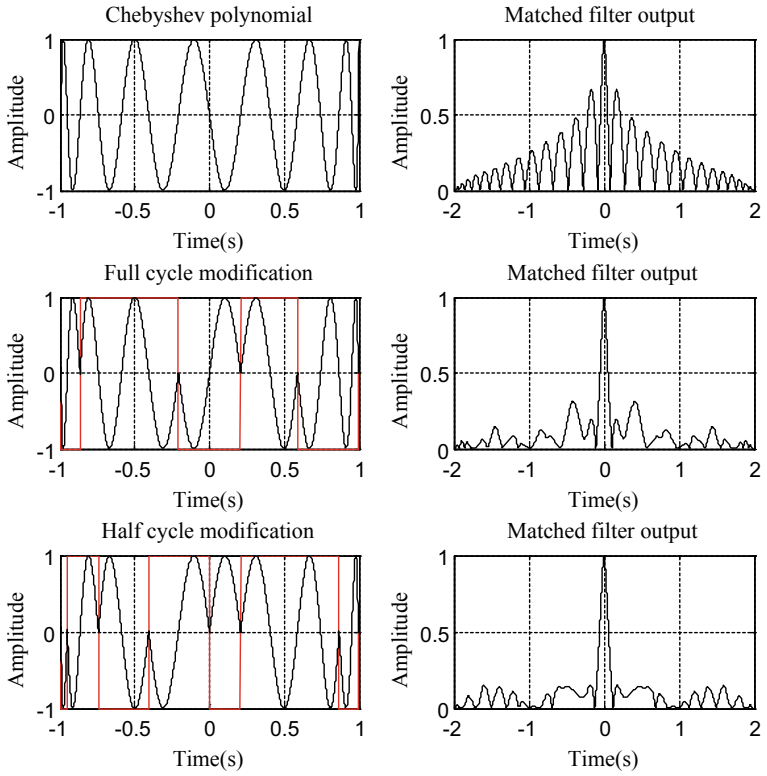


Fig. 3 Time domain and matched filter response for $n = 15$

row, full-cycle modification are done for this specific polynomial, red signal (multiplying signal) is used for polarity iteration with a view to multiply with the above Chebyshev polynomial as a way to alternate the cycles of the polynomial and this modified full-cycle polynomial-matched filter response is analyzed, it is observed that it have less PSL than simple Chebyshev polynomial. Here, $n = 15$ is taken because full- and half-cycle modifications are possible for odd number, moreover, it gives better PSL as observed in Fig. 2.

In the same manner, half-cycles are modified and matched filter response is analyzed, it is perceived that this half-cycle-modified Chebyshev polynomial have improved PSL as represented in row three of Fig. 3. Usually in pulse compression when we control one parameter, then other is disturbed [14–17]. Here also in column two, side lobes are repressed but relative main lobe width is enlarged somewhat that affects the range resolution.

3.2 Discussion with Bandwidth

The product of the signal duration and its spectral width is known as TB product. In line with Fourier concept, short-duration signal have trendy spectral contents and vice versa. Utilizing this product, we will become aware of the available bandwidth for a given channel and it will have to be more in the pulse compression for just right detection as well as for excellent resolutions properties [18–20].

In pulse compression, lengthy period signal are transmitted whose bandwidth is improved by means of an angle modulation, therefore, this transmitted lengthy signal is responsible for goal detection and increased bandwidth is liable for velocity resolution. For this motive bandwidth or spectral contents, analyses are done for different polynomials and are represented in Fig. 4. It is analyzed that the altered

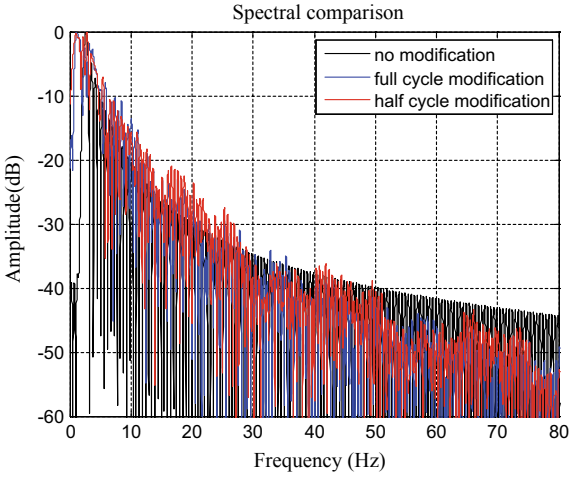


Fig. 4 Spectral comparison for $n = 15$

partial cycle polynomial has substantially broader frequency pattern than other polynomials. These frequency spectrums are got with the help of Fast Fourier Transform (FFT) [21].

3.3 Ambiguity Function Observations

In Fig. 5a–c, the AF for Chebyshev polynomials and modified polynomials are analyzed. From the simulation, it's discovered that Chebyshev polynomials delay–Doppler plane have no longer uniform energy spreading which may lead to the unwanted goal detection. Its decision isn't good for velocity and range selection, and it means these polynomials cannot be used for resolution purposes. Now, the half-cycle and full-cycle-modified Chebyshev polynomial's ambiguity evaluation are carried out and their delay–Doppler plane becomes more concentrated on the center, which tells concerning the goal characteristics with no ambiguity in velocity and range.

It is analyzed that the half-cycle-modified Chebyshev polynomials are good for range as good as velocity resolution given that their response is extra concentrated at the origin. Despite the fact that the AF offers higher shape at the zero Doppler cut, if Doppler shift exists then the side lobe level raises greatly. Further, with the use of delay and Doppler cut, range, and velocity can be determined for the designed polynomials. Also, range–Doppler coupling effects can be observed which is used for the clutter suppression in radar system.

4 Conclusion

The specified evaluation is applied to the resulting signal with appreciate to PSL, TB product, and ambiguity analysis. It is analyzed that half-cycle improved polynomials provide higher results than complete cycle-modified polynomials. Together with this, there may be somewhat improvement within the TB product and its ambiguity function is extra targeted toward the center, which gives excellent resolution properties. So, these modified polynomials will also be utilized for the sensible applications in the radar system for side lobe reduction. Limitation of this generated signal is that relative main lobe width raises slightly that impacts the velocity resolution. In summary, the procedure has been proven to be actual side lobe reduction method having usefulness in microwave radars. The extension of this work is related to the use of these generated polynomials in pulse compression radar. As pulse compression radar required the waveform, which gives less PSL values in order to avoid the target masking which is present near to main target.

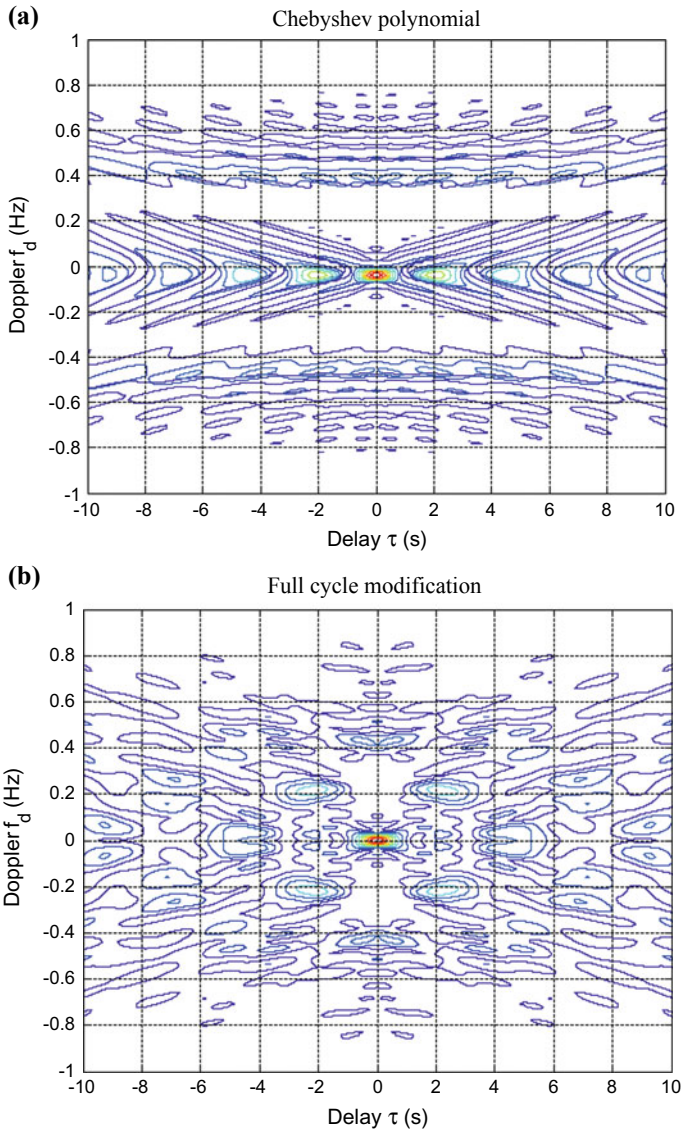


Fig. 5 **a** Ambiguity function analysis for $n = 15$ of Chebyshev polynomial, **b** ambiguity function analysis for $n = 15$ of full cycle, **c** ambiguity function analysis for $n = 15$ of half-cycle

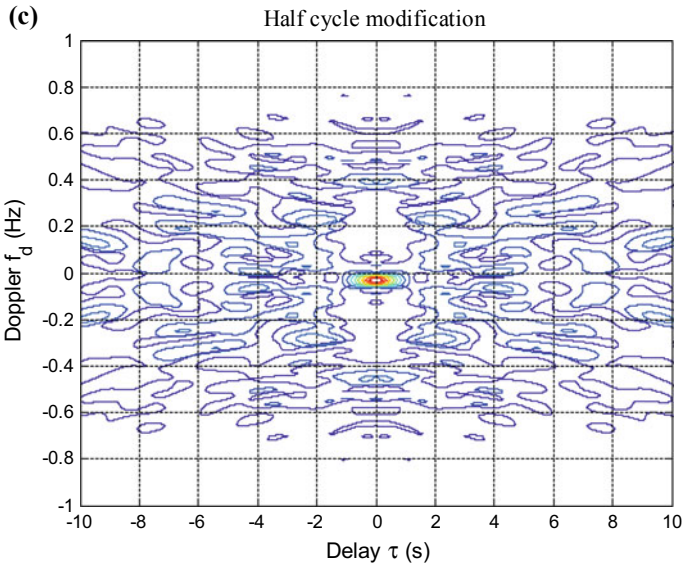


Fig. 5 (continued)

References

1. M.I. Skolnik, *Introduction to Radar System* (McGraw-Hill, New York, 2001)
2. R.M. Mahafza, *Radar Systems Analysis and Design using Matlab* (CRC Press, Boca Raton, FL, 2000)
3. A. Mark, *Richards: Fundamentals of Radar Signal Processing* (McGraw-Hill, New York, 2014)
4. B.L. Lewis, F.F. Kretschmer, Linear frequency modulation derived polyphase pulse compression codes. IEEE Trans. Aerosp. Electron. Syst. **17**, 637–641 (1982)
5. W.K. Lee, A pair of asymmetrical weighting receivers and polyphase codes for efficient aperiodic correlations. IEEE Commun. Lett. **10**, 387–389 (2006)
6. D. Yunkai, H. Yinghui, G. Xupu, Hyper chaotic logistic phase coded signal and its sidelobe suppression. IEEE Trans. Aerosp. Electron. Syst. **46**, 672–686 (2010)
7. I.C. Vizitiu, Some aspects of sidelobe reduction in pulse compression radars using NLFM signal processing. Prog. Electromagn. Res. Lett. **47**, 119–129 (2014)
8. A. Thakur, S.R. Talluri, R.K. Panigrahi, Sidelobe reduction in pulse compression having better range resolution. Comput. Electron. Eng. **74**, 520–532 (2019)
9. Y.K. Chan, M.Y. Chua, V.C. Koo, Side-lobes reduction using simple two and tri-stages nonlinear frequency modulation. Prog. Electromagn. Res. Lett. **65**, 33–52 (2009)
10. A. Thakur, S.R. Talluri, Comparative analysis on pulse compression with classical orthogonal polynomials for optimized time-bandwidth product. Ain Shams Eng. J. **9**(4), 1791–1797 (2018)
11. V. Baghel, G. Panda, Development of an efficient hybrid model for range sidelobe suppression in pulse compression radar. Aerosp. Sci. Technol. **27**, 156–162 (2013)
12. B.G. Zakeri, M.R. Zahabi, S. Alighale, Side lobes level improvement by using a new scheme used in microwave pulse compression radars. Prog. Electromagn. Res. Lett. **30**, 81–90 (2012)
13. M. Kumaria, K. Rajarajeswari, K. Krishna, Low sidelobe pattern using Woo filter concepts. AEU Int. Journal of Electron. Commun. **59**, 499–501 (2005)
14. A. Thakur, S.R. Talluri, D.S. Saini, Signal generation employing Chebyshev polynomial for pulse compression with small relative side-lobe level. in *IEEE International Conference on Sensing, Signal Processing and Security (ICSSS)* (2017) (pp. 114–118)

15. A. Thakur, S.R. Talluri, A novel pulse compression technique for side-lobe reduction using woo filter concepts. in *IEEE International Conference on Communication Signal Processing (ICCSP)* (2017) (pp. 1086–1090)
16. J.N. Blakely, N.J. Corron, Ambiguity in Range-Doppler determination using waveforms of a solvable chaotic oscillator. Signal Process. **104**, 136–142 (2014)
17. F.Y. Lin, J.M. Liu, Ambiguity functions of laser-based chaotic radar. IEEE J. Quantum Electron. **40**, 1732–1738 (2004)
18. A. Gil, J. Segura, N. Temme, *Numerical Methods for Special Functions*. Siam (2007)
19. W. Diego, D. Cristallini, Optimal trajectories for range resolution improvement in multi-PCL SAR. AEU-Int. J. Electron. Commun. **73**, 173–182 (2017)
20. Q.A. Farhan, T.F. Adly, Doppler tolerant and detection capable polyphase code sets. IEEE Trans. Aerosp. Electron. Syst. **51**(2), 1123–1135 (2015)
21. T.S. Chihara, *An Introduction to Orthogonal Polynomials* (Dover, Gordon and Breach, 2011)

Wideband MIMO Antenna with Diverse Polarization and Radiation Pattern for 5G Applications



Dinesh Sharma and E. Kusuma Kumari

1 Introduction

With the advent of the 5G communications, the need for the communication systems with high data transfer rate will be essential and to serve the need of the high data rates of future wireless communication systems, the need for antennas with wide bandwidth has become a necessity. Any frequency in the range of 400–90 GHz can be used for the 5G applications but for the initial commercial purpose, frequency band less than 6 GHz is been given priority all over the world, especially, the frequency of 3.5 GHz [1]. This particular frequency has got keen interest as the present-day communications systems available with the current service providers can support this frequency with some minor modifications and is useful for the implementation of the 5G services in the early stages.

Present-day communications systems require multimode and multi-band antennas to satisfy the multifunctional requirements of the users. For an antenna to serve multiple applications techniques of polarization diversity and radiation pattern diversity have come into limelight. But the major limitation of the MIMO antennas is the mutual coupling effect, which can be eliminated by using some decoupling structures [2]. In [3], a space coupling technique has been proposed to reduce the coupling but it will increase the size of the antenna. In [4], a separate built-in decoupling structure has been used in the antenna to reduce the coupling but with the introduction of the additional structure, the design complexity has been increased. Similarly, different techniques were been proposed by different researchers in [5–13] but all of them have either antenna size increasing problem or increase in design complexity.

D. Sharma
ECE Department, CCET, Chandigarh, India
e-mail: sharma82dinesh@gmail.com

E. Kusuma Kumari (✉)
ECE Department, Sri Vasavi Engineering College, Tadepalligudem, Andhra Pradesh, India
e-mail: kusuma.madhu.ece@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_68

To reduce the mutual coupling in between the MIMO antenna, in [14], a resonator has been placed in between the adjacent antennas which will isolate the antenna elements from one another. A similar work has been carried out in [15] where a resonator is been used which is made up of microstrip lines with slots, which is used to reduce the coupling in antenna array. Passive resonators are been used in [16] for reduction of the mutual coupling in printed inverted F antenna [17, 18] show that the use of defective ground structures and additional resonating elements in terms of parasitic patches can be used for reducing the mutual coupling by increasing the isolation in between the antenna elements. In a similar way, we can also use electromagnetic bandgap structures for reduction of mutual coupling [19–21] but all the mentioned techniques presented here are having an involvement of introducing additional conducting elements into the antenna, which will increase the complexity of the antenna construction.

Proposed is a 2 × 2 MIMO antenna which is having a wide bandwidth along with the capabilities of polarization diversity and radiation pattern diversity along with a simple isolator at the antenna center which will be useful for the reduction of the mutual coupling in between the antenna elements. Proposed antenna resonates at the frequency range of 3.2–6 GHz, which has been identified as the operating frequency for the upcoming 5G communications.

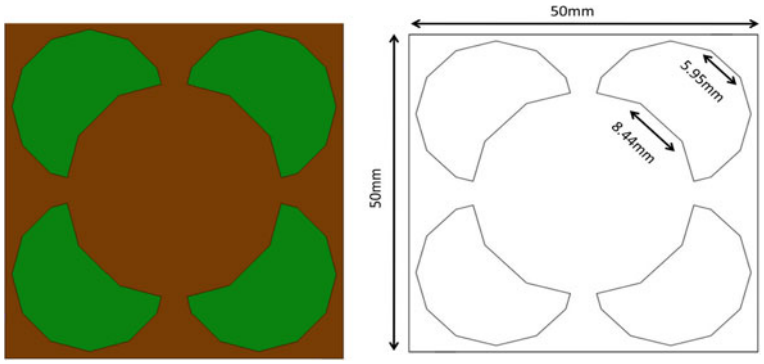
2 Development of Proposed Antenna

A four-port 2 × 2 antenna array has been proposed for the 5G frequency band. The proposed antenna is having wide bandwidth, polarization, and radiation pattern diversity. FR4 glass epoxy substrate has been used to design the antenna. The overall dimension of the antenna will be 50 mm × 50 mm × 1.6 mm. Microstrip line feed has been used to excite the antenna elements individually. The fundamental structure of the radiating element will be a rectangular sheet with dodecagon slots in four quadrants with a dodecagon isolator at the center of the patch as shown in Fig. 1.

The feed is provided at the other side of the substrate as shown in Fig. 2. Initially, a plane rectangle copper patch has been considered in which four dodecagon slots are made such that each slot will be centered at the center point of one of the four quadrants of the rectangular patch. Then, one more dodecagon patch is placed at the center of the slotted rectangle patch which gives us the proposed design of the final patch.

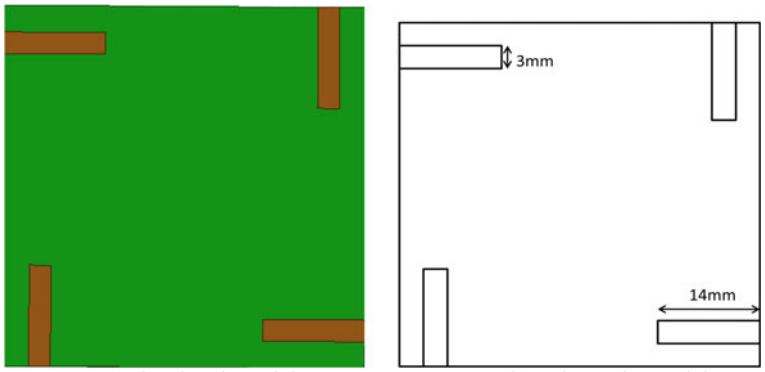
3 Results and Discussion

The excitation will be provided to the feed element and the power is transferred to the dodecagon slot above the feed line in the opposite side of the substrate. As there is no direct contact between the radiating element and feed line, radiation performance of



(a) Simulated Model (b) Schematic Model

Fig. 1 Proposed antenna front view



(a) Simulated Model (b) Schematic Model

Fig. 2 Proposed antenna rare view

the antenna will be high and wide bandwidth can be achieved with proper impedance matching between the feed line and slot. The impedance characteristics of the antenna are shown in Figs. 3 and 4.

Figure 3 represents the return loss plot of the antenna and Fig. 4 represents the VSWR plot of the antenna. From Fig. 3, we can observe that the antenna is radiating from 3.2 to 6 GHz and return loss value is less than −10 dB for the entire frequency band. Similarly, from Fig. 4, we can see that the VSWR value is less than 2 for the entire operating frequency range. Both the plots ensure that the antenna is having good impedance matching.

The gain of the proposed antenna element at two intermediate frequencies is shown in Figs. 5 and 6. The two frequencies selected were 3.5 and 5.3 GHz. Observed a gain of 2.97 dB at 3.5 GHz and a gain of 3.93 dB at 5.3 GHz. The frequencies were selected depending upon the frequency allocation for the 5G communication in near future.

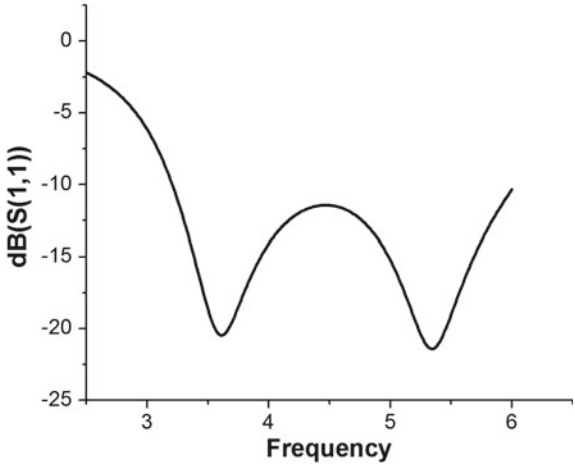


Fig. 3 Return loss

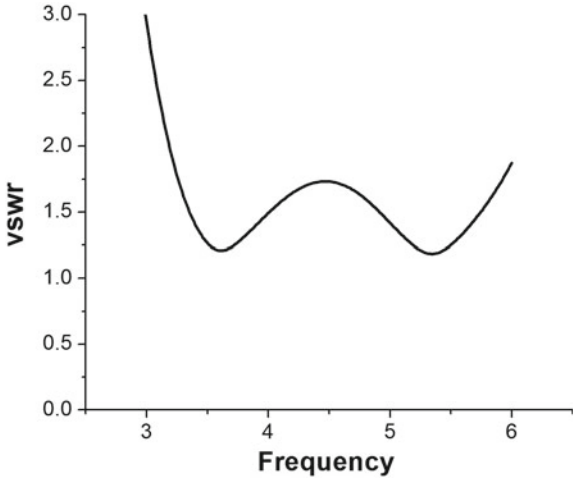


Fig. 4 VSWR

The polarization of the antenna is very important in the present communication systems and a lot of care has been taken while designing the antenna in this aspect. Our proposed antenna is having polarization diversity. If anyone of the antenna elements is excited alone, then the polarization of the electromagnetic waves generated will be linear polarization which can be seen from the plots 7 and 8 below. The axial ratio value of the antenna with single antenna element excitation is given as 16 and 8 dB for the frequencies 3.5 and 5.3 GHz, which represents linear polarization (Figs. 7 and 8).

But if two or more than two antenna elements were excited at the same time, then the resultant field generated is having circular polarization which is depicted in Figs. 9, 10. An axial ratio value of 2.55 and 1.35 dB has been observed with two or

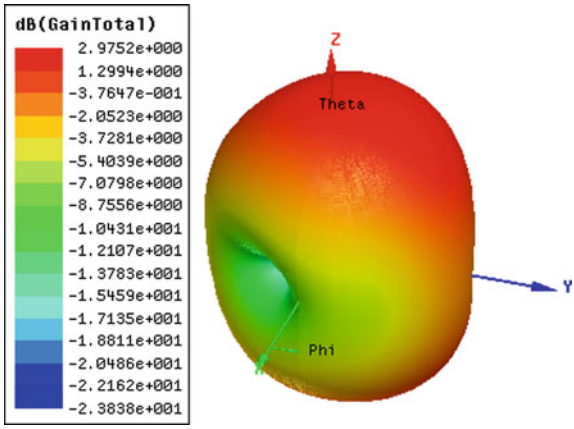


Fig. 5 Gain at 3.5 GHz

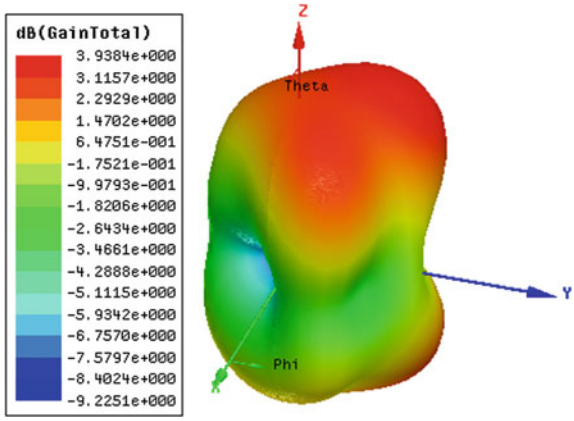


Fig. 6 Gain at 5.3 GHz

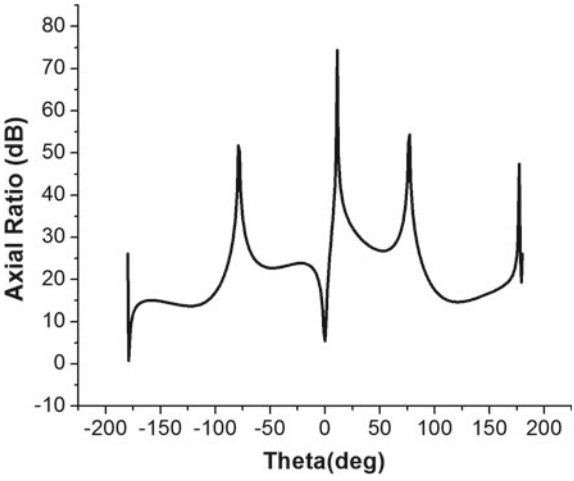


Fig. 7 Axial ratio at 3.5 GHz

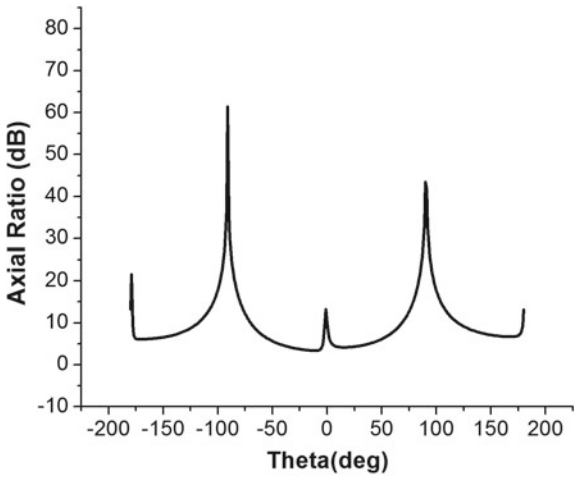


Fig. 8 Axial ratio at 5.3 GHz

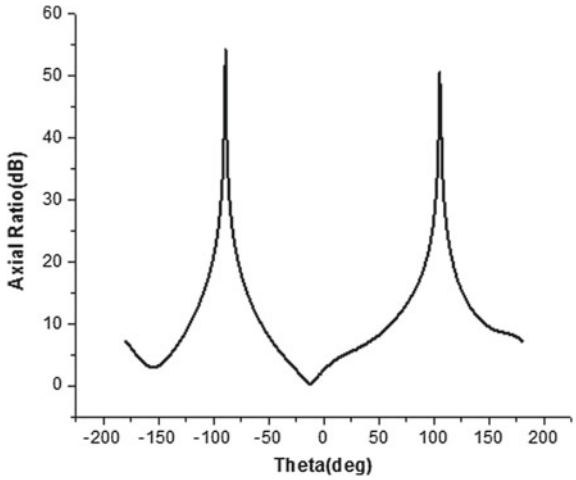


Fig. 9 Axial ratio at 3.5 GHz

more than two antenna elements excitation, which represents circular polarization at the frequencies of 3.5 and 5.3 GHz. This is possible because of the dodecagon slots, which act as radiating elements and their placement in the antenna with respect to each other and the isolator used in the center of the patch, which aligns the current directions orthogonally for getting circular polarization.

Figures 11and 12 depict the radiation pattern of the proposed MIMO antenna. Both the elevation plane and azimuthal plane patterns were studied for the different modes of operation of the MIMO antenna. By exciting different antenna elements, the coverage area of the proposed antenna can be varied such that the antenna can cover a large area. The same has been shown in the below figures.

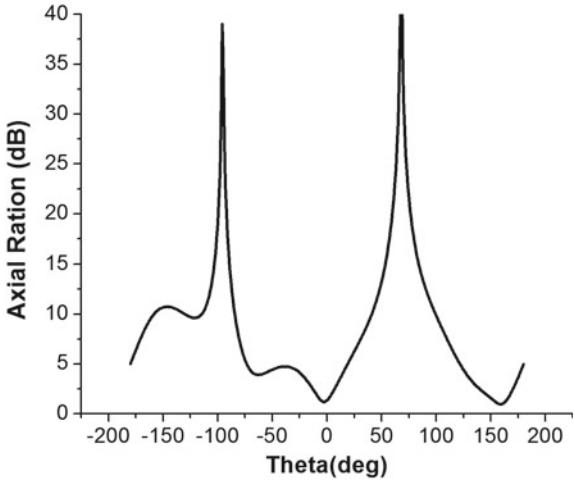


Fig. 10 Axial ratio at 5.3 GHz

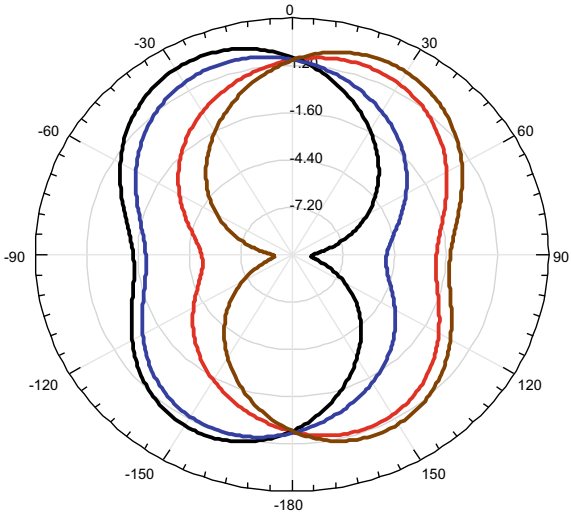


Fig. 11 Elevation plane patterns

Figures 13 and 14 show the current distributions of the antenna with single antenna element excitation and two antenna element excitation, which are diagonal to each other.

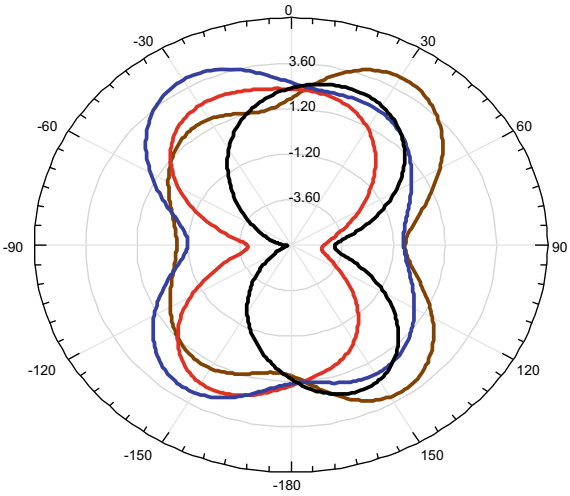


Fig. 12 Azimuthal plane patterns

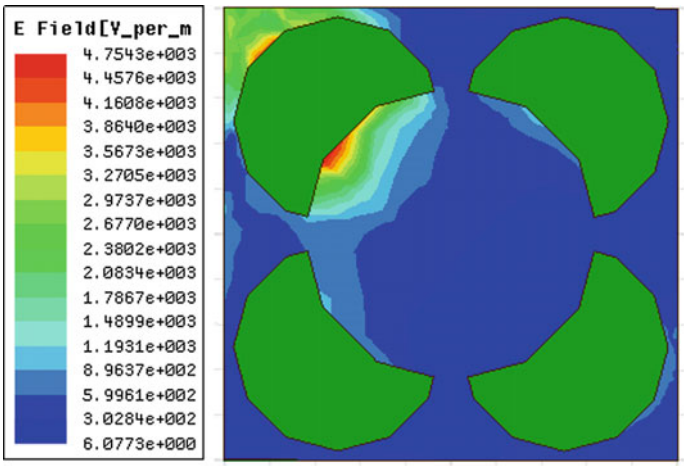


Fig. 13 Current distribution with single antenna excitation

We have observed earlier that when a single antenna element is excited the proposed antenna is having linear polarization but when two orthogonal elements were excited, we got circular polarization. The same can be depicted from the current distributions shown in the below figures.

One of the major limitations for MIMO antennas is the mutual coupling effect in between the antenna elements. Here, in the proposed antenna design, a dodecagon isolator has been placed in the center of the four elements which acts as an isolator in between the antenna elements.

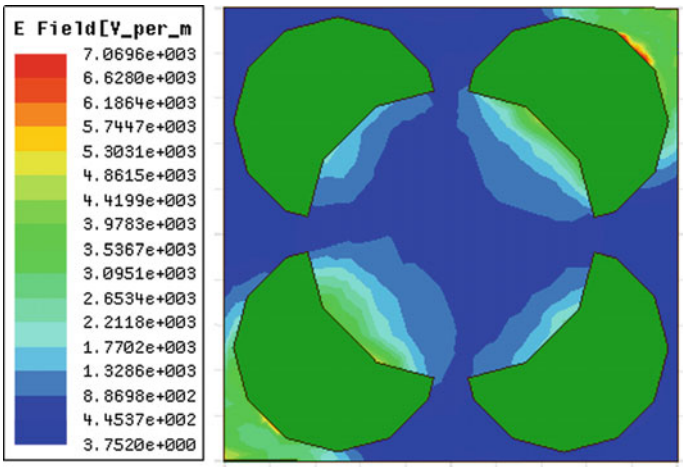
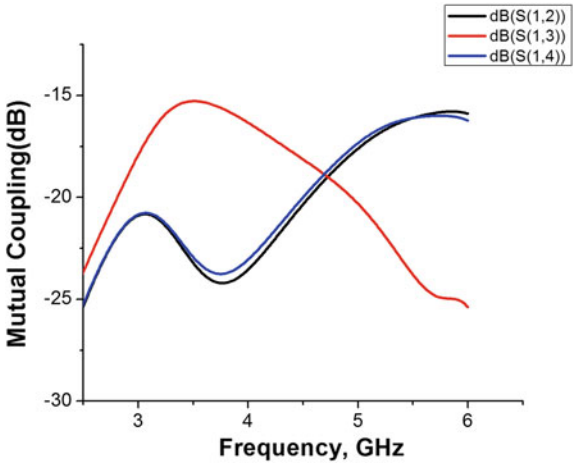


Fig. 14 Current distribution with orthogonal antenna excitations

From Fig. 15 above we can observe that the maximum mutual coupling effect is around −16 dB which is observed at the higher frequency range of the bandwidth and minimum coupling is around −24 dB which is observed at the low-frequency range of the operating bandwidth of the antenna.

Fig. 15 S_{12}, S_{13}, S_{14} plots



4 Conclusion

In this letter, a 2 × 2 MIMO antenna for the 5G application has been proposed and investigated. The proposed antenna is having polarization diversity and radiation pattern diversity which is possible by exciting different antenna elements. Radiating element is a dodecagon slot, which is radiated by a simple stripline feed which helped in achieving wide bandwidth. Mutual coupling effect, which is a major limitation of the MIMO antenna has also been addressed properly by placing a dodecagon isolator at the center of the antenna elements. From the simulation results, we can say that the proposed antenna is best suited for the 5G applications which require high data transfer rate.

References

1. Nokia white paper, '5G deployment below 6 GHz'. https://resources.ext.nokia.com/asset/201315. Accessed Aug 2017
2. S. Rajkumar, N.V. Sivaraman, S. Murali et al., Heptaband swastika arm antenna for MIMO applications. Microw. Antennas Propag. **11**(9), 1255–1261 (2017). https://doi.org/10.1049/iet-map.2016.1098
3. A. Ramachandran, S. Mathew, V. Rajan et al., A compact triband quad-element MIMO antenna using SRR ring for high isolation. Antennas Wirel. Propag. Lett. **16**, 1409–1412 (2017). https://doi.org/10.1109/LAWP.2016.2640305
4. S.M. Wang, L.T. Whang, C.J. Lee et al., MIMO antenna design with built-in decoupling mechanism for WLAN dual-band applications. Electron. Lett. **51**(13), 966–968 (2015). https://doi.org/10.1049/el.2014.4352
5. C.A. Balanis, *Advanced Engineering Electromagnetics* (Wiley, New York, 2012)
6. M. Sekhar, P. Siddaiah, Triple frequency circular patch antenna. in *2014 IEEE International Conference on Computational Intelligence and Computing Research*, Park College of Engineering and Technology, ISBN: 978-1-4799-1594-1
7. E. Kusuma Kumari, A.N.V. Ravi Kumar, Wideband high-gain circularly polarized planar antenna array for L band radar. in *2017 IEEE International Conference on Computational Intelligence and Computing Research* (Tamilnadu College of Engineering), ISBN: 978-1-5090-6620-9
8. E. Kusuma Kumari, A.N.V. Ravi Kumar, Development of an L band beam steering cuboid antenna array. in *2017 IEEE International Conference on Computational Intelligence and Computing Research* (Tamilnadu College of Engineering), ISBN: 978-1-5090-6620-9
9. S. Sreenu, V.S. Rao, Stacked microstrip antenna for global positioning system. in *2017 IEEE International Conference on Computational Intelligence And Computing Research* (Tamilnadu College of Engineering), ISBN: 978-1-5090-6620-9
10. N.A. Rao, S. Kanapala, Wideband circular polarized binomial antenna array for L-band radar. in ed. by G. Panda, S. Satapathy, B. Biswal, R. Bansal. *Microelectronics, Electromagnetics and Telecommunications*. Lecture Notes in Electrical Engineering, vol. 521 (Springer, Singapore)
11. S. Kanapala, N.A. Rao, Beam steering cuboid antenna array for L Band RADAR. in ed. by G. Panda, S. Satapathy, B. Biswal, R. Bansal. *Microelectronics, Electromagnetics and Telecommunications*. Lecture Notes in Electrical Engineering, vol. 521 (Springer, Singapore)
12. K. Sehgal et al., Recent advances in networks and data security survey on various mobile operating systems. in *Lecture Notes in Networks and Systems International Conference on Innovative Computing and Communications Proceedings of ICICC 2018*, vol. 1

13. M.P. Agrawal et al., TRAX: smart mobile application to improve the safety and security of automobile vehicles. in *Lecture Notes in Networks and Systems International Conference on Innovative Computing and Communications Proceedings of ICICC 2018*, vol. 1
14. L.Y. Zhao, K.L. Wu, A dual-band coupled resonator decoupling network for two coupled antennas. IEEE Trans. Antennas Propag. **63**(7), 2843–2850 (2015)
15. M.G.N. Alsath, M. Kanagasabai, B. Balasubramanian, Implementation of slotted meander-line resonators for isolation enhancement in microstrip patch antenna arrays. IEEE Antennas Wirel. Propag. Lett. **12**, 15–18 (2013)
16. L. Minz, R. Garg, Reduction of mutual coupling between closely spaced PIFAs. Electron. Lett. **46**(6), 392–394 (2010)
17. W. Han, X. Zhou, J. Ouyang, Y. Li, R. Long, F. Yang, A six-port MIMO antenna system with high isolation for 5-GHz WLAN access points. IEEE Antennas Wirel. Propag. Lett. **13**, 880–883 (2014)
18. C.C. Hsu, K.H. Lin, H.L. Su, Implementation of broadband isolator using metamaterial-inspired resonators and a T-shaped branch for MIMO antennas. IEEE Trans. Antennas Propag. **59**(10), 3936–3939 (2011)
19. Y.F. Lu, Y.C. Lin, Electromagnetic band-gap based corrugated structures for reducing mutual coupling of compact 60 GHz cavity-backed antenna arrays in low temperature co-fired ceramics. IET Microw. Antennas Propag. **7**(9), 754–759 (2013)
20. J.Y. Lee, S.H. Kim, J.H. Jang, Reduction of mutual coupling in planar multiple antenna by using 1-D EBG and SRR structures. IEEE Trans. Antennas Propag. **63**(9), 4194–4198 (2015)
21. J.S. Sun, H.S. Fang, P.Y. Lin, C.S. Chuang, Triple-band MIMO antenna for mobile wireless applications. IEEE Antennas and Wirel. Propag. Lett. **15**, 500–503 (2016)

Study the Effect of Dielectric Constant on Notch-Loaded U-Shaped Microstrip Patch Antenna

Saurabh Singh, Sudhanshu Verma and Rishabh Kumar Baudh

1 Introduction

In the recent years, wireless communication is essential for everybody in routine life like mobile, Wi-Fi, and WiMax. This advancement in a wireless system is due to the integration of microwave circuit components with the same circuit board such as resistor, inductor, capacitor and microstrip patch antenna. For better performance of wireless communication, wideband antenna will be required to design [1–3]. The dielectric materials change the parameter of antenna which affects the performance characteristics of antenna in the manufacturing process. In antenna especially in integrated antenna dielectric constant lessens the superficial losses. It improves the reflection coefficient of the signal. For the analysis of the performance of the proposed antenna, the glass epoxy, RT/Duroid 5880, and textile material are used that have the dielectric constant [4–6]. The textile material is flexible, compact, and light. Textile antennas are used in portable health care, tracking, navigation, sports monitoring, etc. [7–10].

For enhancing the bandwidth of the proposed antenna, notch on U-shaped microstrip patch antenna and microstrip feed is used. Comparative analysis of reflection coefficient, gain, efficiency, radiation pattern, and directivity is done.

S. Singh · S. Verma
ECE Department, MMMUT, Gorakhpur, UP, India
e-mail: singh.saurabh047@gmail.com

S. Verma
e-mail: svermaecec@gmail.com

R. K. Baudh (✉)
ECE Department, BIET, Jhansi, UP, India
e-mail: rbaudha@yahoo.in

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_69

2 Geometrical Specifications Estimate Dimensions of MPA

The numerical equation [11] is utilized to compute the elements of a ground plane and microstrip patch as length and width.

The equation of determining the designed specification of the geometry given as

$$W = \left(\frac{c}{2f_r}\right)\left(\frac{\varepsilon_r + 1}{2}\right)^{-1/2}$$

where $c = 3 \times 10^8$ ms^{-1}, ε_r = dielectric constant, $f_r = 2.45$ GHz

Specification of the effective dielectric is given as

$$\varepsilon_{\text{eff}} = \left(\frac{\varepsilon_r + 1}{2}\right) + \left(\frac{\varepsilon_r - 1}{2}\right)\left(1 + \frac{12h}{W}\right)^{-1/2}$$

Wherever: h as a height of the substrate

Specification of extension in length given as

$$\Delta L = 0.412h\left(\frac{\varepsilon_{\text{eff}} + 0.3}{\varepsilon_{\text{eff}} - 0.258}\right)\left(\frac{\left(\frac{W}{h}\right) + 0.264}{\left(\frac{W}{h}\right) + 0.8}\right)$$

The specification of determining the length of patch is given as

$$L\left(\frac{c}{2f_r\sqrt{\varepsilon_{\text{eff}}}}\right) - 2\Delta L$$

3 Antenna

For calculation of dimension several parameters, for example, designing frequency (f_r), dielectric constant of the material (ε_r), thickness of the material and loss tangent and microstrip line used as feed into patch are required. The antenna specifications that are required are illustrated in Table 1.

Table 1 Design specification

S. no.	Antenna specification	Glass epoxy	RT/Duroid 5880	Textile
1	Resonant frequency (f_r) (GHz)	2.45	2.45	2.45
2	Thickness (h) (mm)	1.6	1.6	1.6
3	Dielectric constant (ε_r)	4.4	2.2	1.7
4	Loss tangent (tan δ)	0.0013	0.0009	0.025

4 Designing Process

Using the above equations and design specification, the dimensions of the geometry is calculated. In designing of proposed geometry, first form a rectangular patch with calculated dimensions of patch and made a U-shaped design on patch by loading notch. Then create a notch of length L1 and width W1 symmetrically on both sidearm of U-shape design near the microstrip feed of 50 Ω. U-shaped design has been modified by increasing notch width and reducing notch length stepwise. By doing this procedure, the bandwidth has been enhanced and resonant peak shifted towards the lower frequency side. The notches are symmetrically designed, and different dimensional characteristics of the proposed antenna are given in Table 2.

A 50 Ω feed used at the middle of the extent of the geometry. The state of feed is used on the real axis from the origin.

The geometry of calculated notch loaded U-shaped MPA for all-dielectric material is shown in Fig. 1.

5 Comparative Analysis and Discussion

In this paper, the notch-loaded U-Shaped patch antenna is designed on different types of material such as glass epoxy, RT/Duroid 5880 and textile. Due to different dielectric constant characteristics, the performance of the antenna has been varied. Table 3 shows the comparative analysis of the characteristics of the antenna on this type of material.

Table 2 Design specifications

S. no.	Dimension	Glass epoxy data (mm)	RT/Duroid 5880 data (mm)	Textile data (mm)
1	Ground length (L_g)	38.4	50.1	55.4
2	Ground width (W_g)	47.4	58	62.3
3	Patch length (L_p)	28.8	40.5	45.8
4	Patch width (W_p)	37.2	48.4	52.7
5	L1	2.55	3.57	4
6	W1	10.14	13.18	14.4
7	L2	4.78	6.7	7.6
8	W2	14.5	18.85	20.6
9	L3	1.91	2.7	3
10	W3	2.17	2.8	3
11	Ls	3	3	3
12	s	4.8	4.8	4.8

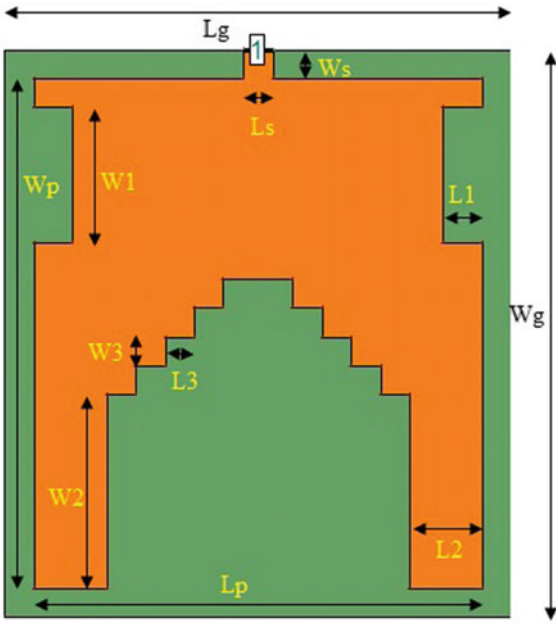


Fig. 1 Design of projected antenna

Table 3 Comparative analysis of characteristics of the antenna on different types of material

S. no.	Material	Bandwidth (MHz)	Fractional bandwidth (%)	Directivity (dBi)	Gain (dBi)
1	Glass epoxy	321	12.3	4.353	4.102
2	RT/Duroid 5880	489	20	4.919	4.871
3	Textile	962	37.68	5.059	4.904

Comparative analysis of bandwidth of notch-loaded U-Shaped microstrip antenna on different materials is shown in Fig. 2. Comparative analysis of directivity of notch-loaded U-Shaped microstrip antenna on different materials is shown below.

Comparative analysis of gain of notch-loaded U-Shaped microstrip antenna on different materials is shown in Fig. 4.

From Fig. 2, it is noted that the bandwidth of an introduced antenna is increased as dielectric constant decreased. It is more recognized that the resonance frequency of the antenna is 2.64 GHz, 2.41 GHz, and 2.29 GHz using glass epoxy, RT/Duroid 5880, and textile material, respectively, which is shifted to the left towards lower frequency as dielectric constant decreased. From Figs. (3 to 4), the directivity of the projected antenna is increased as dielectric constant decreased, the gain of proposed antenna is increased as dielectric constant decreased respectively. The overall discussion is that by enhancing the dielectric constant the parameters of the antenna is decrease and vice versa.

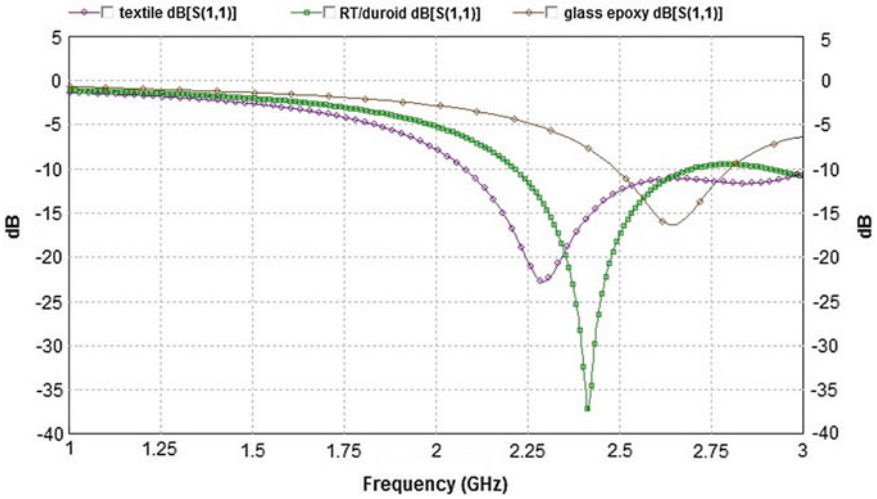


Fig. 2 Comparative BW impedance of the projected antenna

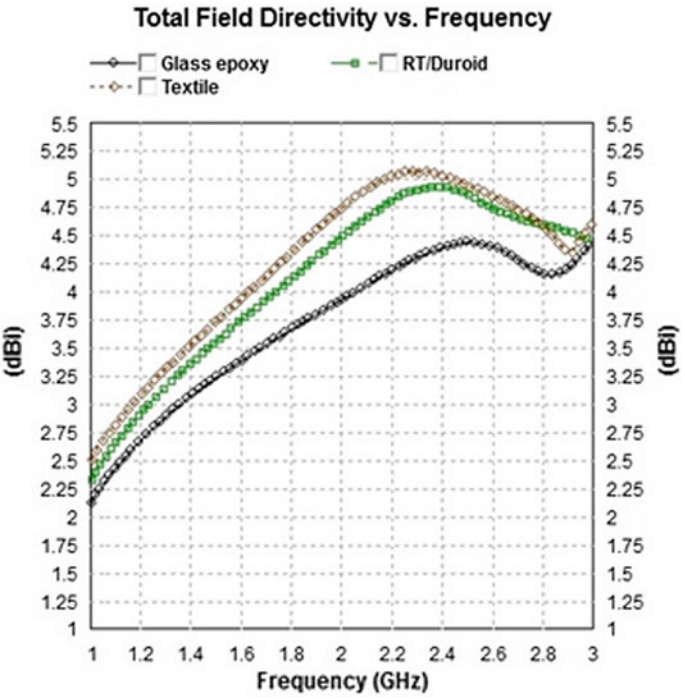


Fig. 3 Directivity of projected antenna

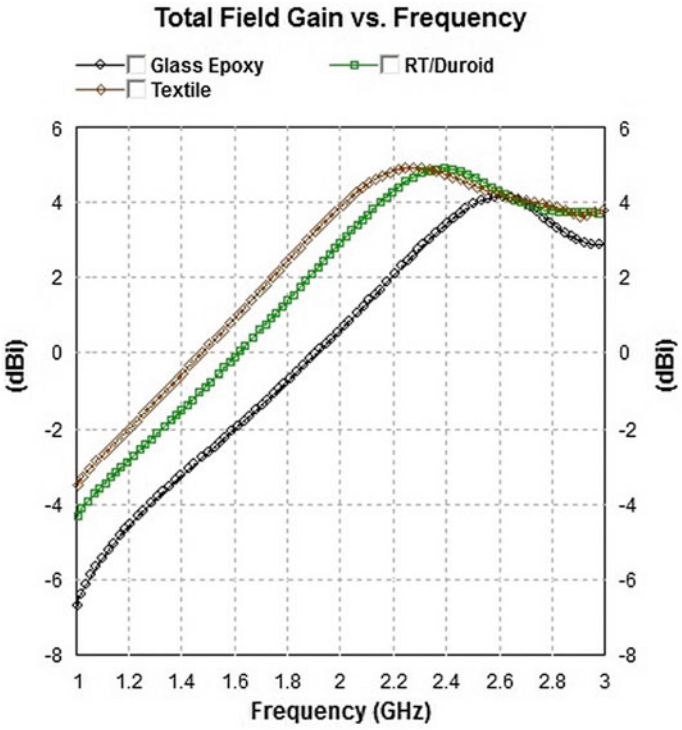


Fig. 4 Total gain impedance

6 Conclusion

A comparison between the different dielectric materials with different dielectric constant has been studied. It noted that the bandwidth of the antenna is increased as a dielectric constant has been decreased. It observed that operating frequency shifted to the left towards lower frequency as dielectric constant decreased. The gain of the antenna is increased as a dielectric constant has been decreased and the directivity of the antenna is also increased as a dielectric constant has been decreased. With the above observation, it can be concluded that by increasing the dielectric constant the parameters of antenna are decreases and vice versa. The proposed studied antenna can be utilized in the wireless communication system.

References

1. M.V. Rooyen, J.W. Odendaal, J. Joubert, High-gain directional antenna for WLAN and WiMAX applications. IEEE Antennas Wirel. Propag. Lett. **16**, 286–289 (2017)
2. Y.J. Chou, G.S. Lin, J.F. Chen, L.S. Chen, M.P. Houng, Design of GSM/LTE multiband application for mobile phone antennas. Electron. Lett. **51**(17), 1304–1306 (2015)
3. N. Herscovivi, N.H. Nashua, E. Dziadek, Omni directional antennas for wireless communication. in *Antennas and Propagation Society International Symposium* (1999) (pp. 556–559)
4. G.B. Waghmare, A.J. Nadaf, P.M. Korake, M.K. Bhandarkar, Comparative analysis of FR4 and RT-duroid materials antenna for wireless application. Int. J. Ser. Eng. Sci. (IJSES) **2**(2) (2016)
5. M. Klemm, G. Troester, Textile UWB antennas for wireless body area networks. IEEE Trans. Antennas Propag. **54**(11), 3192–3197 (2006)
6. P.B. Samal, P.J. Soh, G.A.E. Vandenbosch, UWB all-textile antenna with full ground plane for off-body WBAN communications. IEEE Trans. Antennas Propag. **62**(1), 102–108 (2014)
7. P. Salonen, Y. Rahmat-Samii, K. Kamardin, P.S. Hall, Wearable antennas: advances in the design characteristics and application. *Antenna and Propagation for Body-Centric Wireless Communication*, 2nd edn. (2012) (pp. 209–270)
8. Z. Wang, L. Zhang, D. Psychoudakis, J.L. Volakis, Flexible textile antennas for body-worn communication. in *IEEE International Workshop on Antenna Technology (iWAT)*, Tucson, USA, 5–7 Mar 2012 (pp. 205–208)
9. N. Amaro, C. Mendas, P. Pinho, Bending effects on a textile microstrip antenna. in *IEEE International Symposium on Antennas and Propagation (APSURSI)*, Spokane, USA, 3–8 July 2011 (pp. 282–285)
10. I. Locher, M. Klamm, T. Kirstein, G. Trster, Design and characterization of purely textile patch antennas. IEEE Trans. Adv. Packag. **29**(4), 777–788 (2006)
11. C.A. Balanis, *Antenna Theory-Analysis and Design*, 3rd edn. (Wiley, INC., Publication)

Triple-Band Reconstruct Circular Aperture Loaded CDRA Deploying $HE_{11\delta}$ and $HE_{21\delta}$ Modes for Wireless Applications



Ajay Kumar Dwivedi, Anand Sharma and Ashutosh Kumar Singh

1 Introduction

In today's contemporary world of radio communication, Dielectric resonator as an antenna is predominately considered and implemented in the past decades because of its variant properties as compared to other planar structured antennas such as minimal metallic loses, outspread impedance bandwidth, high radiation efficiency, and ability to get excited with different arrangements [1]. The researchers S. A. Long, M. W. McAllister and L. C. Shen initiated the capability of DR as a radiator way back in 1983 [2]. Referring to open literature, it reveals that the different DRA structures are accessible (hemispherical, cylindrical, and rectangular) but the most calibrated one is the cylindrical shape owing to the fact that its surface area is compact, its ease of manufacturing and commercially availability [3]. CDRA has the potential to bear three nonidentical mode field patterns (TE_{mnp}, TM_{mnp}, HE_{mnp}) thereby accomplishing the distributed radiation pattern [4]. In the contemporary field of wireless communication, the use of dual/multiband antenna along with circular polarization characteristics has been drawing the attention of the researchers. There are valid reasons to use such antenna: (i) a single antenna can well be deployed for different applications (ii) capacity to overpower the interferences in the signal (iii) reduces the direction, dependency between transmitter and receiver [1]. In the current scenario various techniques are implemented to acquire dual band characteristics

A. K. Dwivedi (✉) · A. K. Singh
Department of Electronics & Communication Engineering, Indian Institute of Information Technology, Allahabad, India
e-mail: rse2018505@iiita.ac.in

A. K. Singh
e-mail: ashutosh_singh@iiita.ac.in

A. Sharma
Department of Electronics Engineering, Rajkiya Engineering College, Sonbhadra, India
e-mail: anandsharma2312@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_70

(linearly/circular polarization) (i) with the use of parasitic element loaded in the DRA [5] (ii) combination of DRA with other resonating element (hybrid DRA) [6, 7], and (iii) higher order mode generation along with the elementary mode of DRA [8]. The third approach is utterly constructive but at the same time generation of higher modes in DRA is quite difficult. The research team of D. Guha had done remarkable work in the area of above-mentioned techniques [9, 10]. In this manuscript, cylindrical DRA is excited with ring shape patch along with four different segments etched out of the copper surface which in turn generates three different frequency bands for wireless applications. First (2.14 GHz) is due to the slot at the ground plane, second (3.68 GHz) and third resonant peak (4.76 GHz) is due to $HE_{11\delta}$ and $HE_{21\delta}$ mode in CDRA, respectively. The next Sections describe the antenna geometry and antenna analysis. Sections 4 and 5 explain the final results and the final outcome of it.

2 Antenna Geometry

This section discusses the geometry and dimensions of the antenna proposed that are depicted in Fig. 1. Figure 1a shows the feeding structure of the proposed antenna

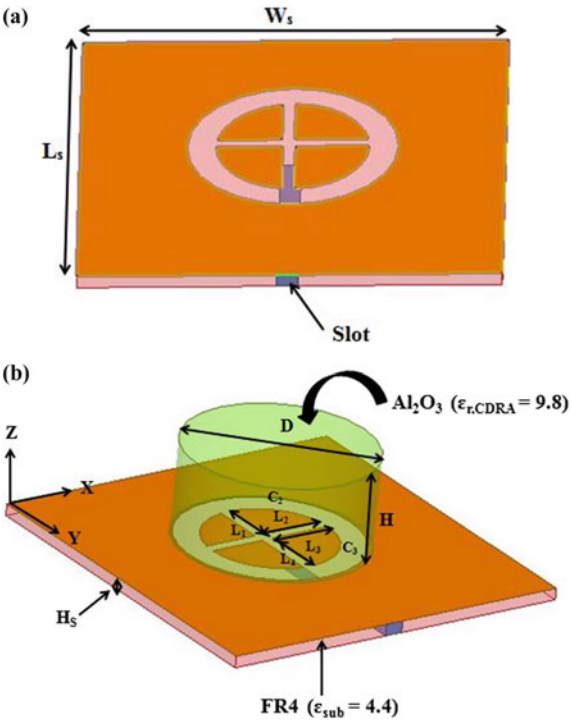


Fig. 1 Profile of the proposed antenna. **a** Feeding procedure, **b** side view

Table 1 Parameters measurements

S.no.	Parameters	Dimensions	Material used
1.	Substrate	$L_S \times W_S \times H_S$ $50 \times 50 \times 1.6$ mm^3	FR4 ($\varepsilon_r = 4.4$) tanδ = 0.025
2.	Slot line	$Ws \times Ls$ 2.5×26 mm^2	Copper
3.	Cylindrical DRA	D = 26 mm Height = 13 mm	Alumina ($\varepsilon_r = 9.8$) tanδ = 0.001
4.	Outer circular ring	D = 25 mm	Copper
5.	Inner circular ring	D = 18 mm	Copper
6.	Inverted *L*-strips and circular cones	$L_1 = 7.49$ mm $L_2 = 8.67$ mm $L_3 = 8.7$ mm $L_4 = 8.99$ mm $C_2 = 12.43$ mm $C_3 = 13.87$ mm	Copper

and the side view is shown in Fig. 1b. The measurements of the presented prototype are tabulated in Table 1.

3 Antenna Analysis

The presented antenna has been analyzed in detail using ANSOFT HFSS-16 electromagnetic solver. Figure 2 depicts the near electric field scattering on CDRA at 2.14 GHz; 3.68 GHz and 4.76 GHz, respectively. Figure 2 illustrates the impact of CDRA and the origination of higher order all along with the basic mode. The field at the lower frequency is due to the using of slot aperture, $HE_{11\delta}$ and $HE_{21\delta}$ originates due to CDRA.

4 Results and Discussion

The optimized results are discussed in this section. Figure 3 shows the$|S_{11}|$ over the wide range of frequency curve in the figure the variation has been shown by considering two different conditions (i) model with CDRA (ii) model without CDRA, it is clearly perceive from the figure that using CDRA the different resonating modes gets generated with appreciable gain and efficiency nearly 90%.

Figure 4 shows variation in VSWR over the frequency of the proposed module. It depicts that due to the use of CDRA the VSWR goes close to ~1, at the three resonating frequencies which shows better impedance matching characteristic. Figure 5 shows peak gain and radiation efficiency over the wider range of frequency bands. It is

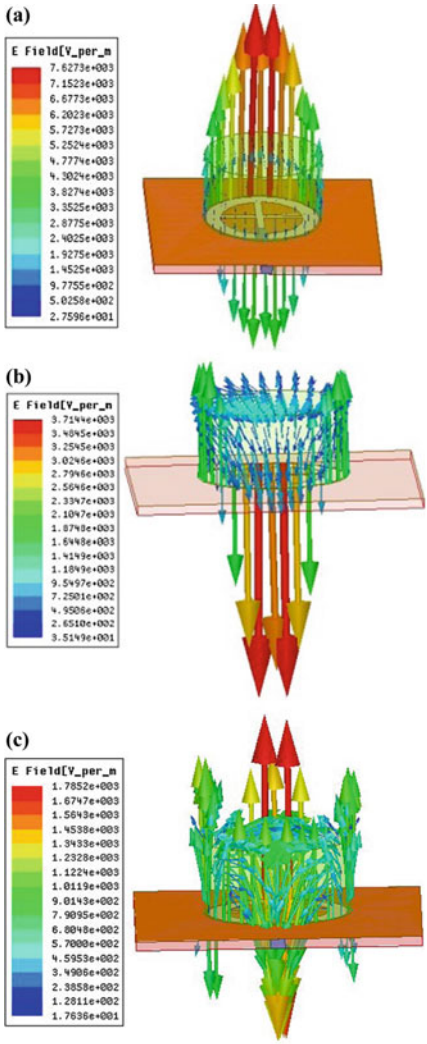


Fig. 2 Electric-field distribution at **a** 2.14 GHz **b** 3.68 GHz **c** 4.76 GHz

observed clearly from Fig. 5 that at st and 2nd order of resonant frequencies higher peak gain obtained.

Figure 6 exhibits the far-field radiation design in the XZ-YZ planes at 2.14 GHz, 3.68 GHz, and 4.76 GHz, respectively. Figure 6 represents that the radiation process at all the obtained resonant peaks has a stable radiation pattern.

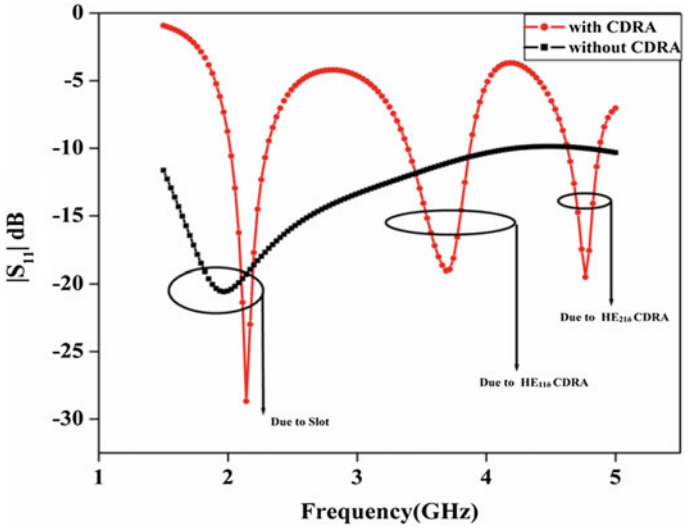


Fig. 3 Return loss of simulated structure over the frequency band

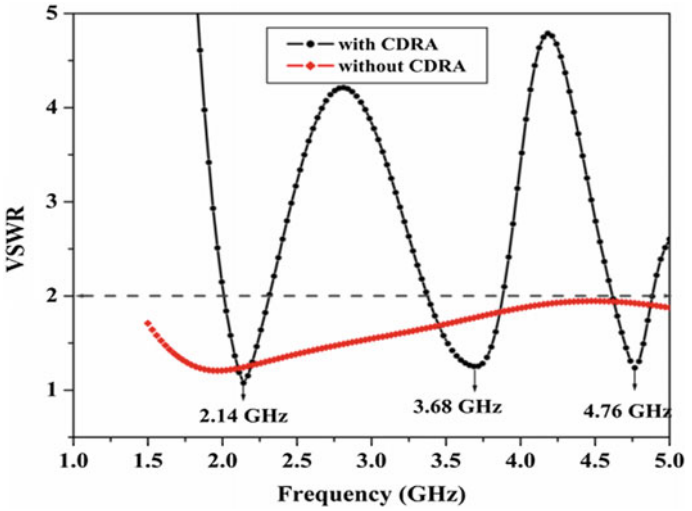


Fig. 4 VSWR versus frequency curve

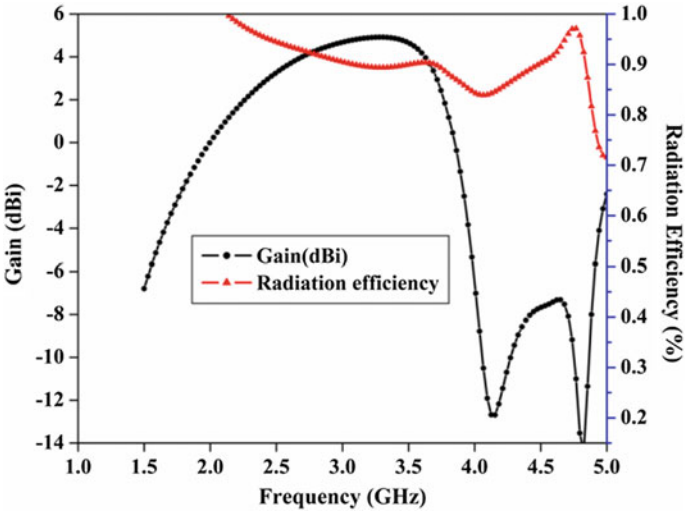


Fig. 5 Peak gain/radiation efficiency for the frequency band

5 Conclusion

This article manifests a triple-band hybrid Cylindrical DRA with a modified circular ring shape along with etched sectors inside. The presented hybrid radiator provides triple-band operation (2.01–2.29, 3.39–3.87, and 4.62–4.87 GHz) self sufficiently handled by two different resonating formations, i.e., cylindrical shapes DRA and the usage of the slot. The first band obtained among the three frequency bands is due to the use of slot and the rest two are due to the use of CDRA. The proposed prototype has an appreciable gain and radiation efficiency capability which make it a good entrant in the modern world of the communication system.

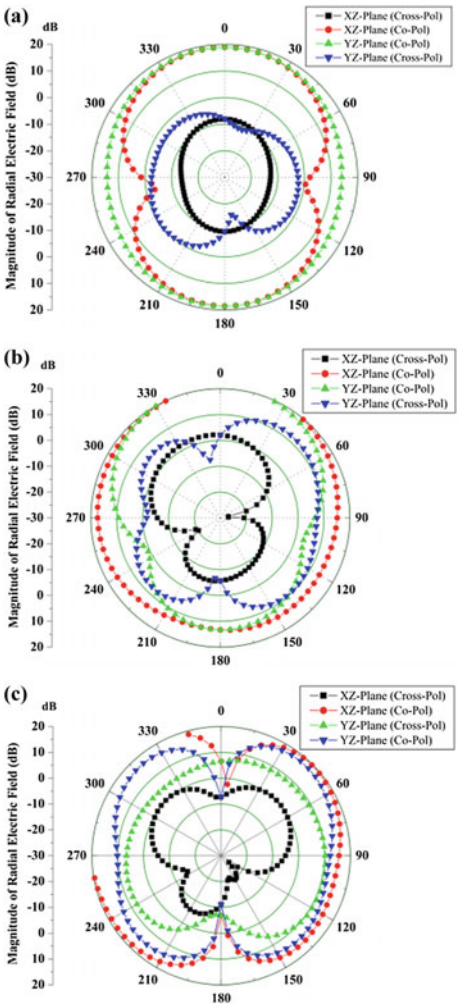


Fig. 6 Far field pattern representation of XZ-YZ planes at **a** 2.14 GHz **b** 3.68 GHz **c** 4.76 GHz

References

1. A. Petosa, *Dielectric Resonator Antenna Handbook* (Artech House, Norwood, MA, USA, 2007)
2. S.A. Long, M.W. McAllister, C.S. Liang, The resonant cylinder dielectric cavity antenna. IEE Trans. Antennas Propag. **31**, 406–412 (1983)
3. K.W. Luk, *Dielectric Resonator Antenna*. (Research Studies Press Ltd, Baldock, Hertfordshire. England, 2003)
4. J.M. Lee, S.J. Kim, G. Kwon, C.M. Song, Y. Yang, K.Y. Lee, K.C. Hwang, Circularly polarized semi-eccentric annular dielectric resonator antenna for X-band applications. IEEE Antennas Wirel. Propag. Lett. **14**, 1810–1813 (2015)
5. M. Khalily, M.R. Kamarudin, M. Mokayef, M.H. Jamaluddin, Omnidirectional circularly polarized dielectric resonator antenna for 5.2-GHz WLAN applications. IEEE Antennas Wirel.

Propag. Lett. **13**, 443–446 (2014)
6. A. Sharma, G. Das, R.K. Gangwar, Dual-band dual-polarized hybrid cylindrical dielectric resonator antenna for wireless applications. in *2017 Progress in Electromagnetic Research Symposium* (2017) (pp. 2190–2194)
7. S.-C. Lin, S.-C. Pan, Y.-K. Wang, Y.-F. Lin, H.-M. Chen, A compact dual-band dielectric resonator antenna using a parasitic slot. IEEE Antennas Wirel. Propag. Lett. **8**, 173–176 (2009)
8. H.S. Ngan, X.S. Fang, K.W. Leung, Design of dual-band circularly polarized dielectric resonator antenna using a higher-order mode, in *Proceedings of the 2012 IEEE-APS Topical Conference Antennas and Propagation in Wireless Communications APWC'12* (2012) (pp. 424–427)
9. D. Yang, M. Zou, J. Pan, A single-point-fed wideband circularly polarized rectangular dielectric resonator antenna. Int. J. Antennas Propag. **2016**, 1–9 (2016)
10. Y.M. Pan, S.Y. Zheng, W. Li, Dual-band and dual-sense omnidirectional circularly polarized antenna. IEEE Antennas Wirel. Propag. Lett. **13**, 706–709 (2014)

Tumor Detection in Multilayer Brain Phantom Model by Symmetrical-Shaped DGS Rectangular Microstrip Patch Antenna



Hemant Kumar Gupta, Raghavendra Sharma and Vandana Vikas Thakre

1 Introduction

Microwave imaging is gaining interest now as the famous research topic to find the inner objects using an electromagnetic wave in the microwave frequency range 300 MHz–300 GHz, by analyzing the transmitted and reflected wave from the inner object one can find the inner object with the help of variation in the performance parameter, i.e., return loss, current density, specific absorption rate of the electromagnetic wave [1, 2]. The discussed method is the same as the MRI (Magnetic resonance imaging) that uses cross-sectional image of the body to detect objects. Microwave imaging is advantageous as compared to the X-ray and MRI because of using low-power signals and repetition of microwave imaging test is not harmful to the body. [3] Ultra-wideband antenna is designed to interface with the brain with and without tumor moreover and the antenna is placed in contact with the brain so that the signal is reflected only from the affected part of the brain. [4–6]. The smart antenna array is also suggested to produce ultra wideband for a microwave imaging application. [7, 8]. Multilayer brain phantom model consists of five different layers starting from the outermost layer skin, fat, muscle, bone, brain with their respective dielectric constant and conductivity. [9] All literature reveals the importance of the ultra-wideband antenna in microwave imaging [10]. Most of the designed antenna produced low gain and high losses but microwave imaging requires antenna with high bandwidth and very low losses. [11]. In this paper designed antenna is ultra-wideband rectangular microstrip patch antenna (RMPA). Initially, RMPA is designed to further slot in the

H. K. Gupta (✉) · R. Sharma
Electronics Department, Amity School of Engineering and Technology, Amity University Gwalior, Gwalior, MP, India
e-mail: hmnt_gpt@yahoo.co.in

V. V. Thakre
Electronics Engineering Department, Madhav Institute of Technology and Sciences, Gwalior, MP, India

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_71

front plane and a defect in the ground plane of the antenna is applied to make it ultra-wideband. The proposed antenna shows a very good bandwidth about 04 GHz, bandwidth is high enough to use in microwave imaging.

The proposed antenna is interfaced with multilayer brain phantom model and all the performance parameter of the antenna has been calculated. Now the proposed antenna is interfaced with multilayer brain phantom model with tumor and all performance parameter of the antenna has been calculated. Tumor in the brain is detected by the variation in the performance parameter of the antenna with and without tumor.

2 Antenna Design

Rectangular microstrip patch antenna has different types of advantages like moderate gain, low size. The bandwidth of these planner antennas is enhanced by applying slots in the front plane and defects in the ground plane of the antenna. In this paper, initially simple RMPA at particular frequency has been designed and converted into ultra-wideband so that becomes applicable for microwave imaging. Simple RMPA with microstrip feeding is shown in Fig. 1.

Figure 1 reveals the length of the RMPA is L and width of the RMPA is W, whereas h is used dielectric substrate height. Basic L and W are finding by the electromagnetic equations [14]. Slots in the front plane defect in ground plane dimensions have been calculated by the hit and trial methods on CST-MWS simulation software [15] for lowest losses and maximum bandwidth. The antenna is said in ultra-wideband region if its bandwidth is at least 500 MHz or greater. Microstrip line feeding and coaxial port are used in the proposed antenna. The proposed antenna with front plane and ground plane and their respective dimensions are shown in Fig. 2. Table 1 also gives all the dimensions of the proposed antenna.

As mentioned earlier Length and width of the proposed planner Antenna resonating at the frequency 5.8 GHz is calculated by relevant electromagnetic relationships.

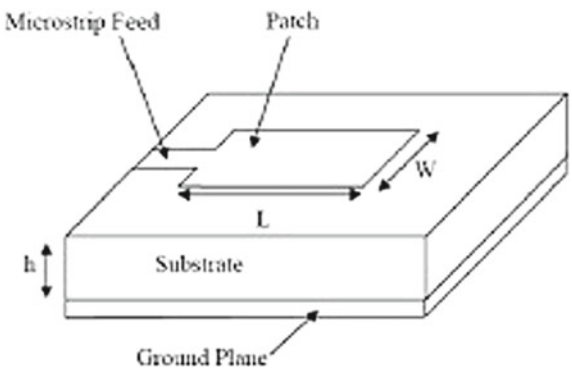


Fig. 1 Rectangular planner antenna [13]

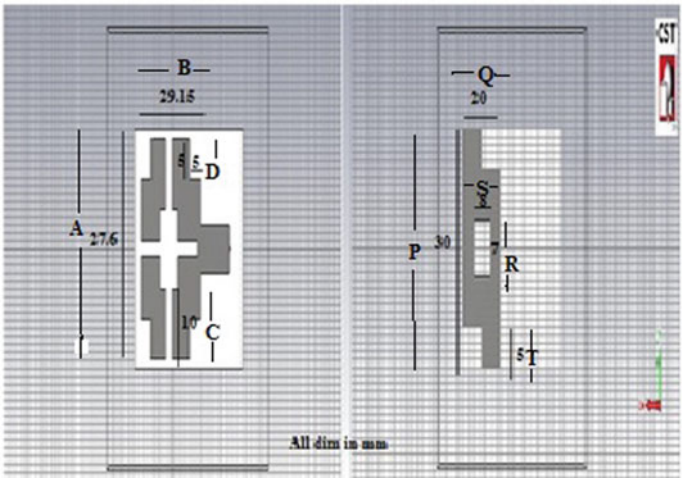


Fig. 2 Proposed dimensions for the ultra-wideband planner antenna

Table 1 Dimensions of the front and ground plane of the antenna

S. no.	Antenna front plane dimensions (mm)	Antenna ground plane dimensions (mm)
1	$A = 27.6$	$P = 30$
2	$B = 29.15$	$Q = 20$
3	$C = 10, D = 5$	$R = 7, S = 8, T = 5$

As the figure shows the parameters A and B have been calculated with help of relevant electromagnetic relationships whereas other dimensions of the front plane of antennas like C and D are determined by the hit and trial method on simulation software [15] for the lowest reflection losses and maximum bandwidth. The ground plane of the rectangular microstrip patch antenna is also modified by introducing different rectangular-shaped DGS whose dimension is given by the parameters P, Q, R, S, T as shown in Fig. 2. Table 1 indicates front plane dimensions and ground plane dimensions of the proposed ultra-wideband planner antenna.

2.1 Multilayer Brain Phantom Model

Simulation software is used to create the multilayer brain phantom model by considering the different layers of the brain with their respective electrical parameters. It is realized by considering the five layers of the brain, viz., CSF, Dura, bone, fat, skin with their relative electrical parameter. The dielectric constant of the brain layer is 45.2 and conductivity is 0.77 s/m. [12] Signal is transmitted into the brain by designed ultra-wideband RMPA (Rectangular Microstrip patch Antenna) and reflected signal

Table 2 Electrical properties of the different layers of the brain [12]

Name of the layer	Dielectric permittivity (f/m)	Electrical conductivity (s/m)	thickness (mm)
Skin layer	40.7	0.65	11
Fat layer	10	0.17	10
Bone layer	20.9	0.33	9
Dura layer	40.7	0.65	8
CSF layer	79.1	2.14	7
Brain model	45.8	0.77	6

from the brain has been compared with a transmitted signal based on the performance parameter, viz., return loss, current density, SAR (Specific absorption rate). The spherical-shaped tumor is placed inside the multilayer brain phantom model. The electrical parameters of the different layers of the brain are displayed inTable 2.

Multilayer brain model is realized by CST-MWS simulation software by taking different types of concentric spheres with their respective dielectric properties as given in Table 2. Later the designed model is interfaced with proposed UWB RMPA shown in Fig. 3.

The antenna performance parameter reflection coefficient with a multilayer brain phantom model is shown in Fig. 4.

Return loss of the RMPA with a multilayer brain phantom model is 41.22 dB at 7.69 GHz and 28.59 dB at 6.42 GHz as shown in Fig. 4. The other two performance parameters shown in Figs. 5 and 6. The observed value of the current density is 2706 A/m^2 and the specific absorption rate is 4.511e+06 W/m^3.

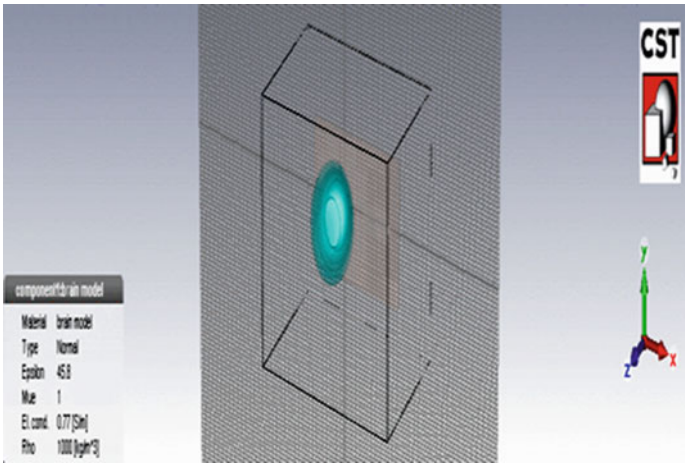


Fig. 3 Multilayer brain phantom model interfaced with UWB antenna

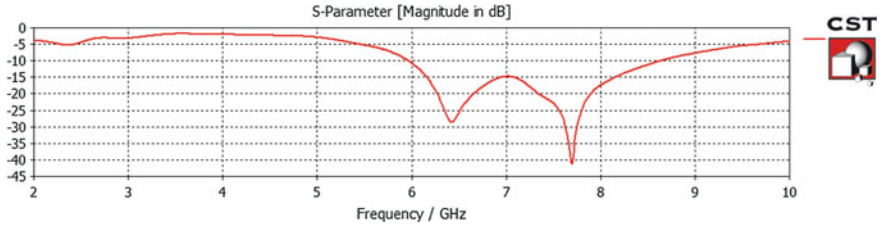


Fig. 4 Reflection parameter of multilayer brain model with UWB antenna

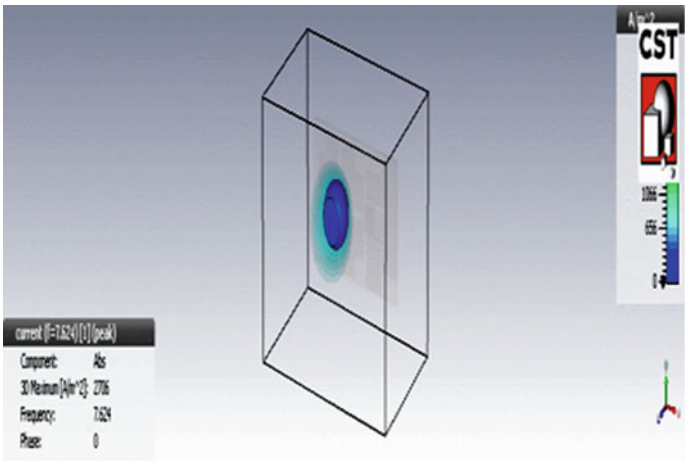


Fig. 5 Simulated current density of multilayer brain model with ultra-wideband antenna

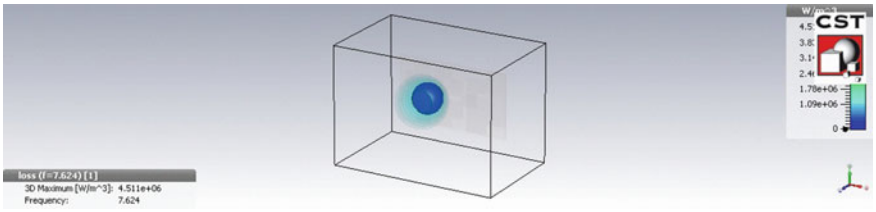


Fig. 6 SAR of the multilayer brain model with proposed UWB antenna

2.2 Multilayer Brain Phantom Model with Cancer Tumor

Generally, tumors cells have high dielectric permittivity and conductivity and these properties of tumors cells use to detect them by microwave imaging. Designed spherical-shaped tumor is shown in Fig. 7. It is designed by the sphere of radius 05 mm with respective electrical parameters. The used dielectric constant of the tumor is 54.2 and conductivity is 2.62 s/m.

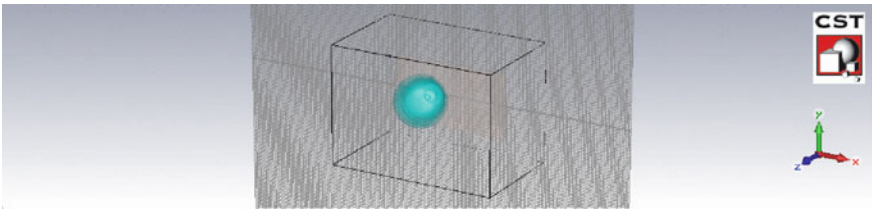


Fig. 7 Multilayer brain model with spherical-shaped tumor

Now the multilayer brain phantom model with the tumor is interfaced with proposed UWB RMPA and their performance parameter are displayed by the figures [7–9].

Figure 8 shows the reflection coefficient produced by RMPA with a multilayer brain phantom model and tumor. Figs. 9 and 10, respectively, show the current density

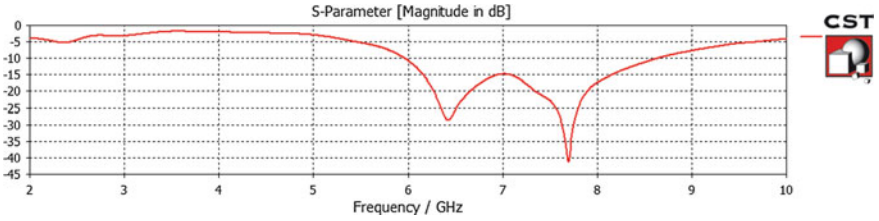


Fig. 8 S_{11} parameter of the multilayer brain phantom model with the tumor

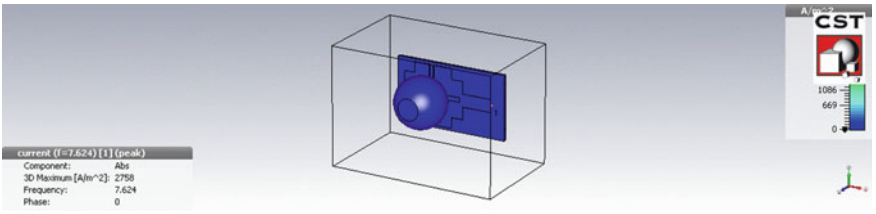


Fig. 9 The current density of the multilayer brain phantom model with tumor

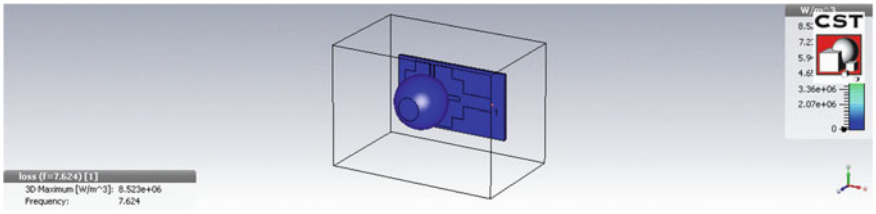


Fig. 10 SAR of multilayer brain phantom model with tumor

Table 3 Antenna performance parameter with and without tumor

Performance parameter	Without tumor	With tumor
Return loss (dB)	41.22	44.42
Current density (A/m^2)	2706	2758
Specific absorption rate ($w/m^{3)}$	4.511e+06	8.523e+06

and SAR of rectangular microstrip patch antenna with a multilayer brain phantom model and tumor.

The values of the reflection coefficient of the multilayer brain phantom model with a tumor when interfaced with the UWB antenna are 44.42 dB at 7.68 GHz and 29.19 dB at 6.42 GHz shown by Fig. 8. The other observable parameters, viz., current density and specific absorption rate for the given case are shown in Figs. 9 and 10. It is observed that the achieved value of current density is 2758 A/m^2 whereas the value of the specific absorption rate with UWB antenna and the tumor is $8.523e*10^6$ w/m^3. All performance parameter is shown in Table 3.

3 Results and Discussion

Antenna performance parameter as reflection parameter, current density, specific absorption rate has been observed and shown in Table 3. It has been shown that the return loss decreases with the tumor due to the tumor in the brain model fewer signals get reflected from the brain. However, the current density with tumor increases as the tumor cells have a very high dielectric permittivity and conductivity. The specific absorption rate is also increased with tumors as due to high dielectric constant and conductivity more signals are absorbed by the brain.

4 Conclusion

Rectangular Microstrip patch antenna with slots and DGS has been designed. The proposed planner UWB antenna is placed in contact with the designed multilayer phantom model with and without a tumor. Tumor in the multilayer brain phantom model is detected by the variation in the performance parameter of the proposed UWB rectangular microstrip patch antenna (RMPA) with and without tumor. Microwave imaging may be used as an effective tool in the field of medical science to detect the tumors in living beings.

References

1. M.E. Bialkowski, Y. Wang, A. Abbosh, UWB microwave monopulse radar system for breast cancer detection, in *2010 4th International Conference on Signal Processing and Communication Systems (ICSPCS)* (13–15 Dec. 2010), pp. 1–4
2. X. Li, S.C. Hagness, B.D. Van Veen, D. van der Weide, *Experimental Investigation of Microwave Imaging via Space-Time Beam-forming for Breast Cancer Detection* (IMS: Philadelphia, PA, USA, 2003), pp. 379–382
3. S.C. Hagness, A. Taflove, J.E. Bridges, Two-dimensional FDTD analysis of a pulsed microwave confocal system for breast cancer detection: fixed-focus and antenna-array sensors. IEEE Trans. Biomed. Eng. **45**(12), 1470–1479 (1998)
4. F. Thiel, O. Kosch, F. Seifert, Contrast agent based tumor detection by ultra-wideband radar: a model approach, in *2010 Proceedings of the Fourth European Conference on Antennas and Propagation (EuCAP)* (12–16 April 2010), pp. 1–3
5. D. Gibbins, M. Klemm, I.J. Craddock, J.A. Leendertz, I.J. Preece, R. Benjamin, Comparison of a wide-slot and a stacked patch antenna for the purpose of breast cancer detection. IEEE Trans. Antennas Propag. **58**(3), 665–674 (2010)
6. E. Pancera, Medical application s of the ultra wideband technology, in *2010 Loughborough Antennas and Propagation Conference (LAPC)* (8–9 November 2010), pp. 52–56
7. H. Zhang, A.O. El-Rayis, N. Haridas, N.H. Noordin, A.T. Erdogan, T. Arslan, A smart antenna array for brain cancer detection, in *Loughborough Antennas & Propagation Conference* (Loughborough, UK, 14–15 November 2011)
8. X.N. Low, Z.N. Chen, Terence S.P. See, A UWB dipole antenna with enhanced impedance and gain performance, IEEE Trans. Antennas Propag. **57**(10) (2009)
9. A. Lak, H. Oraizi, Evaluation of SAR distribution in six-layer human head model. Int. J. Antenna Propag. (2013)
10. M. O'Halloram, M. Glavin, E. Jones, Rotating antenna microwave system for breast cancer detection. Prog. Electromagn. Res. **107** (2010)
11. N. Ojaroudi, M. Ojaroudi, F. Geran, Sh. Amiri, Ominidirectional antenna with multi resonance performances for microwave imaging system. Telfor J. **05** (2013)
12. D. Ireland, M. Bialkowski, Microwave head imaging for stroke detection. Prog. Electromagn. Res. M **21**, 163–175 (2011)
13. R. Inum, M. Rana, K.N. Shushama, M. Quader, EBG based microstrip patch antenna for brain tumor detection via scattering parameters in microwave imaging system. Int. J. Biomed. Imaging Hindawi Publ. (Article ID 8241438 Volume) (2018)
14. C. Balinies, Antenna Theory (Wiley, 1997) (2nd addition chapter-14 ISBN 0-471-59268-4)
15. http://www.cst.com/content/products/mws/overview.aspx © (2014) CST Computer Simulation Technology

Design and Optimization of S-Shape Multi-resonant Microstrip Antenna



Mohit Gaharwar, Shivam Tyagi and Aditya Boggavarapu

1 Introduction

Satellite and wireless correspondence have been produced quickly in the previous decades and it has as of now an emotional effect on human life. Present-day media transmission frameworks require antennas with more extensive data transmissions and littler measurements than customarily conceivable. This has started reception apparatus inquire about in different ways, one of which is by utilizing fractal molded antenna components. As of late, a few fractal geometries have been presented for antenna applications with differing degrees of accomplishment in enhancing antenna qualities. A portion of these geometries has been especially valuable in diminishing the extent of the antenna, while different plans go for joining multi-band attributes [1, 2]. However, no huge advancement has been made in proving fractal properties of these geometries with qualities of reception apparatuses. The exploration work exhibited here is basically expected to break down geometrical highlights of fractals that impact the execution of antennas utilizing them.

A few antenna arrangements dependent on fractal geometries have been accounted for as of late [3–7]. These are low profile antennas with moderate addition and can be made operational at different recurrence groups and henceforth are multi-utilitarian. In this work, the multi-band part of antenna plans is investigated further with extraordinary accentuation on distinguishing fractal properties that sway reception apparatus

M. Gaharwar (✉) · S. Tyagi
Sachdeva Institute of Technology, Farah, Mathura, India
e-mail: mohitgaharwarcoe@gmail.com

S. Tyagi
e-mail: tyagishivam123.st@gmail.com

A. Boggavarapu
Bhilai Institute of Technology, Raipur, CG, India
e-mail: b.aditya@hotmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_72

Table 1 Data sheet of designed antenna

Ground	Material used	Copper
	Length (mm)	47.39
	Width (mm)	60.85
Substrate	Material used	FR4-lossy
	Length (mm)	50
	Width (mm)	36
Patch	Material used	Copper
	Length (mm)	46
	Width (mm)	36
Thickness between ground and substrate (mm)		0.1
Thickness between substrate and patch (mm)		1.6

multi-band qualities [8–10]. Reception apparatuses with diminished size have been gotten utilizing Hilbert bend fractal geometry.

2 Antenna Design

This segment portrays the plan of S-shape antenna which is built utilizing emphasis. It has been discovered that the multi-band nature of the reception apparatus can be changed into wideband qualities by utilizing an extremely high dielectric consistent substrate and appropriate engrossing materials [11–14]. The following material and their qualities are considered for making the ideal radio wire. A wide range of dielectric types might be utilized for this structure; however, the main simple accessible material for PCB printing is FR4 lossy. The dielectric substrate FR4 lossy is basically an epoxy and glass texture overlay, with a dielectric steady $\varepsilon_r = 4.2$ at 1.5 GHz recurrence. An information sheet is given in Table 1. The thickness of the dielectric accessible is 1.6 mm.

3 Results and Discussions

A ground plane was imprinted on the contrary side of the substrate from the fractal design, as is required in a microstrip patch structure. For the most part, bigger ground planes permit better radiation execution, however, an exchange off must be made against the final size of the reception apparatus permitted. What's more, bigger regions of PCBs cost more to create. There are various strategies to nourish a flag to

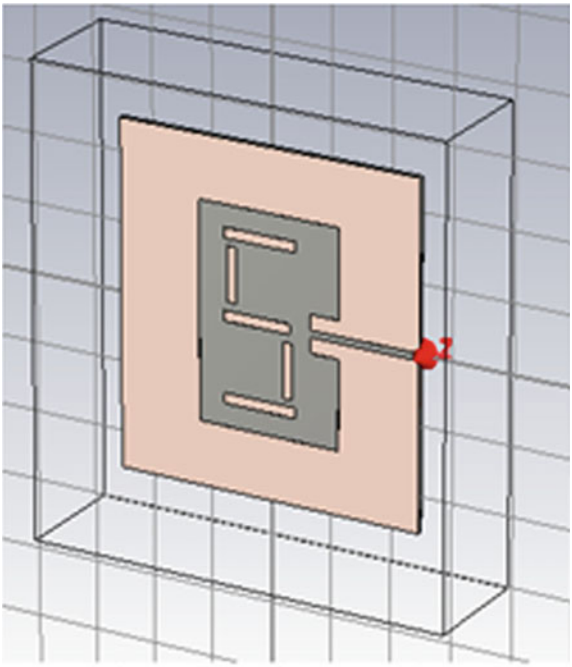


Fig. 1 Structure of anticipated antenna

a microstrip patch, in spite of the fact that the most favorable for this plan are edge and test sustaining as appeared in Fig. 1.

Figure 1 shows the s-shaped perspective view of our desired antenna which is finally designed in CST software at 1.5 GHz cut-off frequency. Figure 2 shows the return loss of the desired antenna. Return loss is the segment of a flag that is lost because of an impression of intensity at line brokenness. It is like VSWR and is commonly favored in the link business to a VSWR determination. Since it is a logarithmic estimation, it is valuable while showing extremely little reflections. Figure 3 describes 3D radiation pattern at 1.5 GHz with the directivity of 6.275 dBi.

Figure 4 shows Smith's diagram of wanted radio wire. The Smith Chart is a phenomenal instrument for envisioning the impedance of a transmission line and receiving wire framework as an element of recurrence. Smith Charts can be utilized to build comprehension of transmission lines and how they carry on from an impedance perspective. The Smith Chart is the graphical portrayal of a complex numerical condition. It is the roundabout plot of the attributes of microwave parts. The Smith Chart is the most utilized instrument for microwave architects to envision complex-esteemed amounts and figure the mapping between them. It comprises of two arrangements of circles for plotting different parameters of jumbled transmission lines. One is the arrangement of complete circles whose fixates lie on the straight line and the other one is the arrangement of two curve circles that lie on either sides

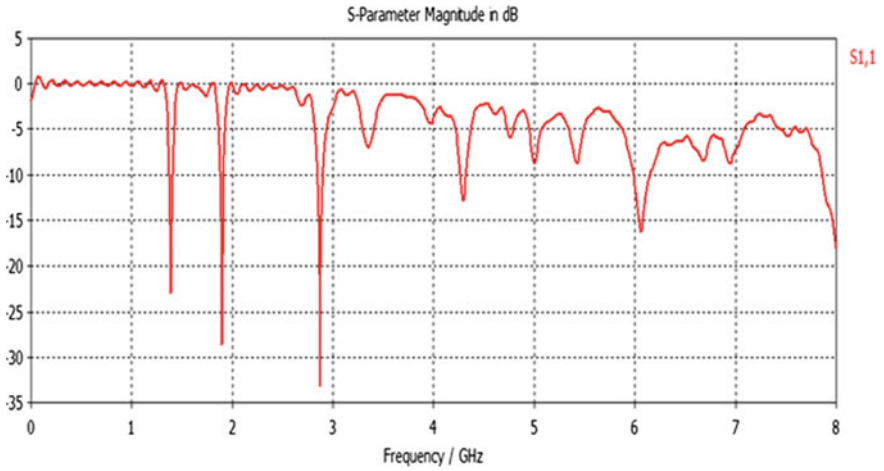


Fig. 2 S_{11} parameter (return loss) of the proposed antenna

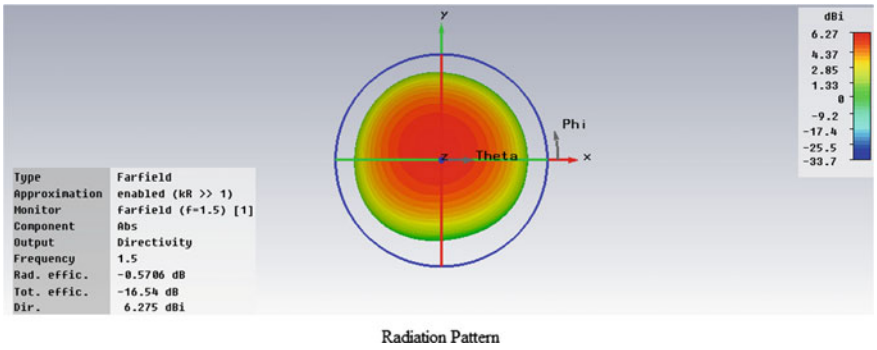


Fig. 3 3D Radiation pattern at 1.5 GHz

of the straight line. Figure 5 shows the hardware of the proposed S-Shape antenna. Figure 6 describes the measured S_{11} parameter of the proposed S-Shape antenna.

4 Conclusion

The reduced, S-shape configuration exhibited in this report is an incredible option in contrast to conventional antenna frameworks in versatile remote recipients. The S-shape monopole performs satisfactorily at 1.5, 1.8, and 2.8 GHz recurrence frequencies and shows space-proficiency through its self-comparative fractal structure. The S-shape configuration shows great execution at 1.5 GHz and has radiation properties about indistinguishable to that of conventional, straight-wire monopoles at that

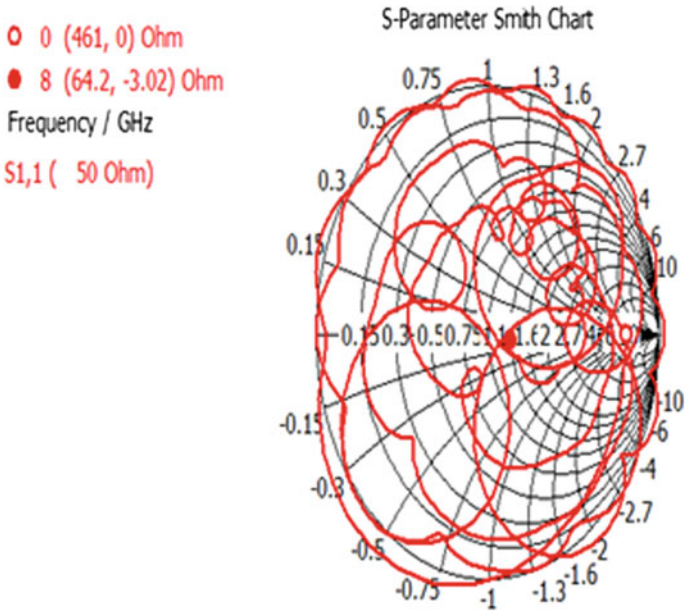


Fig. 4 Smith chart of the proposed antenna

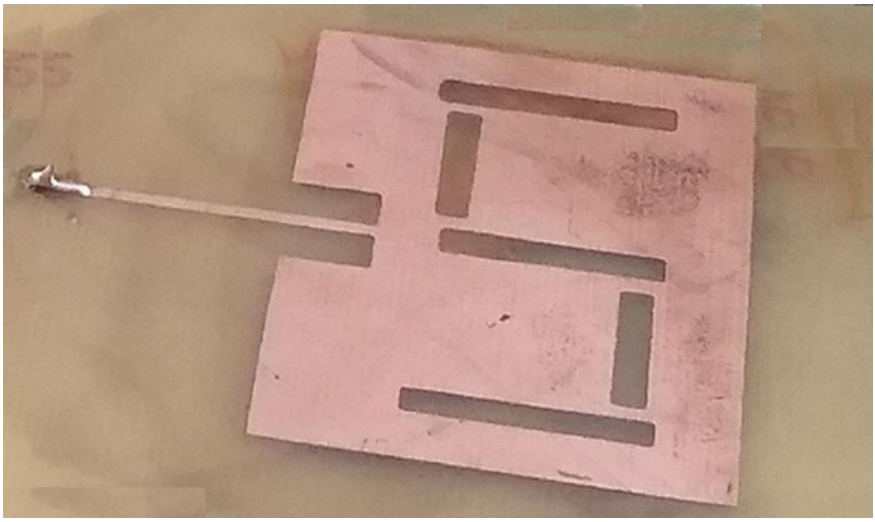


Fig. 5 Hardware proposed S-shape antenna

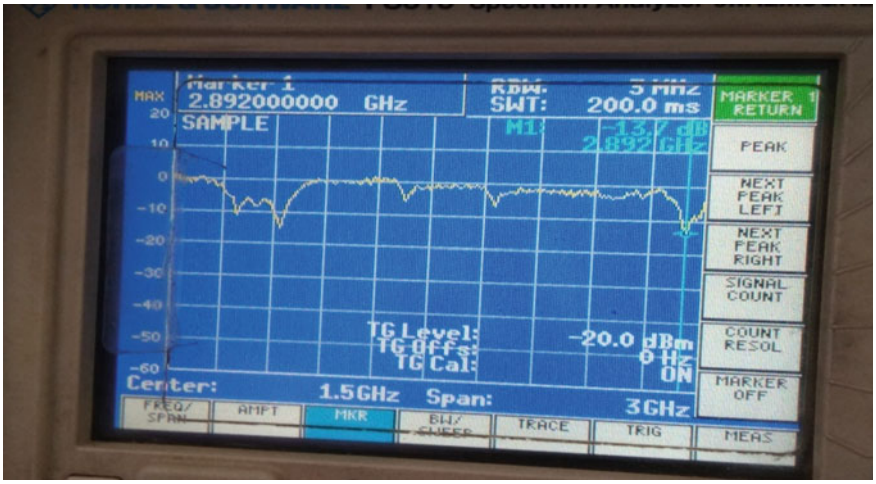


Fig. 6 Measured S_{11} parameter of proposed S-shape antenna

recurrence. The objectives of this structure venture were along these lines effectively practiced.

References

1. S. Kakkar, T.S. Kamal, A.P. Singh, Design and analysis of I-shaped fractal antenna for emergency management. IETE J. Res. (Taylor & Francis), ISSN: 0377-2063 (2018)
2. A.A. Deshmukh, D. Singh, M. Gala, Broadband slot cut S-shaped microstrip antenna. in *International conference in Computing Communication Control and automation (ICCUBEA)* (IEEE, 2017)
3. R. Kumar, P. Malathi, On the design of CPW-feed diamond shape fractal antenna for UWB applications. J. Cogent Eng. **4**(1) (2017)
4. A. Behera, S. Satrusallya, M.N. Mohanty, Design of fractal antenna based on Koch-Sierpinski combination for UWB application. in *IEEE International Conference on Communication and Signal Processing* (India), ISBN No. 978-1-4673-8549-7, 6–8 Apr 2016
5. Y.S. Mezaal, New microstrip semi-fractal antenna: design and simulation results. in *IEEE 24th Signal Processing and Communication Application Conference (SIU)* (Zonguldak, Turkey, 2016)
6. S. Singhal, T. Goel, A.K. Singh, Hexagonal tree shaped ultra–wideband fractal antenna. IETE J. Res. (Taylor & Francis) 335–348 (2016)
7. J.M. Patel, S.K. Patel, F.N. Thakkar, Design of S-shaped multiband microstrip patch antenna. in *International Conference NUiCONE* (IEEE, 2012)
8. K. Kumar, N. Gunasekaran, A novel wideband slotted microstrip patch antenna. in *International Conference on Signal Processing, Communication, Computing and Networking Technologies (ICSCCN)* 21–22 July 2011 (pp. 10–14)
9. M.A.R. Osman, M.K.A. Rahim, M. Azfar, N.A. Samsuri, F. Zubir, K. Kamardin, Design, implementation and performance of ultra-wideband textile antenna. Progress Electromagn. Res. B **27**, 307–325 (2011)

10. V.K. Singh, Z. Ali, S. Ayub, A.K. Singh, A wide band Compact Microstrip Antenna for GPS/DCS/PCS/WLAN applications. in *Intelligent Computing, Networking, and Informatics*, vol. 243. (Springer, 2014) pp. 1107–1113. ISBN: 978-81-322-1664-3
11. R. Srivastava, V.K. Singh, S. Ayub, Comparative analysis and bandwidth enhancement with direct coupled c slotted microstrip antenna for dual wide band applications. in *Frontiers of Intelligent Computing: theory and Applications*, vol. 328. (Springer, 2015), pp. 449–455. ISBN: 978-3-319-12011-9
12. V.K. Singh, Z. Ali, S. Ayub, A.K. Singh, Bandwidth optimization of compact microstrip antenna for PCS/DCS/bluetooth application. Cent. Eur. J. Eng. **4**(3), 281–286 (Springer, 2014)
13. N. Singh, A.K. Singh, V.K. Singh, Design and performance of wearable ultra wide band textile antenna for medical applications. Microw. Opt. Technol. Lett. **57**(7), 1553–1557 (2015)
14. S. Srivastava, V.K. Singh, Z. Ali, A.K. Singh, Duo triangle shaped microstrip patch antenna analysis for WiMAX lower band application. in *International Conference on computational Intelligence: Modelling Techniques and Applications (CIMTA-2013), Procedia Technology Elsevier*, vol. 10 (2013), pp. 554–563

Sigma-Structured Microstrip Antenna for Harvesting Energy for Low-Power Devices



Bharat Bhushan Khare, Vinod Kumar Singh and Anurag Saxena

1 Introduction

The utilization of sustainable power sources to control hardware gadgets is certifiably not another idea. The way toward extracting energy from the surrounding condition to create power is named as energy scavenging [1, 2]. This energy can be gathered from different sources accessible in the encompassing condition, for example, thermal energy, mechanical energy, and radiant energy [3–7]. In Table 1, properties of generally used encompassing vitality sources are presented. Among the accessible encompassing vitality sources, RF vitality has incredibly increased because of the dominance of remote signs, for example, portable base stations [8] and Wi-Fi systems [9], radio and TV transmitters [10–12], and microwave radios and cell phones [13]. This strategy could be particularly helpful for driving up the remote systems sent in cruel conditions, charging batteries, or putting away vitality in supercapacitors. In this paper, we center on surrounding RF energy reaping innovation that will have a vital potential to affect sensors situated in cruel conditions or remote spots, where other vitality sources as wind or sun-based sources are impracticable. This method pulled in huge consideration and various RFEH systems including antenna, coordinating circuits, and rectifier circuits which have been produced for the green supply of low-utilization gadgets.

B. B. Khare
Department of Electrical Engineering, UIT, Bhopal, India
e-mail: bharatbhushankhare1996@gmail.com

V. K. Singh (✉) · A. Saxena
Department of Electrical Engineering, SRGI, Jhansi, India
e-mail: singhvinod34@gmail.com

A. Saxena
e-mail: anurag.saxena.anurag@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_73

Table 1 parameter of designed antenna

S. no.	Antenna parameter	Value
1	Relative permittivity	1.7 F/meter
2	Substrate thickness	1 mm
3	Length of sigma l_1	80 mm
4	Width of sigma l_2	100 mm
5	Side width l_3/l_4	19.5 mm
6	Left-sided middle width l_5	24 mm
7	Inclined arm length l_6	65 mm
8	Right-sided arm length l_7/l_8	10 mm
9	Ground plane dimension (l_9 * l_{10})	50 * 5 mm*mm

The rectenna is the combination of antenna and rectifier circuit so in this paper, the designed antenna is flexible in nature or can be fabricated on clothes because it is a textile [13, 14]. Nowadays, many types of wearable antennas are available. Here the sigma-structured antenna is used which is designed on CST Microwave studio. Then this antenna is connected to the rectifier circuit, which is designed on Pspice software then the theoretical analysis is performed to harness the power which is available in cell phone tower in the range of 935–960 MHz [15–20]. This range of frequency is falling in the category of GSM 900 according to the report on cell tower radiation, which is submitted to the secretary, DOT, Delhi by the electrical engineering department of IIT Bombay, Mumbai.

2 Antenna Design

The sigma-structured microstrip antenna is used for harvesting the energy. It is the combination of ground, substrate, and patch. The antenna designed on the textile substrate by using copper tape so it is flexible in nature and having wearable property. The parameters of this antenna are given in Table 1 and the structure of the ground and patch is shown in Fig. 1.

3 Calculation of Received Power and Voltage of Antenna

According to the report on cell tower radiation, submitted to the DOT, Delhi by IIT Bombay, the transmission frequency range of cell tower transmitter antenna is 869–894 MHz (CDMA), 935–960 MHz (GSM-900) and 1810–1880 MHz (GSM-1800). So the designed sigma-structured antenna has a gain of 2.639 dBi at 960 MHz and 5.443 dBi at 2.12 GHz. Here the theoretical analysis is conducted on 960 MHz frequency only because this is the range of GSM-900. On the behalf of the report of

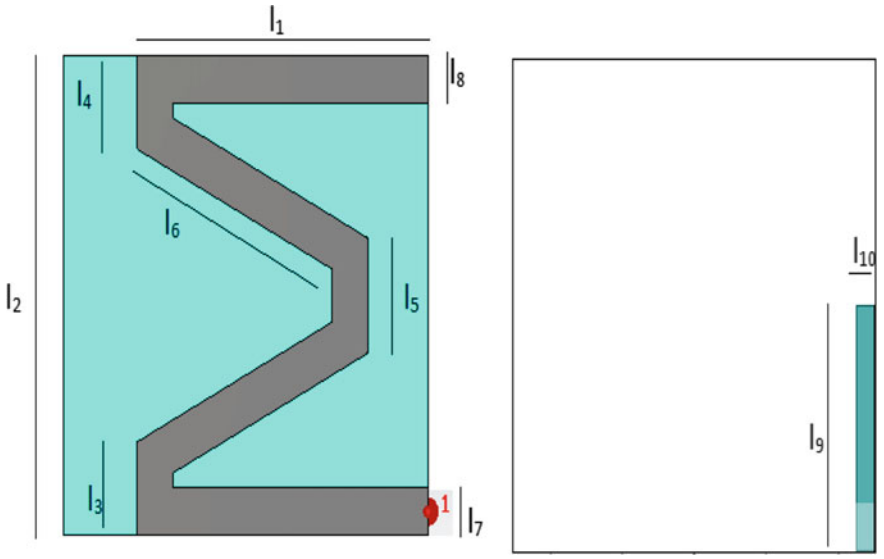


Fig. 1 Structure of antenna patch and ground

IIT Bombay, the transmitted power of transmitting antenna is 20 W and the gain of transmitter is 17 dBi (50), then the calculation of power density at the distance of 1 m can be calculated by

$$P_d = \frac{P_e * G_e}{4\pi R^2}$$

where

P_d is the power density
P_e is the minimum transmitted power from cell phone transmitter
G_e is gain of the transmitter
R is the distance between transmitter and receiver

By putting the value $P_e = 20$, $G_e = 50$ and $R = 1$ m we find the value of $P_d = 79.57747$ W/m^2. Then we can calculate the received power of the proposed antenna if the gain of that antenna is 2.639 dBi which can be gathered by microwave studio software. The received power can be easily calculated by

$$P_r = \frac{P_d * G_m * \lambda^2}{4\pi}$$

where

G_m is the gain of the proposed antenna
λ is the wavelength ($\lambda = C/f$) in meter

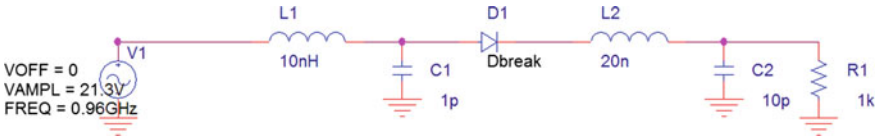


Fig. 2 Rectifier circuit utilizing anticipated antenna as a source

By putting the value of $P_d = 79.577$, $G_m = 1.8361$ (2.639 dBi) and $\lambda = 0.3125$ m, the value of $P_r = 1.1354658$ W.

4 Rectifier Circuit

The rectifier circuit is designed on Pspice software in which Schottky diode is used. Some inductors and capacitors are connected in series and parallel combinations to reduce the distortion which is shown in Fig. 2. The calculation of received voltage at the rectifier can be calculated by

$$V_s = 2 * \sqrt[2]{2 * R_s * P_r}$$

where

R_s is the source impedance. P_r is received power from the antenna.

Substitution the values $R_s = 50\ \Omega$ and $P_r = 1.1354658$ W, the value of $V_s = 21.3$ V. the figure of rectifier circuit is shown in Fig. 3.

5 Illustration of Results

The received power at the distance of 1 m from cell phone tower is 1.1354658 W at sigma-structured antenna. This received power is easily rectified by using a rectifier circuit that is designed in Pspice software and the power output after the rectifier at the load of 1 kΩ is approx 0.6 W (according to the graph of Pspice). So this power can be utilized to operate small sensors and microelectronic gadgets. The gain of antenna and return losses versus frequency plot is shown in Figs. 3 and 4. The graph of voltage, current and power is shown in Figs. 5, 6 and 7 respectively.

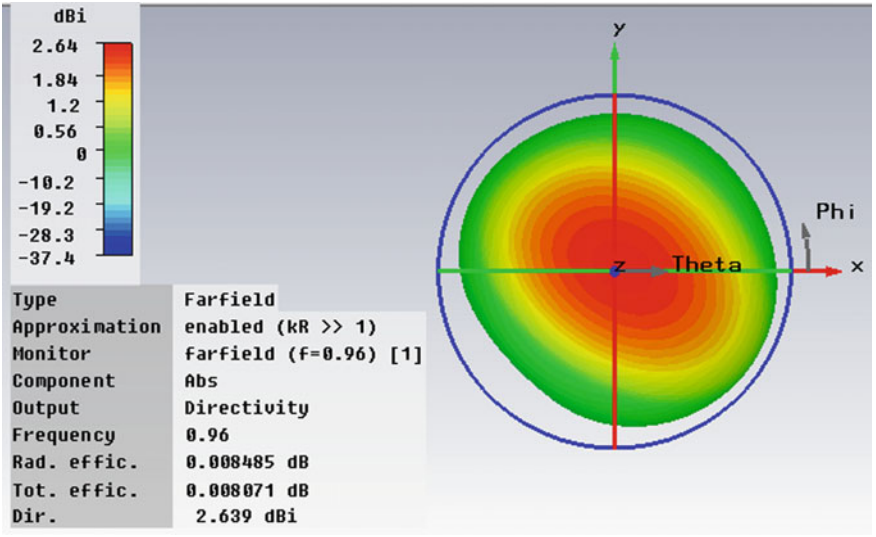


Fig. 3 Gain of proposed antenna

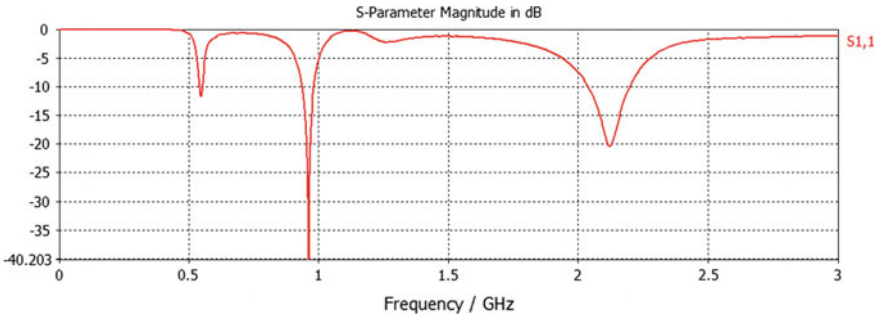


Fig. 4 Return losses versus frequency plot

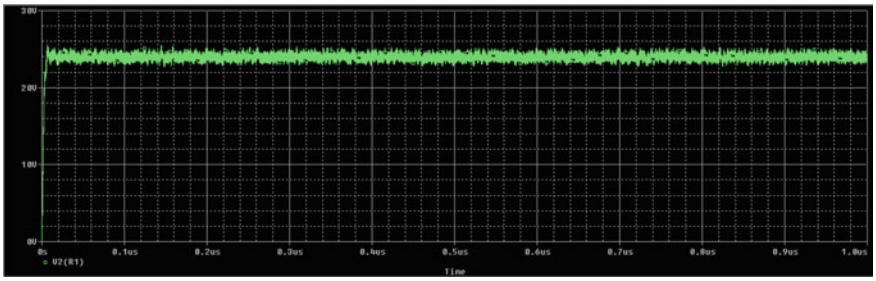


Fig. 5 Output voltage of rectenna circuit at load

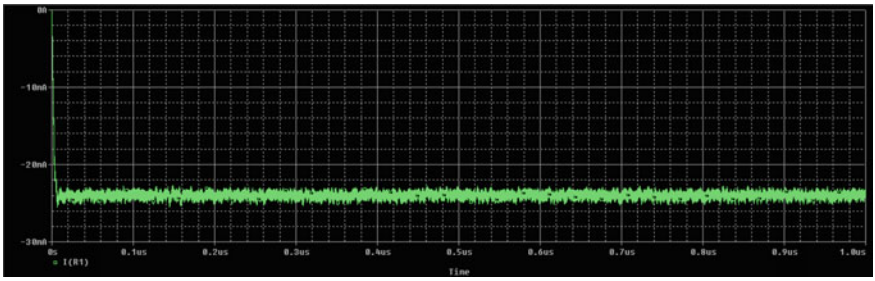


Fig. 6 Output current of rectenna circuit at load

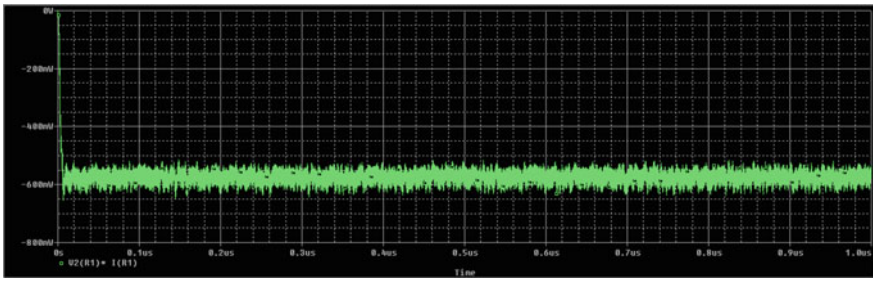


Fig. 7 Output power in mW of rectenna circuit at load

6 Conclusion

This rectenna is theoretically designed and it can be used to harvest the energy and this energy can be utilized to operate some microelectronic gadgets. Generally, the received power is inversely proportional to the square of the distance. If this antenna is utilized less than one-meter distance then the received power will be definitely enhanced. For the enhancement of received power, the antenna array is the best alternative.

References

1. S. Chalasani, J.M. Conrad, A survey of energy harvesting sources for embedded systems, in *Proceedings of the IEEE Southeastcon* (IEEE, Huntsville, Ala, USA, 2008), pp. 442–447
2. S. Sudevalayam, P. Kulkarni, Energy harvesting sensor nodes: survey and implications. IEEE Commun. Surv.
3. L. Mateu, C. Codrea, N. Lucas, M. Pollak, P. Spies, Energy harvesting for wireless communication systems using thermo generators, in *Proceedings of the 21st Conference on Design of Circuits and Integrated Systems (DCIS '06)*, Barcelona, Spain, November 2006

4. H. Bottner, J. Nurnus, A. Gavrikov, et al., New thermoelectric components using microsystem technologies. J. Microelectromechanical Syst. **13**(3), 414–420 (2004)
5. Y.K. Tan, K.Y. Hoe, S.K. Panda, Energy harvesting using piezoelectric igniter for self-powered radio frequency 8 journal of sensors (RF) wireless sensors, in *Proceedings of the IEEE International Conference on Industrial Technology (ICIT '06)*, Mumbai, India, December 2006, pp. 1711–1716
6. A. Hande, T. Polk, W. Walker, D. Bhatia, Indoor solar energy harvesting for sensor network router nodes. Microprocess. Microsyst. **31**(6), 420–432 (2007)
7. C. Alippi, C. Galperti, An adaptive system for optimal solar energy harvesting in wireless sensor network nodes. IEEE Trans. Circuits Syst. **55**(6), 1742–1750 (2008)
8. L.M. Borges, N. Barroca, H.M. Saraiva, et al., Design and evaluation of multi-band RF energy harvesting circuits and antennas for WSNs, in *Proceedings of the 21st IEEE International Conference on Telecommunications*, Lisbon, Portugal, May 2014, pp. 308–312
9. M. Ali, L. Albasha, N. Qaddoumi, RF energy harvesting for autonomous wireless sensor networks, in *Proceedings of the 8th International Conference on Design & Technology of Integrated Systems in Nanoscale Era (DTIS '13)* (IEEE, Abu Dhabi, United Arab Emirates, 2013), pp. 78–81
10. T. Sogorb, J. Vicente Llario, J. Pelegri, R. Lajara, J. Alberola, Studying the feasibility of energy harvesting from broadcast RF station for WSN, in *Proceedings of the IEEE Instrumentation and Measurement Technology Conference Proceedings (IMT'08)* (IEEE, Victoria, Canada, 2008), pp. 1360–1363
11. T.S. Salter, Low power smart dust receiver with novel applications and improvements of an RF power harvesting circuit. Ph.D. thesis (University of Maryland, 2009)
12. B. Naresh, V.K. Singh, V.K. Sharma, Flexible hybrid energy harvesting system to power wearable electronics, in *4th International Conference on Advances in Electrical, Electronics, Information, Communication and Bio-Informatics (AEEICB-18)*, Prathyusha Engineering College, Chennai, Tamil Nadu, India, 27–28 February 2018
13. C. Mikeka, H. Arai, Sustainable energy harvesting technologies—past, present and future, in *Design Issues in Radio Frequency Energy Harvesting System*, ed. by Y.K. Tan (In Tech, Vienna, Austria, 2011)
14. Report on cell tower radiation submitted to secretary, DOT, Delhi by electrical engineering department IIT Bombay, in December 2010
15. A. Gopinath, *All About Transferring Power Wirelessly (PDF), Electronics for You Ezine* (EFY Enterprises Pvt. Ltd., 2015), pp. 52–56
16. S. Ashley, *Wireless Power Transfer via Radio Waves* (Wiley, 2015), pp. 9–13
17. B. Naresh, V.K. Singh, V. Bhargavi, Low power circularly polarized wearable rectenna for RF energy harvesting, in *Advances in Power Systems and Energy Management. Lecture Notes in Electrical Engineering*, vol. 436, ed. by A. Garg, A. Bhoi, P. Sanjeevikumar, K. Kamani (Springer, Singapore, 2018), pp. 131–138
18. B. Naresh, V.K. Singh, Dual band RF energy harvester for wearable electronic technology, in *3rd International Conference on Advances in Electrical, Electronics, Information, Communication and Bio-Informatics (AEEICB-17)*, Prathyusha Engineering College, Chennai, Tamil Nadu, India (978-1-5090-5434-3), 27–28 February 2017
19. X.H. Wang, L. Zhang, Y. Xu, Y.F. Bai, C. Liu, X.W. Shi, A tri-band impedance transformer using stubbed coupling line. Prog. Electromagn. Res. 141, 33–45 (2013)
20. B. Naresh, V.K. Singh, V. Bhargavi, A. Garg, A.K. Bhoi, Dual-band wearable rectenna for low-power RF energy harvesting, in *Advances in Power Systems and Energy Management. Lecture Notes in Electrical Engineering*, vol. 436, ed. by A. Garg, A. Bhoi, P. Sanjeevikumar, K. Kamani (Springer, Singapore, 2018), pp. 13–21

Nodal Pricing Analysis of Distribution System with Wind Power and D-STATCOM for Realistic ZIP and RIC Loads



Banothu Sridhar and Ashwani Kumar

Nomenclature

S_r	Number of reactive power sources
S_l	Number of branches
i, j, l &t	Index variables for buses, lines and time
S_g	Number of generators
S_b	Number of buses
a_i, b_i, c_i	Coefficients of cost (0.01, 40 and 9)
d_k, e_k, f_k	Coefficients of cost of distributed generation (dg)
CF	Cost factor of D-STATCOM
U_i^{max}, U_i^{min}, U_i	Voltage maximum and minimum limit and magnitude.
U_0	Nominal voltage (1p.u)
$P_{wind(k}^{max}$	Maximum limit of wind power
$P_{wind(k)}^{min}$	Minimum limit of wind power
P_{gi}^{max}	Active power generation maximum limit
P_{gi}^{min}	Reactive power generation minimum limit
$Q_{D_{STATCOM(z)}}^{min}$	D-STATCOM reactive power minimum limit
$Q_{D-STATCOM(z)}^{max}$	D-STATCOM reactive power maximum limit.
P_{gi}	Active power generations
Q_{gi}	Generation of reactive power
D-STATCOM	Distribution static synchronous compensator
P_{di}	Real power load at each bus i
Q_{di}	Reactive power load demand at each node i

B. Sridhar (✉) · Ashwani Kumar
National Institute of Technology, Kurukshetra 136119, Haryana, India
e-mail: sridhar.banothu95@gmail.com

Ashwani Kumar
e-mail: ashwaks@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_74

P_{dgi}	Real power generation of wind dg
P_i	Real power injected at bus i
Q_i	Reactive power injected at bus i
L_{net}	Net power loss of the system
P_{di0}	IEEE 69 bus base load active power
Q_{di0}	IEEE 69 bus base load of reactive power
NLP	Nonlinear programing
ρ_{Pi}	Real power marginal loss coefficient
ρ_{Qi}	Marginal loss coefficient of reactive power
TP_{dg}	Cost function of wind dg
G_{ij}	Line conductance
B_{ij}	Line susceptance
θ_i	Angle of bus voltages
θ_{ij}	Angle of node voltages ij

1 Introduction

In recent decades, the electricity sector was a single entity managed by the government. As a vertically integrated public service, consumers have only one power purchase option in the public sector. Over time, due to social welfare, the electricity market has emerged as competitors of energy suppliers. Renewable Energy sources integration into the distribution system has gained a lot of importance due to green energy which has lower emission of greenhouse gases in the environment. However, with the distributed generation there is a need to determine the Nodal pricing in the distribution system. The analysis of the distribution system has changed greatly in recent years. Node pricing is a cost-effective pricing method commonly used in transmission systems and which can also be used in the distribution system according to current national plans [1]. Nodal pricing which is also called Locational marginal pricing implemented first and improved in [2] gives the technique of pricing to determine the transmission operation for a very short period of time. Nodal pricing is usually related to congestion in [3], the distribution system should avoid overloading and ensure that service obligations are not compromised. However, another aspect of nodal pricing is price variation for the limit loss. In addition to low voltage operation using smaller conductors, many distribution systems have lower customer density and longer wiring. In general, in many systems where the distance is relatively short relative to transmission, the marginal loss of the distribution can be large. In [4], it was proposed to use nodal pricing in the distribution system to deliver the appropriate price signals to place DG resources and appropriately reward DG resources to reduce line losses. Power systems restructuring will promote the use of distributed generations in power distribution networks [5]. The different financial and technical benefits of integrating DGs into the energy systems have been extensively studied in the past [6]. Considering the nodal price behavior of ZIP load and RIC load, a

formula based on load flow is considered in [7]. For various load conditions including seasonal fluctuation, the optimum current including the readjusted marginal loss factor and price fluctuation is shown [8]. In [9], the optimum position of D-STATCOM was studied to increase the voltage and reduce the loss of RDS. The effects of wind power on the price value of the hourly distribution system by using the normal power flow method are in [10].

In this paper, the impact of D-STATCOM and Wind deployment, on the nodal prices using proposed an Optimal Power Flow approach in GAMS of 69 bus radial distribution system at every bus for each hour is studied. Optimal location and sizing found using sensitivity based methods.

Other sections of this paper are presented as: Sect. 2 Nomenclature. Section 3 Problem Formulation. Section 4 D-STATCOM and Wind models along with load model. Section 5 Steps to calculate nodal prices. Section 6 Results of the work done. Section 7 Conclusion of the paper.

2 Problem Formulations

2.1 Formulation of Optimal Load Flow

Mathematical modeling of optimum load flow is done by simultaneously optimizing the net operational cost and power loss since GAMS software having simultaneous optimizing feature and the equations are given from (1)–(11) as follows:

Minimizing

$$\sum\nolimits_i a_i P_{gi}^2 + b_i P_{gi} + c_i + TQ + TPdg \tag{1}$$

$$\mathrm{TQ} = cost\ of\ DSTATCOM\ Q = CF * \sum\nolimits_z Q_{\mathrm{D_STATCOM(z)}}$$

$$\mathrm{CF} = \text{cost factor}$$

$$\mathrm{TP_{dg}} = \sum\nolimits_k d_k P_{dgk}^2 + e_k P_{dgk} + f_k \sum\nolimits_{ij} P_{\mathrm{loss}ij} \tag{2}$$

Subjected to the following constraints:

Power balance Eqs. (3)–(6)

$$\mathrm{P_i} = P_{gi} - P_{di} \tag{3}$$

$$\mathrm{Q_i} = Q_{gi} - Q_{di} \tag{4}$$

$$P_i = U_i \sum_{j=1}^{n} U_j\left(G_{ij} * cos\theta_{ij} + B_{ij} * sin\theta_{ij}\right),\ i \in S_b \tag{5}$$

$$Q_i = U_i \sum_{j=1}^{n} U_j\left(G_{ij} * sin\theta_{ij} - B_{ij} * cos\theta_{ij}\right),\ i \in S_b \tag{6}$$

Real and reactive power generation limits,

$$P_{gi}^{\min} \le P_{gi} \le P_{gi}^{\max},\ i \in S_g \tag{7}$$

$$Q_{ri}^{\min} \le Q_{ri} \le Q_{ri}^{\max},\ i \in S_r \tag{8}$$

Bus voltage limits and limits of voltage angle,

$$U_i^{\min} \le U_i \le U_i^{\max},\ i \in S_b \tag{9}$$

$$\delta_i^{\min} \le \delta_i \le \delta_i^{\max},\ i \in S_b \tag{10}$$

$$|P_l| = \left|P_{ij}\right| = \left|U_i U_j\left(G_{ij} cos\theta_{ij} + B_{ij} sin\theta_{ij}\right) - U_i^2 G_{ij}\right| \le P_l^{\max}, l \in S_l \tag{11}$$

2.2 Determination of Nodal Prices Using MLCs

In the RDS, nodal prices or locational marginal prices can be calculated by the method given below in which to find prices at each buses of distribution network, marginal loss coefficients should be determined prior. Marginal loss coefficients are the coefficients which indicates marginal deviation in the total loss of real power due to variation in reactive and real power injections at a specific bus of the network [11].

$$\rho_{Pi} = \frac{\partial L_{\text{net}}}{\partial P_i} \tag{12}$$

$$\rho_{Qi} = \frac{\partial L_{\text{net}}}{\partial Q_i} \tag{13}$$

The supply point or substation is a point where the interconnection of two networks, i.e., distribution and transmission. The cost at this supply point is given by λ (USD/MWh) and found by using following equation;

$$\frac{\partial F}{\partial P_i} = \lambda = 2a_i P_{gi} + b_i \tag{14}$$

Equations for real and reactive power nodal price calculations are expressed as

$$N_{Pi} = (\lambda + \lambda * \rho_{Pi}) \tag{15}$$

$$N_{Qi} = \lambda * \rho_{Qi} \tag{16}$$

But, the cost of reactive power taken zero at substation point for the analysis purpose. For marginal loss coefficients calculation, Jacobean matrix is used in which the intermediary variables, node voltage magnitudes, and voltage angles from load flow outputs are utilized and it is given as

$$\begin{bmatrix} \rho_{Pi} \\ \rho_{Qi} \end{bmatrix} = \left[\begin{array}{ccc|ccc} \frac{\partial P_1}{\partial \theta_1} & \frac{\partial P_2}{\partial \theta_1} & \frac{\partial P_n}{\partial \theta_1} & \frac{\partial Q_1}{\partial \theta_1} & \frac{\partial Q_2}{\partial \theta_1} & \frac{\partial Q_n}{\partial \theta_1} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial P_1}{\partial \theta_n} & \frac{\partial P_2}{\partial \theta_n} & \frac{\partial P_n}{\partial \theta_n} & \frac{\partial Q_1}{\partial \theta_n} & \frac{\partial Q_2}{\partial \theta_n} & \frac{\partial Q_n}{\partial \theta_n} \\ \hline \frac{\partial P_1}{\partial U_1} & \frac{\partial P_2}{\partial U_1} & \frac{\partial P_n}{\partial U_1} & \frac{\partial Q_1}{\partial U_1} & \frac{\partial Q_2}{\partial U_1} & \frac{\partial Q_n}{\partial U_1} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial P_1}{\partial U_n} & \frac{\partial P_2}{\partial U_n} & \frac{\partial P_n}{\partial U_n} & \frac{\partial Q_1}{\partial U_n} & \frac{\partial Q_2}{\partial U_n} & \frac{\partial Q_n}{\partial U_n} \end{array}\right]^{-1} \left[\begin{array}{c} \frac{\partial L}{\partial \theta_1} \\ \vdots \\ \frac{\partial L}{\partial \theta_n} \\ \hline \frac{\partial L}{\partial U_1} \\ \vdots \\ \frac{\partial L}{\partial U_n} \end{array}\right] \tag{17}$$

$$\frac{\partial P_i}{\partial \theta_j} = U_i * U_j\left[G_{ij} * sin\theta_{ij} - B_{ij} * cos\theta_{ij}\right] \tag{18}$$

$$\frac{\partial P_i}{\partial \theta_i} = -B_{ii}U_i^2 - \sum_{j=1}^{n} U_i U_j\left[G_{ij}sin\theta_{ij} - B_{ij}cos\theta_{ij}\right] \tag{19}$$

$$\frac{\partial P_i}{\partial U_j} = U_i\left[G_{ij} * cos\theta_{ij} + B_{ij} * sin\theta_{ij}\right] \tag{20}$$

$$\frac{\partial P_i}{\partial U_i} = G_{ii}U_i + \sum_{j=1}^{n} U_j\left[G_{ij}cos\theta_{ij} + B_{ij}sin\theta_{ij}\right] \tag{21}$$

$$\frac{\partial Q_i}{\partial \theta_j} = -U_i U_j\left[G_{ij}cos\theta_{ij} + B_{ij}sin\theta_{ij}\right] \tag{22}$$

$$\frac{\partial Q_i}{\partial \theta_i} = -G_{ii}U_i^2 + \sum_{j=1}^{n} U_i U_j\left[G_{ij}cos\theta_{ij} + B_{ij}sin\theta_{ij}\right] \tag{23}$$

$$\frac{\partial Q_i}{\partial U_j} = U_i\left[G_{ij}sin\theta_{ij} - B_{ij}cos\theta_{ij}\right] \tag{24}$$

$$\frac{\partial Q_i}{\partial U_j} = -B_{ii}U_i + \sum_{j=1}^{n} U_j\left[G_{ij}sin\theta_{ij} - B_{ij}cos\theta_{ij}\right] \tag{25}$$

The distribution network total loss is as follows:

$$L_{net} = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} G_{ij}\left[U_i^2 + U_j^2 - 2U_iU_jcos\theta_{ij}\right] \tag{26}$$

Partial derivative of the Eq. (26) with U_i *and* θ_i are as follows:

$$\frac{\partial L}{\partial \theta_i} = 2\sum_{j=1}^{n} U_iU_jG_{ij}sin\theta_{ij} \tag{27}$$

$$\frac{\partial L}{\partial U_i} = 2\sum_{j=1}^{n} G_{ij}\left[U_i - U_jcos\theta_{ij}\right] \tag{28}$$

Reconciliation of MLCs

From the mentioned above analysis, losses are greater than the actual with MLCs as in [12]. Consequently, marginal loss coefficients are reconstructed to account price of actual loss of network with the multiplication factor, i.e., Reconciliation factor (R.F) and is given as

$$R.F = \frac{L_{\text{net}}}{L_{\text{approx}}} \tag{29}$$

Whereas L_{approx} is given by

$$L_{\text{approx}} = \sum_{i=1}^{n}\left[\rho_{Pi} * P_i + \rho_{Qi} * Q_i\right] \tag{30}$$

Modified nodal price equation for real and reactive power with reconciliation is as follows:

$$N_{Pi} = \lambda * (1 + Rf * \rho_{Pi}) \tag{31}$$

$$N_{Qi} = \lambda * Rf * \rho_{Qi} \tag{32}$$

3 Wind and D-STATCOM, Load Modeling

D-STATCOM and Wind optimal location in 69 bus system is obtained by using sensitivity based methods, they are as follows:

3.1 Index Vector

The vector index of bus i is given as

$$\text{Index}[i] = \frac{1}{U_i^2} + \frac{I_q[l]}{I_p[l]} + \frac{Q_{\text{eff}}[i]}{Q_{\text{total}}} \tag{33}$$

Whereas, $I_p[l]$ and $I_q[l]$ are real and reactive current component of branch l, $Q_{\text{eff}}[i]$ and Q_{total} are the effective reactive load at node i and the total reactive load of system. The normalized voltage $\left(V_i^{\text{norm}} = U_i/0.95\right)$, the node which is having high index vector and normalized voltage less than 1.01 is selected to be an optimal location for the D-STATCOM placement (Figs. 1, 2).

From the above plots, bus 61 has highest index and is less than normalized voltage, i.e.,1.01p.u. So, it is selected as optimal location for D-STATCOM.

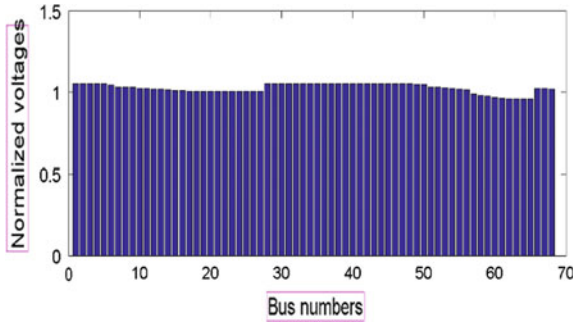


Fig. 1 Normalized voltages

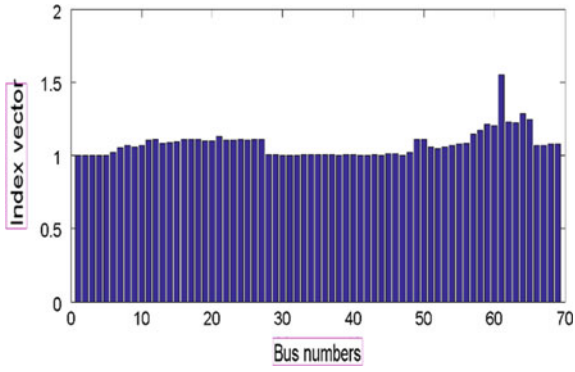


Fig. 2 Index vector

3.2 Combined Power Loss Sensitivity

In [13], the node which is having higher sensitivity factor and below the 1.01 normalized voltage is selected as the optimum location for the wind DG placement which is having unity power factor (UPF) is given as

$$\frac{\partial P_{loss}}{\partial Q_i} = \frac{2 * Q_i * R_l}{U_i^2} \tag{34}$$

$$\frac{\partial Q_{loss}}{\partial Q_i} = \frac{2 * Q_i * X_l}{U_i^2} \tag{35}$$

$$\frac{\partial P_{loss}}{\partial P_i} = \frac{2 * P_i * R_l}{U_i^2} \tag{36}$$

$$\frac{\partial Q_{loss}}{\partial P_i} = \frac{2 * P_i * X_l}{U_i^2} \tag{37}$$

Partial derivative of combined power loss with respect to reactive power is

$$\frac{\partial S_{loss}}{\partial Q_i} = \frac{\partial P_{loss}}{\partial Q_i} + j\frac{\partial Q_{loss}}{\partial Q_i} \tag{38}$$

Partial derivative of combined power loss with respect to real power is

$$\frac{\partial S_{loss}}{\partial P_i} = \frac{\partial P_{loss}}{\partial P_i} + j\frac{\partial Q_{loss}}{\partial P_i} \tag{39}$$

60 branch, i.e., 61st bus is chosen as an optimal place for the wind distributed generation from the above-given plots.

After optimal location selection of wind and D-STATCOM, optimum size selection can be done by using GAMS CONOPT NLP solver (Fig. 3).

The modified power balance equations with wind and D-STATCOM are modeled in GAMS software as follows:

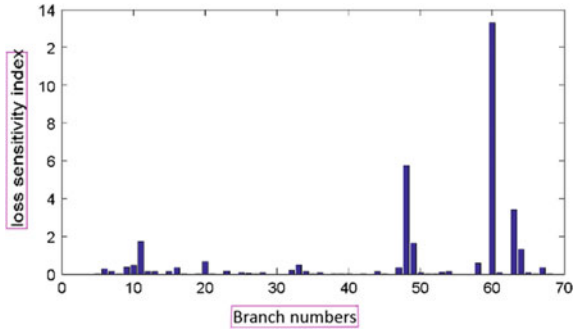


Fig. 3 Combined power loss sensitivity factor

$$P_i = P_{gi} + \sum_k P_{wind(k)} - P_{di} \tag{40}$$

$$Q_i = Q_{gi} + \sum_z Q_{D_STATCOM(z)} - Q_{di} \tag{41}$$

$$P^{min}_{wind(k)} \leq P_{wind(k)} \leq P^{max}_{wind(k)} \tag{42}$$

$$Q^{min}_{D_STATCOM(z)} \leq Q_{D_STATCOM(z)} \leq Q^{max}_{D-STATCOM(z)} \tag{43}$$

Here, the lower and upper limits of the wind real power and D-STATCOM reactive power is taken has (30–40) % and two-thirds of the total real and reactive power load of the network to be considered.

3.3 Modeling of Load

For the optimum load flow, here two types of loads are considered namely ZIP and RIC.ZIP load is combination of constant impedance, constant current, constant power, whereas RIC is a Residential, Commercial and Industrial loads. These loads are modeled depending on voltage, real, and reactive powers at their buses and it changes with time

$$P_{di}(t) = P_{di0}(t)\left[\alpha(t)\left(\frac{U_i(t)}{U_0}\right)^{npr} + \beta(t)\left(\frac{U_i(t)}{U_0}\right)^{npc} + \gamma(t)\left(\frac{U_i(t)}{U_0}\right)^{npi}\right] \tag{44}$$

$$Q_{di}(t) = Q_{di0}(t)\left[\alpha(t)\left(\frac{U_i(t)}{U_0}\right)^{nqr} + \beta(t)\left(\frac{U_i(t)}{U_0}\right)^{nqc} + \gamma(t)\left(\frac{U_i(t)}{U_0}\right)^{nqi}\right] \tag{45}$$

$P_{di0}(t)$, $Q_{di0}(t)$ are the base case real and reactive power load for each node i.

U_0 is a base voltage, generally considered as1p.u. $\alpha(t)$, $\beta(t)$ and $\gamma(t)$ are the fractions of RIC load at node *i* for the period *t* whereas nqr, nqc and nqi, npr, npc, and npiare variable to have various loads as Constant Impedance (CZ), Constant Current (CI), and Constant Power (CP) in [14]. And the sum of these fractions must be one.

$$\alpha(t) + \beta(t) + \gamma(t) = 1 \tag{46}$$

4 Steps to Calculate Nodal Prices

The following steps considered to determine the real and reactive power nodal prices, they are as follows:

Step1: System bus data and line data read from the MATLAB.
Step2: With the help of MATLAB and GAMS interfacing, run the optimal power flow and get the values of voltage magnitudes and angles, etc. from GAMS.
Step3: By using marginal loss coefficients (MLCs) expressions in MATLAB nodal prices are calculated.
Step4: Graphical results are also obtained from MATLAB.
Step5: Finally obtained results are compared.

5 Results and Discussions

In this section a case study of IEEE 69 bus RDS with a base of 100MVA, 12.66 kV is considered. The total load on this system is 3.8013 + j*2.6936MVA [15, 16]. Simulation results are obtained in MATLAB7.0.4 environment and optimization problem is solved in GAMS23.4 CONOPT solver and also interfacing of above software were made using matout.gms and gams.dil files for the nodal price calculations [17]. The active and reactive power losses obtained by the OPF method are 224.834 kw and 102.993KVAR for normal load respectively. Whereas for ZIP load, 173.9912 kw real power loss to 20.7316 kw is reduced with D-STATCOM and Wind. But the operational cost 900000(₹/MWh) is increased to 3.0094E + 008(₹/MWh). For the comparison purpose, any one of the buses can be picked up randomly. So, in this paper bus 13 is considered for analysis (Table 1).

From the above table, it is clearly noticed that nodal prices with optimal power flow got better than the normal load flow approach i.e., Forward–backward sweep method and also observed that the maximum and minimum Active power nodal prices are 40.95801(USD/MWh) at 8 h and 40.86685(USD/MWh) at 20 h, respectively, since it is having peak and lower loads. Whereas maximum and minimum reactive power nodal prices are 0.361155(USD/MWh) at 5 h and 0.332181(USD/MWh) at 18 h, respectively, for 13 bus ZIP load with Wind and D-STATCOM. Likewise, it is observed for every bus of the 69 bus RDS system for both the loads RIC and ZIP, respectively. So, it is clearly noticed here that the nodal prices of active power and Reactive power are reduced by using the wind power sources and D-STATCOM individually as well as considering them at a time in the test network (Table 2).

Graphical representations of nodal prices with reconciliation for24 h of 69 bus RDS is as follows,
For ZIP load (Fig. 4, 5, 6, 7, 8 and 9).
For RIC load (Fig. 10, 11, 12, 13, 14 and 15).

Table 1 Active and reactive power nodal prices comparison for ZIP load

Hours	ZIP load									
	Without D-STATCOM and wind				Optimal load flow					
	Normal load flow		Optimal load flow		D-STATCOM		Wind		Wind and D-STATCOM	
	Npi1	Nqi1	Npi1	Nqi1	Npi1	Nqi1	Npi1	Nqi1	Npi1	Nqi1
1	44.2997	0.6091	41.47256	0.624904	41.05656	0.065827	41.04095	0.423932	40.89471	0.339995
2	44.2992	0.6096	41.48511	0.630547	41.06416	0.067016	41.04887	0.427378	40.89864	0.338203
3	44.2995	0.6082	41.49553	0.635234	41.06977	0.067545	41.05546	0.430243	40.90192	0.34282
4	44.2981	0.6121	41.49957	0.637053	41.07522	0.070261	41.05802	0.431357	40.9032	0.356012
5	44.2944	0.6231	41.4846	0.630318	41.06457	0.067518	41.04857	0.427246	40.89856	0.361155
6	44.2930	0.6274	41.5233	0.647729	41.08746	0.070803	41.07301	0.437883	40.91053	0.361144
7	44.2929	0.6275	41.63413	0.697662	41.15476	0.081963	41.14334	0.468533	40.94472	0.360989
8	44.2930	0.6273	41.67716	0.717065	41.1812	0.086608	41.1708	0.480507	40.95803	0.359285
9	44.2934	0.6259	41.67698	0.716983	41.18145	0.086822	41.17069	0.480459	40.95801	0.354276
10	44.2948	0.6217	41.67482	0.71601	41.1769	0.084623	41.16931	0.479858	40.95759	0.358233
11	44.2937	0.6250	41.66101	0.709776	41.16926	0.083584	41.16053	0.476022	40.95318	0.357698
12	44.2938	0.6246	41.61805	0.690371	41.14649	0.081332	41.13321	0.464102	40.94023	0.355319
13	44.2945	0.6226	41.65208	0.705747	41.16439	0.082968	41.15484	0.47354	40.95046	0.355084
14	44.2945	0.6224	41.64745	0.703656	41.16118	0.082268	41.1519	0.472255	40.94907	0.353241
15	44.2950	0.6209	41.62687	0.69437	41.15059	0.081328	41.13882	0.466549	40.94907	0.345716
16	44.2971	0.6146	41.62511	0.693576	41.15071	0.082008	41.13768	0.466054	40.94232	0.336395
17	44.2996	0.6070	41.60943	0.686506	41.13778	0.078188	41.12773	0.46171	40.93754	0.326
18	44.3025	0.5982	41.54621	0.658024	41.10279	0.074094	41.08762	0.44423	40.91801	0.332181

(continued)

Table 1 (continued)

Hours	ZIP load									
	Without D-STATCOM and wind				Optimal load flow					
	Normal load flow		Optimal load flow		D-STATCOM		Wind		Wind and D-STATCOM	
	Npi1	Nqi1	Npi1	Nqi1	Npi1	Nqi1	Npi1	Nqi1	Npi1	Nqi1
19	44.3022	0.5983	41.46939	0.623473	41.05503	0.065824	41.039	0.423075	40.89387	0.332187
20	44.3007	0.6036	41.41075	0.597536	41.0169	0.060326	40.99993	0.40646	40.86685	0.339995
21	44.3007	0.6036	41.4351	0.608074	41.03541	0.063638	41.01736	0.413672	40.88295	0.338203
22	44.3006	0.6040	41.43543	0.608223	41.0337	0.062073	41.01755	0.413758	40.88296	0.34282
23	44.3001	0.6054	41.43955	0.610072	41.03618	0.062451	41.02014	0.414884	40.88427	0.332693
24	44.3008	0.6053	41.45393	0.616531	41.04559	0.064333	41.0292	0.418822	40.88885	0.334457

Npi1* is Active power nodal price with reconciliation in USD/MWh
Nqi1* is Reactive power nodal price with reconciliation in USD/MWh

Table 2 Active and reactive power nodal prices comparison for RIC load

Hours	RIC load									
	Without D-STATCOM and wind				Optimal load flow					
	Normal load flow		Optimal load flow		D-STATCOM		Wind		D-STATCOM and Wind	
	Npi1	Nqi1	Npi1	Nqi1	Npi1	Nqi1	Npi1	Nqi1	Npi1	Nqi1
1	46.1216	1.3316	41.47256	0.624904	41.05656	0.065827	41.04095	0.423932	40.89471	0.336718
2	44.3357	0.5037	41.48511	0.630547	41.06416	0.067016	41.04887	0.427378	40.89864	0.338234
3	44.3353	0.5014	41.49553	0.635234	41.06977	0.067545	41.05546	0.430243	40.90192	0.339498
4	44.3349	0.5014	41.49957	0.637053	41.07522	0.070261	41.05802	0.431357	40.9032	0.339995
5	44.3328	0.5074	41.4846	0.630318	41.06457	0.067518	41.04857	0.427246	40.89856	0.338203
6	44.3377	0.4921	41.5233	0.647729	41.08746	0.070803	41.07301	0.437883	40.91053	0.34282
7	44.3481	0.4607	41.63413	0.697662	41.15476	0.081963	41.14334	0.468533	40.94472	0.356012
8	44.3488	0.4584	41.67716	0.717065	41.1812	0.086608	41.1708	0.480507	40.95803	0.361155
9	44.3484	0.4596	41.67698	0.716983	41.18145	0.086822	41.17069	0.480459	40.95801	0.361144
10	44.3483	0.4596	41.67482	0.71601	41.1769	0.084623	41.16931	0.479858	40.95759	0.360989
11	44.3464	0.4654	41.66101	0.709776	41.16926	0.083584	41.16053	0.476022	40.95318	0.359285
12	44.3410	0.4818	41.61805	0.690371	41.14649	0.081332	41.13321	0.464102	40.94023	0.354276
13	44.3456	0.4679	41.65208	0.705747	41.16439	0.082968	41.15484	0.47354	40.95046	0.358233
14	44.3448	0.4702	41.64745	0.703656	41.16118	0.082268	41.1519	0.472255	40.94907	0.357698
15	44.3418	0.4794	41.62687	0.69437	41.15059	0.081328	41.13882	0.466549	40.94292	0.355319

(continued)

Table 2 (continued)

Hours	RIC load									
	Without D-STATCOM and wind				Optimal load flow					
	Normal load flow		Optimal load flow		D-STATCOM		Wind		D-STATCOM and Wind	
	Npi1	Nqi1	Npi1	Nqi1	Npi1	Nqi1	Npi1	Nqi1	Npi1	Nqi1
16	44.3424	0.4773	41.62511	0.693576	41.15071	0.082008	41.13768	0.466054	40.94232	0.355084
17	44.3405	0.4831	41.60943	0.686506	41.13778	0.078188	41.12773	0.46171	40.93754	0.353241
18	44.3342	0.5023	41.54621	0.658024	41.10279	0.074094	41.08762	0.44423	40.91801	0.345716
19	44.3297	0.5157	41.46939	0.623473	41.05503	0.065824	41.039	0.423075	40.89387	0.336395
20	44.3294	0.5167	41.41075	0.597536	41.0169	0.060326	40.99993	0.40646	40.86685	0.326
21	44.3283	0.5201	41.4351	0.608074	41.03541	0.063638	41.01736	0.413672	40.88295	0.332181
22	44.3290	0.5180	41.43543	0.608223	41.0337	0.062073	41.01755	0.413758	40.88296	0.332187
23	44.3294	0.5167	41.43955	0.610072	41.03618	0.062451	41.02014	0.414884	40.88427	0.332693
24	44.3306	0.5130	41.45393	0.616531	41.04559	0.064333	41.0292	0.418822	40.88885	0.334457

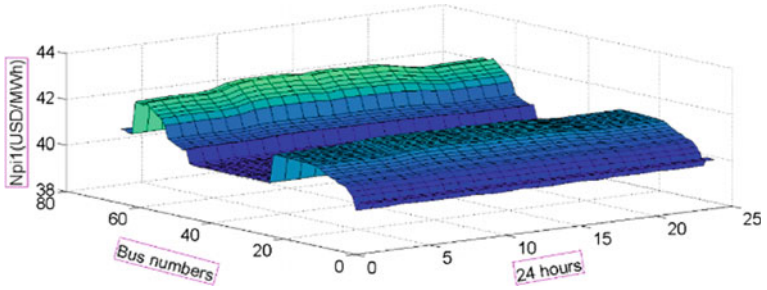


Fig. 4 Active power nodal price with D-STATCOM

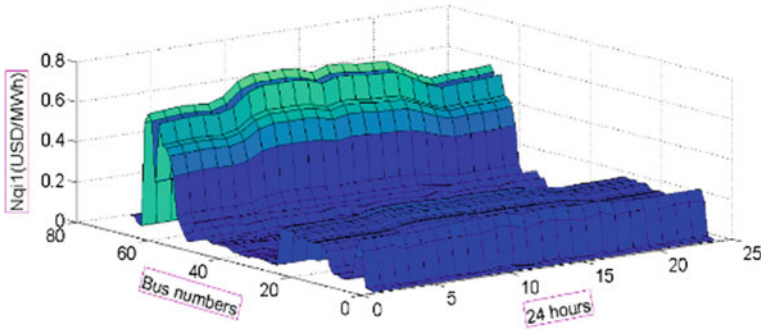


Fig. 5 Reactive power nodal price with D-STATCOM

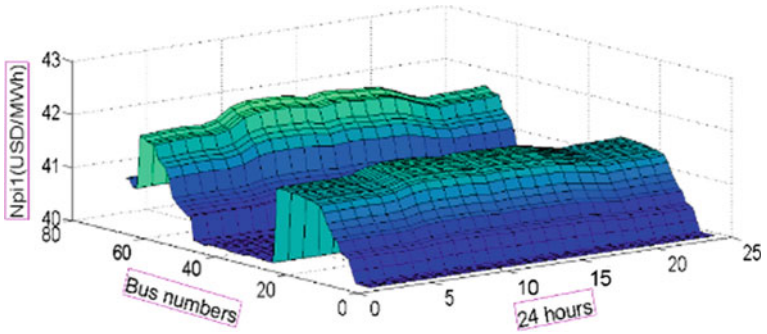


Fig. 6 Active power nodal prices with wind

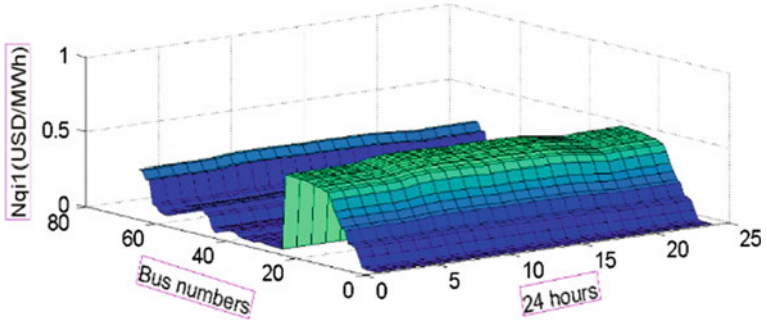


Fig. 7 Reactive power nodal prices with wind

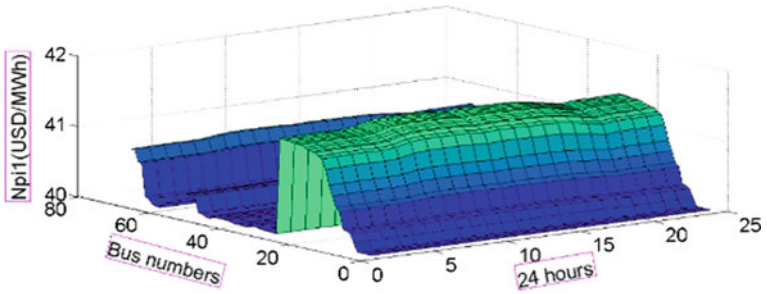


Fig. 8 Active power nodal price with wind and D-STATCOM

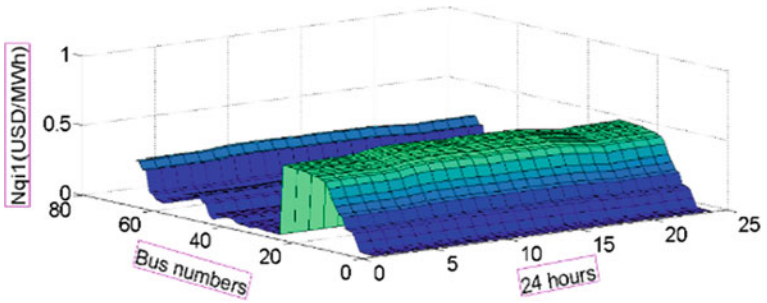


Fig. 9 Reactive power nodal price with wind and D-STATCOM

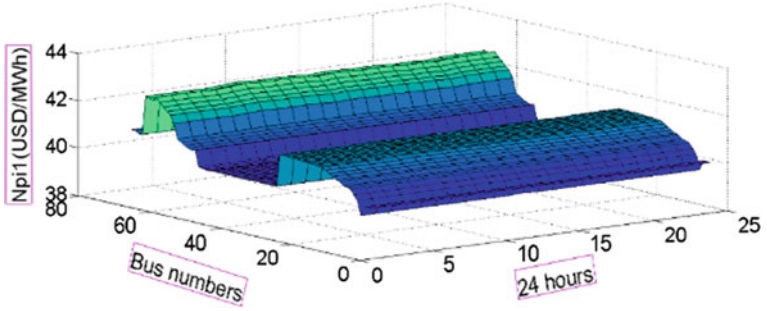


Fig. 10 Active power nodal prices with D-STATCOM

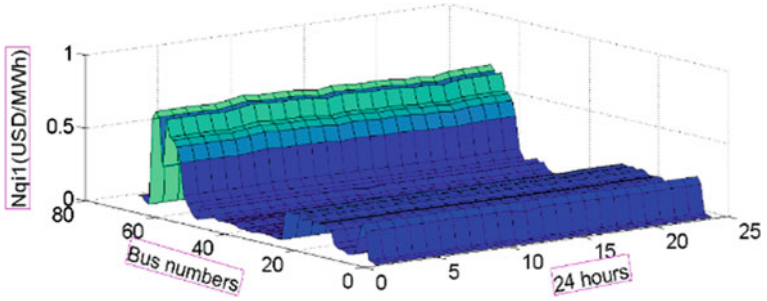


Fig. 11 Reactive power nodal prices with D-STATCOM

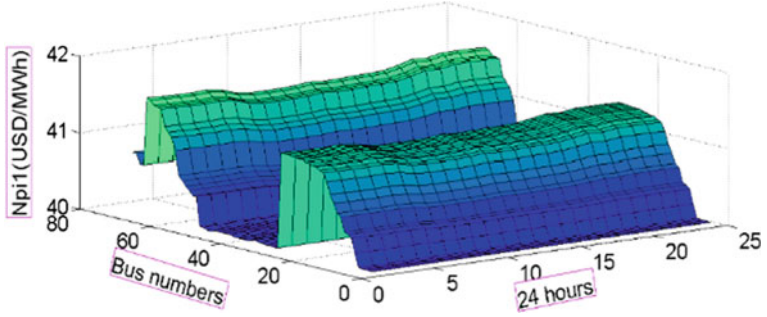


Fig. 12 Active power nodal prices with wind

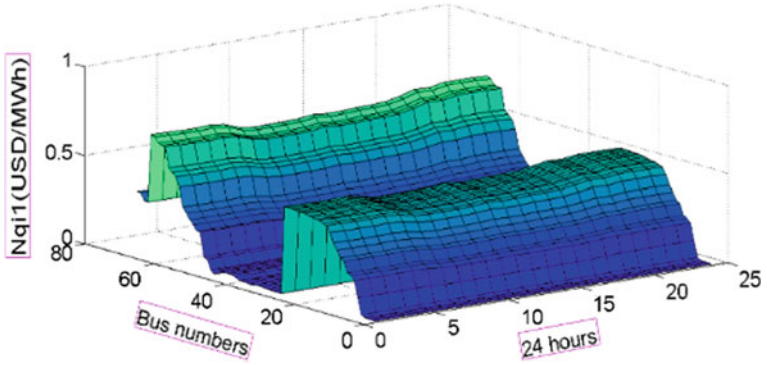


Fig. 13 Reactive power nodal prices with wind

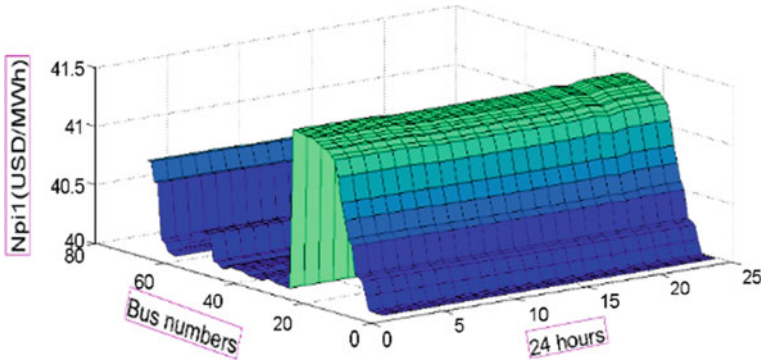


Fig. 14 Active power nodal prices with wind and D-STATCOM

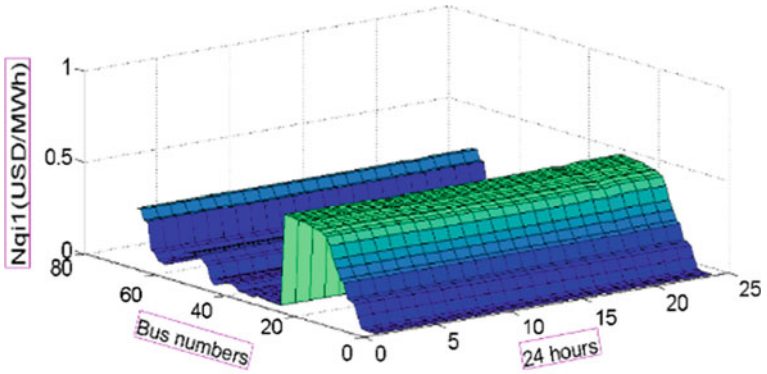


Fig. 15 Reactive power nodal prices with wind and D-STATCOM

6 Conclusion

In Restructured Electricity Market, the determination of distribution systems pricing is a really great challenging task for the electrical professionals. So, in this work Active and Reactive power node prices of the IEEE 69 bus RDS system are obtained and results are compared with Wind, D-STATCOM individually placed in the system and with their simultaneous placement. The prices obtained are lower with DG and D-STATCOM as compared to prices when DG and D-STATCOM were placed individually. The total power loss reduction was higher with DG and D-STATCOM as compared to with DG and D-STATCOM placed separately, i.e., reduced to 20.7316 kw and the node voltages have been improved by using D-STATCOM, but economically ineffective and also nodal prices with optimum power flow gave the better results over the normal load flow.

References

1. M. Ghayeni, R. Ghazi, Transmission network cost allocation with nodal pricing approach based on ramsay pricing concept. IET Gener. Transm. Distrib. **5**, 384–392 (2011)
2. F.C. Schweppe, M.C. Caramanis, R.D. Tabors, R.E. Bohn, Spot Pricing of Electricity (1988)
3. W.W. Hogan, Nodes and zones in electricity markets: seeking simplified congestion pricing, in *Designing Competative Electricity markets*, ed. H.-P. Chao, H.G. Huntington (1998)
4. P.M. Sotkiewicz, J.M. Vignolo, Nodal pricing for distribution networks: efficient pricing for efficiency enhancing DG. Power. Syst. IEEE Trans. **21**, 1013–1014 (2006)
5. R.C. Dugan, T.E. Mcdermont, Distributed generation. IEEE Ind. Appl. Mag. (2002), pp. 19–25
6. L. F. Ochoa, A. Padilha-Feltrin, G.P. Harrison, Evaluating distributed generation impacts with a multiobjective index. IEEE Trans. Power Delivery **21**(3), 1452–1458 (2006)
7. K. Sooraj Narayan, A. Kumar, Distribution system nodal pricing analysis with realistic ZIP load and variable wind power source. IEEE India International Conference (INDICON) (2015)
8. K. Polisetti, A. Kumar, Distribution system nodal prices determination for realistic ZIP and Seasonal loads: an optimal power flow approach **25**, 702–709 (2016)
9. M Subbaramiah, S.M. Suhail Hussain, An Analytical Approach for Optimal Location of DSTATCOM in RDS
10. P.M. Sotkiewicz, J.M. Vignolo, Towards a cost causation – based tariff for distribution network with DG. IEEE Trans. Power Syst. **22**, 1051–1060 (2007)
11. T.E. Mcdermont, Distributed generation. IEEE Ind. Appl. Mag. 19–25 (2002). R.C. Dugan
12. S. Gopiya Naik, D.K. Khatod, M.P. Sharma, Optimal allocation of distributed generation in distributed system for loss reduction, in *2012-IACSIT Coimbatore Conferences,* pp. 42–46
13. D.Q. Hung, N. Mithulananthan, K.Y. Lee, Optimal placement of dispatchable and nondispatchable renewable DG units in distribution networks for minimizing energy loss. Int. J. Electr. Power Energy Syst. **55**, 179–186 (2014)
14. V.V.S.N. Murthy, A. Kumar, Mesh distribution system analysis in presence of distributed generation with time varying load model. Int. J. Elect. Power Energ. Syst. **62,** 836–854 (2014)
15. R.D. Zimmerman, C.E. Murillo-Sánchez, R.J. Thomas, MATPOWER: steady-state operations, planning and analysis tools for power systems research and education. IEEE Trans. Power Syst. **26**(1), 12–19 (2011)

16. R. Parasher, Load Flow Analysis of Radial Distribution Network using Linear Data Structure (2014) arXiv:14034702
17. P. Vijay Babu, S.P. Singh, Optimal placement of DG in distribution network for power loss minimization using NLP and PLS technique, in *ICAER* (2015), pp. 15–17

Enhancement of Bandwidth and Gain of a Slotted E-Shaped Patch Antenna



Yatharth Shankar Misra and Ramesh Kumar Verma

1 Introduction

Microstrip patch antennas have been intensively studied and developed since the concept was first introduced in the 1950s. The study related to the microstrip patch antenna has been progressive over the years. With the fast development of wireless communication, there is a need for compact, simple, and light microstrip antenna. The microstrip patch antenna possesses some disadvantages like narrow bandwidth and low gain which has many effects when using the antenna in communication. So to reduce this effect we are advancing the design of microstrip patch antenna so that it can achieve a wide range of bandwidth, high gain, high performance, and these are compatible with communication devices.

The typical substrate which is used for antenna designing has the dielectric constant (ϵ_r) typically ranges [1] from $2.2 \leq \epsilon_r \leq 12$. The proposed antenna is designed by using a substrate called glass epoxy or FR-4 with dielectric constant ($\epsilon_r = 4.4$). The dielectrics are used for better electrical stability. It helps in reducing the size of the antenna making it compact and simple [2]. It also helps to produce displacement current which produces time-varying magnetic field and electric field and a propagating EM field. Hence, a substrate can enhance antennas radiation capability. The design frequency used over here is 2.38 GHz. In this paper, the bandwidth of the antenna is enhanced by using the technique of redesigning the rectangular patch by cutting an E-shaped patch with a rectangular hole. After the simulation of the antenna design, the antenna has a frequency band that lies between 3.273 GHz and 5.444 GHz

Y. S. Misra (✉) · R. K. Verma
ECE Department, Bundelkhand Institute of Engineering & Technology, Jhansi, UP, India
e-mail: mishra.yatharth194@gmail.com

R. K. Verma
e-mail: ramesh85.ec@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_76

which is suitable for wireless communication devices. This antenna design has been given a 50 Ω microstrip feed line.

Zakir Ali [3] et al. represented a patch that has been designed as E-shaped and its operating frequency is from 5 to 6 GHz. The bandwidth achieved in this paper is 14.90%. The E-shaped designed is used for the increment of bandwidth. Aug [4] et al., designed an antenna selecting an E-shaped patch that can work in multiple frequencies ranging from 5.05 to 5.88 GHz. Tong [5] et al. proposed a special U-shaped patch not only to increase impedance but also the axial ratio bandwidth. Over here 8 dB gain is achieved. Tong [6] et al. designed a U-shaped patch to increase bandwidth with a thick substrate. The achieved bandwidths are 25 and 27%. Maci [7] et al. designed a new antenna and analyzed it. This antenna provides a dual-frequency feature. The frequency of the antenna can be controlled between 1.6 GHz and 2 GHz. Mandal [8] et al. introduced a new concept that has been introduced by selecting U shaped and the substrate sleeted over here is polytetrafluoroethylene. The impedance BW that achieved is 86.79% and gain is 4.1 dB. In Zhang [9]'s paper, a wideband aperture coupled patch antenna is designed that is U shaped for the increment of bandwidth. In Khodai [10] et al.'s paper, a special patch has been designed in U-shaped to increase bandwidth using probe feed.

2 Formula Used for Antenna Design

The formula used for calculation of the width of the antenna is given below

$$\mathrm{W} = \frac{c}{2fr}\sqrt{\frac{2}{\in r+1}} \tag{1}$$

Here c stands for velocity of light whose value is (3×10^3), where f_r is the frequency at which antenna is being designed, over here W is used for describing the width of the patch, and ϵ_{reff} is defined as the effective dielectric constant given as [11]

$$\epsilon_{reff} = \frac{\in r+1}{2} + \frac{\in r-1}{2}\left(1+12\frac{h}{w}\right)^{-1/2} \tag{2}$$

The increment in the length which is used in the antenna is defined as ΔL and the formula is given below [11]

$$\frac{\Delta L}{h} = 0.412\frac{(\varepsilon reff+0.8)(W/h+0.264)}{(\varepsilon reff-0.258)(W/h+0.8)} \tag{3}$$

After solving the above equation, we can define the length of the patch as [11]

Table 1 Antenna parameters and dimensions

S. No.	Parameters	Values
1	Frequency used for design	2.38 GHz
2	Dielectric constant of substrate	4.4
3	Height of the substrate	1.6 mm
4	Loss tangent (tan δ)	0.025
5	Width taken for patch design (W_p)	38.30 mm
6	Length taken for the patch (L_p)	29.65 mm
7	Calculated ground width (W_g)	47.90 mm
8	Calculated ground length (L_g)	39.25 mm

$$L = \frac{c}{2fr\sqrt{\varepsilon reff}} - 2\Delta L \tag{4}$$

It is important to calculate the length and width of the antenna ground and the formula for finding it is given below [11]

$$L_g = 6h + L \tag{5}$$

$$W_g = 6h + W \tag{6}$$

3 Demonstration of Antenna Design

The proposed antenna is designed by using a substrate called glass epoxy (FR-4) which has a dielectric constant 4.4 and is made with a constant height of 1.6 mm and loss tangent (tan δ) 0.025. The length and width of the antenna are calculated with Eqs. (1)–(4). The length and width of the ground are calculated with the formula (5) and (6). The patch length and width have been taken out by using the formula which is given above so after putting values, the width and length are 38.3 mm and 29.65 mm, respectively. The length and width of the ground are simply calculated by adding 9.6 mm in the patch length and width which gives our desired values as 39.25 mm and 47.9 mm, respectively.

Antenna parameters and dimensions are given in Table 1.

4 Procedures for Designing of Antenna

For designing of microstrip antenna at 2.38 GHz frequency, IE3D software is used for designing and simulation [12]. The design process is starting from lower left values of x and y at (0, 0). The microstrip feed line of 50 Ω is placed at lower left corner of patch. The designed antenna is feeding with 50 Ω microstrip line feed. Step

by step designing of antenna geometry is shown in Fig. 1. Parameters of slot and notches are given in Table 2.

Initially in the first step, antenna (a) as shown in Fig. 1a is designed by cutting a notch of size 12.03 × 4.18 mm at the right lower edge of the patch. The bandwidth of this antenna is 27.30% with −20.84 dB return loss. In the second step, by cutting another notch of size 12.3 × 5 mm and a slot of size 10.75 mm × 3.6 mm in the patch. The bandwidth is obtained 21.37% with −23.56 dB return loss. Finally, a notch of size 12.03 × 6.18 mm is cut at the right middle of the patch. The bandwidth of the antenna is obtained 49.81% between 3.273 and 5.444 GHz resonating at 4.889 GHz with high return loss of −29.49 dB. Comparative result is shown in Table 3.

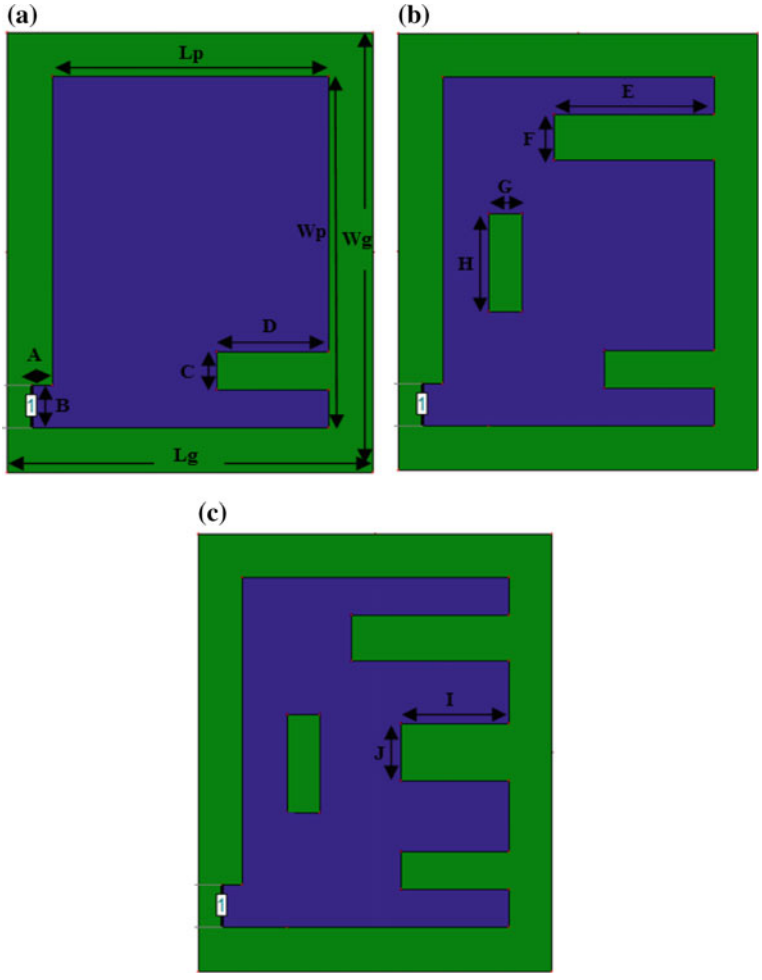


Fig. 1 Proposed design of E-shaped patch antenna

Table 2 Dimensional values of slotted patch

S. No.	Variables	Values (mm)
1	A	2.23
2	B	4.65
3	C	4.18
4	D	12.03
5	E	12.13
6	F	5
7	G	3.6
8	H	10.75
9	I	12.03
10	J	6.18

Table 3 Comparative results of three antenna designs

Antenna	Resonant frequency (GHz)	Return loss (dB)	Bandwidth (%)
Antenna (a)	4.737	−20.84	27.30
Antenna (b)	4.788	−23.56	21.37
Antenna (c)	4.889	−29.49	49.81

5 Discussions of Simulation Results

Microstrip patch antenna has some drawbacks like low efficiency, low gain, and low power handling capability. In the proposed antenna design, we are trying to remove the above drawbacks. The proposed antenna is simulated from 1 to 6 GHz. All the simulated results of the antenna are shown in Figs. 2, 3, 4, 5, 6, 7, 8, 9. After the simulation process, the achieved bandwidth of the slotted antenna is 49.81% at 4.89 GHz resonant frequency with −29.49 dB return loss. A comparison between the three designed antennas graph is shown in Fig. 2. In Fig. 3, it can be seen that VSWR of the antenna at resonant frequency 4.89 GHz is 1.068. Figure 4 shows the calculated gain of antenna which is 7.43 dB. The directivity of the antenna is 7.44 dB as shown in Fig. 5. Figure 6 shows the efficiency of the antenna in percentage. The efficiency of the antenna is increased and achieved 99.89% at 4.89 GHz resonant frequency. Smith chart of the slotted antenna is presented in Fig. 7. 2D and 3D radiation patterns of the slotted antenna are shown in Figs. 8 and 9, respectively.

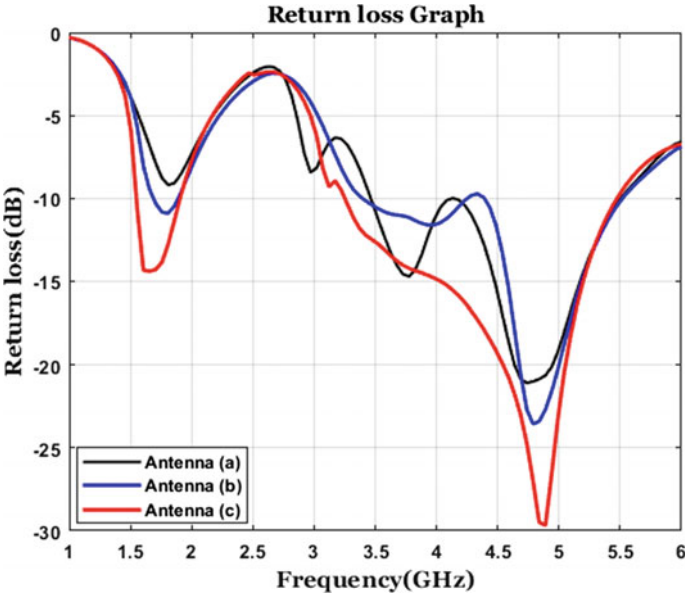


Fig. 2 Return loss versus frequency graph

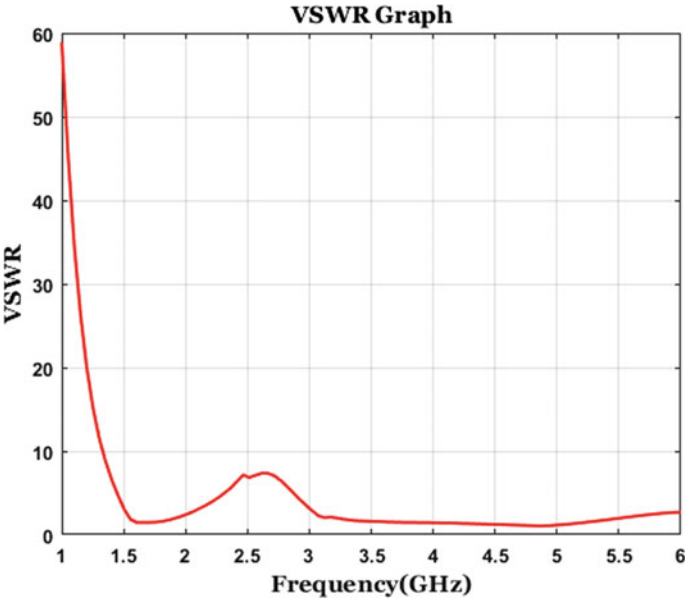


Fig. 3 Graph showing VSWR of the slotted antenna

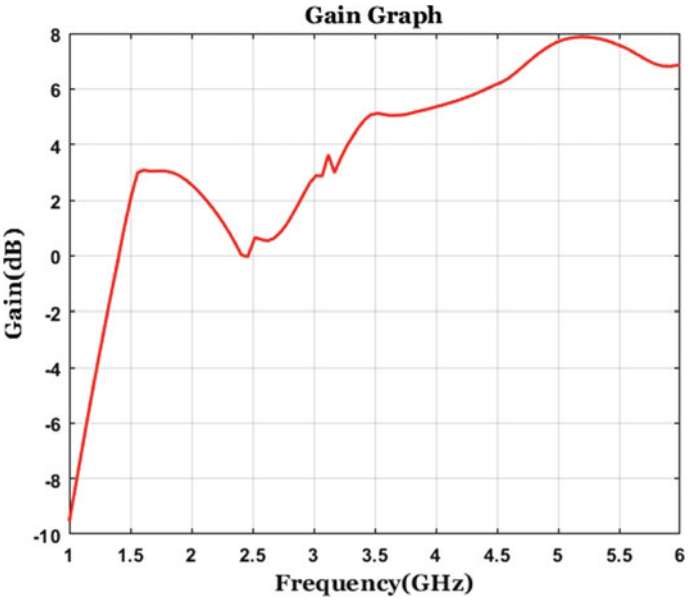


Fig. 4 Gain versus frequency graph of slotted antenna

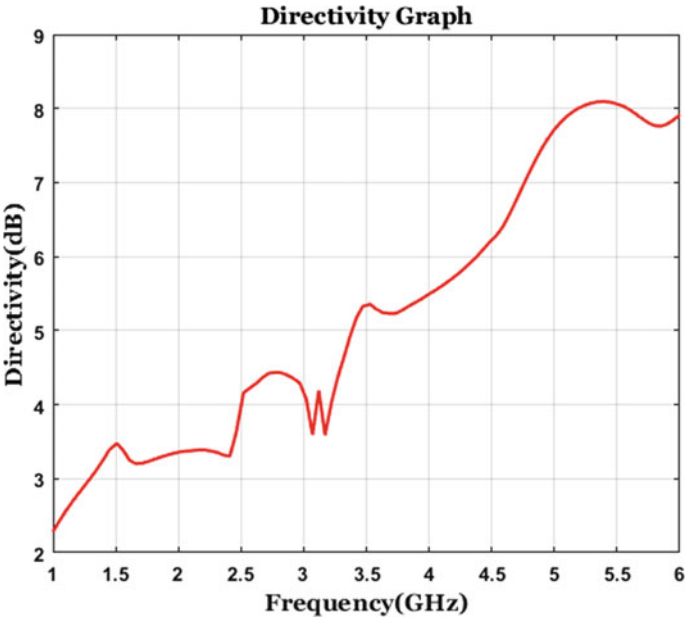


Fig. 5 Directivity versus frequency of the slotted antenna

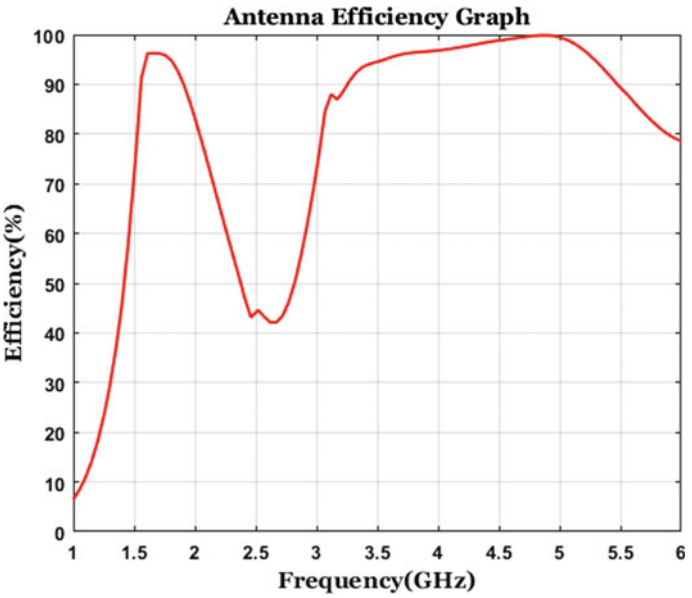


Fig. 6 Graph of efficiency versus frequency

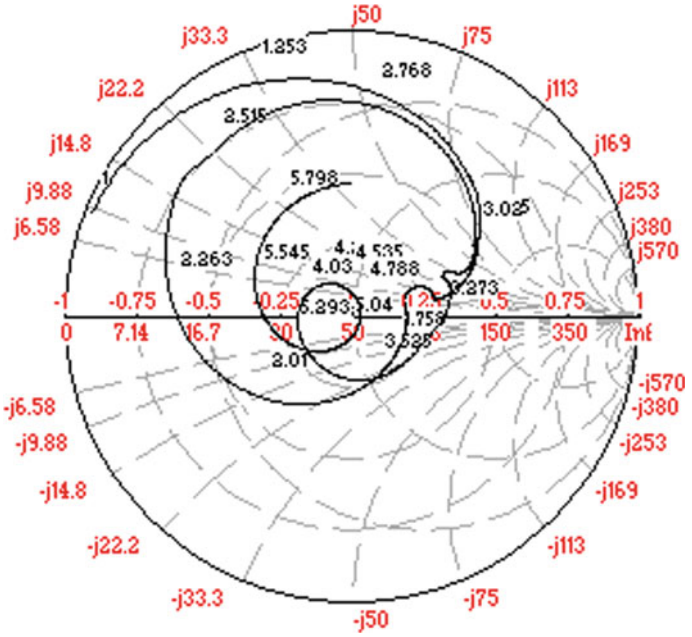


Fig. 7 Smith chart of slotted antenna

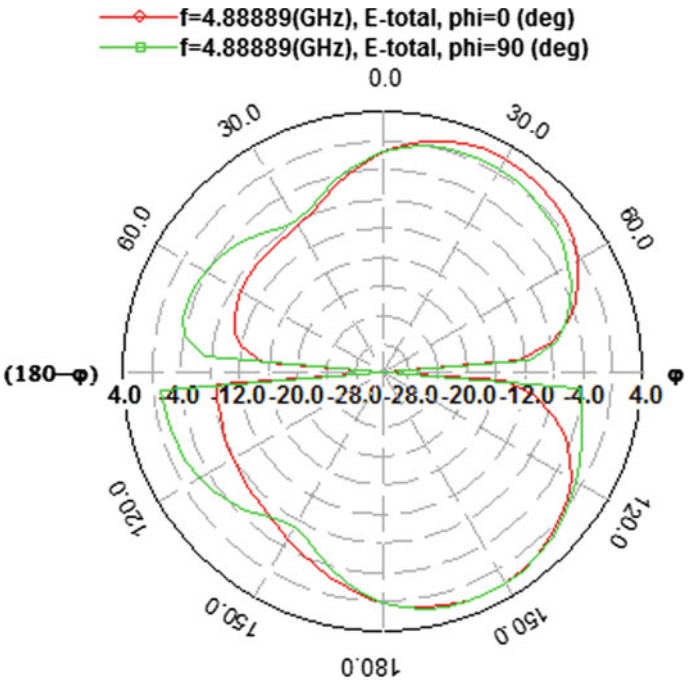


Fig. 8 2D radiation pattern of the designed antenna

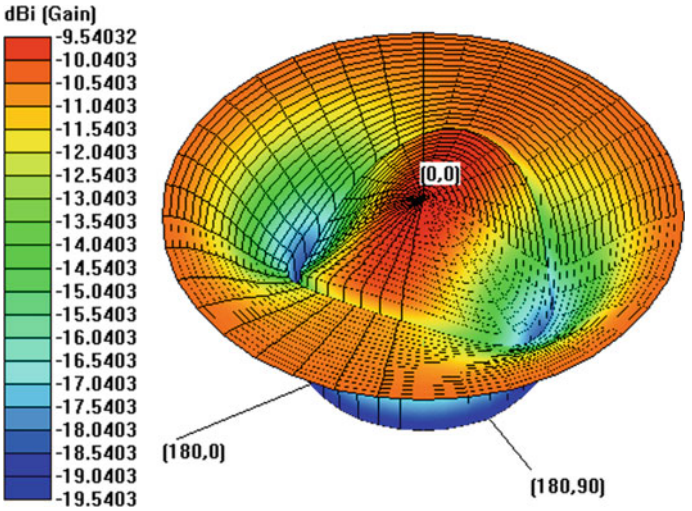


Fig. 9 3D radiation pattern of the slotted antenna

6 Conclusion

Over here we discussed the design and bandwidth increment. This design provided the best technique to solve the narrow bandwidth problem of the microstrip patch antenna. After the simulation process, the bandwidth of the antenna is increased from 27.30 to 49.81% with high gain of 7.43 dB at resonant frequency 4.889 GHz range from 3.273 to 5.444 GHz. This frequency band is suitable for wireless devices. The return loss of antenna also increased up to −29.49 dB. VSWR of this designed antenna is 1.068 and efficiency is 99.89% at 4.889 GHz.

References

1. C.A. Balanis, *Antenna Theory, Analysis and Design* (Wiley & Sons, New York, 2005)
2. D.M. Pozar, Microstrip antennas. Proc. IEEE **40**(1) (1992)
3. Z. Ali, V.K. Singh, A.K. Singh, Ayub, S., E-shaped microstrip antenna on Rogers substrate for WLAN applications, in *International Conference on Computational Intelligence and Communication Networks, Gwalior* (2011), pp. 342–345
4. B.K. Ang, B.K. Chung, A wideband E shaped microstrip patch antenna for 5–6 GHz wireless communications. Progress Electromag. Res. B **75**, 397–407 (2007)
5. K.F. Tong, T.P. Wong, Circularly polarized U-slot antenna. IEEE Trans. Antennas Propag. **55**(8), 2382–2385 (2007)
6. K.F. Tong, K.M. Luk, K.F. Lee, R.Q. Lee, A broadband U-slot rectangular patch antenna on a microwave substrate. IEEE Trans. Antennas Propag. **48**(6), 954–960 (2000)
7. S. Maci, G.B. Gentili, P. Piazzesi, C. Salvador, Dual-band slot-loaded patch antenna. IEEE Proc. Microwav. Antennas Propag. **142**(3), 225–232 (1995)
8. K. Mandal, P.P. Sarkar, High gain wide-band U-shaped patch antennas with modified ground planes. IEEE Trans. Antennas Propag. **61**(4), 2279–2282 (2013)
9. W.X. Zhang, C.S. Pyo, S.I. Jeon, S.P. Lee, N.H. Myung, A new type of wideband slot-fed U-slotted patch antenna. Microwave Opt. Technol. Lett. **22**, 378–381 (1999)
10. G.F. Khodaei, J. Nourinia, C. Ghobadi, A practical miniaturized U-slot patch antenna with enhanced bandwidth. Progress Electromag. Res. B. **3**, 47–62 (2008)
11. R.K. Verma, D.K. Srivastava, A compact inverted Y slot rectangular microstrip patch antenna for Bluetooth applications. Indonesian J. Electr. Eng. Comput. Sci. **11**(2), 413–418 (2018)
12. IE3D electromagnetic simulation and optimization package, Version 9.0

Circularly Polarized Compact Monopole Antenna with an Offset Microstrip Feedline for C-Band Applications



Akansha Yadav, Sudhanshu Verma and Saurabh Singh

1 Introduction

Presently, with the increase in demands for communication systems, small antenna's research with different radiating elements, polarizations, and various feeding designs has attracted immense attention. Circularly polarized antennas are earning favored due to its various supremacy of good mobility and weather penetration in a communication system as compared to linearly polarized antennas. In present days, many of the researches because of broad bandwidth and low profile designed wide slot circularly polarized antenna and circularly polarized wide slot antenna can also be integrated with MMIC, i.e. microwave monolithic integrated circuits. Circularly polarized antennas operate on the principle of exciting two modes have equal amplitudes and orthogonal to each other. Various techniques have been proposed to introduce circularly polarized radiation and various circularly polarized antennas have been described in the literature. These involve the insertion of two L-shape strips at the ground plane of the antenna with square slot and both L–strip are placed at the slot's opposite corners [1, 2] and ARBW is enhanced with the addition of the third L-strip on the ground plane in the square slot [3]. Various techniques for exciting circular polarization mode have been executed and mostly based on the modification of the ground plane and antenna resonator [4, 5] and in planer antennas wide axial ratio bandwidth (ARBW) is obtained with the modifications in the various feedline

A. Yadav (✉) · S. Verma · S. Singh
Department of Electronics and Communication Engineering, Madan Mohan Malaviya University of Technology, Gorakhpur, India
e-mail: yadavakansha1994@gmail.com

S. Verma
e-mail: svermaeced@gmail.com

S. Singh
e-mail: singhsaurabh047@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_77

[5, 6]. Circularly polarized wide ARBW has been proposed with the utilization of the perturbations through feedlines [7–10], with the help of slot structures [11–14] and through array configurations [15, 16]. CP antenna can also be obtained through compact monopole antenna [17]. Although these designs of the antenna are very complicated that they induce difficulties in designs and fabrications of antenna.

In the present paper, an elementary and compact wideband CP microstrip antenna has been proposed. The proposed antenna has a microstrip feedline, a wide rectangular slot at the ground and rectangular stub at the upper right side of the ground plane. This proposed antenna has 3.5 GHz –10.15 GHz wide impedance bandwidth and 3 dB ARBW is about 43.94% (4.67–7.3 GHz) for WLAN and other C-band wireless systems.

2 Design Procedure

A. *Antenna Model and Configuration*

The presented antenna is fed by 50Ω offset microstrip feedline at the upper side of FR4 dielectric. FR4 dielectric is used with $\varepsilon_r = 4.4$, and height $h_{sub} = 1.6$ mm. The width of the microstrip feedline is about 3 mm. The rectangular gap of length L_{g2} and width W_{g2} at the GND plane just below the feedline is created to decrease coupling b/w the ground plane and feedline. To create CP, microstrip feedline is positioned at the right fringe of antenna underneath the extended stub. The parameter of the presented antenna is depicted in Table 1 (Fig. 1).

B. *Progression of Proposed Design*

Antenna 1 has a wide slot at the ground of length (L_S), width (Ws) and microstrip feed line with length F_L and width F_W at the center generates primary resonance close to central frequency. Rectangular gap of length lg_2 and width Wg_2 at the GND plane just below the feedline is created to minimize the coupling b/w the ground plane and feedline (Fig. 2).

Antenna 1 shows a completely linear polarization (LP). In Antenna 2 microstrip feedline is shifted towards the right edges of the antenna with an offset distance

Table 1 Parameters and optimized value of proposed antenna

Parameters	Optimized value(mm)	Parameters	Optimized value(mm)	Parameters	Optimized value(mm)
L_g	25	Ws	17	L_2	5.75
W_g	25	L_1	7.75	F_{L1}	5.77
Ls	16	W_1	7	W_2	5
L_3	5.15	L_4	3.15	S_w	1.8
W_3	4	W_4	2	F_w	3

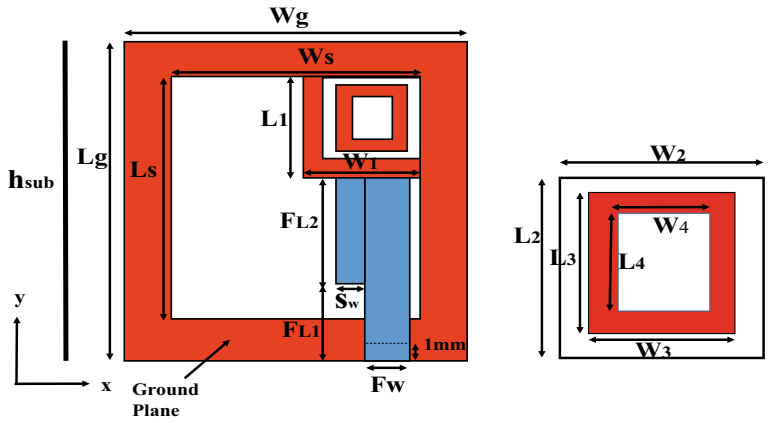
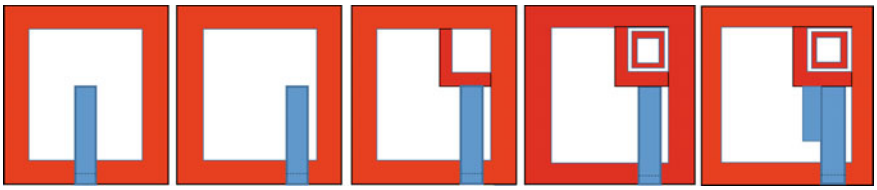


Fig. 1 Geometry of circularly polarized compact monopole antenna



(a) Antenna1 (b)Antenna2 (c) Antenna 3 (d) Antenna 4 (e) Antenna 5

Fig. 2 Evalution of CP compact monopole Antenna, **a** Antenna 1 with wide slot at ground plane and feedline at center **b** Antenna 2 with offset feedline **c** Antenna 3 includes L-shape protruded stub **d** Antenna 4 with rectangular stub bounded from protruded stub **e** Antenna 5 has microstrip feedline with small rectangular stub

d where $d = 5.4$ mm. In Antenna 3 there is L-shape protruded stub of length L1 (along –y direction) and Width W1(along –x direction) to improve the bandwidth and there is a rectangular stub bounded with the L-shape protruded stub in Antenna 4. In Antenna 5, a small stub with length L_3 (along –y direction) and width W_3 (along –x direction) is positioned on the tip of the feedline. Reflection coefficient and ARBW of the antennae 1-5 is depicted in Figs. 3 and 4 respectively. Description of the outcomes of Antennas 1−5 will be in Sect. 3

3 Results and Discussion

CP microstrip antenna has been designed on the basis of measurements given in Table 1 and with the facilitate of ANSYS Version 16.2 presented antennas have been simulated. The performance progress of proposed antennas has been clarified from

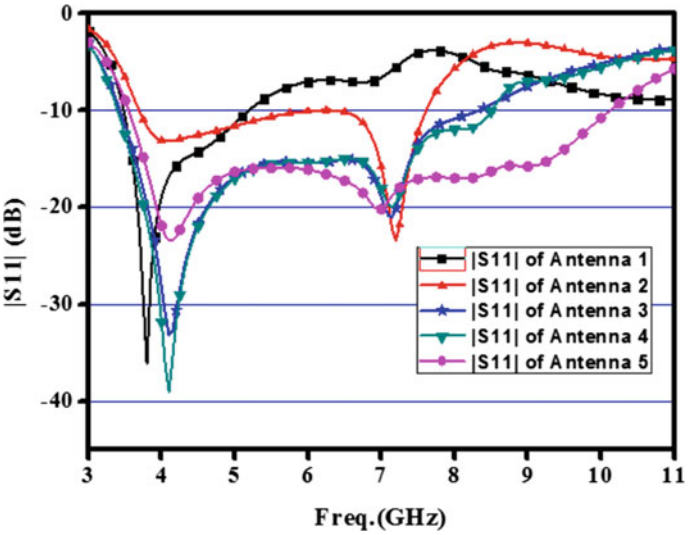


Fig. 3 Simulated $|S_{11}|$ of Antenna 1, Antenna 2, Antenna 3, Antenna 4 and Antenna 5

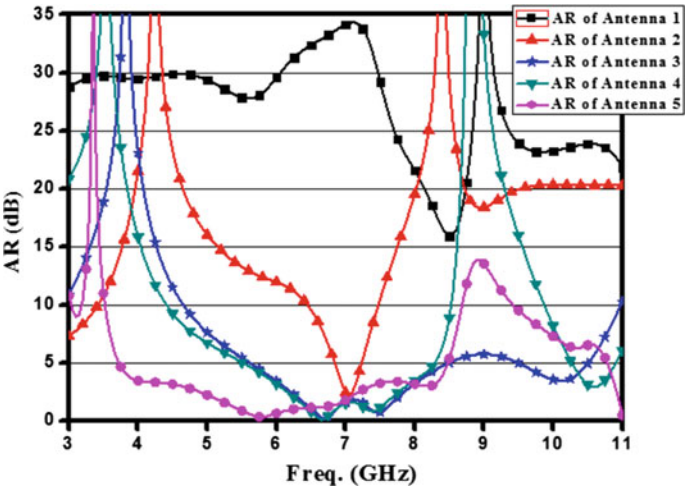


Fig. 4 Simulated AR of Antenna 1, Antenna 2, Antenna 3, Antenna 4, and Antenna 5

Fig. 3 which shows the comparison of the AR and impedance bandwidth of Antennas 1–5. It might be observed from Fig. 3a that |S11| < 10 dB of Antenna 1 is about 3.47–5.18 GHz but it is LP. The bandwidth of the AR in almost frequencies of |s11| < 10 dB is higher than 28 dB is depicted in Fig. 3(b). Two orthogonal E vectors $E_{Horizontal,}$ i.e., E_{Hor} (Horizontal planes complex voltage), and $E_{vertical,}$ i.e., E_{ver} (Vertical planes complex voltage) of similar amplitude with 90° phase difference, CP is effectuated.

When the microstrip feedline is shifted with the same dimensions rectangular gap just below the feedline toward the right edge of the antenna shown as Antenna 2 then CP is effectuated from 6.95 GHz to 7.13 GHz. Protruded stub at the ground plane depicted in Antenna 3 improves the antenna ARBW and impedance bandwidth. Strategically, AR tremendously reaches to 0.1 dB at 6.6 GHz. Yet, the antenna is scarcely circularly polarized and still has not wideband CP characteristics. In order to improve impedance bandwidth significantly rectangular stub is introduced which is bounded with protruded stub as shown in Antenna 4 and obtained impedance bandwidth of Antenna 4 is 5.12 GHz (3.38–8.5 GHz) and ARBW is 1.76 GHz (6.04–7.8 GHz).

Impedance bandwidth with ARBW is improved by using a small stub with length L3 (along –y direction) and width W3 (along –x direction) which is positioned on the tip of the feedline as depicted in Antenna 5.Antenna 5 introduces wide impedance bandwidth is about 97.43% (3.5–10.15 GHz) and ARBW is about 43.94% (4.67–7.3 GHz) and AR < 3 dB is also obtained between 3.5 and 10.15 GHz.

Simulated $|S_{11}|$ of the presented antenna 5 is illustrated in Fig. 5 and ARBW of the presented antenna 5 is shown in Fig. 6. The simulated surface current distribution of the presented antenna 5 at 4.05 GHz and 6.9 GHz is illustrated in Fig. 7. The simulated current distribution of Antenna 5 at 4.05 GHz is shown in Fig. 7a clarify that surface current distribution at the feedline and at the edge of the rectangular slot just below the feedline is high which is responsible to create resonance peak at lower frequency 4.05 GHz and Fig. 7b manifests surface current distribution of presented antenna 5 at 6.9 GHz which clarify that surface current density at the feedline and at the L-shape protruded stub is high which is responsible for generating resonant peak at 6.9 GHz and due to presence of dense current at the feedline at frequency 4.05 GHz and 6.9 GHz and this dense current will increase coupling and hence, bandwidth will

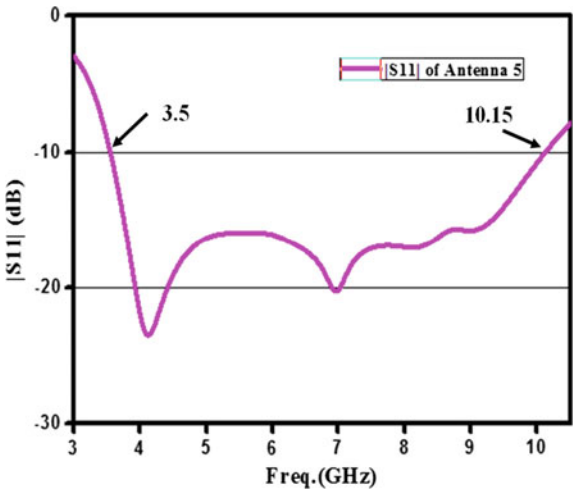


Fig. 5 Simulated $|S_{11}|$ of the proposed CP monopole antenna

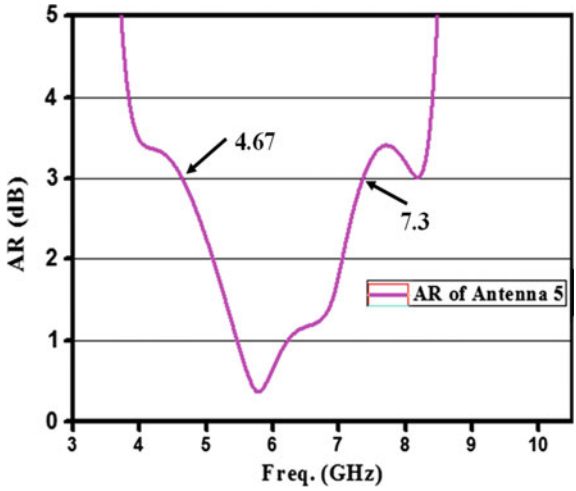


Fig. 6 Simulated AR of the proposed CP monopole antenna

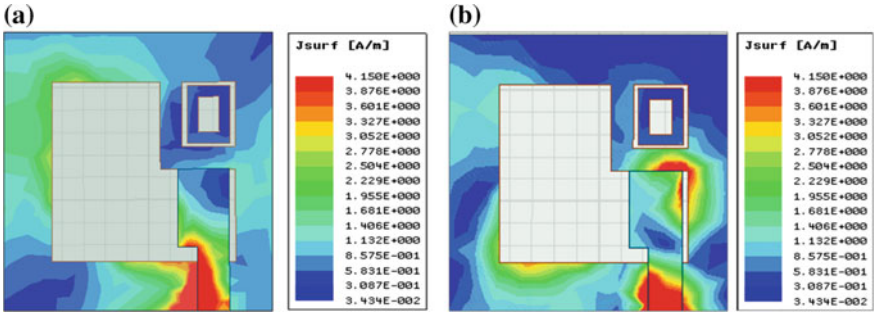


Fig. 7 **a** Surface current density of proposed Antenna 5at 4.05 GHz, **b** Surface current density of proposed Antenna 5 at 6.9 GHz

get enhanced. Simulated realized a total gain of presented Antenna 5 is depicted in Fig. 8. The obtained peak gain of the presented antenna is around 4.43 dB.

Simulated radiation pattern of the proposed antenna 5 at 9 (a) and (b) at 4.05 GHz and at 9 (c) and (d) at 6.9 GHz. From the diagrams of radiation pattern, it is observed that LHCP and RHCP at (phi = 0 and phi = 90) for 4.05 GHz and 6.9 GHz is the mirror image of each other which clarify that proposed Antenna 5 is circularly polarized (Table 2) (Fig. 9).

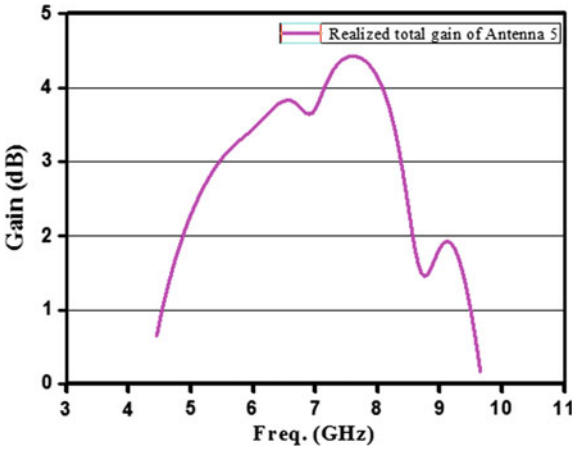


Fig. 8 Simulated gain of proposed Antenna 5

Table 2 Comparison of the dimensions and performance of the proposed antenna with other CP microstrip antenna

Antenna	Method	*fc* (Mz), BW (%)	ARBW (%)	Ant. Area (mm × mm)
Wide slot [1]	Pair of inverted L grounded strip	2.423, 52	28.8	60 × 60
Wide slot [8]	Semicircular protruded stub	2.5, 16	16	30 × 46
Ring slot [12]	Open radiation slot	3.7, 111	27	50 × 50
Wide slot [17]	protruding a horizontal stub from the ground plane	6.375, 90.2	40	25 × 25
Proposed	L-shape protruded stub from the ground	6.65, 97.43	43.94	25 × 25

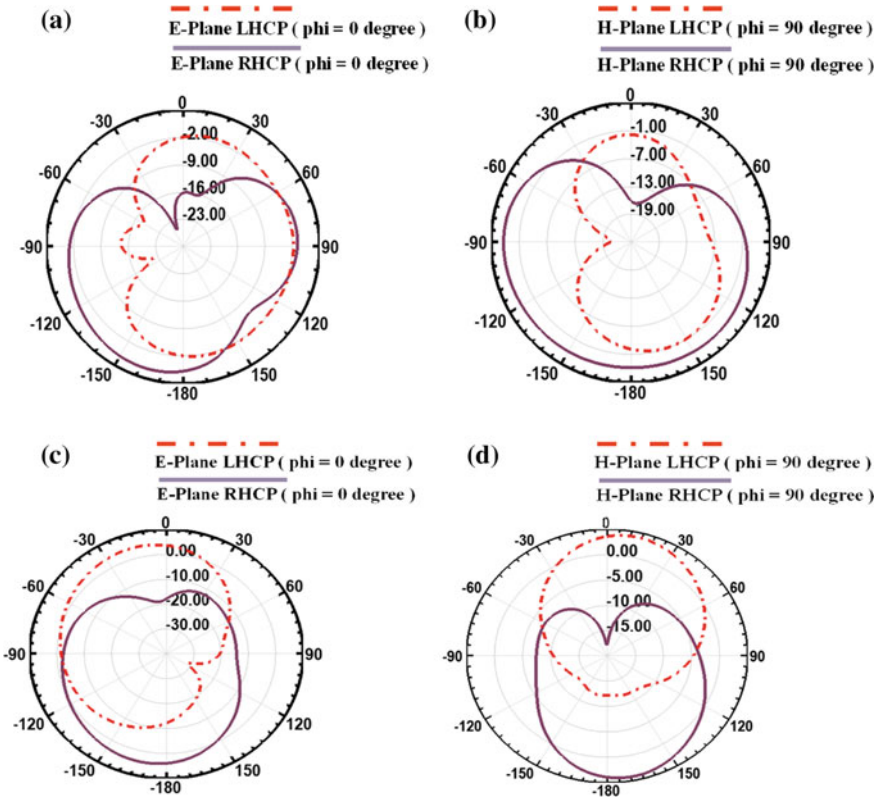


Fig. 9 Simulated radiation pattern of proposed Antenna 5 at different frequencies, (**a**) and (**b**) at 4.05 GHz and (**c**) and (**d**) at 6.9 GHz

4 Conclusions

This work manifests a compact circularly polarized antenna. Circular polarization of an antenna is easily accomplished like other wide slit antenna. Simulated results show accomplished impedance bandwidth of the presented antenna is 97.43% (3.5 GHz–10.15 GHz) and obtained ARBW is about 43.94% GHz (4.67 GHz–7.3 GHz). AR < 3 dB is accomplished by merely L-shape protruding stub from the GND plane in direction of the center of the slot, and after then wide slot fed with an offset microstrip feedline at an offset distance d, beneath the L-shape protruded stub. It can be also demonstrated that the addition of rectangular stub which is bounded with L-shape protruded stub at the GND plane and short stub at the feedline improved the ARBW of the proposed antenna. The proposed CP compact monopole antenna may be used for C-band in wireless communication, including WLAN (5.2 GHz/5.8 GHz) and WiMAX (5.5 GHz). My future work will be to cover a lower frequency range with wide ARBW.

References

1. J.-Y. Sze, C.-C. Chang, Circularly polarized square slot antenna with a pair of inverted-L grounded strips. IEEE Antennas Wirel. Propag. Lett. **7**, 149–151 (2008)
2. J. Pourahmadazar, C. Ghobadi, J. Nourinia, N. Felegari, H. Shirzad, Broadband CPW-fed circularly polarized square slot antenna with inverted L-strip for UWB applications. IEEE Antennas Wirel. Propag. Lett. **10**, 369–372 (2011)
3. N. Felegari, J. Nourinia, C. Ghobadi, J. Pourhmadazar, Broadband CPW-Fed circularly polarized square slot antenna with three inverted L-shaped grounded strips. IEEE Antennas Wirel. Propag. Lett. **10**, 274–277 (2011)
4. R.C. Han, S.-S. Zhong, Broadband circularly-polarised chifreshaped monopole antenna with asymmetric feed. Electron. Lett. **52**(4), 256–258 (2016)
5. T. Kumar, A.R. Harish, Broadband circularly polarized printedslot-monopole antenna. IEEE Antennas Wirel. Propag. Lett. **12**, 1531–1534 (2013)
6. G. Feng, L. Chen, X. Xue, X. Shi, Broadband circularly polarized crossed-dipole antenna with a single asymmetrical cross-loop. IEEE Antennas Wirel. Propag. Lett. **16**, 3184–3187 (2017)
7. K. Ding, C. Gao, T. Yu, D. Qu, Broadband C-shaped circularly polarized monopole antenna. IEEE Trans. Antennas Propag. **63**(2), 785–790 (2015)
8. T.N. Chang, Circular polarized antenna for 2.3–2.7 GHz WiMAX band. Microw. Opt. Technol. Lett. **51**(12), 2921–2923 (2009)
9. J.Y. Sze, J.C. Wang, C.C. Chang, Axial-ratio bandwidth enhancement of asymmetric-CPW-fed circularly-polarised square slot antenna. Electron. Lett. **44**(18), 1048–1049 (2008)
10. J.Y. Sze, S.-P. Pan, Design of cpw-fed circularly polarized antenna with a miniature configuration. IEEE Trans. Antennas Propag. **10**, 1465–1468 (2011)
11. S.H. Yeung, K.F. Man, W.S. Chan, A bandwidth improved circular polarized slot antenna using a slot composed of multiple circular sectors. IEEE Trans. Antennas Propag. **59**(8), 3065–3070 (2011)
12. J.-Y. Jan, C.-Y. Pan, K.-Y. Chiu, H.-M. Chen, Broadband cpw-fed circularly polarized slot antenna with an open slot. IEEE Trans. Antennas Propag. **61**(3), 1418–1422 (2013)
13. Q.X. Chu, S. Du, A CPW-fed broadband circularly polarized square slot antenna. Microw. Opt. Technol. Lett. **52**(2), 409–412 (2013)

14. T.N. Chang, Wideband circularly polarised antenna using two linked annular slots. Electron. Lett. **47**(13), 737–739 (2011)
15. S. Fu, S. Fang, Z. Wang, X. Li, Broadband circularly polarized slot antenna array fed by asymmetric CPW for L-band application. IEEE Antennas Wirel. Propag. Lett. **8**, 1014–1016 (2009)
16. J. Pourahmadazar, V. Rafii, Broadband circularly polarised slot antenna array for L and S-band applications. Electron. Lett. **48**(10), 542–543 (2012)
17. M.S. Ellis, Z. Zhao, X. Ding, Z. Nie, Q.H. Liu, A novel simple and compact microstrip-fed circularly-polarized wide slot antenna with wide axial ratio bandwidth. IEEE Antennas Wirel. Propag. Lett. **64**, 1552–1555 (2016)

An H-Shaped Reconfigurable Slot Antenna Using SIW Technology



Rishi Raj Singh and Akhilesh Mohan

1 Introduction

In today's wireless communication technology, there is an increasing demand for systems that are capable of being multifunctional and compact at the same time. The reconfigurable antenna is the most promising technology which can adapt to any utility depending upon the application, network capability, and other parameters. Reconfigurable systems are able to tune antenna characteristics such as frequency, radiation patterns, and polarization. The frequency tuning feature allows effective utilization of the spectrum. Adding the tunability of polarization and radiation pattern features to the antenna increases the usability of antenna like changing antenna direction to avoid noise or interferences hence increases the channel capacity. Many applications like radar or mobile devices involve the use of different antennas for scanning purpose in different directions, reconfigurable antennas provides a good solution to this problem by using a single antenna integrated with radiation pattern tunability.

Because of the capability of the antenna to tune different antenna characteristics like frequency, polarization and radiation pattern individually or the combination of above, reconfigurable antennas find ample opportunities in the field of wireless communication technologies. These applications include radar and space communications, interference rejection, multiple-input multiple-output (MIMO) systems, cognitive radio systems, wireless devices, wireless network security, ultra-wideband systems, etc. One of the essential components while designing the reconfigurable antenna is the type of switching mechanism used. In the past few years, various

R. R. Singh (✉) · A. Mohan
Indian Institute of Technology Kharagpur, Kharagpur, West Bengal, India
e-mail: rishis874@gmail.com

A. Mohan
e-mail: am@ece.iitkgp.ernet.in

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_78

methods of switching techniques have been evolved in remarkable ways ranging from tuning via posts [1] to pin diodes/MEMS switches [2, 3]. Due to recent trends in the mechanism of switching, hardware complexity of the system is reduced either by using diodes or MEMS switches, although, they individually possess their own pros and cons depending upon the application. In addition, there is a growing trend in integrating the substrate integrated waveguide technology with the circuit design due to its simplicity. Substrate-integrated waveguide (SIW) is constructed by using periodic rows of metallic vias connecting the upper and lower plane of the dielectric substrate. It shows advantages in terms of high power handling capability, low loss, high-quality factor, complete shielding, and easy incorporation with planar circuits. They are becoming highly popular in microwave and millimeter-wave domain nowadays [4, 5].

For the past three decades, there have been many reported designs of reconfigurable antennas with one configuration or two or all three. But combining all the three characteristics in one single antenna without compromising with antenna performance is still a challenging task for antenna designers. Antenna designs are proposed where one of the antenna characteristics is tuned one at a time [6–14] or two of the antenna properties like frequency and polarization reconfiguration [15–19] or frequency and radiation pattern reconfigurability in [20–22] or pattern and polarization tunability in [23–25] where most of the designs are realized by changing either the shape of the slot or positions of feed. One of the SIW based slot antennae was realized in [3] in which the author was able to provide frequency shift along with dual linear polarization and dual circular polarization with the help of a single slot structure on the two different antennas etched above and below the plane respectively. Though it achieved much-improved radiation performance along with other antenna parameters including high gain, high FTBR, and low CPL but at the expense of increased size which is one of the main important factors in wireless communication systems. Another design of a single feed reconfigurable antenna for polarization and frequency diversity for wireless applications is studied in [19]. This antenna has achieved frequency and polarization with the use of open stubs at the feed line and pin diodes with a maximum gain of about 3 dB which is much less than compared to the proposed structure.

In most of the designs, the main basic concept behind frequency shifting is feeding antenna at different locations [23] or the use of pin diodes/varactor diodes [2] connected across the slot which eventually changes the effective length of the slot and it resonates at a different frequency.

However, there are still very few designs that are successful in combining all three characteristics in one single structure without degrading the antenna performance. The paper studied in [26] is the modified version of [3], in which instead of using two different antennas, one single antenna is used, on which same crossed slot structure is etched on both sides of the planes of the substrate and diodes are located on the slot at various positions which are responsible for switching different frequencies, polarization and radiation pattern configurations. This antenna is able to switch between different linearly polarized (LP) and circularly polarized states independently with gain approximately around 5 dBi. The main drawback of this antenna was narrow

bandwidth which is improved in the proposed structure along with improved gain, front to back ratio (FTBR), and antenna efficiency of the system.

In this paper, a substrate-integrated waveguide-based reconfigurable antenna is presented. An H-shaped slot is etched on both sides of the substrate along with diodes which are responsible for reconfiguring frequency and polarization states. Radiation reconfiguration is achieved by changing the slots above and below the plane by switching the diodes ON/OFF. On switching diodes ON/OFF the effective length of the slot varies and it resonates at the different frequencies and consequently, it gives rise to different distinct polarizations corresponding to each frequency.

2 Antenna Design

Figure 1 shows a simplified construction of a reconfigurable slot antenna using a substrate-integrated waveguide (SIW) technology. The upper and lower cross-sectional plane of the structure consists of H-shaped slot etched on both sides with p-i-n diodes connected across the slot at various locations. The dimensions of the design are shown in Table 1. This antenna consists of Rogers 5880 RT/Duroid substrate (thickness = 0.5 mm, $\varepsilon_r = 2.2$) and metallized via arrays popularly known as substrate-integrated waveguide (SIW) technology around the slot as shown. In order to make the rows of vias equivalent to the conventional rectangular waveguide, the condition of $d/d_p \geq 0.5$ and $d/\lambda_0 \leq 0.1$, is maintained, where λ_0 is the free-space wavelength. To feed the antenna, grounded coplanar waveguide (GCPW) is used because of its wider bandwidth, less dispersion and less radiation loss as compared to microstrip.

For switching purposes, five diodes D_{B1}, D_{B2}, D_{B3}, D_{B4}, and D_{B5} are connected across the slot at various locations in the lower plane of the substrate and in similar fashion, five more diodes D_{T1}, D_{T2}, D_{T3}, D_{T4}, and D_{T5} connections are made on the

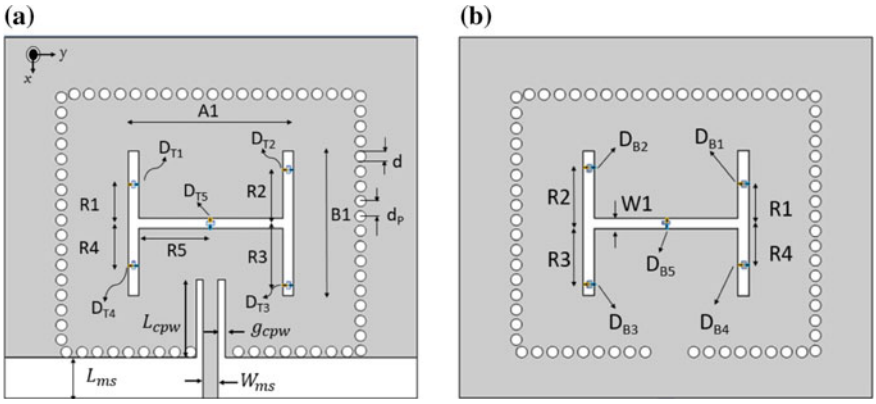


Fig. 1 Geometry of the reconfigurable antenna, **a** top view, **b** bottom view

Table 1 Optimized geometrical parameters of the reconfigurable antenna

Parameters	Values (mm)	Parameters	Values (mm)
L_{ms}	4.0	A_1	16.0
L_{cpw}	7.5	B_1	14.0
g_{cpw}	0.7	d	1.0
W_{ms}	1.45	d_p	1.5
R1	3.8	R2	5.2
R3	6	R4	4
R5	7	W1	1.0

top plane. By switching the diodes ON/OFF, different states of antenna characteristics are realized. To model diodes in Ansoft HFSS simulation [27], the equivalent circuit of diodes characteristics is used as shown in Fig. 2. The inductor is used as an RF choke to bias the DC voltages to p-i-n diodes. When the diode is ON, it behaves as 4.7 Ω resistor and when it is OFF, it behaves as 0.017 pF in the equivalent circuit. During fabrication, to model diode during fabrication, metal strip is used along with 100 pF capacitor and p-i-n diode as shown in Fig. 3. The function of the capacitor is to block dc and pass only RF signals to the diode. By switching to different states of

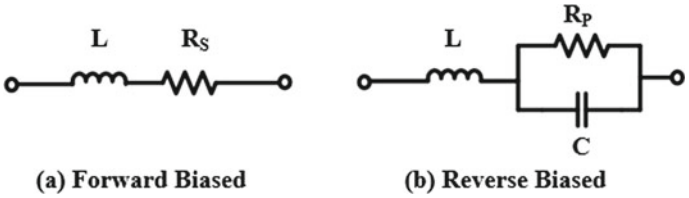


Fig. 2 Equivalent circuit of PIN diode

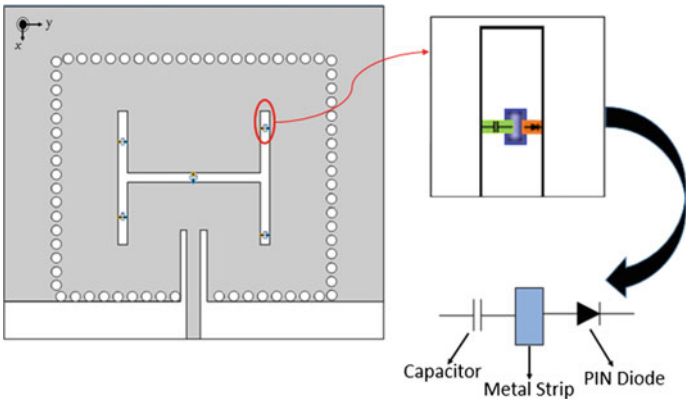


Fig. 3 Proposed reconfigurable H-shaped slot antenna with diode connections

diodes, the antenna is able to tune distinct states of frequency and polarization. To obtain the radiation pattern reconfiguration, upper or lower plane diodes are switched ON/OFF to direct the pattern in +z or −z direction.

3 Frequency and Polarization Reconfigurable Antenna

To obtain the frequency tunability, the effective length of the slot is changed by switching diodes placed on the slot. Two identical H-shaped slots are etched on both sides of the substrate and diodes are placed on it at various strategic locations. The diodes are switched ON/OFF and correspondingly it leads to the generation of five different states of antenna configurations as shown in Table 2. For simplicity, the diodes on the bottom plane is demonstrated, for top plane the results follows in a similar fashion with a slight variation due to the presence of feed line. When the diodes are switched ON, the two ends of the slot are shortened and effective length of slot changes. The five different states of diodes give rise to five different frequency bands as shown in Fig. 4.

Table 2 States of switches for different antenna configurations

	D_{B1}	D_{B2}	D_{B3}	D_{B4}	D_{B5}	Frequency (GHz)
Ant. I	OFF	OFF	OFF	OFF	OFF	4.59
Ant. II	ON	OFF	ON	OFF	OFF	4.63
Ant. III	ON	OFF	OFF	ON	OFF	4.82
Ant. IV	OFF	ON	ON	OFF	OFF	4.65
Ant. V	ON	ON	ON	ON	OFF	4.92

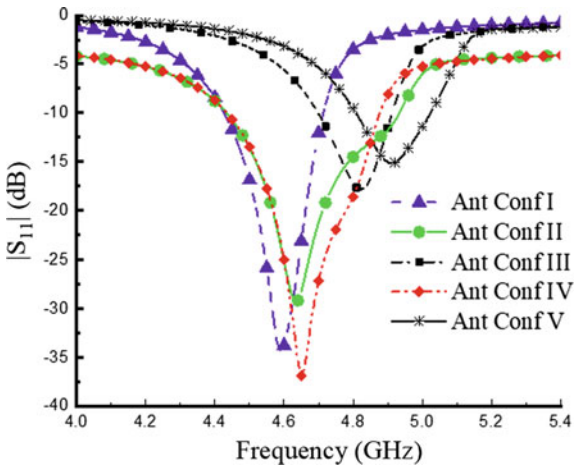


Fig. 4 Frequency reconfiguration states for different antenna configurations

This antenna shows an improved bandwidth with an impedance bandwidth of approximately 4–6%.

To further improve the operation of the antenna, the feature of changing polarization is integrated within the structure. Diodes on the slot are switched simultaneously to generate different polarization states. They are used to switch among different linearly polarized (LP) states and from linearly polarized (LP) to circularly polarized (CP) state or vice versa. The E-field distribution for LP and LHCP states are shown in Fig. 5. In the figure, the resultant direction of the fields is shown by the big arrow and the other fields are canceled. From Fig. 5b, it clearly shows that E fields are rotating anti-clockwise direction and since the slot etched on the lower metal plane of the substrate is active so wave is radiated in the –z axis direction thus we can say that the field is left hand circularly polarized for the given combination of state of switches at 4.63 GHz. Antenna configurations I, III, IV, and V show linearly polarized states at different frequencies while antenna configuration II is circularly polarized at 4.63 GHz with an axial ratio of 1.52 at 4.75 GHz.

In antenna configuration II, the diagonally opposite diodes D1 and D3, positioned on the H-shape slot are switched ON and rest all other diodes are OFF, such that the amplitudes of the orthogonal modes generated in the structure are equal and

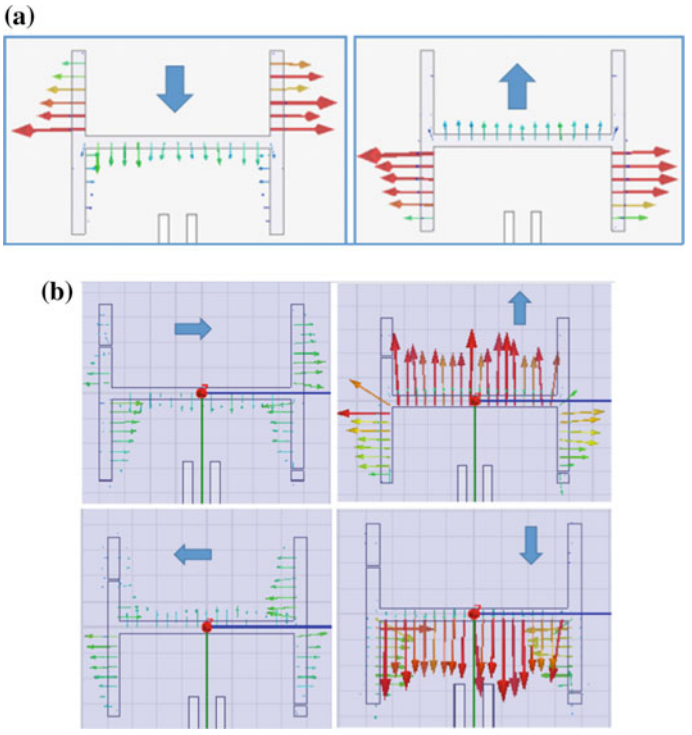


Fig. 5 Electric field distribution of the reconfigurable antenna, **a** LP state at (i) at t = 0 (ii) t = T/4, **b** LHCP state at (i) at t = 0 (ii) t = T/4 (iii) t = T/2 (iv) t = 3T/4

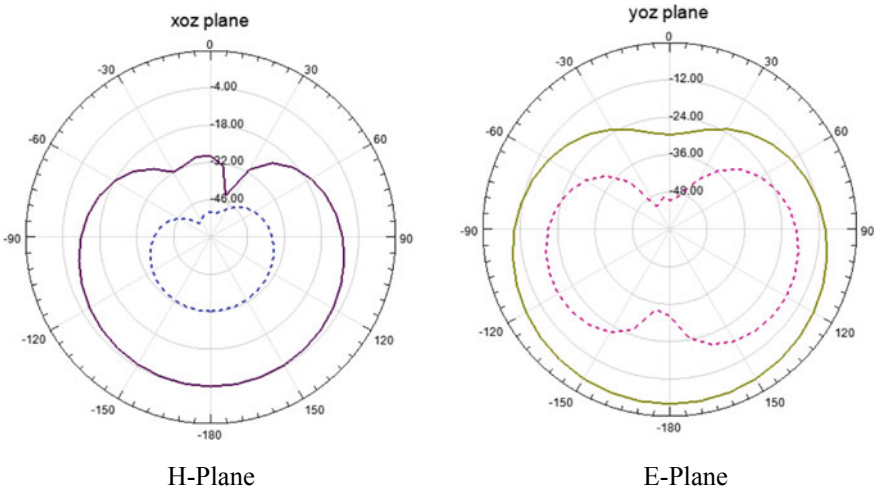


Fig. 6 Simulated co-polarization (dark line) and cross-polarization (dashed line) states of the proposed antenna

produces phase shift of 90° which eventually, results in circular polarization. Other configurations like I and IV lie in the same band as configuration II, thus it is possible for the antenna to switch from CP to LP or vice versa by switching the diodes ON/OFF.

The simulated co-polarization and cross-polarization states of antenna configuration are demonstrated by Fig. 6. The co-polarization state of an antenna is the desired polarization and cross-polarization is the polarization orthogonal to the desired one which is expected to be as minimum as possible. It is impossible to completely avoid the cross-polarization in practical antennas.

4 Radiation Reconfigurable Antenna

Reconfiguring the radiation pattern involves either changing gain, direction, beam angle, etc. In the given antenna structure, the direction of the radiated beam is altered by switching the lower and upper plane of the substrate, such that on switching diodes, the radiated beam is directed in either +z direction or –z direction, respectively. The design presented in this paper demonstrated two H-shaped slots etched on the upper and lower metal plane of the substrate. A total of 10 diodes from which five diodes D_{T1}, D_{T2}, D_{T3}, D_{T4}, D_{T5} are connected on the upper plane and the rest five diodes D_{B1}, D_{B2}, D_{B3}, D_{B4}, D_{B5} on the lower plane of the metal plane are connected.

When all the diodes connected on the lower plane of the substrate are turned ON, there is no radiation from the lower plane such that the radiation pattern is directed toward the +z direction. Similarly, when all the diodes connected on the upper plane of the substrate are turned ON, it behaves as a single metal plane and there is no

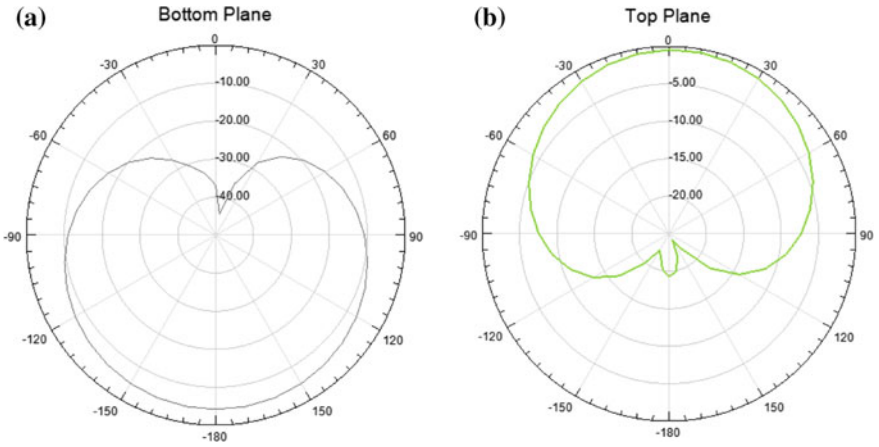


Fig. 7 Radiation pattern along xz plane, **a** lower plane, **b** upper plane

radiation from the upper metal plane. Thus, the radiation is directed toward −z. The radiation pattern plot for the upper and lower plane is illustrated in Fig. 7a and b respectively for antenna configuration I, where all diodes are OFF. The rest other configurations of diodes show an identical pattern for their individual frequency and polarization states.

5 Simulated Results

A fully functional reconfigurable antenna using a substrate integrated waveguide (SIW) technology is successfully analyzed. The simulation was accomplished by using Ansoft HFSS [27].

The antenna parameters for different antenna configurations are compared in Table 3 as shown. The antenna illustrated in this paper shows a broadside gain of approximately 6 dBi. The antenna is able to perform reconfiguration of frequency by switching diodes ON/OFF covering band from 4.5 to 5 GHz The cross-polarization

Table 3 Measured results of different states with backward radiation towards −z direction

	Frequency (GHz)	Gain (dBi)	FTBR (dB)	Ant. efficiency (%)
Ant. I	4.59	5.8	33.4	94.78
Ant. II	4.63	5.6	29.8	92.44
Ant. III	4.82	5.7	33.5	92.47
Ant. IV	4.65	5.8	38.3	94.34
Ant. V	4.92	5.8	37.9	91.56

level (CPL) is achieved as minimum as possible for all the five configurations. The errors due to fabrication and dc bias are neglected during the simulation of the design.

6 Discussions

Of all, the two most important standards for defining tech in today's evolution of antenna world is minimization and performance, the reconfigurable antenna is the most promising answer for revolutionizing the wireless world. A single antenna is able to switch frequency according to the user's requirement, alter directions of radiation and change polarization, just by switching diodes ON/OFF. This reduces the use of several antennae. Thus the antenna size is miniaturized without compromising the performance of the antenna.

In the recent years of development, there have been many reported designs of the reconfigurable antenna with variations either based on feed, the shape of slots, switches like p-i-n diodes/MEMS switches. Although, most of the previous designs were able to reconfigure all of the three characteristics of antenna simultaneously in a single structure but at the expense of short bandwidth compared to the proposed design.

In the proposed design, the placing of diodes on the slot was a crucial job. Since the slot shape is symmetric so in order to generate a circularly polarized wave, the diodes are positioned on the slot in such a way that the slot becomes asymmetric and it generates a phase shift of 90° due to its geometry and hence generates the circularly polarized wave.

The proposed antenna design is able to alter the frequency, radiation pattern and polarization with increased bandwidth, improved gain and antenna efficiency by switching p-i-n diodes ON/OFF.

References

1. D.H. Schaubert, F.G. Farrar, A. Sindoris, S.T. Hayes, Microstrip antennas with frequency agility and polarization diversity. IEEE Trans. Antennas Propag. AP **29**(1) (1981)
2. S. Sam, H. Kang, S. Lim, Frequency reconfigurable and miniaturized substrate integrated waveguide interdigital capacitor (SIW-IDC) antenna. IEEE Trans. Antennas Propag. **62**(3) (2014)
3. G.Q. Luo, Z.F. Hu, Y. Liang, L.Y. Yu, L.L. Sun, Development of low profile cavity backed crossed slot antenna for planar integration. IEEE Trans. Antennas Propag. **57**(10), 2972–2979 (2009)
4. M. Bozzi, L. Perregrini, K. Wu, P. Arcioni, Current and future research trends in substrate integrated waveguide technology. Radio-Eng. **18**(2) (2009)
5. F. Xu, K. Wu, Guided wave and leakage characteristics of substrate integrated waveguide. IEEE Trans. Microw. Theory Tech. **53**(1) (2005)
6. B. Avser, G.M. Rebeiz, Tunable dual-band antennas for 0.7–1.1-GHz and 1.7–2.3-GHz carrier aggregation systems. IEEE Trans. Antennas Propag. **63**(4), 1498–1504 (2015)

7. A. Khidre, F. Yang, A.Z. Elsherbeni, A patch antenna with a varactor-loaded slot for reconfigurable dual-band operation. IEEE Trans. Antennas Propag. **63**(2), 755–760 (2015)
8. A.P. Saghati, K. Entesari, A reconfigurable SIW cavity backed slot antenna with one octave tuning range. IEEE Trans. Antennas Propag. **61**(8) (2013)
9. L. Ge, K.-M. Luk, A band-reconfigurable antenna based on directed dipole. IEEE Trans. Antennas Propag. **62**(1), 64–71 (2014)
10. C.-Y.-D. Sim, Y.-J. Liao, H.-L. Lin, Polarization reconfigurable eccentric annular ring slot antenna design. IEEE Trans. Antennas Propag. **63**(9), 4152–4155 (2015)
11. J.-S. Row, W.-L. Liu, T.-R. Chen, Circular polarization and polarization reconfigurable designs for annular slot antennas. IEEE Trans. Antennas Propag. **60**(12), 5998–6002 (2012)
12. J.-S. Row, M.-J. Hou, Design of polarization diversity patch antenna based on a compact reconfigurable feeding network. IEEE Trans. Antennas Propag. **62**(10), 5349–5352 (2014)
13. L.Y. Ji, Y.J. Guo, P.Y. Qin, S.X. Gong, R. Mittra, A reconfigurable partially reflective surface (PRS) antenna for beam steering. IEEE Trans. Antennas Propag. **63**(6), 2387–2395 (2015)
14. L. Ge, K.-M. Luk, Linearly polarized and dual-polarized magneto-electric dipole antennas with reconfigurable beamwidth in the H-plane. IEEE Trans. Antennas Propag. **64**(2), 423–431 (2016)
15. B. Liang, B. Sanz-Izquierdo, E.A. Parker, J.C. Batchelor, A frequency and polarization reconfigurable circularly polarized antenna using active EBG structure for satellite navigation. IEEE Trans. Antennas Propag. **63**(1), 33–40 (2015)
16. P.-Y. Qin, Y.J. Guo, Y. Cai, E. Dutkiewicz, C.-H. Liang, A reconfigurable antenna with frequency and polarization agility. IEEE Antennas Wireless Propag. Lett. **10**, 1373–1376 (2011)
17. N. Nguyen-Trong, L. Hall, C. Fumeaux, A frequency- and polarization-reconfigurable stub-loaded microstrip patch antenna. IEEE Trans. Antennas Propag. **63**(11), 5235–5240 (2015)
18. L.-R. Tan, R.-X. Wu, Y. Poo, Magnetically reconfigurable SIW antenna with tunable frequencies and polarizations. IEEE Trans. Antennas Propag. **63**(6), 2772–2776 (2015)
19. A. Sedghara, Z. Atlasbaf, A novel single feed reconfigurable antenna for polarization and frequency diversity. Int. J. Microw. Wirel. Technol. 1–7
20. H.A. Majid, M.K.A. Rahim, M.R. Hamid, M.F. Ismail Frequency and pattern reconfigurable slot antenna. IEEE Trans. Antennas Propag. **62**(10), 5339–5343 (2014)
21. S. Nikolaou et al., Pattern and frequency reconfigurable annular slot antenna using PIN diodes. IEEE Trans. Antennas Propag. **54**(2), 439–448 (2006)
22. G.H. Huff, J. Feng, S. Zhang, J.T. Bernhard, A novel radiation pattern and frequency reconfigurable single turn square spiral microstrip antenna. IEEE Microw. Wireless Compon. Lett. **13**(2), 57–59 (2003)
23. A.C.K. Mak, C.R. Rowell, R.D. Murch, C.L. Mak, Reconfigurable multiband antenna designs for wireless communication devices. IEEE Trans. Antennas Propag. **55**(7) (2007)
24. W. Cao, B. Zhang, A. Liu, T. Yu, D. Guo, K. Pan, A reconfigurable microstrip antenna with radiation pattern selectivity and polarization diversity. IEEE Antennas Wireless Propag. Lett. **11**, 453–456 (2012)
25. B. Poussot, J.-M. Laheurte, L. Cirio, O. Picon, D. Delcroix, L. Dussopt, Diversity measurements of a reconfigurable antenna with switched polarizations and patterns. IEEE Trans. Antennas Propag. **56**(1), 31–38 (2008)
26. L. Ge, Y. Li, J. Wang, C.Y.D. Sim, A low-profile reconfigurable cavity-backed slot antenna with frequency, polarization, and radiation pattern agility. IEEE Trans. Antennas Propag. **65**(5) (2017)
27. Ansys Corp. HFSS: high frequency structure simulator based on the finite element method. http://www.ansys.com

Frequency Scanning SIW Leaky Wave Horn Antenna for Wireless Application



Mahesh Kumar Aghwariya, Tanvi Agarwal and Ragini Sharma

1 Introduction

Rectangular waveguide has some disadvantages like high cost and bulky. Moreover, it is not useful for planar Microstrip circuits. Substrate integrated waveguide (SIW) has become the interest of researchers for its characteristics like low loss and smooth incorporation with microwave circuits. SIW consists of a wide microstrip line shorted at the edges with conductive vias, as a rectangular waveguide [1, 2]. Designing of Leaky Wave Antennas (LWAs) depends on geometrical design parameters of the antenna and complex propagation constant of the radiating leaky mode as a function of frequency [3]. SIW leaky wave antennas consist of a simple structure and generate a narrow beam. This narrow beam scans with frequency.

Different kinds of SIW leaky wave antennas have been designed and investigated recently. Transverse slotted leaky wave antenna has beam scanning capabilities and orthogonal polarization [4]. The transverse slotted rectangular waveguide can scan from near broadside to end fire provided waveguide is filled with a dielectric material.

Broadside to end-fire scanning by Leaky wave antenna is given by Eq. (1),

$$\sin(\alpha) = \beta(\omega)/k_0, \tag{1}$$

where $\beta(\omega)$ is the phase constant of propagating wave and k_0 is free space wave number [5]. In [6], a novel substrate integrated waveguide leaky wave antenna is designed having reconfigurable characteristics. The antenna produced both end-fire

M. K. Aghwariya (✉)
THDC-IHET, Tehri, India
e-mail: mahi_wings@ymail.com

T. Agarwal · R. Sharma
KIET Group of Institutions, Ghaziabad, India

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_79

and broadside radiation characteristics at two different frequencies. In [7], a novel SIW MIMO antenna is presented working in wireless frequency range.

In this paper, a SIW Leaky wave antenna horn antenna is presented. The SIW antenna scans from broadside radiation to end fire at different frequencies with a very high gain. The proposed SIW leaky wave antenna consists of the advantages of both rectangular waveguide and SIW. Geometry of this antenna is a simple and microstrip line can be easily used as a feed.

2 Geometry of Proposed Antenna

The proposed configuration of substrate integrated waveguide LWA is shown in Fig. 1. SIW horn antenna is formed by placing metallic vias in form of horn. A substrate integrated waveguide antenna is formed on a substrate with dielectric constant $\in_r = 4.4$ and height 1.76 mm (FR4 Substrate). The metallic sidewalls of the antenna are used to avoid leakage. The feed of this antenna is excited by applying waveguide port as shown in Fig. 1.

First, SIW designates to work for a cutoff frequency of 5 GHz with the following design specifications:

SIW Width $a = 20$ mm,
Spacing between the metallic vias from center-to-center $p = 2$ mm,
Metallic vias diameter $d = 1.5$ mm.

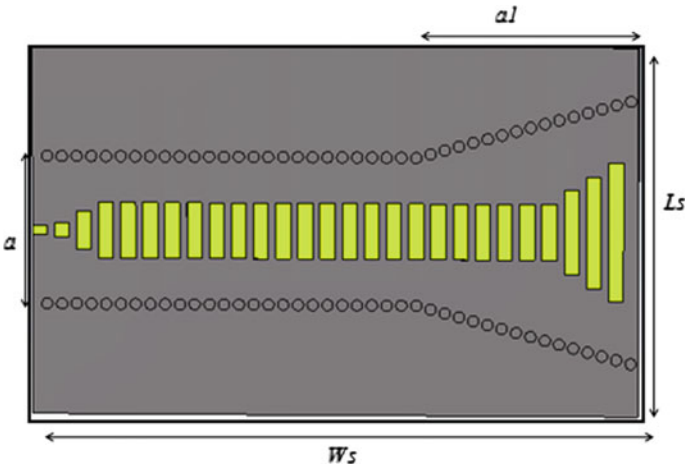


Fig. 1 Design parameters of SIW leaky wave horn antenna: $a = 20$ mm, $Ws = 60$ mm, $a1 = 12$ mm, $P = 2$ mm, $d = 1.5$ mm, height of thickness $= 1.76$ mm

Above values of the proposed design are computed using the following equations for designing of SIW:

A. Diameter of via hole is

$$d < \frac{\lambda_g}{5}$$

B. Distance between the via holes is

$$P \leq 2d$$

C. SIW physical width is

$$w_{eff} = a - 1.08\frac{d^2}{p} + 0.1\frac{d^2}{a},$$

where w_{eff} is the width of the waveguide.

Now, the length of the substrate integrated waveguide is flared to form the waveguide into horn antenna (as shown in Fig. 1). The substrate length 'L' of the antenna is 50 mm and 'W' = 60 mm. The flared length is denoted as '$a1$' which is 12 mm. It is designed for the Ku-band applications. The flare angle is 20.56° and the distance between two parallel rows of vias is 20 mm. Substrate integrated LWA is designed to scan the radiation from broadside to end fire. The frequency range of the antenna is 6–12 GHz.

3 Results

The designed antenna was simulated using software CST MW studio Software. The antenna is excited using a waveguide port indicated in Fig. 1. The antenna has frequency range from 5.9 to 12 GHz. The simulated return loss is value defined by S11 < −10 dB for all working frequencies (as shown in Fig. 2).

Figure 3 shows Electric field distribution of presented SIW Leakey wave horn antenna. It is concluded from Fig. 3 that beam scanning with frequency variations. Pattern of beam scanning is from near broadside to near end fire. Beam starts scanning from near broadside at the frequency 5.9 GHz.

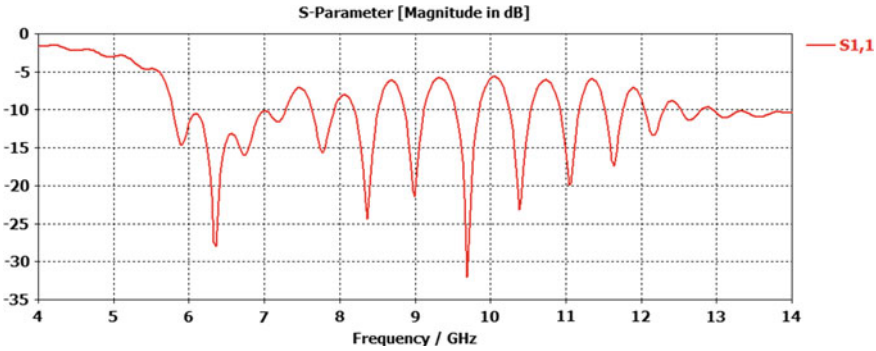


Fig. 2 S11 parameter of SIW leaky wave horn antenna

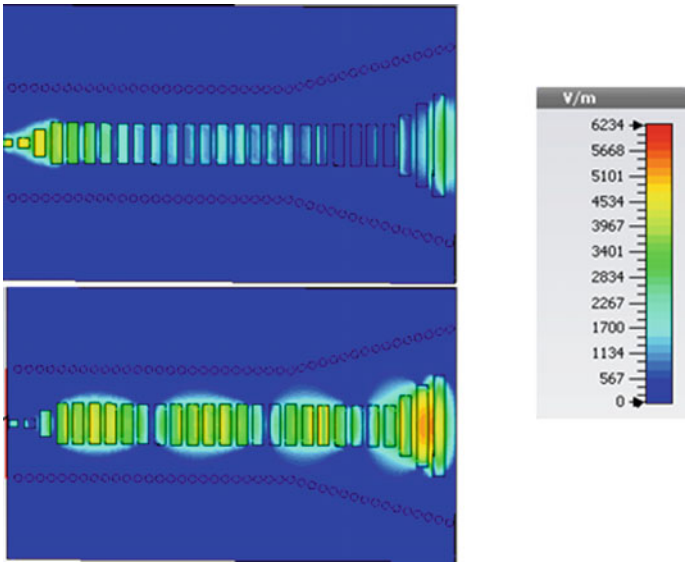


Fig. 3 Electric field distribution of proposed SIW leaky wave horn antenna

Figures 4 and 5 show the field pattern of proposed SIW Leakey wave horn antenna at 5.9 GHz and 6.4 GHz, respectively. The corresponding gain at these frequencies is 7.2 dB and 7.5 dB. Figures 6 and 7 show the field pattern of proposed SIW Leakey wave horn antenna at 8.3 GHz and 11 GHz, respectively. The corresponding gain at these frequencies is 7.9 dB and 8.0 dB.

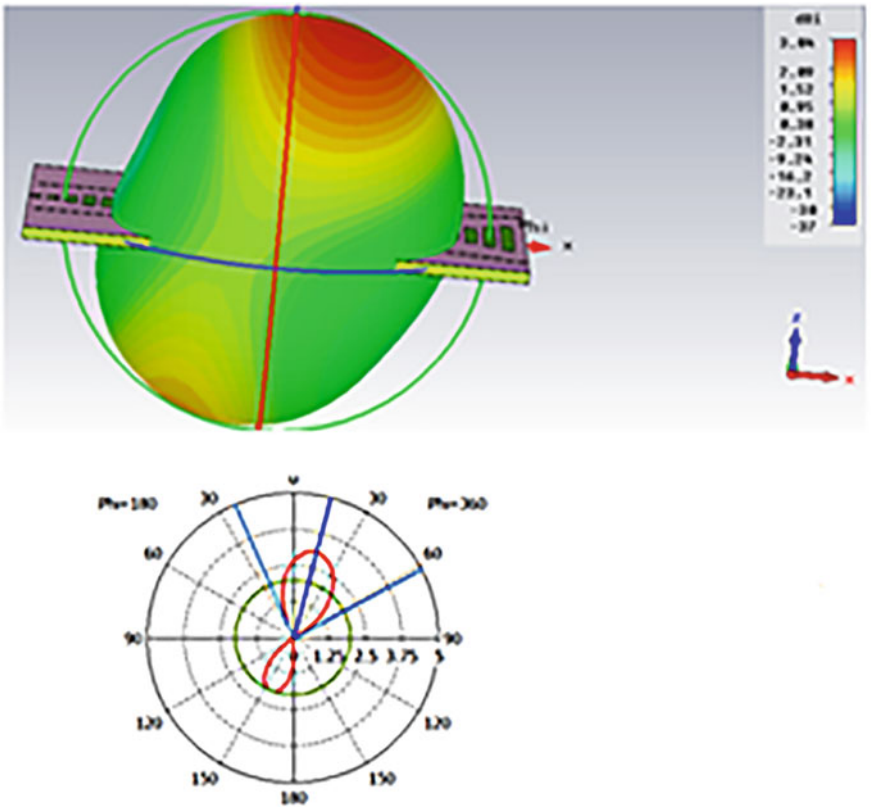


Fig. 4 Polar plot of proposed SIW leaky wave horn antenna at 5.9 GHz

4 Conclusions

A novel SIW leaky wave horn antenna has been designed for multiband application. The proposed prototype of leaky wave antenna efficiently scanned end -fire pattern from broadside pattern with a very high gain. Periodic nonuniform slot causes leakage current since there is an interruption of current flow on the upper wall. The antenna is working in the frequency range from 5.9 to 12 GHz and has an application in Ku-band. The proposed SIW Horn antenna is simulated using CST MW studio Software. The designed antenna has an application in wireless frequency. The presented antenna has a very low loss as it is using SIW technology.

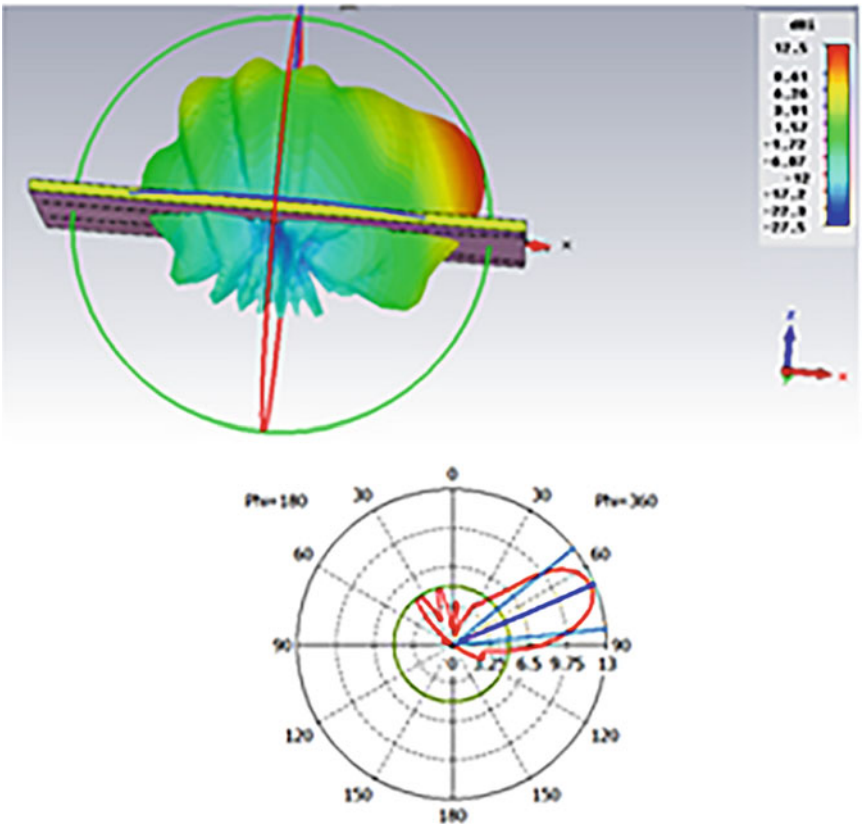


Fig. 5 Polar plot of proposed SIW leaky wave horn antenna at 6.4 GHz

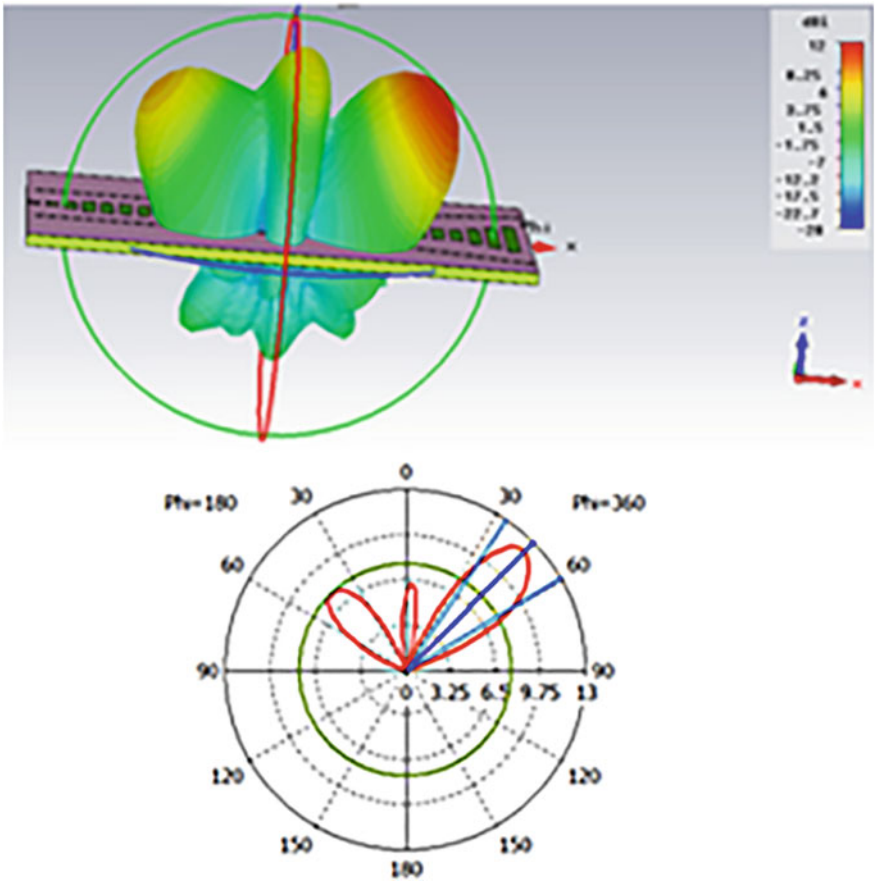


Fig. 6 Polar plot of proposed SIW leaky wave horn antenna at 8.3 GHz

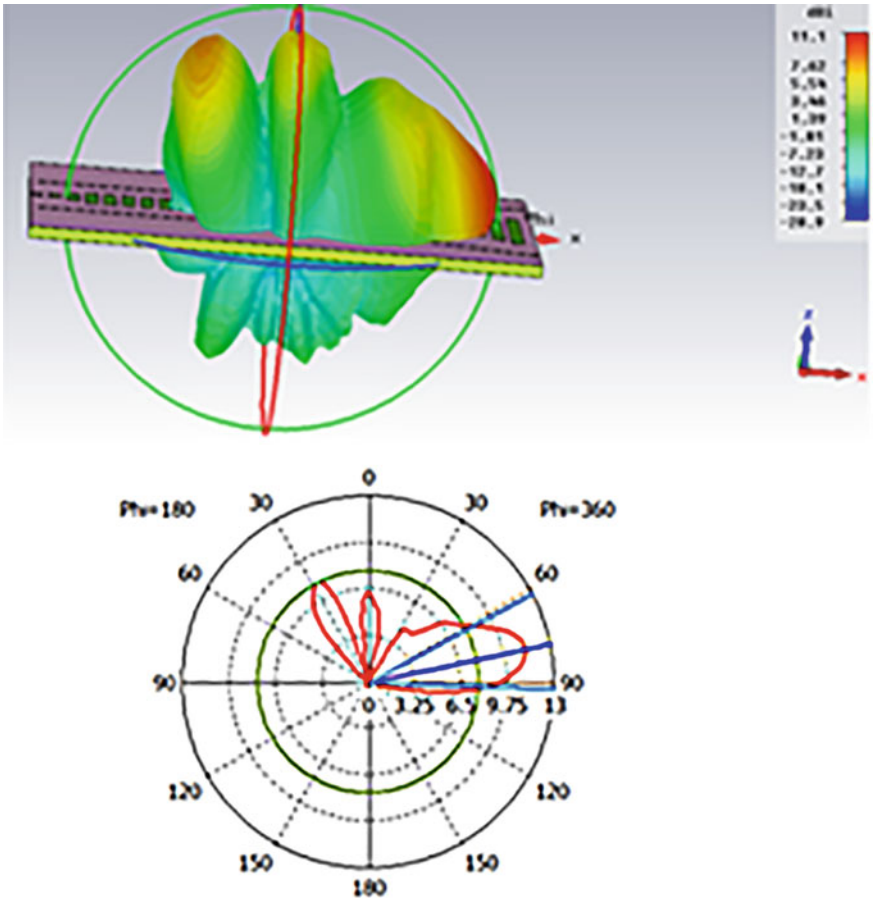


Fig. 7 Polar plot of proposed SIW leaky wave horn antenna at 11 GHz

References

1. D. Deslandes, K. Wu, Design consideration and performance analysis of substrate integrated waveguide components, in *Proceedings of 32nd European Microwave Conference*, vol. 2 (Milan, Italy, Sep. 2002), pp. 881–884
2. R. Sharma, M. kumar et al., Prototype of slotted microstrip patch antenna for multiband applications. Int. J. Eng. Technol. **7**, 1199–1201
3. D. Deslandes, K. Wu, Accurate modeling, wave mechanisms, and design considerations of a substrate integrated waveguide. IEEE Trans. Microw. Theory Tech. **54**(6), 2516–2526 (2006)
4. A.A. Oliner, D.R. Jackson, J.L. Volakis (ed.), Leaky-wave antennas, in *Antenna Engineering Handbook*, 4th ed. (McGraw-Hill, New York, NY, USA, Jun. 2007, ch. 11)
5. A.J. Martinez-Ros, J.L. Gomez-Tornero, G. Goussetis, Planar leaky-wave antenna with flexible control of the complex propagation constant. IEEE Trans. Antennas Propag. **60**(3), 1625–1630 (2012)
6. T. Agrawal, S. Srivastava, Ku band pattern reconfigurable substrate integrated waveguide leaky wave horn antenna. Elsevier AEU—Int. J. Electron. Commun. **87**, 70–75 (2018)
7. T. Agrawal, S. Srivastava, Two element MIMO antenna using substrate integrated waveguide (SIW), in *IEEE International Conference of Signal Processing and Communication* (JIIT Noida, Dec. 2016), pp. 508–511

Randomizing Vigenere Cipher in Polyalphabetic Substitution Scheme



Prabha Elizabeth Varghese and Latha R. Nair

1 Introduction

Cryptography is the tested and proven method to ensure data security. In the past, cryptography was considered to be the science of transforming messages to make them immune to attacks. Cryptography in the present world involves encryption, decryption, and authentication. Cryptanalysis is the science of breaking ciphers to retrieve the ciphertext from plain text, which in turn leads to the development of more secure algorithms. Cryptography and cryptanalysis together form cryptology.

Encryption algorithms are mainly classified as symmetric key algorithms and asymmetric key algorithms. In symmetric key algorithms, same key is used for encrypting the plain text and decrypting the ciphertext. In asymmetric key algorithm, the sender encrypts the plain text with the public key of the recipient and the ciphertext is decrypted with the private key of the recipient.

The basic building blocks of all symmetric methods are substitution and transposition. In substitution ciphers, as the name suggests the letters of the plain text are substituted by other letters, symbols, or numbers. Repositioning of the plain text characters occurs in transposition ciphers. In both cases, the encrypted message looks different from the original message. Another technique is permutation technique where the order of the characters in the plain text is shuffled to get the ciphertext. Symmetric ciphers can be divided into stream ciphers and block ciphers based on how the plain text is processed. Stream ciphers encrypt the plain text one bit or byte at a time. Plain text is encrypted in blocks in block ciphers. Substitution ciphers are further classified as monoalphabetic ciphers and polyalphabetic ciphers. In monoalphabetic ciphers, a character is the plain text that is always replaced by the

P. E. Varghese (✉) · L. R. Nair
SOE, CUSAT, Kochi, India
e-mail: prabha.varghese@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_80

same character in the ciphertext regardless of its position in the plain text. In polyalphabetic ciphers, the same character in the plain text may be replaced by different characters in the ciphertext.

2 Vigenere Cipher

Vigenere cipher is one of the oldest and simplest polyalphabetic ciphers. It was proposed by Blaise de Vigenere in the sixteenth century. The ciphertext is obtained by modular addition of a plain text and a repeating key phrase both of the same length with regard to a 26 * 26 matrix called Vigenere table. Varying shift values are derived from the table and accordingly, substitution is made to get the ciphertext (Fig. 1).

If the letters of the plain text are taken to be numbers 0–25, then the encrypted text can be written as

$$C_i = (P_i + K_i) \bmod 26,$$

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
A	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
B	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A
C	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B
D	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C
E	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D
F	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E
G	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F
H	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G
I	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H
J	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I
K	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J
L	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K
M	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L
N	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M
O	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N
P	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
Q	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
R	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
S	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
T	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
U	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
V	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
W	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V
X	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
Y	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X
Z	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y

Fig. 1 Vigenere Table

and decryption as

$$Pi = (C_i - K_i) \bmod 26,$$

where C_i is the cipher text character, P_i is the plain text character and K_i is the key phrase character.

3 Review on Existing Enhancements of Vigenere Cipher

In [1] more number of encryption keys were used to modify the Vigenere cipher. The ciphertext will have a different encryption pattern as the key is changed automatically after each encryption step. The successive key will depend on the initial key value. If a complex function is used to generate the key, this algorithm will be strong.

To make it more difficult for brute force attack, the key used in Vigenere cipher was modified to include numbers, punctuations, symbols, etc. [2]. A hybrid cryptosystem based on Vigenere cipher was proposed in [3] where the plain text is scrambled using transposition cipher, which in turn is used as the key for Vigenere encryption. In [4] double columnar transposition was proposed for improving security. The plain text is first encrypted using Vigenere cipher then columnar transposition is applied twice to scramble the text further. Computational complexity was found to be more in this case.

In [5] a combination of Vigenere cipher and stream cipher is proposed. The plain text letters in the odd positions are encrypted using Vigenere cipher and the letters in the even positions are converted to its ASCII value and then to its equivalent binary form which is then encrypted with stream cipher using a random binary key value. As the relationship between the plain text and ciphertext is hidden here, it is more secure. The use of repeated random binary key in the case of long messages poses a threat to the security of this algorithm.

Vigenere table handles only English alphabets. Many modifications to Vigenere table were proposed by including numbers 0–9, all the characters in the keyboard [6], a special symbol for blank space [7]. The issue with the repeating nature of the key still remains here.

4 Proposed Work

The main weak point of Vigenere cipher is the repeating nature of its key. Once the key is obtained, the ciphertext can be easily decrypted. In this paper, the conventional Vigenere Table is randomized and a new Randomized Vigenere Table is created. In this approach, even if the key is known, ciphertext cannot be decrypted. For a text encrypted using Randomized Vigenere cipher Table, the random table should be

	U	C	I	K	V	E	J	M	A	P	H	T	N	G	W	Q	F	B	L	D	S	R	Z	O	Y	X
S	N	V	B	D	O	X	C	F	T	I	A	M	G	Z	P	J	Y	U	E	W	L	K	S	H	R	Q
C	X	F	L	N	Y	H	M	P	D	S	K	W	Q	J	Z	T	I	E	O	G	V	U	C	R	B	A
U	P	X	D	F	Q	Z	E	H	V	K	C	O	I	B	R	L	A	W	G	Y	N	M	U	J	T	S
E	Z	H	N	P	A	J	O	R	F	U	M	Y	S	L	B	V	K	G	Q	I	X	W	E	T	D	C
K	U	C	I	K	V	E	J	M	A	P	H	T	N	G	W	Q	F	B	L	D	S	R	Z	O	Y	X
W	R	Z	F	H	S	B	G	J	X	M	E	Q	K	D	T	N	C	Y	I	A	P	O	W	L	V	U
T	O	W	C	E	P	Y	D	G	U	J	B	N	H	A	Q	K	Z	V	F	X	M	L	T	I	S	R
G	B	J	P	R	C	L	Q	T	H	W	O	A	U	N	D	X	M	I	S	K	Z	Y	G	V	F	E
H	C	K	Q	S	D	M	R	U	I	X	P	B	V	O	E	Y	N	J	T	L	A	Z	H	W	G	F
I	D	L	R	T	E	N	S	V	J	Y	Q	C	W	P	F	Z	O	K	U	M	B	A	I	X	H	G
A	V	D	J	L	W	F	K	N	B	Q	I	U	O	H	X	R	G	C	M	E	T	S	A	P	Z	Y
L	G	O	U	W	H	Q	V	Y	M	B	T	F	Z	S	I	C	R	N	X	P	E	D	L	A	K	J
D	Y	G	M	O	Z	I	N	Q	E	T	L	X	R	K	A	U	J	F	P	H	W	V	D	S	C	B
Y	T	B	H	J	U	D	I	L	Z	O	G	S	M	F	V	P	E	A	K	C	R	Q	Y	N	X	W
M	H	P	V	X	I	R	W	Z	N	C	U	G	A	T	J	D	S	O	Y	Q	F	E	M	B	L	K
F	A	I	O	Q	B	K	P	S	G	V	N	Z	T	M	C	W	L	H	R	J	Y	X	F	U	E	D
N	I	Q	W	Y	J	S	X	A	O	D	V	H	B	U	K	E	T	P	Z	R	G	F	N	C	M	L
B	W	E	K	M	X	G	L	O	C	R	J	V	P	I	Y	S	H	D	N	F	U	T	B	Q	A	Z
V	Q	Y	E	G	R	A	F	I	W	L	D	P	J	C	S	M	B	X	H	Z	O	N	V	K	U	T
P	K	S	Y	A	L	U	Z	C	Q	F	X	J	D	W	M	G	V	R	B	T	I	H	P	E	O	N
X	S	A	G	I	T	C	H	K	Y	N	F	R	L	E	U	O	D	Z	J	B	Q	P	X	M	W	V
O	J	R	X	Z	K	T	Y	B	P	E	W	I	C	V	L	F	U	Q	A	S	H	G	O	D	N	M
J	E	M	S	U	F	O	T	W	K	Z	R	D	X	Q	G	A	P	L	V	N	C	B	J	Y	I	H
K	F	N	T	V	G	P	U	X	L	A	S	E	Y	R	H	B	Q	M	W	O	D	C	K	Z	J	I
R	M	U	A	C	N	W	B	E	S	H	Z	L	F	Y	O	I	X	T	D	V	K	J	R	G	Q	P
Q	L	T	Z	B	M	V	A	D	R	G	Y	K	E	X	N	H	W	S	C	U	J	I	Q	F	P	O

Fig. 2 Randomized Vigenere Table

recreated for decryption. Unlike Vigenere cipher, the strength of the cipher is not just dependent on the key, it also depends on the randomness of the table.

A random seed id used to generate the random table. The key is first encrypted using the Randomized Vigenere Table where both the plain text and the key is the keyword itself. The random seed used to generate the random table is appended to the encrypted key. This is then sent to the receiver. The receiver separates the encrypted key and the random seed. The receiver recreates the same Randomized Vigenere Table using the seed supplied to retrieve the plain text from the ciphertext. This method enhances the security of the message (Fig. 2).

5 Vigenere Versus Randomized Vigenere

Conventional Vigenere cipher and Randomized Vigenere cipher were compared based on the Coincidence Index and entropy of the message. The Coincidence Index is an indicator used in cryptanalysis, which makes it possible to evaluate the global distribution of letters in encrypted messages for a given alphabet (Table 1).

Table 1 Comparison of Vigenere and Randomized Vigenere Cipher

	Index of coincidence	Entropy (bits)
Vigenere Cipher	0.02925	3.7216
Randomized Vigenere Cipher	0.02743	3.9812

6 Conclusion

In this paper, we have randomized the conventional Vigenere cipher and the system is found to be very secure. As the table is limited only to alphabets, characters in the plain text were restricted. As an enhancement, the Randomized Vigenere Table can be extended to alpha-numerals and ASCII.

References

1. Q.A. Kester, A cryptosystem based on Vigenère cipher with varying key. Int. J. Adv. Res. Comput. Eng. Technol. (IJARCET) **1**(10), 108–113 (2012)
2. C.R.S. Bhardwaj, Modification of vigenère cipher by random numbers, punctuations and mathematical symbols. IOSR J. Comput. Eng. (IOSRJCE) **4**(2), 35–38 (2010)
3. Q.A. Kester, A hybrid cryptosystem based on vigenère cipher and columnar transposition cipher. Int. J. Adv. Technol. Eng. Res. **3**(1), 141–147 (2013)
4. S. Nishith, B. Kishore, Improving security of vigenere cipher by double columnar transposition. Int. J. Comput. Appl. **100**(14), 6–10 (2014)
5. M.S.A. Fairouz, H.S. Falah, Enhancing security of vigenere cipher by stream cipher. Int. J. Comput. Appl. **100**(1), 1–4 (2014)
6. K. Ravindra Babu, S. Udaya Kumar, A. Vinaya Babu, M. Thirupathi Reddy, An enhanced cryptographic substitution method for information security. Int. J. Math. Arch. **2**(10), 2078–2083 (2011)
7. B.K. Ravindra, K.S. Udaya, B.A. Vinaya, An enhanced polyalphabetic cipher using extended vigenere table. Int. J. Adv. Res. Comput. Sci. **2**(2), 463–465 (2011)

Novel Tool to Determine Kinetic Parameters of Thermoluminescence (TL) Glow Curve—CGCD: $CaZrO_3$: Eu^{3+}, Tb^{3+}

Shubha Tripathi, Vikarm Awate, K. K. Kushwah, Ratnesh Tiwari and Nigama Prasan Sahoo

1 Introduction

CZO:Eu^{3+}, Tb^{3+} ion percentage had been studied for TL glow curve analysis. A novel tool used to determine TL glow curve via computerized glow curve deconvolution (CGCD) technique. Previously, so many authors reported zirconium-based phosphors and its application in phototherapy, display applications, and useful for white light emission [1–20]. Our aim to report the prepared phosphor for TL glow curve analysis and its application in thermoluminescence dosimeter.

2 Experimental

For synthesis of Eu^{3+}, Tb^{3+} activated CZO with varying concentrations of Tb^{3+} ion (0.5–2.5 mol%) were prepared by modified solid-state diffusion method. The stoichiometric amounts starting materials $CaCO_3$, ZrO_2, Eu_2O_3, Tb_2O_3, and H_3BO_3 (as a flux) in molar ratio (0.5–2.5% of Tb) calcination temperature is 900 °C and sintering temperature is 1350 °C were used [21–25]. Characterization techniques are

S. Tripathi · V. Awate
Dr. C.V. Raman University, Bilaspur, India

K. K. Kushwah
Department of Physics, JEC, Jabalpur, India

R. Tiwari (✉)
Department of Physics, Bhilai Institute of Technology, Raipur, India
e-mail: 31rati@gmail.com

N. P. Sahoo
Sri Rawatpura Sarkar Institute of Technology, Raipur, India

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_82

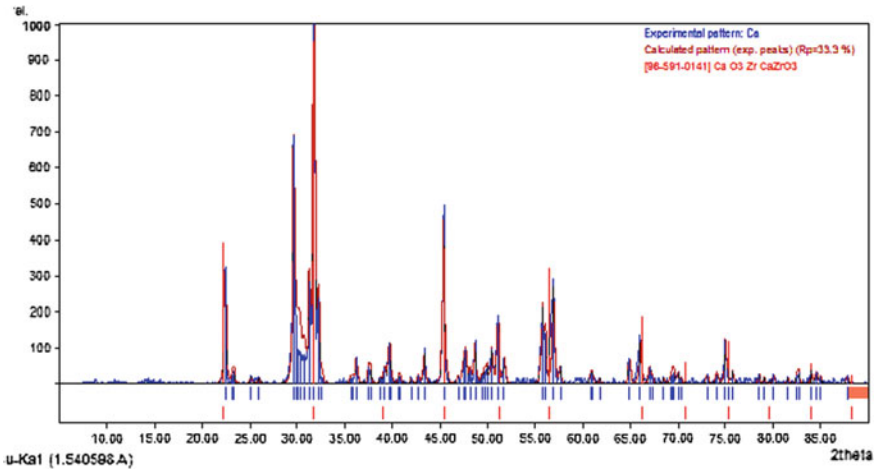


Fig. 1 XRD plot and Rietveld refinement of Eu^{3+}, Tb^{3+} doped $CaZrO_3$ samples

XRD, and TL glow curve analysis using X-ray diffraction meter and TLD reader [26–30].

3 Results and Discussion

3.1 XRD Analysis

The XRD analysis was carried out for comprehensive knowledge of phase formation and structural parameters.

Figure 1 shows XRD analysis of prepared CZO:Eu, Tb doped phosphor and it is refinement pattern which is matched by match software. The refinement pattern compares the experimental and approved data of XRD (Table 1).

3.2 Thermoluminescence Glow Curve Analysis and Trapping Parameters for $CaZrO_3$:Eu^{3+} and Tb^{3+}

It is shown in Fig. 2a that TL glow curve variation with UV exposure time (5–25 min) and it shows such a kind of linearity with dose (Fig. 2b). It is not as such linear response but it is sublinear response with dose. Using CGCD technique, we can compare the results of experimental TL glow curve (Fig. 2c) as well as the theoretically observed TL glow curve [36–42].

Table 1 Refined parameter of X-ray diffraction analysis

Diffraction lines					
d [Å]	Int.	h	k	l	Mult.
3.9900	411.1	1	0	0	6
2.8214	1000.0	1	0	1	12
2.3036	0.1	1	1	1	8
1.9950	395.9	2	0	0	6
1.7844	148.1	2	0	1	24
1.6289	336.5	2	1	1	24
1.4107	197.9	2	0	2	12
1.3300	60.7	2	1	2	24
1.2617	126.1	3	0	1	24
1.2030	3.1	3	1	1	24
1.1518	57.4	2	2	2	8
1.1066	24.9	2	3	0	24
1.0664	143.1	3	2	1	48

The CGCD pattern plotted for the $CaZrO_3$: Eu^{3+}, Tb^{3+} (2 mol%) for the UV 25 min. The figure of merit by CGCD technique is ~3.51% shown in Fig. 2c (Table 2).

For Eu^{3+}, Tb^{3+} (2 mol%) doped $CaZrO_3$ phosphor irradiated by γ Co^{60} dose of 0.5–2 kGy had a peak centered at 156 °C (Fig. 3a). TL glow curve had the general order of kinetic. Figure 3c represents the CGCD pattern of experimental and theoretical TL glow curve irradiated with gamma rays (2 kGy). Four different resolved peaks were determine the kinetic data (Table 3) it is observed that FOM of nearly 2.4%.

Dose versus intensity plot for gamma rays describe its applicability related to patient treatment in medical field (Fig. 3b).

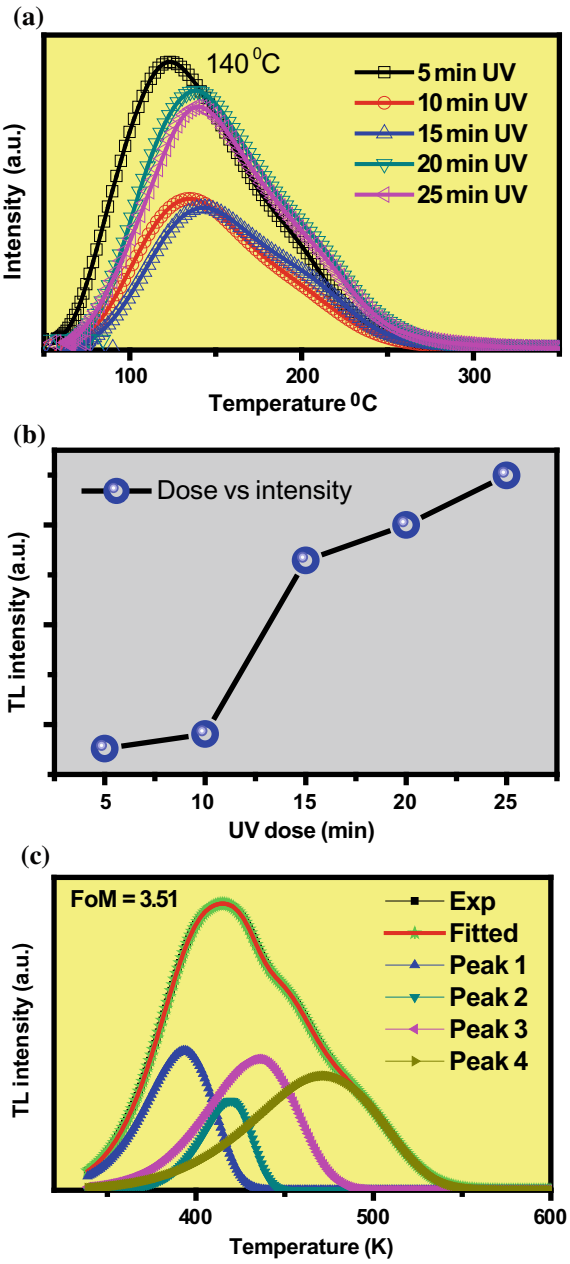


Fig. 2 **a** TL study of UV-ray irradiated sample (5–25 min), **b** dose response of sample, **c** CGCD curve for the $CaZrO_3:Eu^{3+}$, Tb^{3+} (2 mol%) doped phosphor

Table 2 $CaZrO_3:Eu^{3+}$, Tb^{3+} (2 mol%) for the UV 25 min. Figure of merit by CGCD technique

Peak	T_1	T_m	T_2	τ	δ	ω	$\mu = \delta/\omega$	Activation energy	Frequency factor
1	368	392	412	24	20	44	0.45	2.38	2×10^{19}
2	396	420	440	24	20	44	0.45	2.58	1×10^{20}
3	400	436	460	36	24	60	0.4	1.78	5×10^{13}
4	430	471	508	41	37	78	0.47	1.74	6×10^{12}

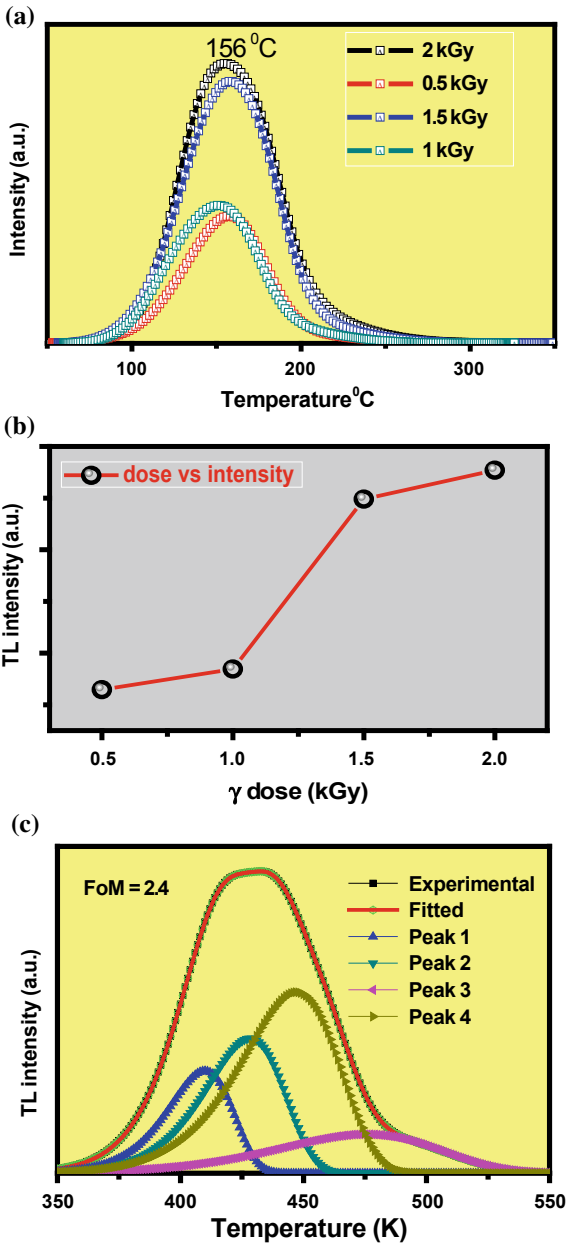


Fig. 3 **a** TL glow curve for $CaZrO_3:Eu^{3+}$, Tb^{3+} 2 mol%) phosphor had γ dose 0.5–2 kGy. **b** Glow peak intensity under 0.5–2 kGy γ dose for $CaZrO_3:Eu^{3+}$, Tb^{3+} (2 mol%). **c** CGCD curve for $CaZrO_3:Eu^{3+}$, Tb^{3+} (2 mol%) phosphor γ 2 kGy irradiated

Table 3 CGCD curve for $CaZrO_3$:Eu^{3+}, Tb^{3+} 2 mol%) doped phosphor γ irradiated sample

Peaks	T_1	T_m	T_2	τ	δ	ω	$\mu = \delta/\omega$	Activation energy	Frequency factor
1	392	409	424	17	15	32	0.46	3.54	4×10^{27}
2	406	428	445	22	17	39	0.43	2.88	1×10^{22}
3	420	446	468	26	22	48	0.45	2.57	2×10^{19}
4	443	481	508	38	27	65	0.41	1.91	7×10^{13}

4 Conclusions

In conclusion, CZO:Eu^{3+}, Tb^{3+} phosphors in are prepared by solid-state reaction method and the various irradiation source are used to compare the TL glow curve analysis here the gamma-irradiated phosphor had the formation of deeper level trapping whereas UV irradiated phosphor shows shallower trapping which was justified by trap depth parameter.

Acknowledgements Authors are very thankful to Prof. S. J. Dhoble, RTM Nagpur University for providing γ irradiation source for our study.

References

1. C. Hoare, A. Li Wan Po, H. Williams, Health Technol. Assess. **4**, 88 (2000)
2. R.P. Sonekar et al., Opt. Mater. **30**, 622 (2007)
3. J.A. Parrish, K.F. Jaenicke, J. Investig. Dermatol. **76**, 359 (1981)
4. H. van Weelden, H. Baart de la Faille, E. Young, J.C. van der Leun, Br. J. Dermatol. **119**, 11 (1988)
5. J. Ferguson, Arch. Dermatol. **135**, 589 (1999)
6. A. Vecht, A.C. Newport, P.A. Bayley, W.A. Crossland, J. Appl. Phys. **84**, 3827 (1998)
7. M.F. Holick, in Vitamin D ed. by D. Feldman, F.H. Glorieux, J.W. Pike (Academic Press, San Diego 1997), pp. 33–39
8. K. Becker, At. Energy Rev. **5**, 43 (1967)
9. J.L. Shie, C.H. Lee, C.S. Chiou, C.T. Chang, C.C. Chang, C.Y.J. Chang, Hazard. Mater. **155**, 164 (2008)
10. L. Scherschun, J.J. Kim, W.H. Lim, J. Am. Acad. Dermatol. **44**, 999 (2001)
11. R.M. MacKie, Radiat. Prot. Dosim. **91**, 15 (2000)
12. C. Sacher, C. Konig, K. Scharffetter-Kochanek, T. Krieg, N. Hunzelmann
13. W.L. Morison, Semin. Cutan. Med. Surg. **18**, 297 (1999)
14. J.A. Parrish, K.F. Jacnicke, J. Invest. Dermatol. **76**, 359 (1981)
15. M. Abdullah, A.W. Khairurrijal, W. Sutrisno, I. Nurhasanah, A.S. Vioktalamo, Powder Technol. **183**, 297 (2008)
16. H. Yoshida, R. Yoshimatsu, S. Watanabe, K. Ogasawara, Jpn. J. Appl. Phys. Part-1 **45**, 146 (2006)
17. S. Okamoto, R. Uchino, K. Kobayashi, H. Yamamoto, J. Appl. Phys. **106**, 013522 (2009)
18. V. Singh, R.P.S. Chakradhar, J.L. Rao, I. Ledoux-Rak, H.-Y. Kwak, J. Mater. Sci. **46**, 1038 (2011)
19. V. Singh, S. Borkotoky, A. Murali, J.L. Rao, T.K. GunduRao, S.J. Dhoble, Spectrochim. Acta Part A: Mol. Biomol. Spectrosc. **139**, 1–6 (2015)
20. A.B. Gawandea, R.P. Sonekara, S.K. Omanwar, Optik **127**, 3925–3927 (2016)
21. V. Dubey, J. Kaur, S. Agrawal, N.S. Suryanarayana, K.V.R. Murthy, Optik—Int. J. Light Electron Opt. (2013). https://doi.org/10.1016/j.ijleo.2013.03.153
22. N. Dubey, V. Dubey, *Reviews in Fluorescence* (Springer, Berlin, 2016) pp. 155–184
23. V. Dubey, R. Tiwari, R.K. Tamrakar, J. Kaur, S. Dutta, S. Das, H.G. Visser, S. Som, J. Lumin. **180**, 169–176 (2016)
24. V.K. Singh, S. Tripathi, M.K. Mishra, R. Tiwari, V. Dubey, J. Disp. Technol. **12**(10), 1224–1228 (2016)
25. NehaTiwari VikasDubey, AIP Conf. Proc. **1728**, 020002 (2016). https://doi.org/10.1063/1.4946052

26. N. Tiwari, V. Dubey, J. Dewangan, N. Jain, J. Displ. Technol. **12**(9), 933–937
27. N. Tiwari, V. Dubey, R.K. Kuraria, J. Fluoresc. **26**(4), 1309–1315
28. Y. Parganiha, J. Kaur, V. Dubey, R. Shrivastava, S.J. Dhoble, Superlattices Microstruct. **88**, 262–270
29. N. Tiwari, V. Dubey, Luminescence **31**, 837–842 (2016)
30. V. Dubey, J. Kaur, S. Agrawal, Res. Chem. Intermed. **41**(7), 4727–4739 (2015)
31. H. Jeong, N. Singh, M.S. Pathak, S. Watanabe, T.K.G. Rao, V. Dubey, V. Singh, Optik-Int. J. Light Electron Opt. **157**, 1199–1206 (2018)
32. V. Singh, N. Singh, M.S. Pathak, S. Watanabe, T.K.G. Rao, P.K. Singh, V. Dubey, Optik-Int. J. Light Electron Opt. (2017). https://doi.org/10.1016/j.ijleo.2017.12.034
33. R.W.G. Wyckoff, *Structure of Crystals*, 2nd edn. (The Chemical Catalog Company, INC, New York, 1931), p. 273, http://jcrystal.com/steffenweber/gallery/StructureTypes/st4.html
34. V. Dubey, J. Kaur, S. Agrawal, N.S. Suryanarayana, K.V.R. Murthy, Optik **124**, 5585 (2013)
35. V. Dubey, N.S. Suryanarayana, J. Kaur, J. Miner. Mater. Charact. Eng. **9**(12), 1101 (2010)
36. V. Dubey, J. Kaur, N.S. Suryanarayana, K.V.R. Murthy, Res. Chem. Intermed. **39**, 3689 (2013)
37. V. Dubey, J. Kaur, Y. Parganiha, N.S. Suryanarayana, K.V.R. Murthy, Appl. Radiat. Isot. **110**, 16 (2016)
38. I.R. Singh, R.S. Ningthoujam, V. Sudarsan, I. Shrivastav, S.D. Singh, G.K. Dey, S.K. Kulshreshtha, Nano Tech. **19**, 1 (2008)
39. G. Baldinozzi, J.-F. Berar, G. Calvarin, Mater. Sci. Forum **278**, 680 (1998)
40. R. Shrivastava, J. Kaur, V. Dubey, B. Jaykumar, Bull. Mater. Sci. **37**(4), 925 (2014)
41. V. Dubey, J. Kaur, N.S. Suryanarayana, S. Agrawal, Int. J. Lumin. Appl. **3**(II), 98 2013
42. S. Som, M. Chowdhury, S.K. Sharma, Radat. Phys. Chem. **110**, 51 (2015)

Determination of Spectroscopic Parameters via Judd–Ofelt Analysis of Eu^{3+} Doped $La_2Zr_2O_7$ Phosphor



Neha Dubey, Jagjeet Kaur, Vikas Dubey and Manish Kumar Mishra

1 Introduction

Lanthanide zirconate with general formula $A_2B_2O_7$ type structure [1–3]. For practical and industrial applications, it is required to prepare homogenous phosphor [4]. Kong et al. (2013) discussed proposed mechanism related to elucidates molecular assembly of this chemical synthesis method [4]. Min et al. (2018) have synthesized Eu^{3+} activated LZO red light emission phosphor for NUV (near UV) (w-LEDs), combined with near-ultraviolet or blue lighting chips [5]. Mohapatra et al. (2016) have discussed effect of Uranium doping level on LZO phosphor and mechanism of multipolar interaction [6]. Singh et al. (2015) synthesized LZO phosphor doped with Gd^{3+} ion and its spectrum at 313 nm range (UV emission) [7]. Tiwari and Dubey (2016) have discussed the basic concept of determination of spectroscopic parameters of zirconium-based phosphor via Judd–Ofelt intensity parameters [8]. Herein synthesized phosphor doped with europium ion and its application in intense red emission for display device applications. The theoretical analysis of spectroscopic parameters is evaluated via Judd–Ofelt intensity parameters.

N. Dubey (✉) · J. Kaur
Department of Physics, Government Vishwanath Yadav Tamaskar Post Graduate Autonomous College, Durg 491001, Chhattisgarh, India
e-mail: tiwarineha1441@gmail.com

V. Dubey
Department of Physics, BIT Raipur, Kendri 493661, Chhattisgarh, India

M. K. Mishra
Career Point University Kota, Kota, Rajasthan, India

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_83

2 Experimental

All phosphor samples were synthesized via solid-state reaction method. Starting materials used for sample preparation are La_2O_3, ZrO_2, and Eu_2O_3 here boric acid used as a flux. All reagents are mixed in agate pastel and mortar by ground it mixed thoroughly using acetone several times up to 3 h. Samples were calcined at 1000 °C for 2 h and then sintered at 1300 °C for 3 h. After that, it was cooled and used for different characterizations.

Structural analysis of prepared phosphor was done by X-ray diffraction analysis using Bruker D8 advance X-ray diffractometer. Photoluminescence analysis (spectra) was recorded using spectrophotoflurometer made by Shimadzu [11–19].

3 Results and Discussion

3.1 XRD Analysis of Prepared Phosphor

Figure 1 shows XRD analysis of prepared phosphor focused only on those angle which clearly mentioned the crystal structure of LZO phosphor (26–38° of 2θ). Herein, the crystallite size was computed using popular Scherer's formula using so many authors [9, 10].

$$D = 0.9\,\lambda/\beta\,\cos\theta \tag{1}$$

The crystallite size found nearly 56 nm.

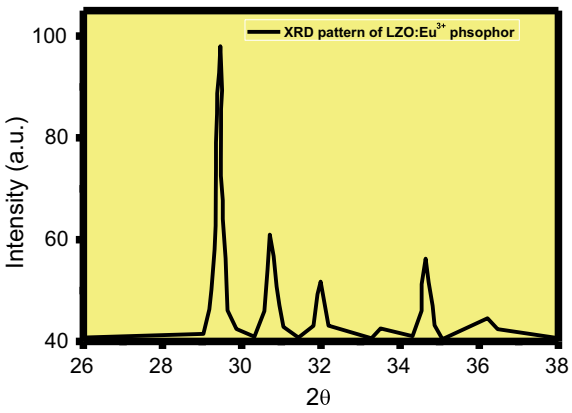


Fig. 1 XRD pattern of $La_2Zr_2O_7$ phosphor

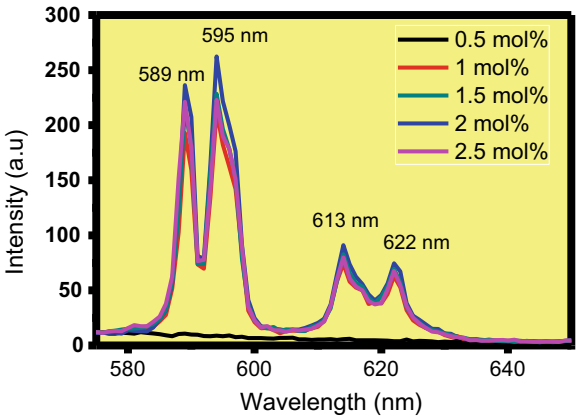


Fig. 2 PL emission spectra of prepared phosphor

3.2 PL Spectra Analysis of LZO Doped with Eu^{3+} Phosphor

It is seen from Fig. 2 the photoluminescence emission spectrum of LZO phosphor doped with various europium ion concentrations, i.e., 0.5–2.5 mol% and found that PL intensity increases with increasing the concentration of doping ions, i.e., Eu^{3+}. It shows linear response up to 2 mol%, optimized concentration for PL spectra after that PL intensity decreases due to concentration quenching phenomenon [8]. Herein, the intensity of orange light dominant over the intensity of red light emission due to magnetic dipole (MD) transition $^5D_0 \rightarrow {}^7F_1$ is dominant the electric dipole (ED) transition $^5D_0 \rightarrow {}^7F_2$, the ration of MD → ED was 3:1 and it is verified by CIE coordinate (Fig. 3). Here the peak depicted at orange light emission ($X = 0.61$ and $Y = 0.38$) it is maybe useful for plasma display panel (PDP) and several display device applications. The Judd–Ofelt analysis was done by our previous work [20]. When we compare our study with previous orange light emitted YBO_3:Eu^{3+} [21] phosphor it is found that LZO:Eu^{3+} has great potential applications and high PL intensity.

Figure 3 shows CIE coordinate for LZO:Eu doped phosphor where **X = 0.61 and y = 0.38** is nearly equal to orange light emission.

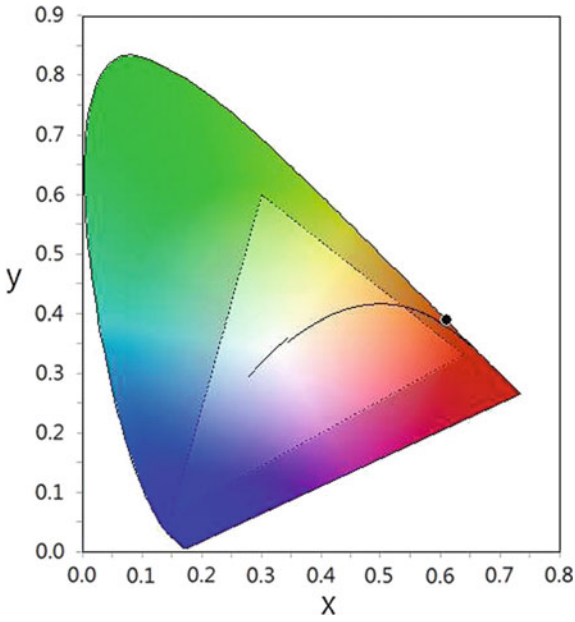


Fig. 3 CIE coordinate for LZO:Eu^{3+} doped phosphor

4 Conclusion

LZO doped with Eu^{3+} phosphors were synthesized and the emission spectra of prepared phosphor show intense orange emission. Here, MD transition dominant over the ED transition. CIE values show nearly ($x = 0.61$ and $y = 0.38$) nearer to commercially available orange luminescence phosphor.

Acknowledgements Authors are very thankful to the Department of Science and Technology funding through WOS-A fellowship. Entitled "Synthesis and Characterization of rare earth activated zirconium based phosphors for biomedical applications" File no. (Project No. SR/WOS-A/PM-14/2018 (G)).

References

1. D.R. Modeshia, R.I. Walton, Solvothermal synthesis of perovskites and pyrochlores: crystallisation of functional oxides under mild conditions. Chem. Soc. Rev. **39**(11), 4303–4325 (2010)
2. R.C. Ewing, W.J. Weber, J. Lian, Nuclear waste disposal—pyrochlore ($A_2B_2O_7$): nuclear waste form for the immobilization of plutonium and "minor" actinides. J. Appl. Phys. **95**(11), 5949–5971 (2004)
3. M.A. Subramanian, G. Aravamudan, G.V. Subba Rao, Oxide pyrochlores—a review. Prog. Solid State Chem. **15**(2), 55–143 (1983)

4. L. Kong, I. Karatchevtseva, D.J. Gregg, M.G. Blackford, R. Holmes, G. Triani, A novel chemical route to prepare $La_2Zr_2O_7$ pyrochlore. J. Am. Ceram. Soc. **96**(3), 935–941 (2013)
5. X. Min, Y. Sun, L. Kong, M. Guan, M. Fang, Y. Liu, W. Xiaowen, Z. Huang, Novel pyrochlore-type $La_2Zr_2O_7$: Eu^{3+} red phosphors: synthesis, structural, luminescence properties and theoretical calculation. Dyes Pigment. **157**, 47–54 (2018)
6. M. Mohapatra, B. Rajeswari, N. S. Hon and R. M. Kadam, Uranium luminescence in $La_2Zr_2O_7$: effect of concentration and annealing temperature. Luminescence (2016)
7. V. Singh, G. Sivaramaiah, J.L. Rao, R.S. Kumaran, P.K. Singh, T.-S. Kim, L.K. Kim, Luminescence and EPR studies of ultraviolet light emitting $La_2Zr_2O_7$:Gd^{3+} phosphor powder. J. Mater. Sci.: Mater. Electron. (2015)
8. N. Tiwari, V. Dubey, Synthesis and characterization of europium doped zirconium based phosphor for display applications. *Reviews in Fluorescence* (2016), pp. 155–184
9. Z. Qi, C. Shi, G. Zhang, Z. Han, H.H. Hung, Phys. Stat. Sol. (a) **195**(2), 311–316 (2013)
10. W.G. Lee, D.H. Lee, Y.K. Kim, J.K. Kim, J.W. Park, J. Nucl. Sci. Technol. **Supplement 4**, 572–574 (2008)
11. N. Thejo Kalyani, S.J. Dhoble, Renew. Sustain. Energy Rev. **16**, 2696 (2012)
12. N.N. Greenwood, A. Earnshaw, *Chemistry of the Elements*, 2nd edn. (Butterworth-Heinemann, Oxford, 1997)
13. A.F. Wells, *Structural Inorganic Chemistry* (Clarendon Press, Oxford, 1984)
14. J.B. Hedrick, Rare Earths In Selected U.S. Defense Applications, 40th Forum On The Geology of Industrial Minerals (2004)
15. J. Kido, K. Nagai, Y. Okamoto, J. Alloy. Compd. **192**, 30 (1993)
16. J. Kido, K. Nagai, Y. Okamoto, T. Skotheim, Chem. Lett. **235**, 1267 (1991)
17. X. Zhang, R. Sun, Q. Zheng, T. Kobayashi, W. Li, Appl. Phys. Lett. **71**, 2596 (1997)
18. Z. Hong, W. Li, D. Zhao, C. Liang, X. Liu, J. Peng, Synth. Met. **104**, 165 (1995)
19. Y. Lan, L. Yi, L. Zhou, Z. Tong, F. Gong, R. Wang, Phys. B: Condens. Matter **405**, 3489 (2010)
20. V. Dubey, R. Tiwari, R.K. Tamrakar, J. Kaur, S. Dutta, S. Das, H.G. Visser, S. Som, J. Lumin. **180**, 169–176 (2016)
21. V. Dubey, J. Kaur, S. Agrawal, N.S. Murthy, K.V.R. Suryanarayana, Superlattices Microstruct. **67**, 156–171 (2014)

CGCD Technique for Thermoluminescence Studies of $Y_2Zr_2O_7:Eu^{3+}$ Phosphor



Seema Chopra, Ratnesh Tiwari, Anshu Gupta and A. K. Beliya

1 Introduction

$Re_2Zr_2O_7$-type phosphors doped with various rare-earth ions are reported earlier by so many authors [1–4]. In view of these important properties, these types of oxides find applications as hosts for fluorescence centers and in gas turbines and diesel engines [1, 5, 6]. Nowadays so many techniques for synthesis are opted by various authors such as sol–gel synthesis, solid-state synthesis, combustion synthesis, etc. [7–11]. In this manuscript YZO:Eu^{3+} phosphor synthesized via modified solid-state reaction method and its application in thermoluminescence studies is reported.

2 Experimental

For synthesis of $Y_2Zr_2O_7$ doped with europium starting reagents used were Y_2O_3, ZrO_2, and Eu_2O_3, whereas boric acid was used as a flux. They were grounded in agate pestle and mortar and put into the alumina crucible for firing at 1000 °C for 1 h for calcination in muffle furnace. Sintering temperature is quite high, i.e., 1300 °C. Thereafter synthesized phosphor was characterized by XRD, SEM, and TL glow curve analysis of the sample irradiated by UV radiation.

S. Chopra · A. Gupta
G. D. Goenka University, Gurgaon, India

R. Tiwari (✉)
Bhilai Institute of Technology, Raipur, India
e-mail: 31rati@gmail.com

A. K. Beliya
Govt. C.V. College, Dindori, MP, India

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_84

3 Results and Discussion

Figure 1 had the X-ray diffraction analysis pattern of YZO:Eu^{3+} doped phosphor with optimized concentration of europium, i.e., 2 mol%.

SEM images of prepared phosphor (Figs. 2a, b) show cubic size formation and some particles agglomerated with each other. Here, the cubic morphology enhances the luminescence efficiency of the phosphor.

The well-resolved peak centered at 178 °C is reported for TL glow curve of YZO:Eu^{3+}. Shallower trapping and deep trapping phenomenon are dependent upon the excitation source and peak temperature of TL glow curve whereas high peak temperature shows deeper trap formation and intensity irradiation source are also for deeper trap. The trap depth as calculated by peak shape method is 2.6 eV which shows

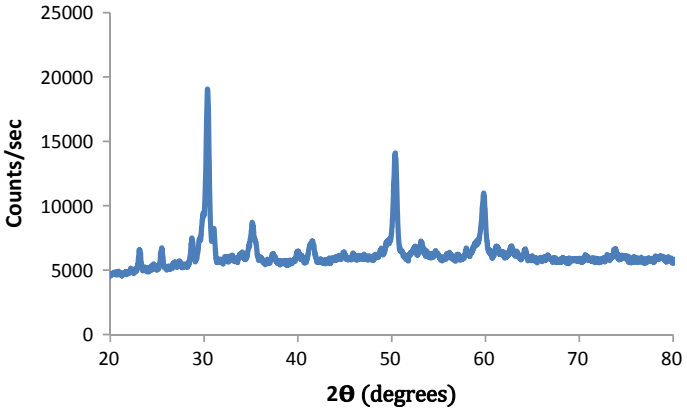


Fig. 1 XRD pattern of YZO:Eu^{3+} doped phosphor

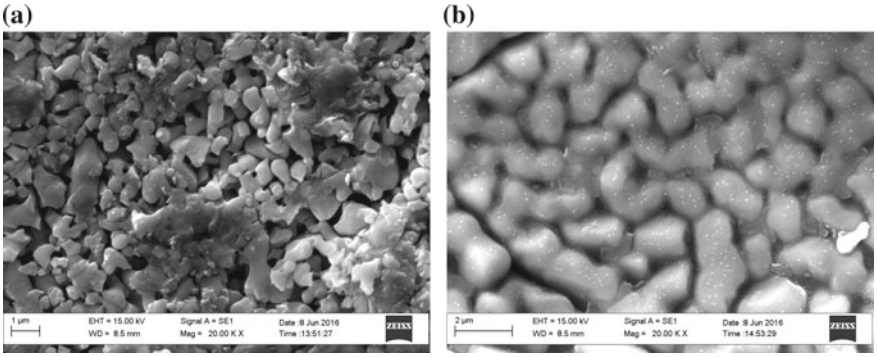


Fig. 2 **a** SEM image of YZO:Eu doped phosphor (1 μm), **b** SEM image of YZO:Eu doped phosphor (2 μm)

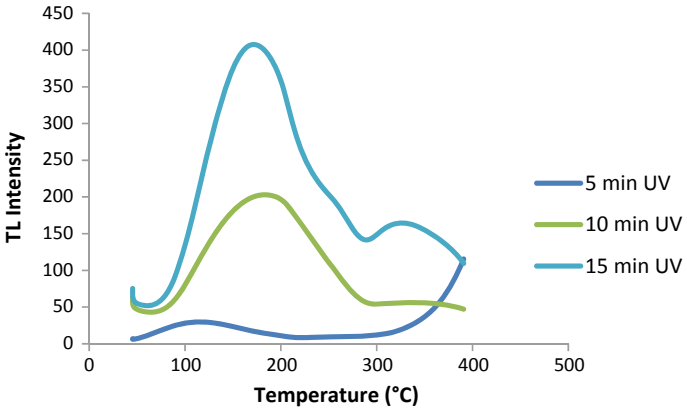


Fig. 3 Thermoluminescence glow pattern of YZO:Eu^{3+} doped phosphor

formation of shallow trapping in the prepared phosphor. Dose-response pattern had the linearity with UV dose (Fig. 3).

4 Conclusion

From TL glow pattern our prepared phosphor by solid-state method had a good response with UV dose as well as it describes shallower level trapping which may be applicable in some practical aspects of dosimeter. Herein morphological point of view due to solid-state method has a good morphology and micron-sized particle distribution. X-ray diffraction analysis shows crystal structure and its lattice parameters as well.

References

1. R. Vassen, X. Cao, F. Tietz, D. Basu, D. Stover, J. Am. Ceram. Soc. **83**, 2023 (2000)
2. J. Cheng, J.J. Li, C.Y. Ma, Z.P. Hao, Catal. Commun. **10**, 1170 (2009)
3. P. Xu, K. Holliday, K.R. Czerwinski, J.C. Nino, J. Nucl. Mater. **394**, 39 (2009)
4. K.E. Sickafus, L. Minervini, R.W. Grimes, J.A. Valdez, M. Ishimaru, F. Li, K.J. McClellan, T. Hartmann, Science **289**, 748 (2000)
5. B. Saruhan, P. Francois, K. Fritscher, U. Schulz, Surf. Coat. Technol. **182**, 175 (2004)
6. O. Otaki, T. Hoshikawa, M. Shimada, K. Koizumi, J. Ceram. Soc. Jpn. **96**, 124 (1998)
7. H. Kido, S. Komarneni, R. Roy, J. Am. Ceram. Soc. **74**, 422 (1991)
8. K. Koteswara Rao, T. Banu, M. Vithal, G.Y.S.K. Swamy, K. Ravi Kumar, Mater. Lett. **54**, 205 (2002)
9. D. Chen, R. Xu, Mater. Res. Bull. **33**, 409 (1998)
10. Y. Matsumura, M. Yoshinaka, K. Hirota, O. Yamaguchi, Solid State Commun. **104**, 341 (1997)
11. M. Yu, J. Lin, Y.H. Zhou, S.B. Wang, Mater. Lett. **56**, 1007 (2002)

A Study of Thermally Induced Vibrations of Circular Plate of Nonuniform Thickness



Narender Kumar Sarswat, Vakul Bansal, Praveen Kumar and Mahesh Kumar Aghwariya

1 Introduction

Tomar, Gupta and Jain have observed the ax-symmetric frequencies of vibrating circular plate of linear thickness non-homogeneity arises by exponential variation of elasticity [1, 2]. For circular plates, Mindlin et al. have observed, in moderately thick plates, that the shearing effects cannot be neglected [3]. Gupta and Sharma have determined vibrations for a rectangular beam with a shear effect [4]. Xiang and Reddy have studied the free vibrations of rectangular plates by leavy type solution [5]. Lee has observed the frequencies of plate by four nodes finite method influence in shear theory [6]. Lee and Cheng have observed the nonlinear free vibrations of bidirectional plates by DQ method in shear effect [7]. Tomar and Tewari analyze the frequencies of linear thermal effected of circular nonuniform plate [8]. Tomar and Gupta have investigated the impact of variable harmonically and variable linear temperature on vibrations for circular plate of varying and uniform thickness [9, 10]. Sarswat et al. have analyzed the temperature effect on damped vibration of circular and square plates by Frobenius method [15–17]. Akkaya et al. observed the transverse vibrations of beams on semi-infinite domains used in toys model and dampers. [16].

N. K. Sarswat (✉)
Department of P.G. Studies & Research in Physics, J.V. Jain College, Saharanpur 247001, UP, India
e-mail: narendersarswat12@gmail.com

V. Bansal
P. G. Studies & Research in Physics, J. V. Jain College, Saharanpur, UP, India

P. Kumar
Department of Physics, Narain College, Shikohabad, UP, India

M. K. Aghwariya
Department of Electronics and Communication Engineering, THDC Institute of Hydropower of Engineering Technology, Chokhala, UK, India

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_85

Dan and vistor have observed frequencies of mobilization of a circular plate for a tapered central hole [18] Yipeng et al. have obtained frequencies of cylindrical shell plate by IFSM method. [19]. Anavariyeh et al. have found the vibrations of vibrating circular plate with air cavity under harmonic excitation by Von Karman theory [20]. Liue et al. have obtained the complex Eigenvalues and Eigenvectors for damaged structures by naval identification technique [21]. This study observed the thermally affected frequencies of nonhomogeneous, thin circular plate of non-variable thickness influence of foundation. For frequency determination, eq"n of motion is analyzed through PSS formulation and frequencies for first four vibration modes at clamped and simply supported edges are computed. The numerical results are shown in graphical form.

2 Mathematical Analysis

Temperature Variation
Let the plate as a function of linear temperature is given by

$$t = t_0(1 - R) \tag{1}$$

where t and t_0 denote the temperature at R and $R = 0$

If the elastic modulii are temperature dependent for most engineering substances [11–13, 15, 17], then we can have

$$\bar{E}(t) = \bar{E}_0(1 - \phi t) \quad and \quad \overline{G}(t) = \overline{G_0}(1 - \phi t) \tag{2}$$

where $\bar{E}_0$, $\overline{G}_0$ are values of Young's and Shear modulus, respectively, at $t = 0$

Now considering t at $R = 1$, as reference temperature, the modulii parameter variation becomes as given below

$$\bar{E}(R) = \bar{E}_0\{1 - \beta(1 - R)\} \; and \; \overline{G}(R) = \overline{G}_0\{1 - \beta(1 - R)\} \tag{3}$$

where $\beta = \phi$, t_0 $(0 \leq \beta < 1)$ is called the thermal gradient, and ϕ is a arbitrary constant.

Equation of Motion
The eq"n of motion for plate variables is given by

$$\left(H^2\overline{E}^I + 3\bar{E}HH^I\right)\left[\overline{\Psi}^I + \left(\frac{\nu}{R}\right)\overline{\Psi}\right] + \left(H^2\bar{E}\right)\left[\overline{\Psi}^{II} + \left(\frac{\overline{\Psi}^I}{R}\right) - \left(\frac{\overline{\Psi}}{R^2}\right)\right]$$
$$- 12\left(1 - \nu^2\right)K^2\bar{G}\left[\overline{\Psi} + \bar{W}^I\right] + \left[A_0a^2\bar{\rho}H^2\right]\overline{\Psi} = 0 \tag{4}$$

and

$$\bar{W} = -\left[\frac{1}{12(1-v^2)\left(p^2a^2\bar{\rho} - \left(\frac{k_f}{H}\right)\right)}\right]\left[\left\{H^2\bar{E}\right\}\overline{\Psi}^{III}\right.$$
$$+\left\{\left(2H^2\bar{E}^I + 6HEH^I\right) + \left(\frac{2H^2\bar{E}}{R}\right)\right\}\overline{\Psi}^{II}$$
$$+\left\{H^2\overline{E}^{II} + 6HH^I\bar{E} + 6\bar{E}(H^I)^2 + 3HH^{II}E + \left(\frac{2+\nu}{R}\right)\left(H^2\bar{E}^I\right) - \left(\frac{H^2\bar{E}}{R^2}\right)\overline{\Psi}^I\right.$$
$$+\left.\left\{\left(\frac{\nu(H^2\bar{E}^{II} + 6HH^I E + 3\bar{H}H^{II}E}{R}\right) - \left(\frac{H^2\bar{E}^I + 3HH^I\bar{E}}{R^2}\right) + \left(\frac{H^2\bar{E}}{R^3}\right)\right\}\overline{\Psi}\right]$$
$$\left.-\left(\frac{1}{12}\right)\left\{H^2\bar{\Psi}^I + \left(\frac{\rho}{\rho^I}H^2 + 3HH^I + H^2/R\right)\overline{\Psi}\right\}\right. \tag{5}$$

Let thickness of the plate varies parabolically given as

$$H = H_0\left(1 - \alpha R^2\right) \tag{6}$$

where $\acute{\alpha}$ is known as taper parameter.

Imposing relation (3) and (6) on Eq. (5), one obtains the equation of motion as

$$\bar{W} = \frac{A_1}{A_2}[B_1^1\overline{\Psi}^{III} + \left\{B_2^1 + \frac{B_2^2}{R}\right\}\overline{\Psi}^{II}$$
$$+\left\{B_3^1 + \left(\frac{(2+\nu)B_3^2}{R}\right) + \frac{B_3^3}{R^2}\right\}\overline{\Psi}^I + \left\{B_4^1 + \frac{\nu B_4^2}{R} + \frac{B_4^3}{R^2} + \frac{B_4^2}{R^3}\right\}\overline{\Psi}] \tag{7}$$

where

$$A_1 = \left(-\frac{H^2}{12}\right),\quad A_2 = (\Omega^2 - D_K/C_0H),\ A_3 = (\Omega^2(1+C_0) - D_K/H),$$
$$A_4 = (\Omega^2C_0 - D_K/H)$$
$$B_1^1 = \left\langle 1 - \alpha R^2\right\rangle^2,\quad B_2^I = \left(\frac{2\beta B_1^{(1)}}{1-\beta+\beta R}\right) - (12\alpha R(1-\alpha R)),\ B_2^2 = 2B_1^{(1)},$$
$$B_3^{(1)} = \left(\Omega^2 B_1^{(1)} + 24\alpha^2R^2 - 6\alpha(3+\nu)\left(1-\alpha R^2\right) - \left(\frac{12\alpha\beta R\left(1-\alpha R^2\right)}{1-\beta+\beta R}\right)\right),$$
$$B_3^{(2)} = \left(\frac{\beta B_1^{(1)}}{1-\beta+\beta R}\right),\ B_3^{(3)} = -B_1^{(1)},\ B_4^{(1)} = \Omega^2\left(\frac{\beta B_1^{(1)}}{1-\beta+\beta R} - 6\alpha R\left(1-\alpha R^2\right)\right),$$
$$B_4^{(2)} = \left(24\alpha^2R^2 - \frac{12\alpha\beta R\left(1-\alpha R^2\right)}{1-\beta+\beta R} - 6\alpha R\left(1-\alpha R^2\right) + \frac{\Omega^2B_1^{(1)}}{\nu}\right),$$
$$B_4^{(3)} = \left(-\frac{\beta\left(1-\alpha R^2\right)^2}{1-\beta+\beta R} - 6\alpha R\left(1-\alpha R^2\right)\right)\quad B_4^{(4)} = \left(\left(1-\alpha R^2\right)^2\right)$$

$$C_0 = \left(\frac{\bar{E}_0}{(1-v^2)\overline{G_0}K^2}\right), \quad D_K = \frac{1}{\overline{G_0}K^2\frac{1}{K_F}}, \Omega^2 = \left(1-v^2\right)\left\{\frac{\bar{\rho}a^2p^2}{\bar{E}_0}\right\}$$

Hence, D_k is foundation modulus, p is circular frequency, and Ω is frequency parameter.

3 Solution

Let the solution for $\bar{W}$ has been assumed in the series form as

$$\overline{\Psi} = \sum_{\lambda=0}^{\infty} a_\lambda R^{C+\lambda} \quad when \quad a_0 \neq 0 \tag{8}$$

where c is the exponent of singularity,

Using Eqs. (3), (4), (6), and (8), one obtains

$$\begin{aligned}
&\sum_{\lambda=0}^{\infty} a_\lambda\left[n_\lambda^{(3)}E_1^{(1)} + n_\lambda^{(2)}E_2^{(1)} + n_\lambda^{(1)}E_3^{(1)} + n_\lambda E_4^{(1)} + E_5^{(1)}\right]R^{C+\lambda-4} \\
&+ \sum_{\lambda=0}^{\infty} a_\lambda\left[n_\lambda^{(3)}E_1^{(2)} + n_\lambda^{(2)}E_2^{(2)} + n_\lambda^{(1)}E_3^{(2)} + n_\lambda E_4^{(2)} + E_5^{(2)}\right]R^{C+\lambda-3} \\
&+ \sum_{\lambda=0}^{\infty} a_\lambda\left[n_\lambda^{(3)}E_1^{(3)} + n_\lambda^{(2)}E_2^{(3)} + n_\lambda^{(1)}E_3^{(3)} + n_n E_4^{(3)} + E_5^{(3)}\right]R^{C+\lambda-2} \\
&+ \sum_{\lambda=0}^{\infty} a_\lambda\left[n_\lambda^{(3)}E_1^{(4)} + n_\lambda^{(2)}E_2^{(4)} + n_\lambda^{(1)}E_3^{(4)} + n_\lambda E_4^{(4)} + E_5^{(4)}\right]R^{C+\lambda-1} \\
&+ \sum_{\lambda=0}^{\infty} a_\lambda\left[n_\lambda^{(3)}E_1^{(5)} + n_\lambda^{(2)}E_2^{(5)} + n_\lambda^{(1)}E_3^{(5)} + n_\lambda E_4^{(5)} + E_5^{(5)}\right]R^{C+\lambda} \\
&+ \sum_{\lambda=0}^{\infty} a_\lambda\left[n_\lambda^{(3)}E_1^{(6)} + n_\lambda^{(2)}E_2^{(6)} + n_\lambda^{(1)}E_3^{(6)} + n_\lambda E_4^{(6)} + E_5^{(6)}\right]R^{C+\lambda+1} \\
&+ \sum_{\lambda=0}^{\infty} a_\lambda\left[n_\lambda^{(3)}E_1^{(7)} + n_\lambda^{(2)}E_2^{7} + n_\lambda^{(1)}E_3^{(7)} + n_\lambda E_4^{(7)} + E_5^{(7)}\right]R^{C+\lambda+2} \\
&+ \sum_{\lambda=0}^{\infty} a_\lambda\left[n_\lambda^{(1)}E_1^{(8)} + n_\lambda E_2^{(8)} + E_3^{(8)}\right]R^{C+\lambda+3} + + \sum_{\lambda=0}^{\infty} a_\lambda\left[E_1^{(10)}\right]R^{C+\lambda+5} \\
&\sum_{\lambda=0}^{\infty} a_\lambda\left[n_\lambda^{(1)}E_1^{(9)} + n_\lambda E_2^{(9)} + E_3^{(9)}\right]R^{C+\lambda+4} + \sum_{\lambda=0}^{\infty} a_\lambda\left[E_1^{(11)}\right]R^{C+\lambda+6} = 0
\end{aligned} \tag{9}$$

Respectively, where

$$n_{\lambda}^{(3)} = (c+\lambda)(c+\lambda-1)(c+\lambda-2)(c+\lambda-3),$$
$$n_{\lambda}^{(2)} = (c+\lambda)(c+\lambda-1)(c+\lambda-2),$$

$$n_{\lambda}^{(1)} = (c+\lambda)(c+\lambda-1), n_{\lambda} = (c+\lambda)\,,\ d = 1-\beta$$

$$E_1^{(1)} = d^2 = 2E_2^{(1)},\ \ E_3^{(1)} = -3E_1^{(1)}\ \ , E_4^{(1)} = 3E_1^{(1)}, E_5^{(1)} = E_3^{(1)},$$

$$E_1^{(2)} = 2\beta E_1^{(1)}, E_2^{(2)} = 6d\beta, E_3^{(2)} = \beta d(\nu-4), E_4^{(2)} = \beta d(3-\nu), E_5^{(2)} = -4\beta d$$

$$E_1^{(3)} = \left(\beta^2 - 2\alpha d^2\right),\ E_2^{(3)} = \left(4\beta^2 - 20\alpha d^2\right),$$

$$E_3^{(3)} = \left(d^2 c - \left(32 + 6\nu\alpha d^2\right) - \beta^2(3-\nu)\right),$$

$$E_4^{(3)} = \left(d^2 c + (4-6\nu)\alpha d^2 + 2\beta^2\right), E_5^{(3)} = E_1^{(1)}(6\alpha\nu - 4\alpha - c),$$

$$E_1^{(4)} = -4\beta\alpha d, E_2^{(4)} = 11E_1^{(4)}, E_3^{(4)} = (2d\beta c - (88+14\nu)\beta d\alpha),$$

$$E_4^{(4)} = (3d\beta c - (6+26\nu)\beta d\alpha),$$

$$E_5^{(4)} = \left(d\beta\nu C_0\Omega^2 + (12\nu-4)\beta d\alpha\right) - 4d\alpha - 2d\beta c$$

$$E_1^{(5)} = \left(d^2\alpha^2 - 2\alpha\beta^2\right), E_2^{(5)} = \left(18d^2\alpha^2 - 24\alpha\beta^2\right),$$

$$E_3^{(5)} = \left(c\left(\beta^2 - 2\alpha d^2\right) + (83+6\nu)\alpha^2 d^2 - \beta^2\alpha(52+8\nu)\right),$$

$$E_4^{(5)} = \left(d^2\alpha^2(77+42\nu) - 2\alpha d^2\Omega^2(6+4C_0) + 2\beta^2 c - \beta^2\alpha(2+22\nu)\right),$$

$$E_5^{(5)} = \left(d^2\Omega^4 C_0 - \Omega^2\beta^2(2+C_0-C_0\nu) - 2\alpha\beta^2\Omega^2(1-9\nu) - 4\alpha\beta\right)$$
$$- \alpha\Omega^2 d^2(16+6C_0\nu-2C_0) - 4d^2\alpha - \frac{12\Omega^2 d^2}{H_0^2} + d^2\alpha^2(30\nu-5)$$

$$E_1^{(6)} = 2\beta d\alpha^2, E_2^{(6)} = 38\beta d\alpha^2, E_3^{(6)} = d\beta\alpha^2(188+13\nu) - 4d\beta\alpha c,$$

$$E_4^{(6)} = d\beta\alpha^2(195+99\nu) - d\beta\alpha\Omega^2(18C_0+26)$$

$$E_5^{(6)} = \left(-d\beta\alpha\Omega^2(36+14C_0\nu+4C_0) + 2\beta d\Omega^4 C_0 - 8\beta d\alpha - 2\alpha^2\beta d(84\nu-10)\right)$$
$$- \frac{24\beta d\Omega^2}{H_0^2},$$

$$E_1^{7)} = \beta^2\alpha^2, E_2^{7)} = 20\beta^2\alpha^2, E_3^{7)} = \beta^2\alpha^2(103+7\nu) + c\left(d^2\alpha^2 - 2\alpha\beta^2\right),$$

$$E_4^{7)} = \beta^2\alpha^2(108+58\nu) - \beta^2\alpha\Omega^2(14+10C_0) + \beta^2\alpha^2\Omega^2(11+7C_0),$$

$$E_5^{7)} = -\beta^2\Omega^2\alpha(18+8\nu C_0+2C_0) - \beta^2\alpha^2(6-42\nu) + d^2\alpha^2\Omega^2(29+C_0 6\nu - C_0)$$
$$+ \Omega^4\beta^2 C_0 - 2\alpha\Omega^4 d^2 C_0 - \frac{12\beta^2\Omega^2}{H_0^2} - 4\beta^2\alpha + 4\beta^2 d^2$$

$$E_1^{8)} = 2\beta d\alpha^2 c, E_2^{8)} = \beta d\alpha^2\Omega^2(23+13C_0), E_3^{8)} = \beta d\alpha^2\Omega^2(62+13C_0\nu-2C_0)$$

$$_4\beta d\alpha\Omega^4 C_0 + 8\beta d\alpha^2,\ E_4^{9)} = \beta^2\alpha^2\Omega^2 c,\ E_2^{9)} = \beta^2\Omega^2\alpha^2(12 + 8C_0),$$
$$E_3^{3)} = \beta^2\alpha^2\Omega^2(32 + 7C_0\nu - C_0) + d^2\alpha^2\Omega^4 C_0 - 2\beta^2\alpha\Omega^4 C_0 + 4\beta^2\alpha^2,$$
$$E_1^{108)} = 2\beta d\alpha^2\Omega^4 C_0,\ E_1^{111)} = \beta^2\alpha^2\Omega^4 C_0,$$

The indicial roots observed for equating to zero, the coefficient of the lowest power of R is $c = 3, 1, 1, -1$. For next higher power of R, $a_1 = 0$ at $c = 1$. For R^{C-2}, a_2 is indeterminate write with a_0. For other higher powers of R, a_3–a_5 is find in terms of a_0 and a_2. The left unknown constants $a_\lambda(\lambda = 6, 7.)$ are observed from the below relation

$$\begin{aligned}
&a_{\lambda+10}\Big[B_1E_1^{(1)} + E_3^{(1)}B_2 + E_4^{(1)}B_3 + E_5^{(1)}\Big] + a_{\lambda+9}\Big[B_4E_1^{(2)} + B_5E_3^{(2)} + B_6E_4^{(2)} + E_5^{(5)}\Big]\\
&+ a_{\lambda+8}\Big[B_7(B_8(B_9E_1^{(3)} + E_2^{(3)}) + E_3^{(3)}) + E_4^{(3)}B_{17} + E_5^{(3)}\Big]\\
&+ a_{\lambda+7}\Big[B_{18}(B_9(B_{11}E_1^{(4)} + E_2^{(4)}) + E_3^{(4)}) + E_4^{(4)}B_{19} + E_5^{(4)}\Big]\\
&+ a_{\lambda+6}\Big[B_{10}(B_{11}(B_{12}E_1^{(5)} + E_2^{(5)}) + E_3^{(5)}) + E_4^{(5)}B_8 + E_5^{(5)}\Big]\\
&+ a_{\lambda+5}\Big[B_{13}(B_{12}(B_{14}E_1^{(6)} + E_2^{(6)}) + E_3^{(6)}) + E_4^{(6)}B_9 + E_5^{(6)}\Big]\\
&+ a_{\lambda+4}\Big[B_{15}(B_{14}(B_{16}E_1^{(7)} + E_2^{(7)}) + E_3^{(7)}) + E_4^{(7)}B_{11} + E_5^{(7)}\Big]\\
&+ a_{\lambda+3}\Big[B_{12}(B_{14}E_1^{(8)} + E_2^{(8)}) + E_3^{(8)})\Big] + a_{\lambda+2}\Big[B_{14}(B_{16}E_1^{(9)} + E_2^{(9)}) + E_3^{(9)})\Big]\\
&+ a_{\lambda+1}\Big[E_1^{(10)})\Big] + a_\lambda\Big[E_1^{(11)})\Big] = 0
\end{aligned} \tag{10}$$

where

$$\begin{aligned}
&B_1 = (c+\lambda+10)(c+\lambda+9)^2(c+\lambda+8),\ B_2 = (c+\lambda+10)(c+\lambda+9)\\
&B_3 = (c+\lambda+10),\ B_4 = (c+\lambda+9)^2(c+\lambda+8)(c+\lambda+7),\\
&B_5 = (c+\lambda+9)(c+\lambda+8),\ B_6 = (c+\lambda+9),\ B_7 = (c+\lambda+8)(c+\lambda+7)\\
&B_8 = (c+\lambda+6),\ B_9 = (c+\lambda+5),\ B_{10} = (c+\lambda+6)(c+\lambda+5),\\
&B_{11} = (c+\lambda+4)\\
&B_{12} = (c+\lambda+3),\ B_{13} = (c+\lambda+5)(c+\lambda+4),),\ B_{14} = (c+\lambda+2),\\
&B_{15} = (c+\lambda+4)(c+\lambda+3),\ B_{16} = (c+\lambda+1),\ B_{17} = (c+\lambda+8),\\
&B_{18} = (c+\lambda+7)(c+\lambda+6),\ B_{19} = (c+\lambda+7)
\end{aligned}$$

If the notations

$$a_\lambda = F_\lambda a_0 + G_\lambda a_2 \quad with \quad \lambda = 0, 1, 2, 3, 4 \ldots\ldots\ldots\ldots \tag{11}$$

are introduced, the solution $\overline{\Psi}$ at $c = 1$ are given as

$$\overline{\Psi} = a_0 x(R, \Omega) + a_2 y(R, \Omega) \tag{12}$$

where

$$x(R,\Omega) = R + \sum_{\lambda=1}^{\infty} F_\lambda R^{\lambda+1} \; and \; y(R, \quad \Omega) = \sum_{\lambda=1}^{\infty} G_\lambda R^{\lambda+1} \Bigg\} \tag{13}$$

where $F_\lambda \; and \; G_\lambda$ are function of β, D_k, $\Omega \; and \; C_0$. The solution at $c = 1$, -1 of the indicial equation being singular at $R = 0$ are omitted, because they are of no interest to any one. The Eqs. (11) and (13) are uniformly convergent $where \quad |\mu < 1|$ according [14].

The solution for $\overline{\Psi}$, in Eq. (6) and the solution for $\bar{W}$ is given as:

$$\bar{W} = a_0 u_1(R,\Omega) + a_2 u_2(R,\Omega) \tag{14}$$

where

$$\begin{aligned} u_1(R,\Omega) = -\left(\frac{A_1}{A_2}\right)\Bigg[& B_1^{(1)} X^{III} + \left(B_2^{(1)} + \frac{B_2^{(2)}}{R}\right) X^{II} \\ & + \left(B_3^{(1)} + \frac{B_3^{(2)}}{R} + \frac{B_3^{(3)}}{R^2}\right) X^I + \left(B_4^{(1)} + \frac{B_4^{(2)}}{R} + \frac{B_4^{(3)}}{R^2} + \frac{B_4^{(4)}}{R^3} + \right) X \Bigg] \\ u_2(R,\Omega) = -\frac{A_1}{A_2}\Bigg[& B_1^{(1)} Y^{III} + \left(B_2^{(1)} + \frac{B_2^{(2)}}{R}\right) Y^{II} + \left(B_3^{(1)} + \frac{B_3^{(2)}}{R} + \frac{B_3^{(3)}}{R^2}\right) Y^I \\ & + \left(B_4^{(1)} + \frac{B_4^{(2)}}{R} + \frac{B_4^{(3)}}{R^2} + \frac{B_4^{(4)}}{R^3}\right) Y \Bigg] \end{aligned} \tag{15}$$

This solution is non-singular at all points of the plate because the substitution for X and Y cancels R from the denominator in the terms involving $1/R$, $1/R^2$, and $1/R^3$

4 Edges Condition

Now employing the given boundary conditions, frequency equations of circular plates have been obtained as

For **clamped edge**:

Now plate clamped at $r = a$, then we get

$$\bar{W}|_{R=1} = 0 = \overline{\Psi}|_{R=1} \tag{16}$$

Imposing (15) to (11) and (13), one gets, after eliminating a0 and a2, frequency equation for clamped- plate as

$$\begin{vmatrix} x(1,\Omega) & y(1,\Omega) \\ u_1(1,\Omega) & u_2(1,\Omega) \end{vmatrix} = 0 \tag{17}$$

where $x\ (1,\ \Omega)$, $y\ (1,\ \Omega)$,$u_1\ (1,\ \Omega)$, and $u_2\ (1,\ \Omega)$ are given by Eqs. (12) and (14), respectively:

For simply supported edge:

Now plate at the edge r = a, then deflection and momentum at this edge should be zero

$$\bar{W}\Big|_{R=1} = 0 = \left[\overline{\Psi}^1 + \frac{\nu}{R}\overline{\Psi}\right]\Big|_{R=1} \tag{18}$$

Again using (11) and (13), one obtains the frequency equation for this boundary condition, the plate as

$$\begin{vmatrix} u_1\ (1,\ \Omega) & u_2\ (1,\ \Omega) \\ u_3\ (1,\ \Omega) & u_4\ (1,\ \Omega) \end{vmatrix} = 0 \tag{19}$$

where

$$u_3(1,\Omega) = \left[\frac{\partial x(R,\Omega)}{\partial R} + \frac{\nu}{R}x(R,\Omega)\right] at \quad R=1$$
$$u_4(1,\Omega) = \left[\frac{\partial y(R,\Omega)}{\partial R} + \frac{\nu}{R}y(R,\Omega)\right]\Big|_{R=1}.$$

5 Results and Conclusion

Many roots can be observed from expressions (16) and (18) for frequency parameters. Here, the vibrations of first two modes of a clamped and simply supported circular plates have been observed on various values of $\acute{\alpha}$, β, and D_K using computer programming and first two modes of vibration here are presented in graphical form. In all the cases, the values of $\nu = 0.4$ and $k2 = 0.85$ have been considered here. The results in Fig. 1 shows the thermal effect of Ω with variation D_k, under shear influences. Observation from these figure is that the Ω increases with D_k for all presented modes of vibration, here for both conditions of plates. It is also noted that the Ω in clamping edges are upper than the corresponding Ω for ss conditions and the frequencies of thermal effective plate are lower than that of normal plate in both the edges. The Fig. 2 represents the thermal variation on the vibrations of a circular plate of uniform thickness at $\acute{\alpha}$ and D_K. It is observed that in both boundary conditions, the Ω decreases against β for nonhomogeneous materials of the plate. One can see that the frequencies of thermal effective plate are greater than that of unheated. Also, the value of Ω in shear theory is less than the corresponding Ω in

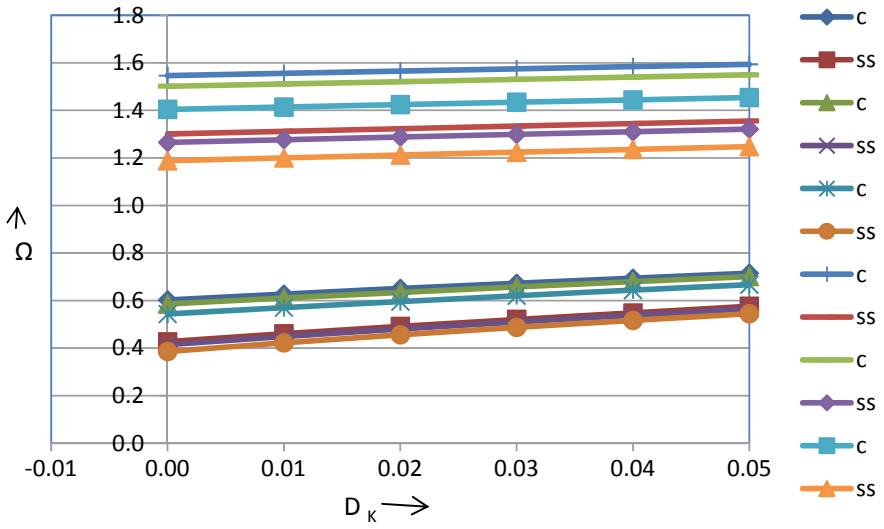


Fig. 1 Frequencies with variations of foundation parameter [$\nu = 0.4$, $\acute{\alpha} = 0.1$, $\beta = 0.0, 0.1, 0.3$]

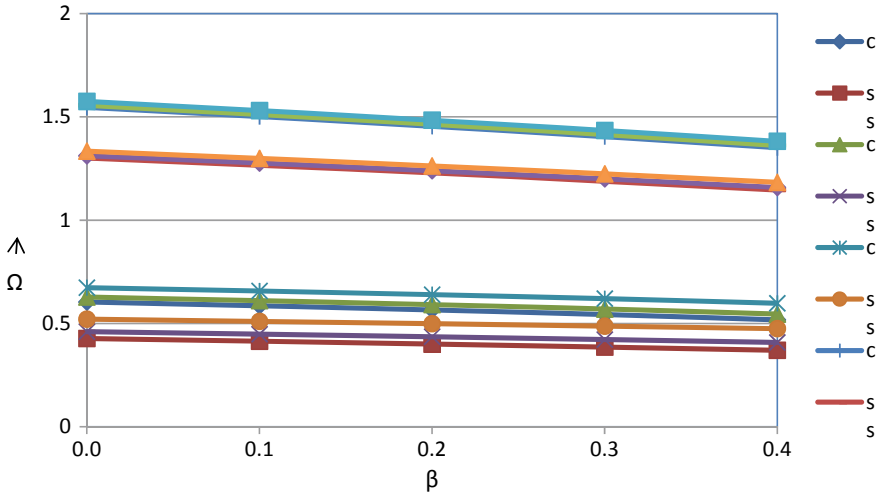


Fig. 2 Frequencies with variations of temperature parameter [$\nu = 0.4$, $\acute{\alpha} = 0.1$, $D_k = 0.0, 0.01, 0.03$]

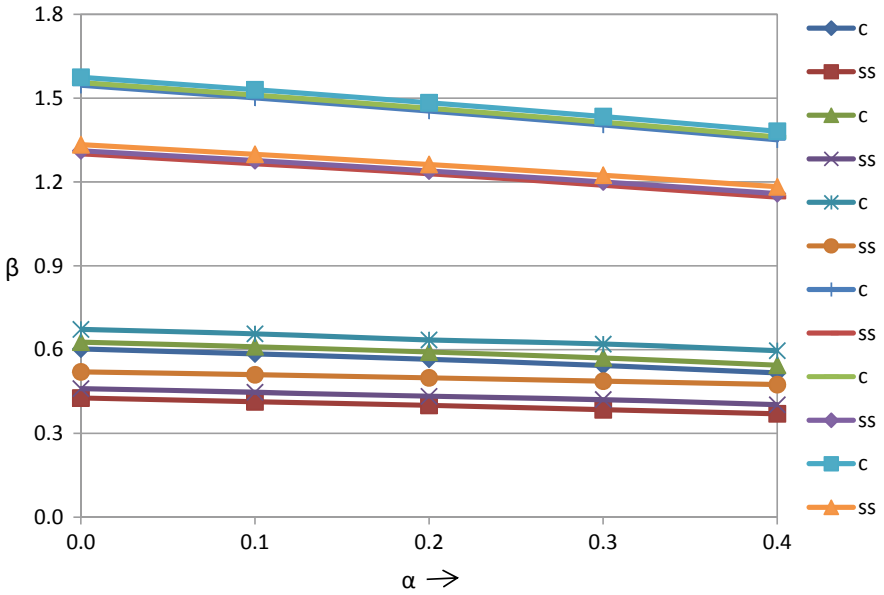


Fig. 3 Frequencies with variations of thickness parameter [$\nu = 0.4$, $\acute{\alpha} = 0.1$, $D_k = 0.01$, $\beta = 0.0$, 0.1, 0.3]

classical plate theory. Its clear that the effects of shear and rotatory inertia cannot be neglected for thin plates when vibrations corresponding to the higher modes are considered (Fig. 3).

References

1. J.S. Tomar, D.C. Gupta, N.C. Jain, Axisymmetric vibrations of an isotropic elastic non-homogeneous circular plate of linearly varying thickness. J. Sound Vib. **85**(3), 365–370 (1982)
2. J.S. Tomar, D.C. Gupta, V. Kumar, Natural frequencies of a linearly tapared non-homogerneous isotropic elastic circular plates resting on an elastic foundation. J. Sound Vib. **III**(1), 1–8 (1986)
3. R.D. Mindlin, H. Deresiewicz, Thickness shear and flexural vibrations of a circular disk. J. Appl. Phys. **25**(10), 1329–1332 (1954)
4. A.P. Gupta, N. Sharma, Effect of transverse shear and rotatory inertia on the forced motion of a stepped rectangular beam: J. Sound Vib. **209**(5), 811–820 (**1998**)
5. Y. Xiang, J.N. Reddy, Natural vibrations of rectangular plates with an internal line hinge using the first order shear deformation plate theory. J. Sound Vib. **263**(2), 285–297 (2003)
6. S.J. Lee, Free vibrations analysis of plates by using a four node finite element formulated with assumed natural transverse shear strain. J. Sound Vib. **278**(3), 657–684 (2004)
7. J.J. Li, C.J. Cheng, Differential quadrature method for non-linear vibrations of orthotropic plates with finite deformation and transverse shear effect. J. Sound Vib. **281**(120), 295–309 (2005)
8. J.S. Tomar, V.S. Tewari, Effect of thermal gradient on frequencies of a circular plates of linearly varying thickness. J. Non-equilib. Thermodyn. **6**, 115–122 (1981)

9. J.S. Tomar, A.K. Gupta, Thermal effect on axisymmetric vibrations of an orthotropic circular plate of variable thickness. J. AIAA **22**(7), 1015–1017 (**1984**)
10. J.S. Tomar, A.K. Gupta, Thermal effect on axisymmetric vibrations of an orthotropic circular plate of parabollicaly varying thickness. J. Pure Appl. Math. **16**(5), 537–545 (**1985**)
11. S. Spinner, Temperature dependence of elastic constant of some cermet specimens, Research National Bureau Standard C. Eng. Inst. **65**, 89–96 (1961)
12. I.E. Garrick, Survey of aerothermoelasticity. J. Aerosp. Eng. **22**, 140–147 (1963)
13. E.H. Mansfield, Bending buckling and curling of a heated elliptical plate. Proc. R. Soc., Ser.—A **28**, 296–417 (**1965**)
14. H. Lamb, Hydrodynamics, Dover Publications, New York, p. 335 (**1945**)
15. N.K. Sarswat, Effect of damping on thermal vibrations of elastic circular plates. IJEARTS **8**(20), 12–21 (**2015**)
16. T. Akkaya, W.T.V. Horssen, On the transverse vibrations of strips and beams on the semi infinite domains. Proc. IUPAM 226–273 (**2016**)
17. N.K. Sarswat, V. Bansal, The study of damping and parabolic temperature on vibrations of isotropic elastic square plate: Asia pacific. JAEEE **1**(2), 1–8 (2017)
18. B.O. Dan, V.K. Vistor, Damping of flexural vibrations in circular plates with tapered central holes. J. Sound Vib. **330**(1), 2220–2238, (**2011)**
19. Y. Cao, R. Zhang, W. Zhang, J. W Vibrations characteristics analysis of cylindrical shell plate coupled structure using an improved fourier series method. Shock Vib. 19 (2018)
20. F.S. Anvariyeh, M.M. Jalili, Nonlinear vibration analysis of a circular plate -cavity system. J. Braz. Soc. Mech. Sci. Eng. (2019)
21. J. Liue, Z. Lu, M. Yu, Damaged identification of non classically damped shear building by sensitivity analysis of complex model parameter. J Sound Vib. **438**(6), 457–475 (2019)

Industrial Motor Bearing Fault Detection Using Vibration Analysis



Rajul Misra, Kshitij Shinghal, Amit Saxena and Alok Agarwal

1 Introduction

In most of the industries, the induction motor plays a very important role in driving conveyor belts, machines, etc. Their unexpected breakdown can have a high financial impact on increased cost of production due to cost overhead during the system shutdown time. According to IEEE motor reliability study, 41% of total faults belong to bearing defects [1–4], followed by stator (37%) and by rotors 10%. Thermal or dynamic mechanical stresses, damages in inner or outer races, misalignment, poor fitting, and insufficient lubricant due to thermal or electrical stresses are the major causes of faults in bearing. Therefore, many research studies have been carried out on fault detection [5, 6]. Schoen et al. on the basis of their work proposed a system in which the bearing fault was detected by the help of rotating eccentricities generated due to the rotation of the faulty bearing [7]. As a faulty bearing shows eccentricity and rotates in erratic manner, this technique proposed by Schoen et al. is being used in many modern machines for detection of bearing faults [8–12]. Although numerous condition monitoring schemes with sensors/transducers have been developed,

R. Misra (✉) · A. Agarwal
Department of Electrical Engineering,
Moradabad Institute of Technology, Moradabad, UP, India
e-mail: rajulmisra71@gmail.com

A. Agarwal
e-mail: alokagarwal2@gmail.com

K. Shinghal · A. Saxena
Department of Electronics and Communication Engineering,
Moradabad Institute of Technology, Moradabad UP, India
e-mail: kshinghal@gmail.com

A. Saxena
e-mail: amitssaksena@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_86

vibration signal based methods are most commonly and widely accepted for monitoring the health of bearing [13–16]. These methods are applied on the principle that if there is an occurrence of single-point fault in the induction motor bearing, then it may cause an occurrence of comparatively more catastrophic failure. Based on the characteristics, a single-point failure in the barrier will have a specific signature vibration frequency. Therefore, the bearing fault can be identified specifically by carefully recording and analyzing the vibration frequency, because each single point bearing fault will result in different motor torques and different current requirements generating a specific signature vibration frequency. Hence, the current harmonics at the specific frequency will be affected by bearing vibration and this current spectrum can be used to detect the bearing faults. The current spectrum includes information such as load oscillation, broken rotor bar, and rotor eccentricity [17, 18]. Muhammad Irfan et al. have developed an intelligent condition monitoring system for identifying outer race bearing faults which utilizes the commonly available motor current and voltage [19]. The developed system continuously tracks real-time outer race defects of the bearing and is capable of determining severity of the faults. But it has certain limitations; first, it requires expert knowledge and full technical specifications of under observed bearing; second, this method is valid for six to twelve balls of bearing only. To overcome these issues in the present work, a vibration signal based bearing faults detection scheme for outer as well as inner race has been developed for any number of bearing balls which does require neither geometry of bearing nor detailed technical specifications of bearing. The rest of the paper is organized as follows: next section illustrates the bearing faults, their possible causes and remedies, Sect. 3 describes the standard, proposed formulas of bearing characteristic frequencies, in Sect. 4, experimental studies have been carried out, and finally conclusions are drawn in last Section.

2 Bearing Faults: Causes and Remedies

Based on industrial best practices, the most common causes and their remedies for bearing failures are as follows:

1. **Improper installation**: For satisfactory operation of bearings, it should be properly installed and must be aligned in all three axes. It should be properly locked to the shaft and shafts must be within tolerance alignment.
2. **Misalignment of drives**: For a prolonged good condition of the drive and permissible vibrations, precise alignment is very important. Misalignment can be avoided by laser alignment or double dial gauges.
3. **Tightness and looseness of belts**: Appropriate tightness in the belt should be maintained at operational time. Time to time inspection of belts to be done.
4. **Improper lubrication**: Lubricant (grease) must be used according to manufacturer compliance and it should be of good quality e.g., EP 1.5–EP 2.

5. **Rotor unbalancing**: This problem should be resolved at start-up. Due to unbalancing, the motor starts moving and vibration level increases. If this suddenly occurs in the motor, then wear, debris buildup, broken, or cracked parts of the rotor should be checked and corrective actions should be taken immediately.
6. **Over (critical) speed**: Sometimes drive runs at beyond full load speed, in such situations, the speed of drive should be properly checked by AMP probe, tachometer, or stroboscope and corrective actions should be taken as earliest as possible.
7. **Poor or no isolation**: Proper isolation should be done and if there is any cracked or broken isolation, it should be replaced immediately.
8. **Contamination**: Extreme care should be taken while installing IM bearings because bearing exposed to moisture, other contaminations may get rusted over time and may get cracked or broken during the operation.
9. **Poor or wrong bearing selection**: New bearing should always be of the same rating as the original one if replaced. Precautions should be taken while driving heavy motor such as their belt drives should have roller bearings in the driving end and fixed bearing in the other end.
10. **Electrically induced bearing damage**: Generally, the motors working on voltages of around 575 V with variable frequency drives often encounter a situation where there is electrically induced voltage building on the rotor of the motor and this building results in milliampere discharge through the bearings back to the stator. Due to this issue, loud noise from bearing, its inner and outer race, rubbing, scratching, riveting, and rolling of balls in races is generated. There are many remedies for avoiding electrically induced bearing damage such as sealed DE/NDE motor bearings packed with electro conductive grease, rotor grounding seal/ring, and use of sine wave filter.

3 Bearing Characteristics Frequencies

A typical induction motor bearing is shown in Fig. 1. The various parts are depicted as (i) inner core of the bearing (ii) outer core of the bearing (iii) rolling balls (iv) its cage.

Analytical models relating to the vibration response of bearing with single-point defects can be found in [20–22]. When a fault occurs on the surface of the bearing, the corresponding fault characteristic frequency is generated. In order to locate the fault, it is mandatory to identify the presence of fault occurrence and its related frequency. These frequencies are a function of the bearing geometry (e.g. pitch diameter, roller diameter, etc.) and relative speed between the two raceways [23–25]. When bearing geometry is known, the fundamental fault frequencies can be calculated.

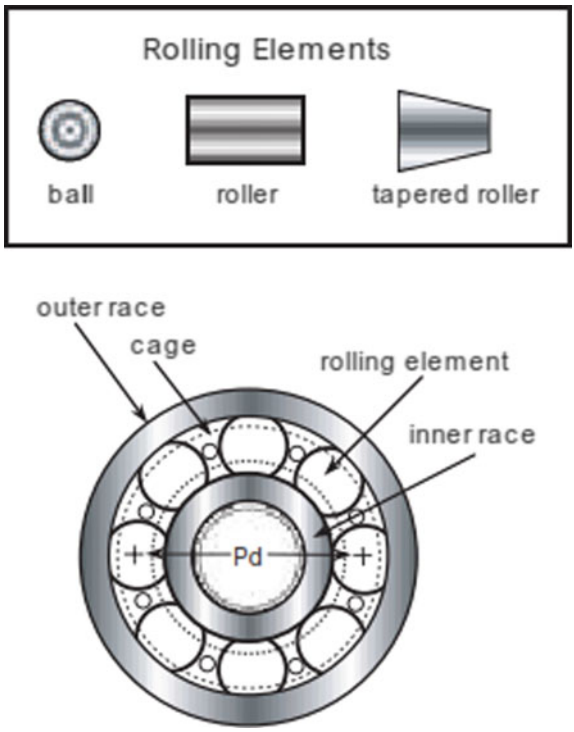


Fig. 1 Indicative diagram of bearing

3.1 Fundamental Train Frequency (FTF)/Cage Defect

When faults occur in the bearing cage, the FTF is shown as

$$\mathrm{FTF} = \frac{\mathrm{S}}{2}\left(1 - \frac{\mathrm{B_d}}{\mathrm{P_d}}\cos\emptyset\right) \tag{1}$$

3.2 Ball Pass Frequency Inner Race (BPFI)

If the inner race of bearing has a crack, the fundamental vibration frequency resulting from ball passing over the fault is called BPFI and given as follows:

$$\mathrm{BPFI} = \frac{\mathrm{N_b}}{2}\mathrm{S}\left(1 + \frac{\mathrm{B_d}}{\mathrm{P_d}}\cos\emptyset\right) \tag{2}$$

3.3 Ball Pass Frequency Outer Race (BPFO)

Similarly, if the bearing outer race has a crack, the fundamental vibration frequency resulting from ball passing over the fault is called BPFO which is given by

$$\text{BPFO} = \frac{N_b}{2} S\left(1 - \frac{B_d}{P_d}\cos \emptyset\right) \tag{3}$$

3.4 Ball Spin Frequency (BSF)/Rolling Element Defect

If the ball has a defect such as pit, the pulse repetition rate occurs each time, the defect is struck, is known as the BSF and given by

$$\text{SF} = \frac{P_d}{2B_d} S\left\{1 - \left(\frac{B_d}{P_d}\right)^2 (\cos \emptyset)^2\right\} \tag{4}$$

where S represents the speed of machine, B_d is ball or roller diameter, N_b is number of balls, P_d is pitch diameter, and $\emptyset$ is ball contact angle (typically equal to 0°).

Besides the above formulation, author proposed the following formulas for characteristics frequencies. All these formulas are deduced from the experience of various experiments performed in Jindal Steel and Power Limited, Raigarh, India. The main feature of these formulas is that only number of balls and speed of observed motor are required to calculate the characteristics frequencies.

$$\text{Proposed BPFO} = \left[\frac{N_b}{2} - 1.2\right] S \tag{5}$$

$$\text{Proposed BPFI} = \left[\frac{N_b}{2} + 1.2\right] S \tag{6}$$

4 Experimental Setup

To validate the proposed formulas, various case studies have been carried out for inner and outer race faults. The snapshot of the experimental setup is shown in Fig. 2.

Case 1. Study of vibration analysis of motor bearing at DE and NDE under outer race fault condition.

The specifications (geometry and defect multipliers for characteristic frequencies @manufacturer: SKF) of under observed bearings are tabulated in Table 1. An experimental study has been performed on two bearings namely 22218 and NU1048.



Fig. 2 Snapshot of experimental setup

Table 1 Specifications of bearing

Case study	Bearing no.	Motor speed (CPM)	N_b	B_d	P_d	Φ	Defect multiplier (BPFO)	Defect multiplier (BPFI)
1	22218	1481	19	0.669	4.939	10	8.23	10.77
2	NU1048	986	25	1.181	11.811	0	11.25	13.75

First, vibrations in horizontal, vertical and axial directions are measured and tabulated in Table 2.

The vibration of fault bearing is given in Fig. 3. From the graph, it is evident that the motor speed is 1481 CPM as indicated by point R and point B (24944 CPM) is the second harmonic of point A (12471 CPM), i.e., frequency at point B is the twice of frequency of point A.

The characteristic frequencies can be calculated by the following steps:

A. Look up the bearing number. Corresponding to this number, find bearing geometry and defect multipliers.
B. Multiply this multiplier by the speed of motor which yields characteristics frequencies. In this case, the speed of motor is 1481 CPM, therefore the characteristics frequencies will be 12188.63 CPM (BPFO) and 15950 CPM (BPFI).
C. If harmonics are present in a vibration pattern, then one can easily classify the types and location of the fault. In this case, it is easily observed that the

Table 2 Measurement of vibrations

Direction	Vibration (22218), mm/s		Remarks
	DE	NDE	
Horizontal	4.7	8.4	Vibration found in increasing trend at NDE bearings and abnormal sound observed from bearings Status: **NOT OK**
Vertical	3.5	6.9	
Axial	3.7	4.2	

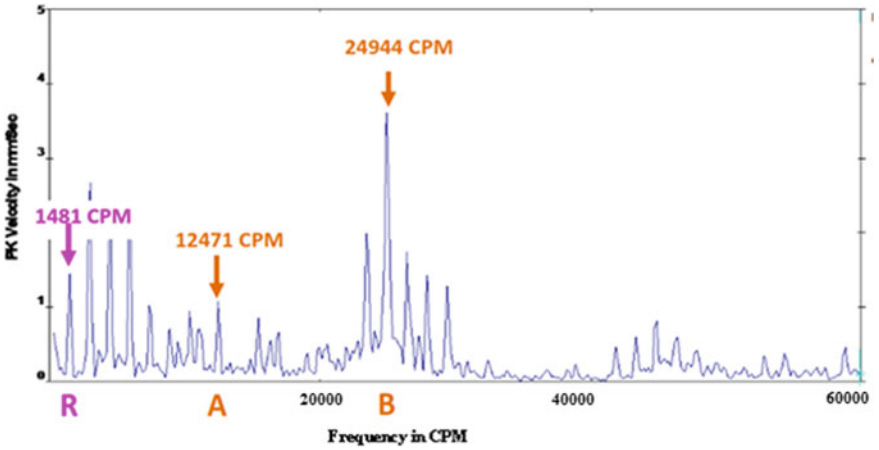


Fig. 3 Vibration patterns (faulty bearing, 22218) at NDE

characteristics frequency of 12188.63 CPM is comparable to 12471 CPM which is measured by vibration analyzer. There is also a second harmonic (24944 CPM) of BPFO presented in vibration pattern which indicates outer race fault of the bearing.

The characteristic frequencies for both bearings are calculated from standard formulas (Eqs. 2 and 3) and proposed formulas (Eqs. 5 and 6) and are tabulated in Table 3 as given above.

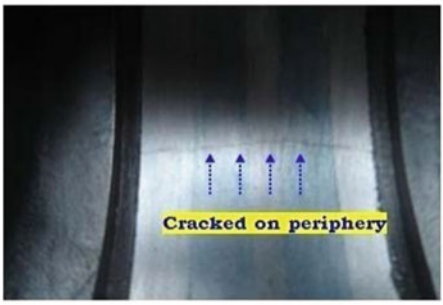
It is observed from Table 3 that without knowing the bearing geometry and technical specifications from the manufacturer, proposed formulas are capable to determine the characteristic frequencies and hence, one can detect the type of fault very easily and efficiently. After inspection of NDE bearing and its adjacent component (i.e., housing), it is found that outer race of the bearing has defect. The blue impressions of the defected bearing are systematically depicted in Fig. 4.

After the replacement of faulty bearing, vibrations have been recorded again which are tabulated in Table 4.

After corrective action, the vibration pattern has been recorded and shown in Fig. 5 and it is observed that dominant peaks at Point A (12471 CPM) and Point B (24944 CPM) have fully vanished after replacement of NDE bearing block.

Table 3 Characteristic frequencies (BPFO)

S. No	Bearing number	Characteristic frequency	Calculated/observed values (CPM)			
			Standard formula	Defect multipliers	Observed from analyzer	Proposed formula
1	22218	BPFO	12192	12188	12471	12292
2	NU1048	BPFO	11536	11092	11687	11635



(a) Periphery



(b) Bottom side



(c) Top side

Fig. 4 Blue impressions of defected NDE bearing

Table 4 Measurement of vibration (Healthy bearing)

Direction	Vibration (22218), mm/s		Remarks
	DE	NDE	
Horizontal	0.7	0.9	Vibration level drastically reduced after the replacement of defected bearing. Status: **OK**
Vertical	1.1	1.2	
Axial	0.9	0.8	

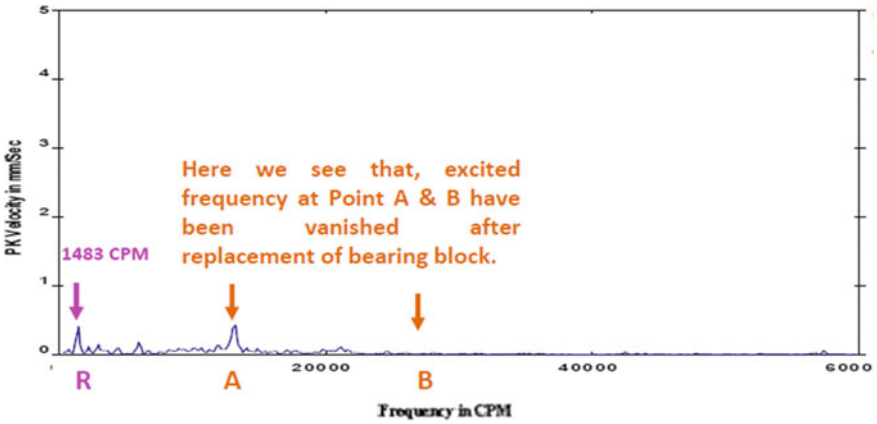


Fig. 5 Vibration patterns (healthy bearing, 22218) at NDE

Similarly, another experimental study on NU1048 has been carried out and all results are found as anticipated as earlier illustrated study.

Case 2. Study of vibration analysis of motor bearing at DE and NDE under inner race fault condition.

The specifications of bearings 22224 and 6309 are tabulated in Table 5.

Vibrations in all axes are measured and tabulated in Table 6.

The vibration pattern is given in Fig. 6. From Fig. 6, it is observed from the pattern that the motor speed is 993 CPM as indicated by point R and point B (20342 CPM) is the second harmonic of point A (10170 CPM).

The characteristic frequencies for both bearings are determined and tabulated in Table 7 as given below.

Table 5 Specifications of bearing

S. No	Bearing no.	Motor speed (CPM)	N_b	B_d	P_d	Φ	Defect multiplier (BPFO)	Defect multiplier (BPFI)
1	22224	993	18	0.905	6.685	10	8.67	11.33
2	6309	1481	8	0.687	2.913	0	3.06	4.94

Table 6 Measurement of vibration

Direction	Vibration (22224), mm/s		Remarks
	DE	NDE	Vibration found in increasing trend at NDE side and abnormal sound is observed Status: **NOT OK**
Horizontal	7.5	14.8	
Vertical	7.0	24.6	
Axial	5.4	9.2	

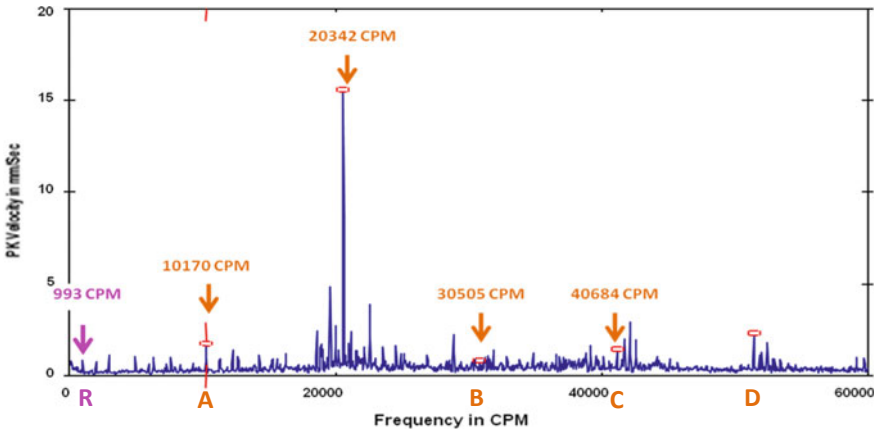


Fig. 6 Vibration patterns (faulty bearing, 22224) at NDE

Table 7 Characteristic frequencies (BPFI)

S. No.	Bearing number	Characteristic frequency	Calculated/observed values (CPM)			
			Standard formula	Defect multipliers	Observed from analyzer	Proposed formula
1.	22224	BPFI	10128	11250	10170	10129
2.	6309	BPFI	7312	7316	7703	7701

It is cleared from Table 7 that proposed formulas are capable to determine the characteristic frequencies. The snapshot of defected bearing is shown in Fig. 7.

After replacing defected bearing, vibrations are recorded and tabulated in Table 8.

After the replacement of defected bearing, vibration pattern is shown in Fig. 8 and it is observed that all dominant peaks have vanished.

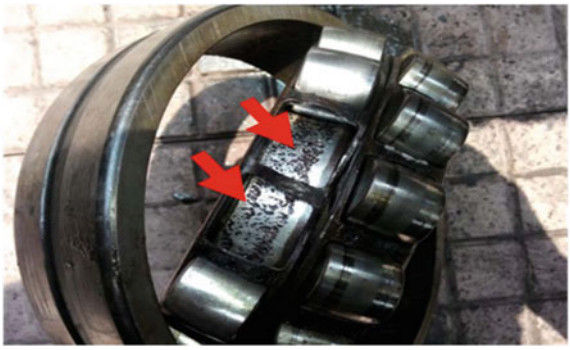


Fig. 7 Snapshot of defected NDE bearing

Table 8 Measurement of vibration (healthy bearing)

Direction	Vibration at NDE (22224), mm/s	Remarks
Horizontal	3.1	Vibration and sound levels drastically reduced after the replacement of bearing. Status: **OK**
Vertical	2.4	
Axial	0.5	

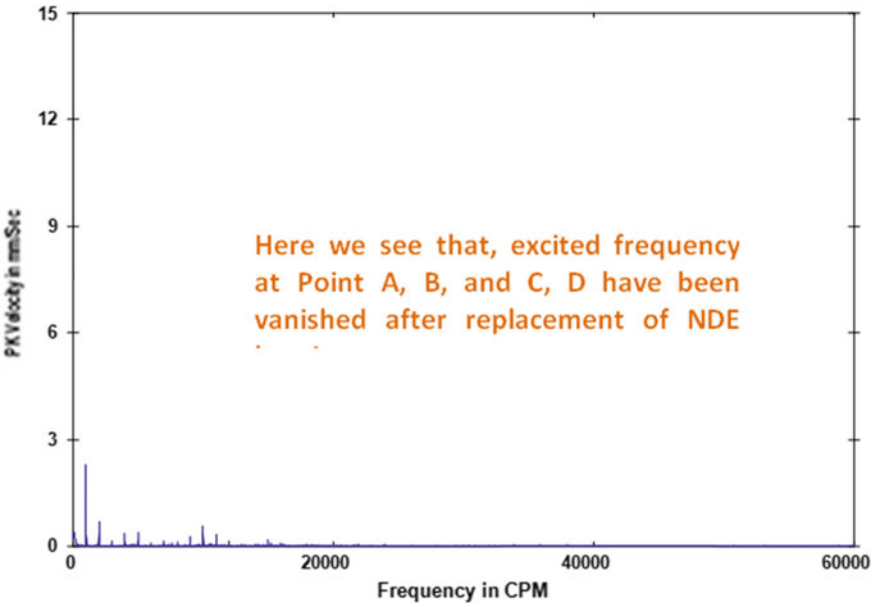


Fig. 8 Vibration patterns (healthy bearing, 22224) at NDE

Similarly, experimental study with 6309 has been performed and all results are verified.

From the Table 9, it is clearly observed that for bearing 22218 and NU1048, the calculated BPFOs from standard formula as well from proposed formulas are much close to measured BPFOs but calculated BPFIs with standard formula and with

Table 9 Summary of experimental results

S. No.	Bearing number	BPFO (CPM)		BPFI (CPM)		Analyzer reading (CPM)
		Standard formula	Proposed formula	Standard formula	Proposed formula	
1.	22218	12192	12292	15946	15847	12471
2.	NU1048	11536	11635	13557	14001	11687
3.	22224	7745	7745	10128	10129	10170
4.	6309	4526	4147	7312	7701	7703

proposed formula are not in the range of measured BPFIs, which clearly indicates existence of outer race fault in bearing. Furthermore for bearing numbers 22224 and 6309, the calculated BPFIs with standard formula as well with proposed formulas are comparable to measured BPFI but calculated BPFOs with standard formula as well with proposed formula are not in the range of measured BPFOs, which clearly indicates the existence of inner race fault in bearing.

5 Conclusions

In this paper, vibration-based fault detection scheme has been presented and based on industrial best practices, new mathematical formulas are proposed. To validate the proposed formulas, experimental studies for inner and outer race faults in Jindal Steel and Power Limited, Raigarh, India have been carried out and a fair comparison has been made with existing standard theory. On the basis of experimental study, it is concluded that without knowing bearing geometry and technical specifications from the manufacturer, proposed formulas are able to determine the characteristic frequencies for higher number of bearing balls and both types of faults efficiently.

Acknowledgements The authors are very much thankful to authorities of Jindal Steel and Power Ltd. especially to Mr. Sanjay Bhansali, Asst. Vice President Central Services, Raigarh, C.G., India, for providing necessary facilities to do this work and MIT, Moradabad for the support in developing this document.

References

1. F. Immovilli, A. Bellini, R. Rubini, C. Tassoni, Diagnosis of bearing faults in induction machines by vibration or current signals: a critical comparison. IEEE Trans. Ind. Appl. **46**(4), 1350–1359 (2010)
2. F. Immovilli, C. Bianchini, M. Cocconcelli, A. Bellini, R. Rubini, Bearing fault model for induction motor with externally induced vibration. IEEE Trans. Ind. Appl. **60**(8), 3408–3418 (2013)
3. M. Blodt, P. Granjon, B. Raison, G. Rostaing, Models for bearing damage detection in induction motors using stator current monitoring. IEEE Trans. Ind. Electron. **55**(4), 1813–1822 (2008)
4. IEEE Motor Reliability Working Group, Report of large motor reliability survey of industrial and commercial installations. IEEE Trans. Ind. Appl. **IA-21**(4), 853–872 (1986)
5. H. Henao et al., Trends in fault diagnosis for electrical machines—a review of diagnostic techniques. IEEE Ind. Electron. Mag. **8**(2), 31–42 (2014)
6. P. Zhang, Y. Du, T.G. Habetler, B. Lu, A survey of condition monitoring and protection methods for medium-voltage induction motors. IEEE Trans. Ind. Appl. **47**(1), 34–46 (2011)
7. R.R. Schoen, T.G. Habetler, F. Kamran, R. Bartheld, Motor bearing damage detection using stator current monitoring. IEEE Trans. Ind. Appl. **31**(6), 1274–1279 (1995)
8. J.H. Jung, J.J. Lee, B.H. Kwon, Online diagnosis of induction motors using MCSA. IEEE Trans. Ind. Electron. **53**(6), 1842–1852 (2006)
9. L. Sun, B. Xu, An improvement of stator current based detection of bearing fault in induction motors, in *Conference Rec. IAS Annual Meeting, Tampa, FL* (2007), pp. 2277–2281

10. L. Eren, K. Teotrakool, M.J. Devaney, Bearing fault detection via wavelet packet decomposition with spectral post processing, in *Proceeding of IEEE Instrument. Measurement Technology, Conference, Warsaw, Poland* (2007), pp. 1–4
11. A. Lebaroud, G. Clerc, Time–frequency classification applied to induction machine faults monitoring, in *Proceedings of 32nd Annual Conference IEEE IECON, Paris, France* (2006), pp. 5051–5056
12. T.C.A. Kumar, G. Singh, V.N.A. Naikan, Effectiveness of vibration and current monitoring in detecting broken rotor bar and bearingfaults in an induction motor, in *IEEE 6th International Conference on Power Systems (ICPS)* (2016), pp. 1–5
13. A. Ibrahim, M. El Badaoui, F. Guillet, F. Bonnardot, A new bearing fault detection method in induction machines based on instantaneous power factor. IEEE Trans. Ind. Electron. **55**(12), 4252–4259 (2008)
14. K.S. Gaeid, H.W. Ping, M.K. Masood, L. Szabó, Survey of wavelet fault diagnosis and tolerant of induction machines with case study. Int. Rev. Electr. Eng. **7**(3), 4437–4457 (2012)
15. E. Ayaz, A. Ozturk, S. Seker, Continuous wavelet transform for bearing damage detection in electric motors, in *Proceeding of IEEE Mediterranean Electrotechnical Conference, Malaga, Spain* (2006), pp. 1130–1133
16. N. Mehala, R. Dahiya, Detection of bearing faults of induction motor using Park's vector approach. Int. J. Eng. Technol. **2**(4), 263–266 (2010)
17. Rajul Misra, G.L. Pahuja, An experimental study of rotor fault detection using motor current signature analysis based on neural networks. Int. J. Adv. Sci. Technol. **79**, 25–36 (2015)
18. Rajul Misra, G.L. Pahuja, Fuzzy logic based rotor health index of induction motor. Int. J. Emerg. Electr. Power Syst. **16**(5), 443–449 (2015)
19. M. Irfan, N. Saad et al., Analysis of bearing outer race defects in induction motor, in *IEEE International Conference on Intelligent and Advanced System, Kuala Lumpur* (2014), pp. 1–6
20. N. Tandon, A. Choudhury, An analytical model for the prediction of the vibration response of rolling element bearings due to a localized defect. J. Sound Vib. **205**(3), 275–292 (1997)
21. P.D. MacFadden, J.D. Smith, Model for the vibration produced by a single point defect in a rolling element bearing. J. Sound Vib. **96**(1), 69–82 (1984)
22. Y.F. Wang, P.J. Kootsookos, Modeling of low shaft speed bearing faults for condition monitoring. Mech. Syst. Signal Process. **12**(3), 415–426 (1998)
23. B. Li, M. Chow, Y. Tipsuwan, J. Hung, Neural-network-based motor rolling bearing fault diagnosis. IEEE Trans. Ind. Electron. **47**(5), 1060–1069 (2000)
24. X. Yu, L. Shunming, W. Jingrui, A novel roller bearing fault diagnosis method based on the wavelet extreme learning machine, in *Prognostics and System Health Management Conference (PHM-Harbin), IEEE Conference* (2017), pp. 1–6
25. A. Boudiaf, A. Djebala; Hocine Bendjma, A. Balaska, A. Dahane A summary of vibration analysis techniques for fault detection and diagnosis in bearing, in *8th International Conference on Modelling, Identification and Control (ICMIC)* (2016), pp. 37–42

A New Modified Form of Murnaghan Thermodynamic Equation of State



Tanveer Ahmad Wani, B. K. Das, Babita Tripathi and Ikhlaq Ahmed Khan

1 Introduction

An equation of state is significant in reviewing the high-pressure properties of solids. A number of attempts have been made by various researchers for the pursuit of a simple form of an equation of state with the least number of parameters, to forecast accurate high-pressure behavior. The parameters are determined by means of existing low-pressure data such as the volume, the isothermal bulk modulus and its pressure derivatives at zero pressure. The importance of equation of states from applied science point of view is no doubt unquestionable. The equation of state offers, using limited quantity of measurements, to forecast and calculate thermodynamic properties of a given solid.

2 Theory

In deriving the equation to an isotherm, one can use simplified notations which avoid partial derivatives. Thus, for example, we can express the isothermal bulk modulus

$$B_T = -V(\frac{\partial P}{\partial V})_T$$

as

$$B = -V(dV/dP) = -V/V' \tag{1}$$

T. A. Wani (✉) · B. K. Das · B. Tripathi · I. A. Khan
Noida International University, Greater Noida 201310, India
e-mail: tanveerur_rahman@yahoo.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_87

where a prime indicates differentiation with respect to pressure. Integration of Eq. (1) with the substitution

$$B = B_0 + B_0'P \tag{2}$$

where the suffix 0 indicates value at zero pressure, yields the Murnaghan equation (ME)

$$V/V0 = (1 + P/P_C) - 1/n \tag{3}$$

where

$$n = B_0'$$

$$P_C = B_0'/n \tag{4}$$

Pc can be treated as a characteristic pressure. The MME is similarly obtained by substituting in (1)

$$B = B_0' + B_0'P + B_0''P^2/2 \tag{5}$$

Integration of (1) then leads to

$$\frac{V}{Vo} = \left[\frac{M - B_0' - B_0''P}{M + B_0' + B_0''P}.\frac{M + B_0'}{M - B_0'}\right]^{-1/M}$$

$$M = \sqrt{\left(B_0'^2 - 2B_0B_0''\right)} \tag{6}$$

In deriving (6), one assumes

$$B_0'^2 > 2B_0B_0'' \tag{7}$$

We do not know of any solid violating condition (7). Differentiation of Eq. (5) gives

$$B' = B_0' + B_0''P \tag{8}$$

Obviously, the MME assumes B′ to be linearly dependent on P while the ME assumes it to be constant. It is a general observation that B′ depends on P rather weakly. This can be also concluded from the fact that the ME presents P-V relationship pretty accurately up to moderately high pressure. One can, therefore, express B′ in more realistic form by

$$B' = n\left(1 + mf\left(\frac{P}{Pc}\right)\right) = n(1 + mf(P)) \tag{9}$$

where n and m are constants and P = P/Pc is a dimensionless pressure variable. In the limit p going to zero, f(p) must go to zero to ensure

$$B'_0 = n'' \tag{10}$$

The various order derivatives of f(p) may, however, go to finite limits which could be connected to the zero pressure values of the derivatives B of the corresponding order. For example, if we choose

$$f(p) = P(1 + bP) \tag{11}$$

We obtain

$$\lim_{P\to 0} f(P) = 0, \lim_{P\to 0} f'(P) = \frac{1}{Pc}, \text{and } \lim_{P\to 0} f''(P) = \frac{2b}{Pc^2}$$

Differentiating Eq. (9) successively and taking the limits as p goes to zero, we obtain

$$nm/Pc = B'_0 \quad \text{and} \quad \frac{2nmb}{Pc^2} = B''_0$$

In view of Eqs. (4–10), we can express these parametric relations as

$$Pc = B_0/B''_0 \tag{12}$$

$$n = B'_0 \tag{13}$$

$$m = B_0 B''_0/B'^2_0 \tag{14}$$

$$b = B'''_0 B_0/2B'B''_0 \tag{15}$$

The choice $B'''_0 = 0$ requires b = 0
Now by differentiating Eq. (1), we obtain

$$B' = -1 + \frac{VV''}{V'^2}$$

Substituting B′ from Eq. (9), we get

$$\frac{VV''}{V'^2} = (1 + n) + nmf(P) \tag{16}$$

Writing

$$n + 1 = 1 \quad \text{and} \quad nm/1 = \lambda$$

We finally express this equation in the form

$$VV''/V'^2 = 1(1 + \lambda f(P)) \tag{17}$$

$\lambda f(p)$ is in general very small as compared to unity. This allows us to use the perturbation method for solving Eq. (17) with λ as a perturbation parameter. We, therefore, express

$$V = V + \lambda V + \lambda V \tag{18}$$

With this substitution, Eq. (17) leads to zero and the first-order perturbation equation as

$$V + \frac{d^2V}{dp^2} = 1\left(\frac{dV}{dp}\right)^2 \tag{19}$$

$$V + \frac{d^2V}{dp^2} + \frac{Vd^2V}{dp^2} = 1\left\{f(p)\left(\frac{dV}{dp}\right)^2 + 2\left(\frac{dV}{dp}\right)\left(\frac{dV}{dp}\right)\right\} \tag{20}$$

Integration of (19) yields the ME with V in place of $\overset{0}{V}$. Equation (20) can be reduced with the help of (19) to the form

$$\left(\frac{d}{dp}\right)\left(\frac{(1+p)^2 du}{dp}\right) = \left(\frac{1}{n^2}\right)f(p) \tag{21}$$

$$u = \overset{1}{V} \Big/ \overset{0}{V} \tag{22}$$

Integrating Eq. (21) twice, we obtain

$$u = \left(\frac{1}{n^2}\right)\int_0^p \left[\int_0^y \left[\frac{\{f(x)dx\}}{(1+y)^2}\right]dy\right. \tag{23}$$

If the integration (23) can be performed, we shall obtain u in an algebraic form. Combining the first-order solution for volume namely $\overset{1}{V}$ with the zeroth-order one, we obtain the resultant volume V or V/V0

$$\frac{V}{V0} = \begin{matrix} 0 \\ (V \end{matrix} + \begin{matrix} 1 \\ \lambda V) \end{matrix} /V0 + \begin{pmatrix} 0 \\ \frac{V}{V_0} \end{pmatrix} (1 + \lambda u) \tag{24}$$

The integrals in (23) can be easily obtained for several convenient forms of f(p). The simplest form can be

$$f(p) = p \tag{25}$$

Substitution of this expression in (23) results in

$$u = \frac{1}{2n^2}\left(p\left(\frac{p+2}{p+1}\right) - 2\ln(p+1)\right) \tag{26}$$

The expression in view of Eq. (24) leads to EOS

$$\frac{V}{V0} = (1+p)^{-1n}\left[1 + \left(\frac{m}{2n}\right)(p(P+2)/(p+1) - 2\ln(p+1))\right] \tag{27}$$

We designate this EOS as the new modified Murnaghan equation (NMME) in view of the choice (25) for f(p). Other forms of f(p) which readily result in algebraic expression for u are

1. $f(p) = p(p+1)$
2. $f(p) = p/(p+1)$
3. $f(p) = \ln(p+1)$
4. $f(p) = \frac{p}{(p+1)^2}$

which also satisfy the required criteria for the functional form of the f(p). The simplest expression for u results from the first of these four functions. The functions u for all the five choices of f(p) given by Eqs. (25) and (27) are given in Table 1.

$$\frac{V}{V0} = (1+p)^{(-1/n)}(1+\lambda u)$$

Table 1 Expression for the function u for different choices of the function f(p)

f(p)	λu
$p(p+1)$	$(m/2n)\left(P^3/3(P+1)\right)$
p	$(m/2n)\{(P(P+2)/(P+1) - 2\ln(P+1)\}$
$\log(p+1)$	$(m/2n)\{\ln(P+1)(\ln(P+1) - 2) + 2P/P + 1\}$
$p/p+1$	$(m/2n)[2\{(P+2)\ln(P+1) - 2P\}/P+1)]$
$p/(p+1)^2$	$\left(\frac{m}{2n}\right)[P(P+2) - \{2(P+1)\ln(P+1)\}/\left(P+1\right)^2]$

The five equations so obtained will be referred to by combining the serial number in Table 1 with Eq. (27). Thus, the EOS (27) corresponding to f(p) = p will be indicated by ((27, 2). In the case of each of the five functions given in Table 1, the leading term in u expressed as a series in increasing powers of p(p < 1) is found to be (1/2n2) (p3/3). If the successive term in increasing powers of p is assumed to be decreasing as fast as they are expected in view of the weak dependence of f(p) assumed, we can conclude that all the EOS ((27, i) i = 1, 2, 3, 4, 5 will yield numerical results in close agreement. The integrals in (23) can also be worked out for the choice (11) for f(p). This leads up to an EOS which is generalized to one further stage as compared to the MME. This equation reads as

$$\frac{V}{V_0} = (1+p)^{-\frac{1}{n}}\left[1+\left(\frac{m}{6n}\right)\left\{b(p+1)^2+3(1-2b)(p+1)+3b-6(1-b)\ln(p+1)-\frac{(3-2b)}{(p+1)}\right\}\right]$$

One can easily check that this equation goes into Eq. (27, 2) for b = 0 and Eq. (27, 1) for b = 1. The parameters of the equation are of course fixed by B_0, B_0', B_0'' and B_0''' according to Eqs. (12)–(15).

3 Applications

An ionic compound NaCl has been used for pressure measurements. A number of experimental studies [1–4] were conducted on the p-v data based on measurements of length under hydrostatic pressure. The highest pressures reached in different studies along with techniques is given in Table 2.

The equation of states used here are Freund and Ingalls equation (FE) [8], Birch equation (BE) [9], and a new modified form of Murnaghan equation (NMME2) and are given, respectively, as:

$$P = \left(3\frac{B_0}{2}\right)\left\{\left(\frac{V_0}{V}\right)^{\frac{7}{3}}-\left(\frac{V_0}{V}\right)^{\frac{5}{3}}\right\}\left\{1+\left(\frac{3}{4}\right)(B_0'-4)\left[\left(\frac{V_0}{V}\right)^{\frac{2}{3}}-1\right]\right\} \quad (28)$$

$$\frac{V}{V_0} = \{1-a\ln(1+bP)\}^c \quad (29)$$

Table 2 P_{max} (GPa) reached with particular technique

Technique	P_{max} (GPa)
Hydrostatic	4.5 [4]
Ultrasonic	27 [5]
Shock wave	23 [6]
X-ray diffraction	25 [7]

Table 3 Best fit values for B_0, B_0', B_0'', and B_0'''

B_0	B_0'	B_0''	B_0'''	Source
25.16	4.50	–	–	BE [9]
23.65	4.82	−0.035	–	NMME1[10]
23.64	4.82	−0.035	−0.014	NMME2 [10]
22.79	5.73	−0.45	–	FIE [8]

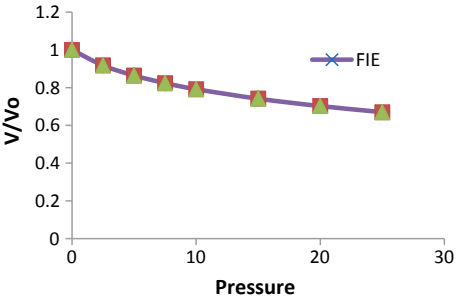


Fig. 1 Pressure versus V/V_0 for NaCl

$$\frac{V}{V_0} = (1+p)^{-\frac{1}{n}}\left[1+\left(\frac{m}{6n}\right)\left\{b(p+1)^2+3(1-2b)(p+1)\right.\right.$$
$$\left.\left.+3b-6(1-b)\ln(p+1)-\frac{(3-2b)}{(p+1)}\right\}\right] \tag{30}$$

For the value of $b = 1$, the above Eq. (30) will become NMME1. The best fit values are given in Table 3.

These values have been plotted against P as shown in Fig. 1. The derived equations stated the V/V_O versus P values of NaCl in agreement with the experimental data [11] and other two EOS's.

References

1. P.W. Bridgeman, Proc. Am. Acad. Arts Sci. **74**, 21 (1940)
2. S.N. Vaidya, G.C. Kennedy, J. Phys, Chem. Solids. **41**, 517 (1971)
3. R. Boehler, G.C. Kennedy, J. Phys, Chem. Solids. **41**, 517 (1980)
4. L.C. Chhabildas, A.L. Ruoff, J. Appl. Phys. **47**, 4182 (1976)
5. J. Frankel, J.F. Rich, C.G. Homan, J. Geophys. Res. **81**, 6357 (1976)
6. J.N. Fritz, S.P. Marsh, W.J. Carter, R.G. Mc Queen, The Hugornioe equation of state of sodium chloride in the sodium chloride structure, in *Accurate Characterization of the High-Pressure Environment*, ed. by E.C. Leoyd (NBS special publication, Washington D.C, 1971), p. 326
7. G. Borelius, *Solid State Physics*, ed. by F. Seitz, D. Turnbull (Academic press, New York, 1958), p. 65
8. J. Freund, R. Ingalls, J. Phys. Chem. Solids **50**, 263 (1989)

9. F.J. Birch, J. Geo. Phys. Res. **57**, 227 (1952)
10. P. Singh, T.A. Wani, M. Mehra, A.S. Khan, Int. J. Phys. Appl. **2**, 57 (2010)
11. H.G. Drickmer, R.W. Lynch, R.L. Clender, E.A. Perez-Albueme, Solid State Phys. **19**, 225 (1966)

A Study of Shear and Temperature on Vibrations of Infinite Plate

Narender Kumar Sarswat

1 Introduction

Tomar et al. have observed the axisymmetric transverse vibrations of a circular plate of linearly nonuniform thickness with exponential variation of elastic modulus and density [1]. Tomar and others analyzed the frequencies of an elastic plate of nonuniform thickness on elastic foundation with different boundary conditions [2]. For uniform circular plates, Mindlin and Deresiewicz have shown that for moderately thick plates, the shearing effects cannot be neglected [3]. Gupta and Sharma have observed the forced vibrations of a stepped rectangular beam with effect of shear and rotatory inertia [4]. Xiang and Reddy have determined the natural vibrations of rectangular plates by leavy type solution [5]. Lee has observed the natural frequencies of plates by finite method under shear theory [6]. Lee and Cheng have observed the nonlinear vibrations of orthotropic plates by DQ method under shear effect [7]. Tomar et al. have studied the effect of linearly thermal field on frequencies of a circular plate of variable thickness [8]. Tomar et al. analyzed the harmonic and linear thermal impart on frequencies of orthotropic circular plates of variable thickness [9, 10]. Sarswat et al. have analyzed the thermal damped vibration of isotropic, square plate by Frobenius method [11, 12]. Shrikant et al. have observed that vibrations of circular plates with holes are extensively used in mechanical components by FEA method [13]. This study shows the effect of the temperature distribution on the frequencies of an isotropic circular plate of non-variable thickness with foundation. The eq'n of motion has been analyzed by Frobenius method and frequencies for first four vibrational modes have been computed at various edges condition. The numerical

N. K. Sarswat (✉)
Department of P.G. Studies & Research in Physics, J.V. Jain College, Saharanpur 247001, UP, India
e-mail: narendersarswat12@gmail.com

C. C. S. University, Meerut, India

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_88

results are shown in graphical form. This study is helpful in designing plate material in structural engineering.

2 Mathematical Formulation

Temperature Variation
The temperature of plate material is subjected in x-direction as linearly and non-homogeneity of the plate material that arises with modulus of elasticity [14–16] is given as

$$\bar{E}(R) = \bar{E}_0\{1 - \eta(1 - X)\}\, and\, \overline{G}(R) = \overline{G}_0\{1 - \eta(1 - X)\} \tag{1}$$

where $\eta = \xi\, T_0$, $(0 \leq \eta < 1)$ is called the thermal gradient and ξ is a arbitrary constant.
Equation of Motion
The fourth-order eq'n of motion is given by

$$\begin{aligned}&\left(\overline{E}\right)\overline{\Psi}^{IV} + \left(3\overline{E}^{I}\right)\Psi^{III} + \left(3\overline{E}^{II} + \{C_2A - A_1\}\overline{E} + A_2\right)\Psi^{II}\\ &+ \left(\overline{E}^{III} + \{C_2A - A_1\}\overline{E}^{I}\right)\Psi^{I} + A_2\Bigg[\left\{C_2A - 12/H^2\right\}\\ &- A_1\left\{1 - (1/A)(12/H^2C_2)\right\}\Bigg]\overline{\Psi} = 0\end{aligned} \tag{2}$$

and

$$\begin{aligned}\overline{W} = &\left[\frac{1}{12\left(1 - v^2\right)\left(-p^2a^2\overline{\rho} + \left(\frac{k_f}{H}\right)\right)}\right]\left[\left\{H^2\overline{E}\right\}\overline{\Psi}^{III} + \left(2H^2\overline{E}^{I}\right)\overline{\Psi}^{II} + \left(H^2\overline{E^{II}}\right)\overline{\Psi}^{I}\right]\\ &+ \left(\frac{H^2p^2a^2\overline{\rho}}{12\left(-p^2a^2\overline{\rho} + \left(\frac{k_f}{H}\right)\right)}\right)\overline{\Psi}^{I}\end{aligned} \tag{3}$$

where a is the width of the plate.

3 Solution

Let the solution for deflection of plate $\bar{W}$ has been assumed in the series form as

$$\overline{\Psi} = \sum_{\lambda=0}^{\infty} a_\lambda X^{C+\lambda} \quad with \quad a_0 \neq 0 \tag{4}$$

where C is the exponent of singularity.

Imposing relation (1) and (4) on Eq. (2), we obtained the equation of motion as

$$\begin{aligned}
&\sum_{\lambda=0}^{\infty} a_\lambda \left[b_\lambda^{(3)} S_1^{(1)} \right] X^{C+\lambda-4} + \sum_{\lambda=0}^{\infty} a_\lambda \left[b_\lambda^{(3)} S_2^{(1)} + b_\lambda^{(2)} S_2^{(2)} \right] X^{C+\lambda-3} \\
&+ \sum_{\lambda=0}^{\infty} a_\lambda \left[b_\lambda^{(3)} S_3^{(1)} + b_\lambda^{(2)} S_3^{(2)} + b_\lambda^{(1)} S_3^{(3)} \right] X^{C+\lambda-2} + \sum_{\lambda=0}^{\infty} a_\lambda \left[S_5^{(5)} \right] X^{C+\lambda} \\
&+ \sum_{\lambda=0}^{\infty} a_\lambda \left[b_\lambda^{(1)} S_4^{(3)} + b_\lambda S_4^{(4)} \right] X^{C+\lambda-1} + \sum_{\lambda=0}^{\infty} a_\lambda \left[S_6^{(5)} \right] X^{C+\lambda+1} = 0
\end{aligned} \tag{5}$$

Respectively, where

$$\begin{gathered}
let\ C = e, b_\lambda^{(3)} = (e+\lambda)(e+\lambda-1)(e+\lambda-2)(e+\lambda-3), \\
b_\lambda^{(2)} = (e+\lambda)(e+\lambda-1)(e+\lambda-2), \\
b_\lambda^{(1)} = (e+\lambda)(e+\lambda-1), b_\lambda = (e+\lambda),
\end{gathered}$$

$$S_1^{(1)} = (1-\eta)^2, S_2^{(1)} = 2\eta(1-\eta), S_2^{(2)} = 3\eta(1-\eta),$$

$$S_3^{(1)} = \eta^2, S_3^{(2)} = 3\eta^2, S_3^{(3)} = (1-\eta)(A_3), S_4^{(3)} = \eta A_4$$

$$S_5^{(5)} = A_4\left(\Omega^2 - A_5(1-\eta)\right), S_6^{(5)} = -A_4 A_5 \eta, A_3 = \Omega^2(1+C_0) - \left(\frac{D_K}{H}\right),$$

$$C_2 = a^2 p^2 / K^2, C_1 = 1/p^2(1-\nu^2),$$

$$A_4 = \left(\Omega^2 C_0\right) - \left(\frac{D_K}{H}\right) and\ A_5 = \left(\frac{12}{H^2 C_0}\right).$$

Here, D_k, P, 'Ω is foundation, circular, frequency parameters, respectively.

The roots observed for lowest power of X are $C = 0, 1, 2, 3$. For next higher power of X for $C = 0$, these give a_1, a_2, a_3 are indeterminate and taken with arbitrary a_0. Similarly obtained a_4 in terms of a_0, a_1, a_2, a_3. The remaining unknown constants $a_\lambda (\lambda = 5, 6, 7 \ldots)$ are got by the following relation.

$$\begin{aligned}
&a_{\lambda+5}\left[B_1 S_1^{(1)} \right] + a_{\lambda+4}\left[B_2 S_2^{(1)} + B_3 S_2^{(2)} \right] + {}_{\lambda+3}\left[B_4 S_3^{(1)} + B_5 S_3^{(2)} + S_3^{(3)} \right] + \\
&a_{\lambda+2}\left[B_6 S_4^{(3)} + S_4^{(4)} \right] + a_{\lambda+1}\left[S_5^{(5)} \right] + a_\lambda \left[S_6^{(5)} \right] = 0
\end{aligned} \tag{6}$$

where

$$B_1 = (e+\lambda+2)(e+\lambda+5)(e+\lambda+4)(e+\lambda+3),$$

$$B_2 = (e+\lambda+4)(e+\lambda+3)(e+\lambda+2)(e+\lambda+1),$$
$$B_3 = (e+\lambda+4)(e+\lambda+3)(e+\lambda+2),$$
$$B_4 = (e+\lambda+3)(e+\lambda+2)(e+\lambda+1)(e+\lambda),$$
$$B_5 = (e+\lambda+3)(e+\lambda+2)(e+\lambda+1),\ B_6 = (e+\lambda+1)(e+\lambda+2)$$

If the notations

$$a_\lambda = f_\lambda^0 a_0 + f_\lambda^1 a_1 + f_\lambda^2 a_2 + f_\lambda^3 a_3 \quad with\, \lambda = 0, 1, 2, 3, 4 \ldots\ldots\ldots \tag{7}$$

are introduced, the solution $\overline{\Psi}$ at $C = 0$ can be written as

$$\overline{\Psi} = a_0 F_0 + a_1 F_1 + a_2 F_2 + a_3 F_3 \tag{8}$$

where

$$\left.\begin{aligned} F_0 &= 1 + \sum_{\lambda=4}^{\infty} f_\lambda^0 X^\lambda, \quad F_1 = X + \sum_{\lambda=4}^{\infty} f_\lambda^1 X^\lambda \\ F_2 &= X^2 + \sum_{\lambda=4}^{\infty} f_\lambda^2 X^\lambda \, and \, F_3 = X^3 + \sum_{\lambda=4}^{\infty} f_\lambda^3 X^\lambda \end{aligned}\right\} \tag{9}$$

where $F_\lambda^i (i = 0, 1, 2, 3)$ are function of η, D_k, $\Omega \, and \, C_0$. The solution (8) is uniformly convergent in the interval $0 \le X \le 1 \, where \quad |\mu < 1|$ as per [17]. Hence, the solution is convergent for all $\left|-(\frac{\eta}{1-\eta})\right| < 1$.

Using solution of $\overline{\Psi}$ from Eq. (8) in (3) and (1), the solution for deflection is given as

$$\overline{W} = a_0 T_0 + a_1 T_1 + a_2 T_2 + a_3 T_3 \tag{10}$$

Respectively, where-

$$T_0 = Q_0 F_0^{III} + Q_1 F_0^{II} + Q_2 F_0^{I} + Q_3 F_0,$$
$$T_1 = Q_0 F_1^{III} + Q_1 F_1^{II} + Q_2 F_1^{I} + Q_3 F_1$$
$$T_2 = Q_0 F_2^{III} + Q_1 F_2^{II} + Q_2 F_2^{I} + Q_3 F_2$$
$$T_3 = Q_0 F_3^{III} + Q_1 F_3^{II} + Q_2 F_3^{I} + Q_3 F_3$$

and

$$\left.\begin{aligned} Q_0 &= \left(1 - \eta + \eta X^2\right),\ Q_1 = (4\eta X) \\ Q_2 &= \left(2\eta + \Omega^2\right) and \, Q_3 = 0, \end{aligned}\right\} \tag{11}$$

4 Edge Conditions and Frequencies Equation

Now employing the following edge conditions and frequency equations of infinite plate plates have been obtained as

Eq'n Clamped-Clamped Plate:

Now plate clamp at $X = 0{,}1$ the deflection and rotation of the plate as

$$\bar{W}\,|_{X=0,1} = 0 = \overline{\Psi}\,|_{X=0,1} \tag{12}$$

Imposing (12) to (8) and (10), we get the frequency equation for this condition as

$$\begin{vmatrix} F_0(0,\Omega)F_1(0,\Omega)F_2(0,\Omega)F_3(0,\Omega) \\ T_0(0,\Omega)T_1(0,\Omega)T_2(0,\Omega)T_3(0,\Omega) \\ F_0(1,\Omega)F_1(1,\Omega)F_2(1,\Omega)F_3(1,\Omega) \\ T_0(1,\Omega)T_1(1,\Omega)T_2(1,\Omega)T_3(1,\Omega) \end{vmatrix} = 0 \tag{13}$$

Eq'n Clamped Simply Supported Plates:

Now plate at the edge $X = 0$ and 1, then W and M_r at this edge should be zero

$$\bar{W}\big|_{X=0,1} = 0 = \left[\overline{\Psi}^1\right]\Big|_{X=1,0} \tag{14}$$

Again using (8), (10), and (13), one obtains the frequency equation for this boundary condition as

$$\begin{vmatrix} F_0(0,\Omega)F_1(0,\Omega)F_2(0,\Omega)F_3(0,\Omega) \\ T_0(0,\Omega)T_1(0,\Omega)T_2(0,\Omega)T_3(0,\Omega) \\ F_0^I(1,\Omega)F_1^I(1,\Omega)F_2^I(1,\Omega)F_3^1(1,\Omega) \\ T_0(1,\Omega)T_1(1,\Omega)T_2(1,\Omega)T_3(1,\Omega) \end{vmatrix} = 0 \tag{15}$$

5 Results and Conclusion

The expressions (13) and (15) of frequency parameters are eq'ns in Ω from which roots are observed for this plate. Observed frequencies for 4- mode of vibrations of c-c and c-ss plates have been analyzed for different values of η and D_k using $c + 2$ and presently shown in graphically form. In all cases, the values of $\nu = 0.3$, $H = 1$ and $k^2 = 0.85$ have been considered here. The results according to Fig. 1 shows the thermally effected variation of Ω parameters with D_k for 24 series at c-c and c-ss conditions, when the shear theory can be considered. As per characteristics, the values of frequencies increases with D_k for all shown modes, here for both conditions

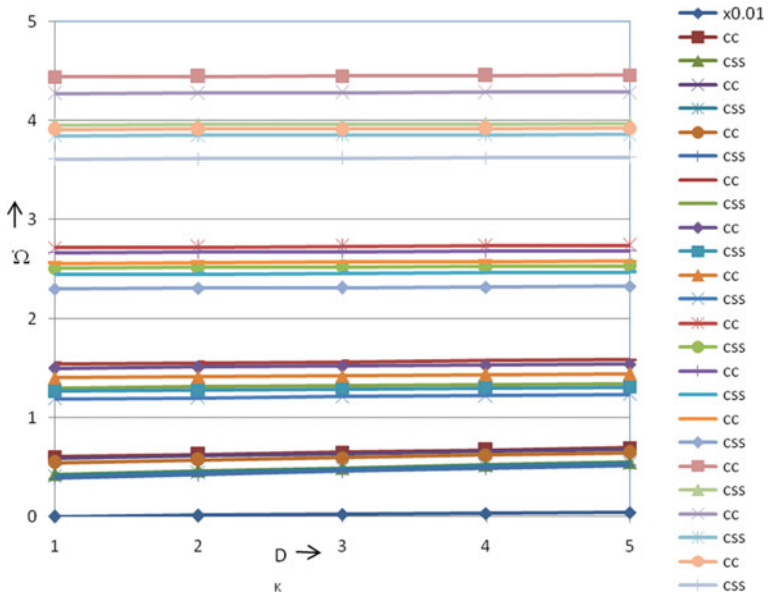


Fig. 1 Variation of ῼ with D_k at $H = 1$, $\acute{\eta} = 0.0, 0.1, 0.3$, and $\nu = 0.3$

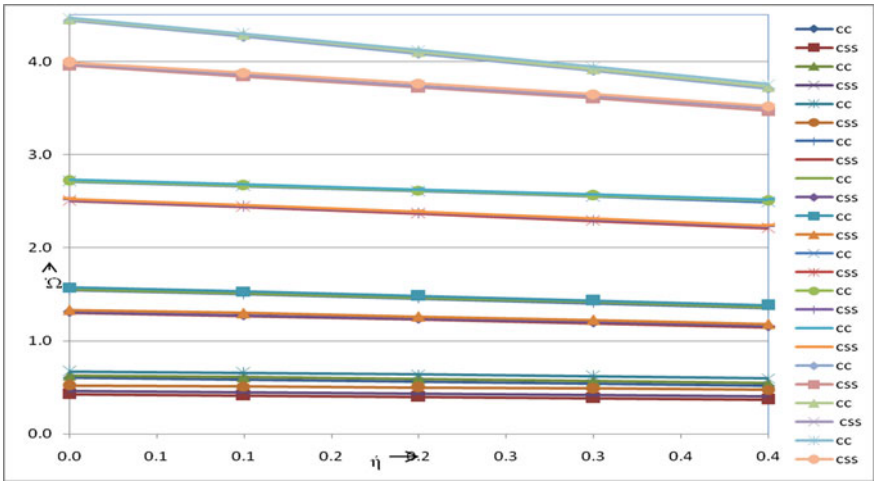


Fig. 2 Variation of ῼ with $\acute{\eta}$ at $H = 1$, $D_k = 0.0, 0.01, 0.03$, and $\nu = 0.3$

of the heated and unheated ($\eta = 0$) infinite plates. It is also noted that the Ω of c-c are greater than the Ω of c-ss plates and the frequency of excited plate are lower than that of normal plate in boundary conditions. Now Fig. 2 in graphical form of 24 series represents the linear thermal impart on the vibrations of plate material of non-variable structure with D_K. It is observed that in both boundary conditions, the

Ω decreases with the increase of η for nonhomogeneous materials of the plate. Also, the value Ω in shear theory is less than the corresponding Ω as per classical plate theory. Thus, the remarkable impact of shear cannot be ignores for this plates, when vibrations corresponding to the higher modes are considered.

References

1. J.S. Tomar, D.C. Gupta, N.C. Jain, Axisymmetric vibrations of an isotropic elastic non-homogeneous circular plate of linearly varying thickness. J. Sound Vib. **85**(3), 365–370 (**1982**)
2. J.S. Tomar, D.C. Gupta, V. Kumar, Natural frequencies of a linearly tapered non-homogerneous isotropic elastic circular plates resting on an elastic foundation. J. Sound Vib. **III**(1), 1–8 (1986)
3. R.D. Mindlin, H. Deresiewicz, Thickness shear and flexural vibrations of a circular disk. J. App. Phys. **25**(10), 1329–1332 (1954)
4. A.P. Gupta, N. Sharma, Effect of transverse shear and rotatory inertia on the forced motion of a stepped rectangular beam. J. Sound Vib. **209**(5), 811–820 (1998)
5. Y. Xiang, J.N. Reddy, Natural vibrations of rectangular plates with an internal line hinge using the first order shear deformation plate theory. J. Sound Vib. **263**(2), 285–297 (2003)
6. S.J. Lee, Free vibrations analysis of plates by using a four node finite element formulated with assumed natural transverse shear strain. J. Sound Vib. **278**(3), 657–684 (2004)
7. J.J. Li, C.J. Cheng, Differential quadrature method for non-linear vibrations of orthotropic plates with finite deformation and transverse shear effect. J. Sound Vib. **281**(120), 295–309 (2005)
8. J.S. Tomar, V.S. Tewari, Effect of thermal gradient on frequencies of a circular plates of linearly varying thickness. J. Non-equilib. Thermodyn. **6**, 115–122 (1981)
9. J.S. Tomar, A.K. Gupta, Thermal effect on axisymmetric vibrations of an orthotropic circular plate of variable thickness. J. AIAA **22**(7), 1015–1017 (1984)
10. J.S. Tomar, A.K. Gupta, Thermal effect on axisymmetric vibrations of an orthotropic circular plate of parabollicaly varying thickness. J. Pure App. Maths. **16**(5), 537–545 (1985)
11. N.K. Sarswat, V. Bansal, The effect of damping on thermal vibrations of isotropic elastic square plates. IOSAR-JAP **9**(6), 31–38 (2017)
12. N.K. Sarswat, V. Bansal, The study of damping and parabolic temperature on vibrations of isotropic elastic square plate. Asia Pac. JAEEE **1**(2), 1–8 (2017)
13. B. Shrikant, L. Thakare, A.V. Damale, Free vibration analysis of circular plates with holes and cutout. IOSR-JMCE **8**(2), 46–54 (2013)
14. S. Spinner, Temperature dependence of elastic constant of some ceramet specimens, Research National Bureau Standard C. Engg. Inst. **65**, 89–96 (1961)
15. I.E. Garrick, Survey of aero-thermo elasticity. J. Aerosp. Engg. **22**, 140–147 (1963)
16. E.H. Mansfield, Bending buckling and curling of a heated elliptical plate. Proc. Roy. Soc. Ser. A **28**, 296–417 (1965)
17. H. Lamb, Hydrodynamics (Dover Pub., New York, 1945), p. 335

Voltage Stability Analysis of Wind Integrated Grid



Taruna Sharma and Omveer Singh

1 Introduction

In recent times, the depletion of fossil fuels has increased to a great extent which further contributes to the greenhouse effect and this will lead to global warming, which adversely affects the environment and climate change, so it is the need of the hour that everyone should look upon to more renewable energy resources which will ensure that there will be no more degradation of the environment. Nowadays, solar energy is being used to a greater extent than the wind energy due to its abundance quantity and decent efficiency, it is much easier in installation. Wind energy lags in some aspects due to various reasons such as non-uniform wind speed, weather conditions, technical problems and some types of instabilities. According to recently organised scotch awards, India has been in fifth position in the production of renewable energy and at fourth position in the production of wind energy. There are several problems due to which the full potential of the wind farm cannot be achieved to remove those instabilities, this paper suggests some of the measures. In this paper, DFIG is used in wind turbine due to following advantages: DFIG are considered because they have a unique feature to operate at variable speed. This property of DFIG allows it to produce electricity even at low speeds. The inverter used in this generator become less expensive because it handles less amount of real power, the amount of reactive power absorbed and supplied to the grid, hence power factor is controlled by it. DFIG is more efficient than conventional induction generators. This type of generators consist the power electronics converters so that they can regulate the Q power, and they can operate on a desired pf, or maintain the voltage of the grid, but PWM converter

T. Sharma (✉) · O. Singh
Electrical Engineering, Gautam Buddha University, Greater Noida 201312, Uttar Pradesh, India
e-mail: tarunasharma111@gmail.com

O. Singh
e-mail: omveers@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_89

has a constrained capacity so, DFIG does not match with the synchronous generator [1–4]. Many factors can affect the voltage stability but it affects the most when DFIG are not able to control the voltage requirement. We can classify voltage stability into two categories, on the basis of the type of disturbance, or on the basis of the time period over which instability can arise. On the basis of disturbances, voltage stability can be divided into small and large disturbances. The former stability occurs when system is not able to regain its original state after being encountered with a very small disturbance such as changing loads, whilst latter stability occurs when system cannot attain its previous conditions after being subjected to large disturbances such as large generation tripping. Related to time, short-time voltage stability is related to the low variations in the fastacting loads whereas long-time voltage stability related to equipment which is very slow in response. Voltage instability can lead to system collapse, by definition system collapse is the sequence of the events which can lead to complete blackout. Due to the reactive power imbalance, various phenomenon can occur such as voltage collapse which can damage the power system equipments such as transmission lines, transformers and loads which accept the Q power. So, the absence of Q power compensator may lead to huge losses and heavy damage in the system and voltage instability issues may occur in the system and it can collapse. In recent times to enhance the quality of the voltage profile of the wind farm system, different types of FACTs devices are used which are called as Flexible AC Transmission devices. These devices belong to power electronics family which is helpful to restrict the flow of P and Q power of the system at large quantity. Voltage profile and loadability of the transmission can be improved or enhanced by providing series or shunt compensation, which can be provided with the help of the FACTs devices. FACTs devices may be used in the future to solve the various problems such as oscillations and variable voltage control and the devices which can be used to provide shunt compensation are SVC and [5–9]. This paper shows the results of the isolated DFIG connected wind farm and wind integrated grid to show how voltage instability occurs in the system due to various reasons which can cause the system failure or reduce the efficiency of the overall wind farm that may affect the performance of the wind farm.

2 Problem Formulation

An analysis is made based on the observations of the Simulink model of the isolated wind farm as well as wind integrated grid. Wind farm instabilities can be checked through various tests. In this paper, symmetrical voltage test, asymmetrical voltage test and dynamic voltage test have been discussed and how it can affect the stability of the wind farm. P - active power Q- Reactive power of the wind turbine system to be observed in this regard by observation values of these both P and Q powers a conclusion is made that to what extent the wind farm is getting affected.

3 Literature Survey

Wind turbine farm system is based on various types of generator such as simple induction generator, permanent synchronous generator, double fed induction generator, and Fig. 1 shows the simple configuration of the wind turbine.

A. Induction Generator Model

The induction generator has two types of rotor; wound rotor and squirrel cage rotor. It has a winding which is short-circuited and not connected to any external voltage source [10–12]. The generators are used to take the mechanical energy and give output in the electrical energy form. Induction generator is cheap, reliable as the construction of these types of generators is based on the squirrel cage induction motor. Induction generators are available in wide ranges and various electrical sizes, these generators are also available in different capacities from fractional horsepower to multimegawatt. These generators can be used for both domestic and commercial purposes with renewable energy wind power applications. Induction generators require reactive power which is usually provided with the help of shunt capacitors in the individual wind turbines. Figure 2 shows the basic configuration of the induction generator connected wind turbine, which consist a gearbox, capacitor banks and various types of loads.

B. Doubly Fed Induction Generator Model

As the name states, the 'doubly fed' refers to the fact that the stator is getting supply from grid and voltage at the rotor is induced with the help of the converter.

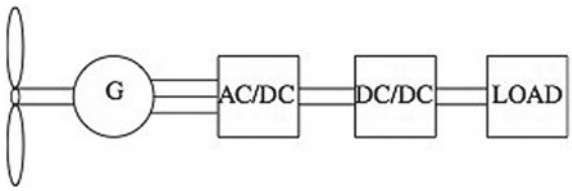


Fig. 1 Configuration of a simple wind turbine system

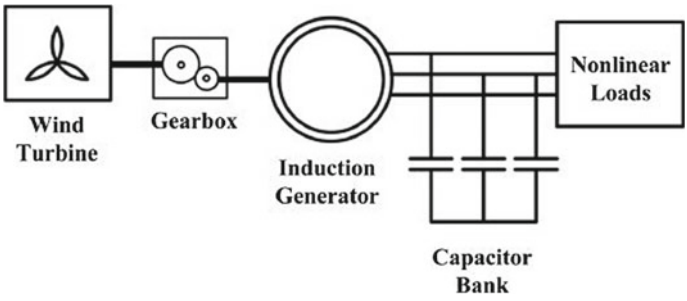


Fig. 2 Configuration of induction generator based wind turbine system

DFIG is considered to perform operations on variable speed performs over a large range but these generators are restricted to a certain range. DFIG by giving input current into the rotor with a variable frequency maintains the difference between the mechanical and electrical frequency. DFIG consist two converters ,the first converter and the second one is grid side converter; these both converters are independently controlled of each other. The first converter is used to control the P and Q power output and grid side converter is used to control the DC link voltages. The main advantage of this type of induction generator is that it can also operate at the variable speeds to produce electricity at even low speed. Figure 3 shows the general layout of the DFIG-based wind farm system. Figure 4 will show the graph associated with DFIG system.

When the DFIG is decoupled by the converter from the transmission system, then by these conditions or theories, the circuit equations of stator and rotor are different from each other and can be written separately. Grid is fed with the per-unit power which is the total sum of the stator and the rotor power when the losses are ignored

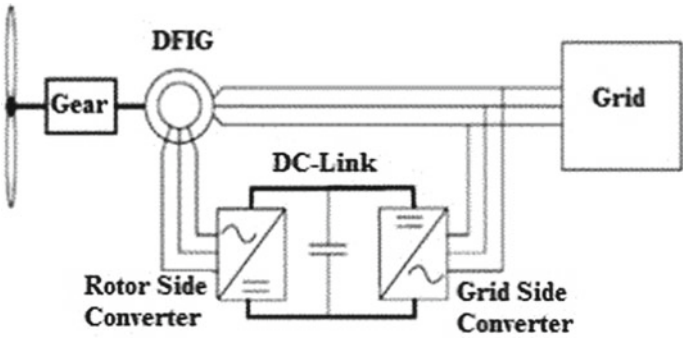


Fig. 3 Configuration of doubly fed induction generator based wind turbine system

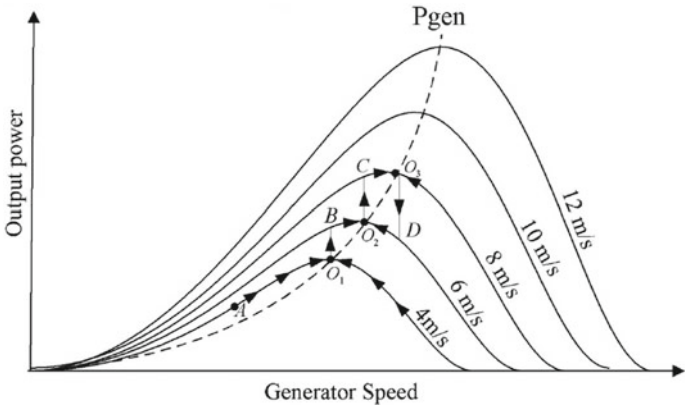


Fig. 4 Graph of DFIG-based wind turbine system

from the system. The reactive power will be absorbed by the grid or released by the grid. Basically, it is used to capture the power or generate the maximum energy at variable wind speed even very low speed so that its full potential can be utilised. Rotor, gearbox and blades of the propeller all are combined together to make mechanical frame of the wind turbine system and these all are elastically connected with each other to provide oscillations in the system.

The wind farm torque which is applied on the blades comes from the values of the pitch angles and rate at which the wind and shaft rotate. The DFIG is used because of its speed control and reduced cost.

Furthermore, DFIG has two control modes

Power Factor Control mode: At the point of connection in load flow studies, P and Q power of the stator should be controlled. DFIG can be modelled as LOAD bus.

Voltage Control mode: In the voltage control mode at the point of common connection, the voltage is kept constant. In the load flow studies, the DFIG is modelled as a GENERATOR bus [13].

C. **Permanent Magnet Synchronous Generator Model**

In PMSG model, the excitation is provided without energy supply that is why its efficiency is higher than the induction generator model. In this model, the excitation is provided with the help of the permanent magnet which further enhances the cost of the overall setup. The stator is wound in this system and rotor has a permanent magnet pole system. Advantage of this system is that at different speeds, power can be generated by the wind turbine system.

A disadvantage of the system is that it is synchronous in nature which can create problems during start-ups and voltage regulations. PMSGs are affected by the temperature so the temperature of the rotor should be kept at safe limit for which cooling system is required in this type of system. Figure 5 shows the general configuration of the PMSG-based turbine system. Figure 6 will show the graph associated with PMSG system.

In PMSG, large no. of poles can be used to eliminate the gearbox from the system. By using large no. of poles, the efficiency of the system will increase as compared to other generators, the PMSG with direct drive is very expensive but still considered as a good option than the other generators because of its high efficiency and low wear and tear [14].

4 Descriptions of the Components and Model

The MATLAB system simulation software is used to analyse the system. The studied system is an isolated DFIG-based wind farm system with a speed of 12 m/s, maximum pitch angle of the wind farm system is 45° and pitch angle of wind turbine system is taken as 2° to calculate the P and Q power output of the system. Further, the wind

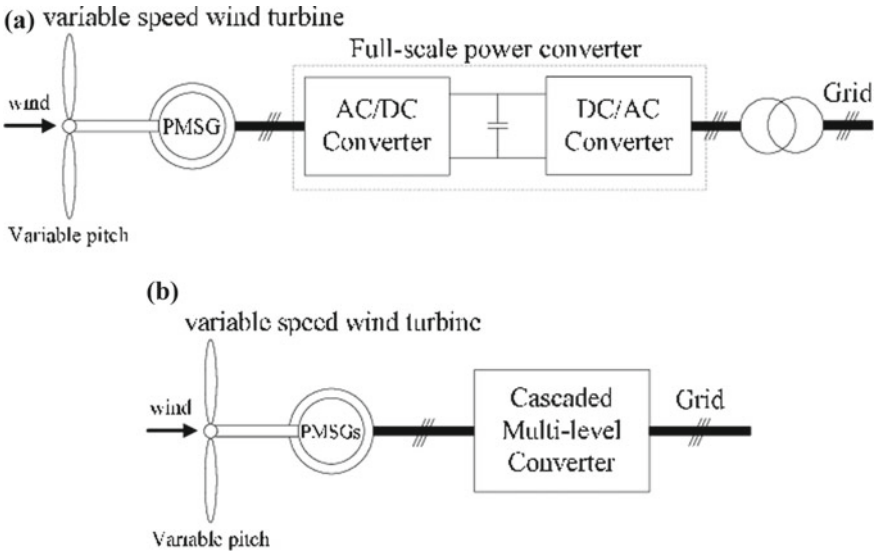


Fig. 5 Configuration of PMSG-based wind turbine system

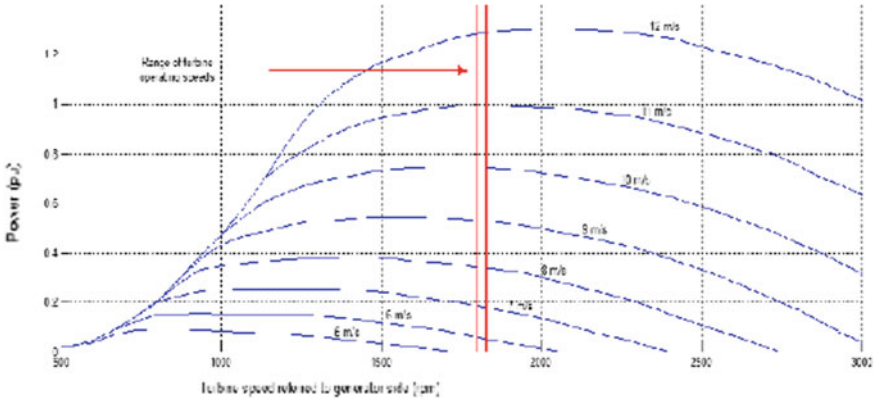


Fig. 6 Graph of PMSG-based wind turbine system

integrated grid system will also be discussed which is consisting of 14 buses system which is connected to the wind turbine system. Figure 7 will show the isolated wind turbine system and Fig. 8 will show the wind integrated grid system.

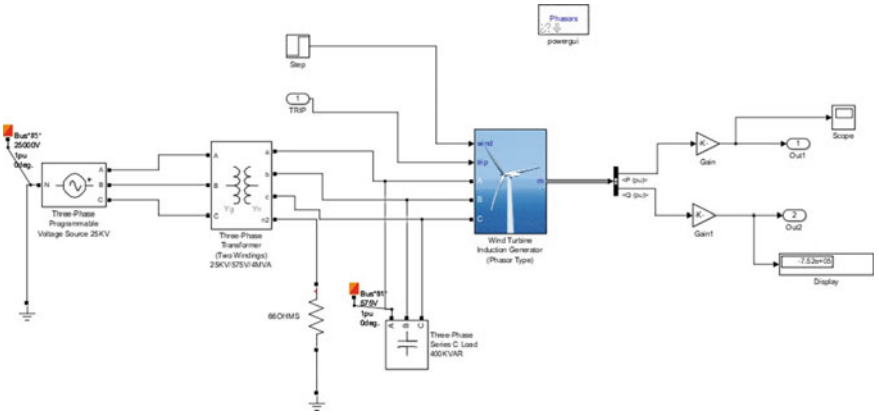


Fig. 7 Wind turbine configuration of the studied system

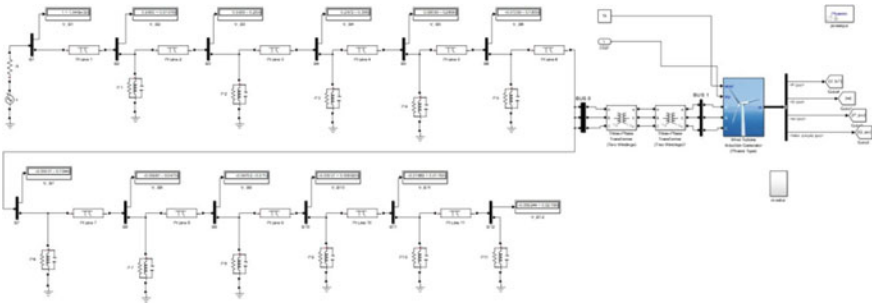


Fig. 8 Wind integrated grid configuration of the studied system

5 Implementation Procedure

Voltage stability analysis can be classified on the basis of two factors such as steady and dynamic analysis. For the steady analysis, the system is studied at different time period in time domain trajectory, this process will be very useful to get information about stability of the voltage level and probability to voltage collapse. To analyse the voltage stability of the system, static techniques are sufficient, dynamic analysis consists of analysis of various differential equations and order of the equations should be first which can be derived and solved by using various integration methods [15]. In this paper, P (active power) and Q (reactive power) of the system of isolated wind farm system and grid integrated wind turbine system are discussed. Figures 9 and 10 will show both P and Q power of the isolated wind turbine system.

P (active power) response of the isolated wind turbine system is 2 MW and the Q (reactive power) output of the isolated wind farm system is −0.789 MVAR, − is the sign which signifies that the power is being absorbed. Figure 11 will show

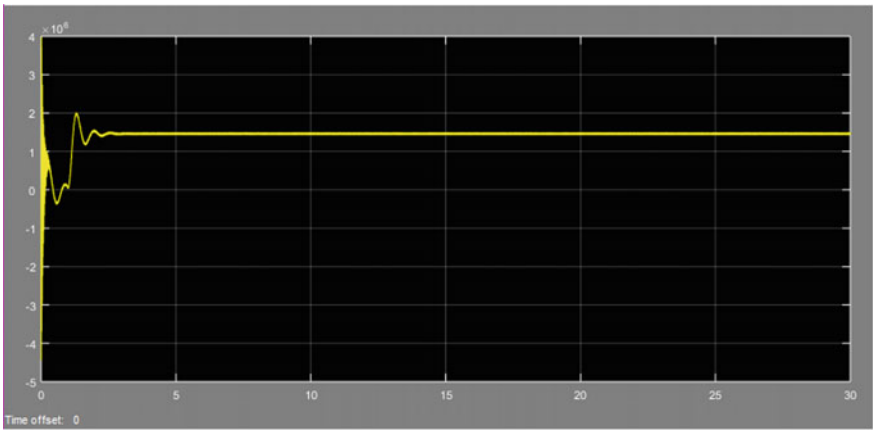


Fig. 9 Active power output of the wind farm

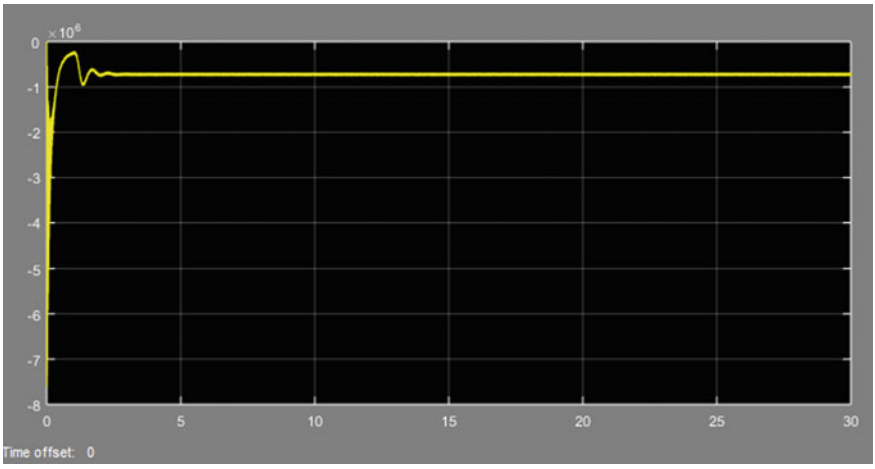


Fig. 10 Reactive power output of the wind farm

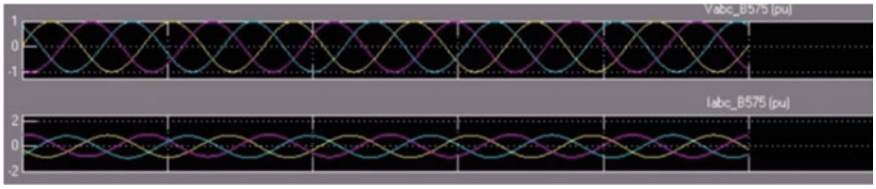


Fig. 11 Input voltage and input current of the wind integrated system



Fig. 12 Active and reactive power output of the wind integrated system

the parameters and sinusoidal wave of input V (voltage) and I (current) of the wind integrated system.

As it is obvious that isolated wind turbine system will show the different output as compared to the integrated grid, so it became important to analyse what will be the response of the power output of the wind integrated system and by analysing the system, it has been observed that system will show some difference in the active and reactive power output response of the system. Figure 12 will show the active and reactive power output of the wind integrated system.

6 Results and Discussions

In the studied system, variations are observed in both real and reactive powers due to various reasons such as sudden wind gusts, electrical loads are not designed to maintain their normal usage behaviour during a voltage dip, load variation and degree of wind power penetration. In isolated wind farm, active power is 2 MW and reactive power is −0.789 MVAR and in wind integrated grid system, these values are somewhat lower than the isolated wind system due to instabilities.

7 Conclusion

For power system, it is very important to eliminate or control the voltage instabilities so that equipment can be prevented from various faults such as excess amount of heating of generators and losses occuring in the transmission lines to make the system capable so that it becomes able to overcome the voltage collapse. So, to maintain voltage stability, reactive power can be decreased to increase the voltage level and reactive power can be increased to decrease the voltage level of the system. If reactive power is in excess, then inductive elements can be used to absorb the reactive power and if reactive power is required, then capacitive elements are required so that there will be a sufficient amount of reactive power for the system. The proper reactive power compensation which is required by the system can be provided with the help of static VAR compensator so that it can help to maintain the voltage level provided by the active power. Compensation is required in the transmission lines by maintaining the voltage levels and the voltage level can be maintained with the help of the shunt connected compensator.

References

1. V. Akhmatov, Analysis of dynamic behavior of electric power systems with large amount of wind power. Ph.D. Dissertation, Department Electric Power Engineering, Orsted-DTU, Denmark, April 2003
2. Y. Lei, A. Mullane, G. Lightbody, R. Yacamini, Modeling of the wind turbine with a doubly fed induction generator for grid integration studies. IEEE Trans. Energy Convers. **21**(1), 257–264 (2006)
3. R. Pena, J.C. Clare, G.M. Asher, Doubly fed induction generator using back- to-back PWM converters and its application to variable speed wind-energy generation. IEE Proc.-Electr. Power Appl. **143**(3), 231–241 (1996)
4. M.A. Poller, Doubly-fed induction machine models for stability assessment of wind farms, in *Power Tech Conference Proceedings, IEEE Bologna*, vol. 3, 23–26 June 2003, 6 p.
5. M. Guleryuz, A. Demiroren, Effects of a wind farm and FACTS devices on static voltage stability of Bursa transmission system in Turkey, in *10th International Conference on Environment and Electrical Engineering (EEEIC)* (2011), pp. 1–5
6. C. Angeles-Camacho, E. Farinas-Wong, F. Banuelos-Ruedas, FACTS: its role in the connection of wind power to power networks, in *Proceedings of the International Symposium on Modern Electric Power Systems (MEPS)* (2010), pp. 1–7
7. N.G. Hingorani, FACTS technology: state of the art, current challenges and the future prospects, in *Power Engineering Society General Meeting* (2007), pp. 1–4
8. I.F. Mohamed, S.H.E. Abdel Aleem, A.M. Ibrahim, A.F. Zobaa, Optimal sizing of C-type passive filters under non-sinusoidal conditions. Energy Technol. Policy **1**, 35–44 (2014)
9. H. Tarafdar, M.A. Roshan, A. Lafzi, Dynamic stability improvement of a wind farm connected to grid using STATCOM, in *5th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)* (2008), pp. 1057–1060
10. K. Nandigam, B.H. Chowdhury, Power flow and stability models for induction generators used in wind turbines. IEEE Power Eng. Soc. Gen. Meet. **2**(6–10), 2012–2016 (2004)
11. P. Kunder, *Power System Stability and Control* (McGrawHill, New York, 1994)
12. F.Q. Zhou, G. Joos, C. Abbey, Voltage stability in weak connection wind farms, in *IEEE Power Engineering Society General Meeting*, vol. 2, 12–16 June 2005, pp. 1483–1488
13. Z. Zeng, X. Li, J. Zhou, Y. Zhang, Investigation of wind farm on power system voltage stability based on bifurcation theory, in *Asia Pacific Power and Energy Engineering Conference (APPEEC)* (2009), pp. 1–4
14. G. Bhuvaneshvari, R. Balasubramanian, H. Ahuja, Performance comparison of DFIG and PMSG based WECS. IIT Delhi
15. V. Ajjarapu, *Computational Techniques for Voltage Stability Assessment and Control* (Springer, New York, 2007); P. Kundur, N.J. Balu, M.G. Lauby, *Power System Stability and Control*, vol. 7 (McGraw-Hill, New York, 1994)

CDM-Based PID Controller with Filter Design for Performance Improvement of Two-Mass Drive System



Subrata Jana and Benjamin A. Shimray

1 Introduction

Various multi-mass Industrial drive systems exists in industries such as textile industries, paper mills, rolling-mills, conveyor belts, robot arm, servo systems, and deep-space antenna drives. There exists torsional vibration which could degrade product qualities. The need for disturbance rejection and vibration suppression originates in such systems. Vibration effect is produced because of backlash, bearing failure, positioning error between motor and load, the heavy weight of the shaft and elastic coupling of the load to the drive motor, etc.

It is very much difficult to control two-mass drive systems when all system variables are not measurable. To overcome the problems, various control strategies have been proposed including resonance ratio control, addition of some feedback loops [1], sliding-mode control [2], PI control, PI-based state-space control, model-based predictive control [3], improved NCTF (nominal characteristic trajectory following) controller [4], explicit MPC (model predictive control)-PI switching control [5], adaptive control, and using Kalman filter [6].

Generally, in the industrial application of a two-mass drive system, the resonance frequency of many drive parts is very low, that is why it is essential to reduce it and improve the damping response. The servo system provides the variable of the motor, but it does not provide the variable of the load. So it is a little bit difficult to design and implement two-mass systems. Using a PID controller, we can obtain improved and effective damping response for speed control. [7]

In this paper, CDM-based PID controller with a low-pass filter is proposed to minimize vibration suppression, faster-tracking load speed with minimum overshoot, effective rejection of load disturbance, and robust control of a two-mass drive system.

S. Jana · B. A. Shimray (✉)
Department of Electrical Engineering, National Institute of Technology Manipur, Imphal, Manipur, India
e-mail: benjaminshimray@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_90

2 Mathematical Analysis of the Two-Mass System

A typical two-mass drive system is as shown in Fig. 1, where T_M is the motor torque, ω_M is the motor angular velocity, T_S is the shaft torsional torque, T_L is the load torque, and ω_L is the load speed. Some of the important parameters are given in Table 1 [8].

The state equations governing two-mass systems are as follows:

$$\frac{\mathrm{d}}{\mathrm{dt}}\omega_L = -\frac{B_L}{J_L}\omega_L + \frac{1}{J_L}T_S - \frac{1}{J_L}T_L \tag{1}$$

$$\frac{\mathrm{d}}{\mathrm{dt}}\omega_M = -\frac{B_M}{J_M}\omega_M - \frac{1}{J_M}T_S + \frac{1}{J_M}T_M \tag{2}$$

$$T_S = B_S(\omega_M - \omega_L) + K_S(\theta_M - \theta_L) \tag{3}$$

$_L =$ Load angular displacement in rad, and
$_M =$ Motor angular displacement in rad.

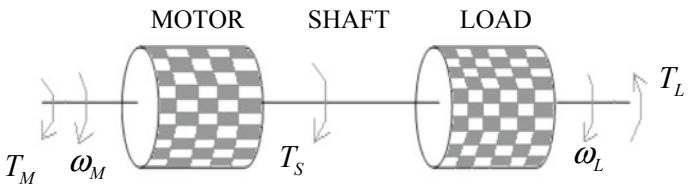


Fig. 1 Two-mass drive system

Table 1 The nominal parameter of a two-mass system

Quantity	Unit	Value
Shaft stiffness (K_S)	Nm/rad	138
Shaft damping coefficient (B_S)	Nm/rad/s	0.1000
Load inertia (J_L)	kg.m2	0.0086
Load viscosity coefficient (B_L)	Nm/rad/s	0.069
Motor inertia (J_M)	kg.m2	0.048
Motor viscosity coefficient (B_M)	Nm/rad/s	0.0013

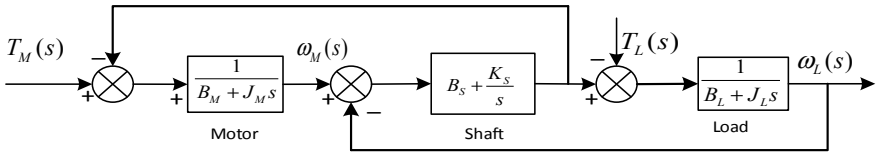


Fig. 2 Block diagram of the two-mass drive system

Figure 2 show the approximate block diagram of a two-mass drive system, where $G_M(s) = \frac{1}{B_M + J_M s}$ (motor transfer function), $G_S(s) = B_S + \frac{K_S}{s}$ (shaft transfer function), and $G_L(s) = \frac{1}{B_L + J_L s}$ (load transfer function).

The transfer function of the load speed, motor speed, and shaft torque are as follows:

$$\omega_L(s) = \frac{G_M(s)G_S(s)G_L(s)}{1 + G_M(s)G_S(s) + G_S(s)G_L(s)} T_M(s) - \frac{G_L(1 + G_S(s)G_M(s))}{1 + G_M(s)G_S(s) + G_S(s)G_L(s)} T_L(s) \tag{4}$$

$$\omega_M(s) = \frac{G_M(s)(1 + G_S(s)G_L(s))}{1 + G_M(s)G_S(s) + G_S(s)G_L(s)} T_M(s) - \frac{G_M(s)G_S(s)G_L(s)}{1 + G_M(s)G_S(s) + G_S(s)G_L(s)} T_L(s) \tag{5}$$

$$T_S(s) = \frac{G_M(s)G_S(s)}{1 + G_M(s)G_S(s) + G_S(s)G_L(s)} T_M(s) - (-\frac{G_S(s)G_L(s)}{1 + G_M(s)G_S(s) + G_S(s)G_L(s)}) T_L(s) \tag{6}$$

To design a closed loop, the transfer function of $\omega_M(s)$ to $T_M(s)$ plays a very important role and is given by [9]:

$$H_{\mathrm{MM}}(s) = \frac{\omega_M(s)}{T_M(s)}|_{T_L=0} = \frac{1}{J_L J_M} \frac{J_L s^2 + (B_L + B_S)s + K_S}{\Delta(s)} \tag{7}$$

where $\Delta(s)$ is the characteristic equation of the system and it is

$$\Delta(s) = s^3 + (\frac{B_M + B_S}{J_M} + \frac{B_L + B_S}{J_L})s^2 + \left(\frac{K_S(J_L + J_M) + B_S(B_L + B_M) + B_L B_M}{J_L J_M}\right)s + \frac{(B_M + B_L)K_S}{J_L J_M} \tag{8}$$

For simplification, we can neglect the friction term as it is very small:

$$H_{\mathrm{MM}}(s) = \frac{1}{J_M} \frac{s^2 + \omega_A^2}{s(s^2 + \omega_R^2)} \tag{9}$$

The antiresonance frequency (ω_A), resonance frequency (ω_R), and inertia ratio (K_J) are given by

$$\omega_A = \sqrt{\frac{K_S}{J_L}}; \tag{10}$$

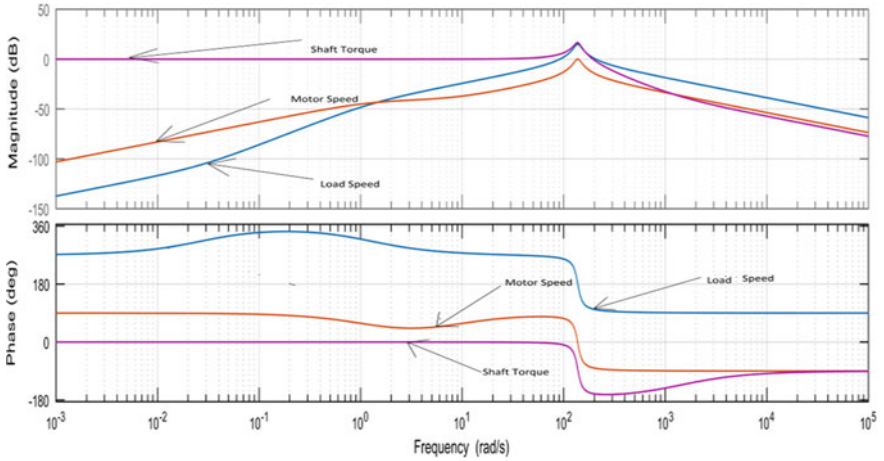


Fig. 3 Frequency domain response of the two-mass system relative to the motor

$$\omega_R = \omega_A \sqrt{1 + \frac{J_L}{J_M}}; \tag{11}$$

$$K_J = \frac{J_L}{J_M}. \tag{12}$$

It is very difficult to suppress of torsional vibration because of the small value of K_J. Also, the control bandwidth in the closed-loop motion control system is limited by the ω_A. The ω_R depends upon the inertia ratio and is affected by the drive type. If motor inertia constant (J_M) is increased, ω_R decreases without affecting ω_A. Conversely, from Eqs. 10 and 11, we can say that ω_A and ω_R increase with increase of Ks and also the mechanical bandwidth in the closed-loop system. So for large systems, it has large inertia and it creates low natural frequencies.

The dominant pole for the open-loop system are $P_1 = -1.25$, and $P_{2,3} = -10.26 \pm j137.14$. The damping factor without controller (η) of system is 0.0498.

The resonant frequency (ω_R) is = 137.6 rad/s and antiresonant frequency (ω_A) is = 126.7 rad/s. Since all poles of the system are on the left-hand side, we can easily observe that the system is stable but highly damped. The frequency domain response relative to the motor is given in Fig. 3.

3 Controller and Filter Design

CDM is an unprecedented technique for analyzing stability, speed response, and robustness of a system. Professor Shunji Manabe first proposed this technique in 1991 [10]. This technique is similar to the indirect pole placement method for finding out the polynomial of a characteristic equation. This technique can design a controller for robust control with desired system response. In this technique, important parameters

called stability index (γ_k) and equivalent time constant (τ) are used. The generalized equation of CDM is

$$P_S\text{target} = a_0[\{\sum_{i=2}^{n}(\sum_{j=1}^{i-1}\frac{1}{\gamma_{i-j}^{j}})(\tau s)^i\} + \tau s + 1] \tag{13}$$

The main objective of CDM is to compare the coefficient of generalized CDM equation to the close characteristic loop equation and derive all the gain values of the PID controller and filter. In a conventional PID controller, $G_C(s) = 0$ and $G_V(s) = K_P + \frac{K_I}{s} + sK_D$.

Figure 4 shows the control structure of the two-mass drive system.

The motor speed, load speed, and shaft torque depend on the disturbance torque $T_L(s)$ and reference command speed ω_C:

$$\omega_M(s) = \frac{G_M(s)G_V(s)G_F(s)(1+G_S(s)G_L(s))}{1+G_M(s)G_V(s)G_F(s)+G_M(s)G_S(s)+G_S(s)G_L(s)+G_L(s)G_S(s)G_M(s)G_V(s)G_F(s)}\omega_C(s) - \frac{G_M(s)G_S(s)G_L(s)}{1+G_M(s)G_V(s)G_F(s)+G_M(s)G_S(s)+G_S(s)G_L(s)+G_L(s)G_S(s)G_M(s)G_V(s)G_F(s)}T_L(s) \tag{14}$$

$$\omega_L(s) = \frac{G_M(s)G_S(s)G_L(s)G_V(s)G_F(s)}{1+G_M(s)G_V(s)G_F(s)+G_M(s)G_S(s)+G_S(s)G_L(s)+G_L(s)G_S(s)G_M(s)G_V(s)G_F(s)}\omega_C(s) - \frac{G_L(1+G_S(s)G_M(s)+G_M(s)G_V(s)G_F(s))}{1+G_M(s)G_V(s)G_F(s)+G_M(s)G_S(s)+G_S(s)G_L(s)+G_L(s)G_S(s)G_M(s)G_V(s)G_F(s)}T_L(s) \tag{15}$$

$$T_S(s) = \frac{G_M(s)G_S(s)G_V(s)G_F(s)}{1+G_M(s)G_V(s)G_F(s)+G_M(s)G_S(s)+G_S(s)G_L(s)+G_L(s)G_S(s)G_M(s)G_V(s)G_F(s)}\omega_C(s) - (-\frac{G_S(s)G_L(s)(1+G_M(s)G_V(s)G_F(s))}{1+G_M(s)G_V(s)G_F(s)+G_M(s)G_S(s)+G_S(s)G_L(s)+G_L(s)G_S(s)G_M(s)G_V(s)G_F(s)})T_L(s) \tag{16}$$

The closed-loop characteristic equation with controller and filter is given by

$$s^5 + \frac{J_M+K_D}{J_MK}s^4 + \frac{J_MK\omega_R^2+K_P}{J_MK}s^3 + \frac{J_M\omega_R^2+K_I+K_D\omega_A^2}{J_MK}s^2 + \frac{K_P\omega_A^2}{J_MK}s + \frac{K_I\omega_A^2}{J_MK} = 0 = P_S \tag{17}$$

Comparing the coefficient of the equation no. 17 and fifth-order generalized CDM equation, the equivalent time constant and gains of PID controller and gain of filter are evaluated as

$$\omega_R^2\tau^4 - \gamma_4\gamma_3^2\gamma_2^2\gamma_1^2\tau^2 + \frac{\gamma_4\gamma_3^2\gamma_2^3\gamma_1^4}{\omega_A^2} = 0; \tag{18}$$

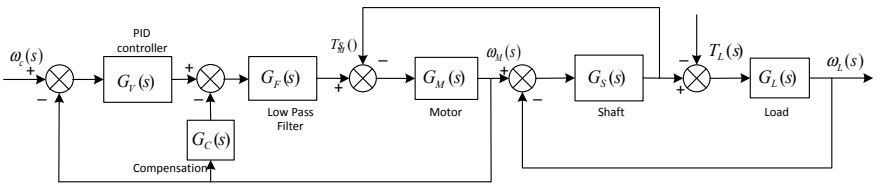


Fig. 4 The control structure of a two-mass drive system

$$a_0 = \frac{\gamma_4 \gamma_3^2 \gamma_2^3 \gamma_1^4}{\tau^5}; \tag{19}$$

$$K_D = \frac{J_M(\omega_A^2 \omega_R^2 \tau^4 + \gamma_3 \gamma_2^2 \gamma_1^3 - \gamma_3 \gamma_2^2 \gamma_1^2 \omega_A^2 \tau^2)}{\gamma_3 \gamma_2^2 \gamma_1^2 \omega_A^2 \tau^2 - \gamma_3 \gamma_2^2 \gamma_1^3 - \omega_A^2 \tau^4}; \tag{20}$$

$$K = \frac{\gamma_1 \omega_A^2 (J_M \omega_R^2 + K_D \omega_A^2)}{a_0 J_M (\omega_A^2 \tau^2 - \gamma_1)}; \tag{21}$$

$$K_P = \frac{a_0 J_M K \tau}{\omega_A^2}; \tag{22}$$

$$K_I = \frac{a_0 J_M K}{\omega_A^2}; \tag{23}$$

To get the proper stability index, we apply Routh–Hurwitz algorithm on fifth-order generalized CDM equation. The condition in Eqs. 24 and 25 should be maintained while designing the controller gain:

$$\gamma_4 \gamma_3 \gamma_2 \gamma_1 > \frac{\gamma_3 \gamma_1 - 1}{\gamma_3 \gamma_2 - 1}; \tag{24}$$

$$\gamma_4^2 \gamma_3^2 \gamma_2^2 \gamma_1 (\gamma_3 \gamma_2 \gamma_1 - \gamma_1 - \gamma_3) + \gamma_4 \gamma_3 \gamma_2 \gamma_1 (2 - \gamma_3 \gamma_2^2 \gamma_1 + \gamma_3 \gamma_2) > 1; \tag{25}$$

Here, Eq. 26 gives the characteristic equation of closed-loop PID controller.

$$s^4 + \frac{K_P}{J_M + K_D} s^3 + \frac{(J_M \omega_R^2 + K_I + K_D \omega_A^2)}{J_M + K_D} s^2 + \frac{K_P \omega_A^2}{J_M + K_D} s + \frac{K_I \omega_A^2}{J_M + K_D} = 0 = P_S \tag{26}$$

Similarly, comparing the coefficient of Eq. 26 and the fourth-order generalized CDM equation, the equivalent time constant and controller's gains are calculated [11].

$$\tau = \frac{\gamma_1 \sqrt{\gamma_2}}{\omega_A}; \tag{27}$$

$$K_I = \frac{K_S \gamma_3}{\gamma_1 \gamma_2 \gamma_3 - \gamma_3 - \gamma_1}; \tag{28}$$

$$K_D = \frac{J_L \gamma_1}{\gamma_1 \gamma_2 \gamma_3 - \gamma_3 - \gamma_1} - J_M; \tag{29}$$

$$K_P = \frac{\gamma_1 \gamma_3 \sqrt{\gamma_2 K_S J_L}}{\gamma_1 \gamma_2 \gamma_3 - \gamma_3 - \gamma_1}; \tag{30}$$

To get the proper stability index, we apply Routh–Hurwitz algorithm on fourth-order generalized CDM equation. The stability index should also maintain the given condition in Eq. 31:

$$\gamma_2 > \left(\frac{1}{\gamma_3} + \frac{1}{\gamma_1}\right); \tag{31}$$

4 Simulation Result

The control structure of two-mass system with a PID controller and filter in MATLAB–Simulink shown in Fig. 5. The controller parameter of the CDM-PID+ filter is obtained using Eqs. (20)–(23) while Eqs. (28)–(30) are used to determine the control parameter for CDM-PID controller. All the controller parameter values are evaluated and presented in Table 2.

Table 3 shows the pole location and equivalent time constant (τ) of the closed-loop systems for different stability indices. All the poles are located on the left-hand side, which indicates that the systems are stable. All simulation results of shaft torque, load speed and motor speed response to a unit-step disturbance torque and command speed are presented in Figs. 6, 7, 8, 9, 10, 11, 12, 13, and 14.

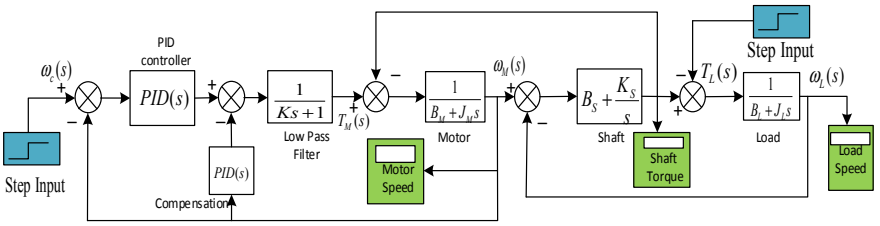


Fig. 5 Control structure in MATLAB of a two-mass system

Table 2 All calculations controller parameter for PID and PID + FILTER

Stability Index (γ_k)	PID			PID with filter			
	K_P	K_I	K_D	K_P	K_I	K_D	K
$\gamma_1 = 2.5$ $\gamma_2 = 2.3$ $\gamma_3 = 1.45$ $\gamma_4 = 1$	1.36	45.6	−0.043	1.47	37.36	−0.0357	0.00119
$\gamma_1 = 3$ $\gamma_2 = 2$ $\gamma_3 = 1.4$ $\gamma_4 = 1.25$	1.61	48.3	−0.041	1.98	46.22	−0.031	0.00143
$\gamma_1 = 2.75$ $\gamma_2 = 2.25$ $\gamma_3 = 1.5$ $\gamma_4 = 1$	1.34	41.1	−0.04	1.39	34.03	−0.038	0.0009

Table 3 Eigenvalue and time constant of the closed-loop system

Stability Index (γ_k)	PID		PID with Filter	
	Pole Location	τ	Pole Location	τ
$\gamma_1 = 2.5$ $\gamma_2 = 2.3$ $\gamma_3 = 1.45$ $\gamma_4 = 1$	$-90.13 \pm j146.2$ $-65.6 \pm j25.4$	0.03	-215.3 $-43.8 \pm j21.5$ $-15.2 \pm j136.6$	0.0394
$\gamma_1 = 3$ $\gamma_2 = 2$ $\gamma_3 = 1.4$ $\gamma_4 = 1.25$	-79.75 -54.40 $-64.98 \pm j145$	0.033	-38.36 $-84.79 \pm j45.8$ $-19.3 \pm j173$	0.043
$\gamma_1 = 2.75$ $\gamma_2 = 2.25$ $\gamma_3 = 1.5$ $\gamma_4 = 1$	$-87.9 \pm j148.6$ $-65.5 \pm j11.53$	0.032	-45.6 $-85.6 \pm j27.5$ $-18.14 \pm$ j189.1	0.041

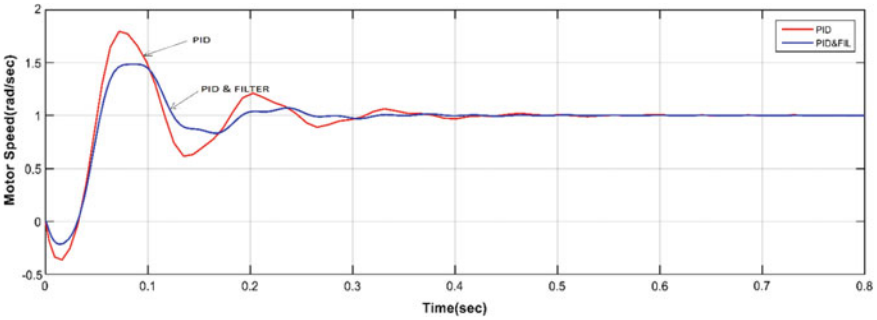


Fig. 6 Motor speed characteristics (with PID controller and filter and only with PID controller)

After using a filter with a PID controller, the peak overshoot and rise time are reduced. The system response is also found faster.

Responses for Stability Index 2.5, 2.3, 1.45, and 1 are shown in Figs. 6, 7, and 8.

Responses for Stability Index 3, 2, 1.4, and 1.25 are shown in Figs. 9, 10, and 11.

Responses for Stability Index 2.75, 2.25, 1.5, and 1 are shown in Figs. 12, 13, and 14.

5 Conclusion

CDM-based PID controller in cascade with a low-pass filter for an industrial drive system was investigated. CDM technique is used to determine the controller parameter. With the increase in complexity of the classical system method of tuning a PID controller, it becomes difficult and time consuming. Many optimization algorithms using

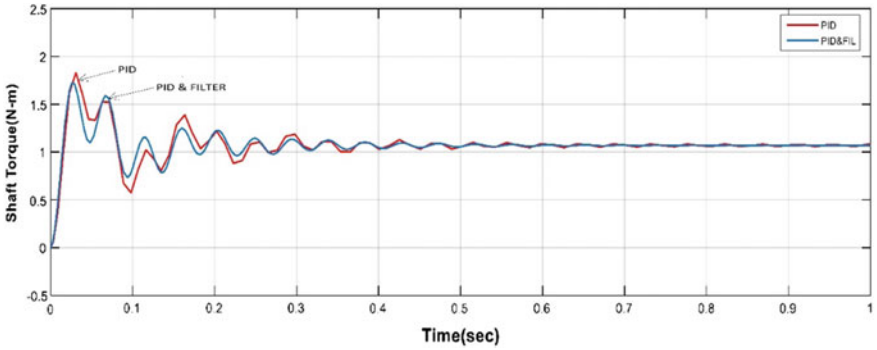


Fig. 7 Shaft torque characteristics (with PID controller and filter and only with PID)

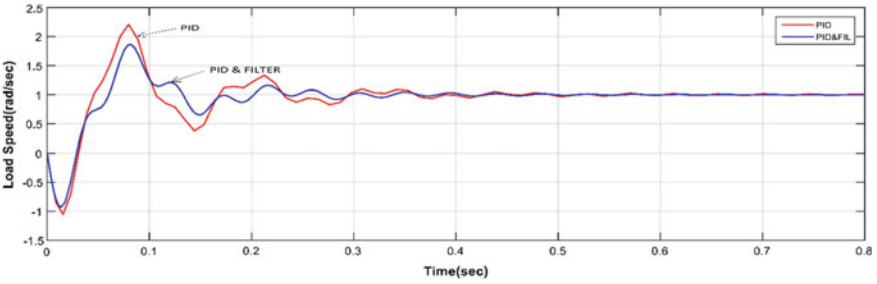


Fig. 8 Load speed characteristics (with PID controller and filter and only with PID controller)

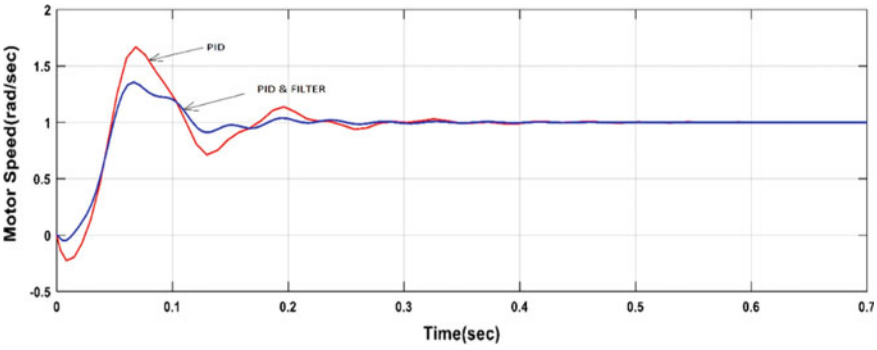


Fig. 9 Motor speed characteristics (with PID controller and filter and only with PID controller)

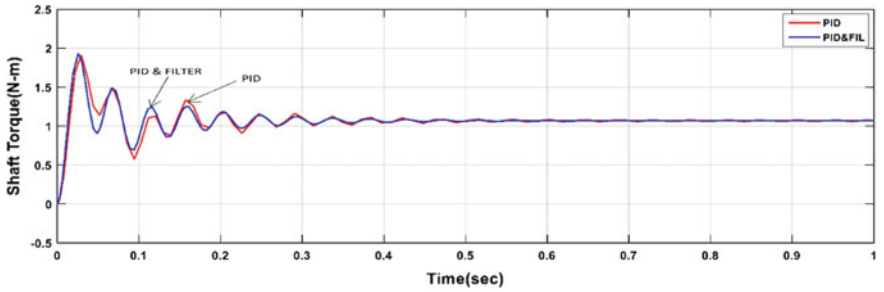


Fig. 10 Shaft torque characteristics (with PID controller and filter and only with PID controller)

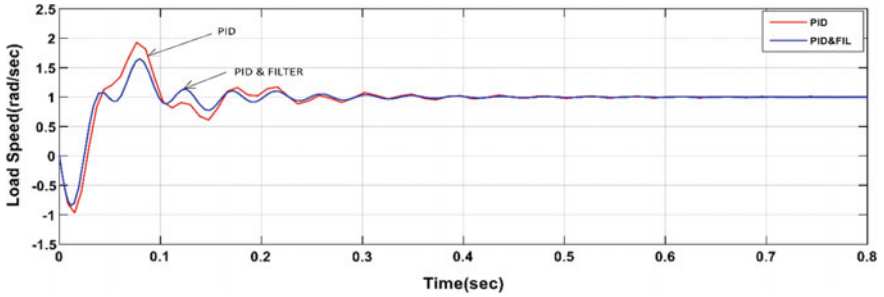


Fig. 11 Load speed characteristics (with PID controller and filter and only with PID controller)

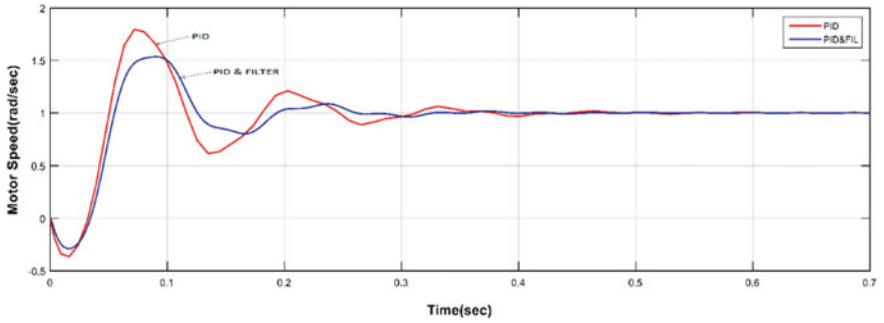


Fig. 12 Motor speed characteristics (with PID controller and filter and only with PID controller)

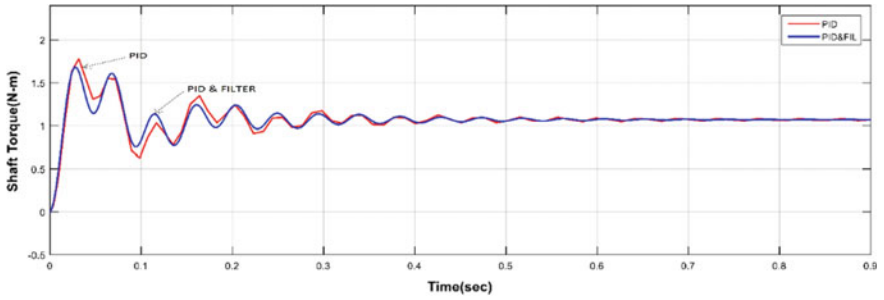


Fig. 13 Shaft torque characteristics (with PID controller and filter and only with PID controller)

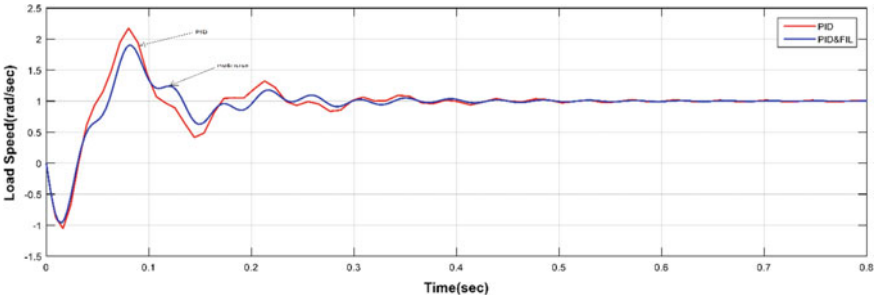


Fig. 14 Load speed characteristics (with PID controller and filter and only with PID controller)

soft-computing techniques were available but CDM-based PID controller parameter optimization is comparatively easier to implement.

The designed controller could effectively improve the system responses. The responses are also observed for different stability indices as proposed in the CDM technique. Responses for Stability Index 3, 2, 1.4, and 1.25 shown in Figs. 9, 10, and 11 present an impressive result. The percentage of overshoot was reduced by almost 40% on the motor side and almost 20% on the load side. Vibration in the shaft is also effectively reduced as indicated in Fig. 10. The results also show that settling time is around 0.3–0.4 s.

Though simple in concept, the controller parameter optimization is time-consuming as the system order increases. In the future, complex industrial drive systems may be modeled as a three-mass system that will give insights on the influence of higher eigenfrequencies, which are important aspects.

References

1. K. Sugiura, Y. Hori, Vibration suppression in 2- mass and 3- mass system based on the feedback of imperfect derivative of the estimated torsional torque. IEEE Trans. Ind. Electron. **43**(1) (1996)
2. K. Erenturk, Nonlinear two-mass system control with sliding-mode and optimized proportional–integral derivative controller combined with a grey estimator. *Published in IET Control Theory and Applications*. https://doi.org/10.1049/iet-cta:20070330
3. S. Thomsen, N. Hoffmann, F.W. Fuchs, PI control, PI-based state space control, and model-based predictive control for drive systems with elastically coupled loads—a comparative study. IEEE Trans. Ind. Electron. **58**(8) (2011)
4. M.F.M. Yakub, W. Martono, R. Akmeliawati, Vibration control of two-mass rotary system using improved NCTF controller for positioning systems, in *2010 IEEE Control and System Graduate Research Colloquium* (2010)
5. C. Wang, M. Yang, W. Zheng, J. Long, D. Xu, Vibration suppression with shaft torque limitation using explicit MPC-PI switching control in elastic drive systems. IEEE Trans. Ind. Electron. **62**(11) (2015)
6. K. Szabat, T. Orlowska-Kowalska, Application of the Kalman filters to the high-performance drive system with elastic coupling. IEEE Trans. Ind. Electron. **59**(11)

7. S. Choudhary, S.K. Sharma, V. Shrivastava, Modelling of speed controller for industrial applications: a two mass drive system, in *2016 IEEE 7th Power India International Conference (PIICON)* (2016)
8. G. Shahgholian, P. Shafaghi, M. Zinali, S. Moalem, State space analysis control design of two-mass resonant system, in *2009 Second International Conference on Computer and Electrical Engineering* (2009)
9. G. Shahgholian, F.M. Tehrani, Development of state space model and control design of two-mass system using standard forms, in *2011 IEEE 3rd International Conference on Communication Software and Networks* (2011)
10. S.S. Mishra, S.K. Mishra, S.K. Swain, Coefficient diagram method (CDM) based PID controller design for magnetic levitation system with time delay, in *2017 IEEE International Conference on Intelligent Techniques in Control, Optimization and Signal Processing* (2017)
11. G. Shahgholian, Modeling and simulation of a two-mass resonant system with speed controller. Int. J. Inf. Electron. Eng. **3**(5) (2013)

Autonomous Vehicle Power Scavenging Analysis for Vehicular Ad Hoc Network



Garima Sharma, Praveen Kumar Singh and Laxmi Shrivastava

1 Introduction

In the fast-growing world, an Autonomous Vehicle has an important role in our life to meet the challenges of safety of a human being. A VANET is a network of intelligent vehicles that can guide itself without human intervention. An Autonomous Vehicle has become a concrete reality in the way of VANET network designing system where computers take over the art of driving. VANET requires various kinds of technologies including GPS sensing knowledge and use of various sensors for efficient communication and equipment to avoid collisions. According to 'National Highway Traffic Safety Administration' (NHTSA) 2016 survey report, about 4.7 million vehicles to vehicle crashes and about 1.44 million crashes had occurred till now, which lead to youngsters' demises; about 1.25 million people die every year worldwide according to World Health Organization (WHO) 2016 report, and 50 million and more injuries were taking place. Road traffic injuries are approximated to be the ninth foremost death sources worldwide and by 2030, it is expected to become the seventh foremost cause of death. Information, vehicle and human safety is the main point of concern in a VANET; two kinds of communication are processed in VANETs—communication among vehicles, communication between vehicle and Infrastructure. Therefore, Autonomous Vehicles are equipped with different types

G. Sharma (✉) · P. K. Singh · L. Shrivastava
Department of Electronics and Communication Engineering, Madhav Institute of Technology & Science, Gwalior, Madhya Pradesh, India
e-mail: sharmagarima2411@gmail.com

P. K. Singh
e-mail: kumarpraveensingh90@gmail.com

L. Shrivastava
e-mail: lselex@mitsgwalior.in

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_91

of sensors which are responsible for safety purposes. For example, obstacle detection sensor, tire pressure monitoring, vehicle speed sensor, and temperature sensor. Presently, battery is the only power source present in a vehicle for powering various sensors. Therefore, this paper uses a preexisting technology called piezoelectricity and focuses on maximum power recovered from the vibrations that decrease our reliance over battery so that a large number of sensors can be deployed for more safety and comfort purposes.

From Table 1, it is clear that the front hood and car battery generate a minimum amount of vibrations and exhaust part produces maximum vibrations followed by the engine, but from the exhaust part, power cannot be extracted because of heat emission and exposed condition. So, the engine is selected as an experimental area in a vehicle. Since vibrations in a vehicle is imperative though contemplated as the research. The aggregated retrieve power is used in Vehicular Ad Hoc Networks by powering different sensor nodes for improving vehicle safety [1]. And, it is also useful for efficient communication by providing continuous power source apart from the battery. By taking concern of crashes worldwide, extensive research has been made which includes vibration signals because it carries large amount of information and is used as a power source for vehicle sensors to acquire real-time information about vehicle to vehicle, vehicle to infrastructure, and road-side unit communication to reduce traffic, accidents, etc. Qingyuan et al. proposed the idea of extracting vibration from different grades of road to power wireless sensors but that can only scavenge 13.3 μw of power [2]. Then, Kabrane et al. proposed that wireless sensor network technology requires inbuilt intelligent sensors for controlling crashes by calculating their communication parameters and each network node exhausts some amount of energy. So, it addresses an alternate energy source because of limited battery capacity. Mouapi et al. use piezoelectric transducer to extract vibrations and get up to 3 μw power [3]. Mohamad et al. given the idea of piezoelectric material to exploit the ambient vibrations in the micro-generating system for low-energy systems by transforming the mechanical energy produced by engines into electrical energy [4]. So nowadays, piezoelectric energy scavenging technology is attracting the attention of researchers. This paper includes SIMSCAPE modeling in MATLAB for converting mechanical vibrations into electrical energy by the use of piezoelectric transducer which is named as piezo stack in MATLAB. It can obtain maximum power at a particular acceleration and frequency. The vibration produced in Kia spectra reported 47.44 μW power in kilohertz range [5]. This paper scavenged maximum power from a vehicle engine vibration which helps in creating alternate power sources

Table 1 Vibration produced by various locations in a vehicle [8]

Location in a vehicle	Vibration (g)
Front hood	0.05
Car battery	0.07
Engine	0.24
Exhaust	0.79

for driving wireless sensors in Vehicular Ad Hoc Network to decrease the threat of life sensor nodes which can effectively communicate with each other.

2 Piezoelectric Transducer Theoretical Background

The piezoelectric material is responsible for producing electrical energy from vibrations since it gives a good repartee to vibrations.

Common piezoelectric material made up of arium Titanite, Lead Titanite and Lead Zirconate Titanite (PZTs) and its function up to 300 °C temperature. A piezoelectric transducer comprises piezoelectric elements which convert mechanical vibrations into electrical energy. A review of a piezoelectric transducer has been proposed in [3]. And the most prominent piezoelectric generation structure is a cantilever beam used for low-frequency applications [2, 6]. There are three elements in a cantilever beam from which the piezoelectric layer is responsible for producing charges across its ends after amplifying the deflection produced by the beam, which produces electrical energy after applying resistive load. Therefore,

$$F = mA_{\text{in}} \tag{1}$$

where m is the value of seismic mass and A_{in} is the peak of measured acceleration.

The piezoelectric layer generates voltage throughout the electrode by the displacement of charges due to stretching or compression. These two effects are determined by piezo stack equation which is used in SIMSCAPE model.

$$D = d \times T + \varepsilon^T \times E_1 \tag{2}$$

$$S = s^E \times T + d \times E_2 \tag{3}$$

where D is the electric flux density, E_1 is the electric field, T is the mechanical stress, ε^T is the permittivity, E_2 is the mechanical strain, and d is the piezoelectric charge coefficient.

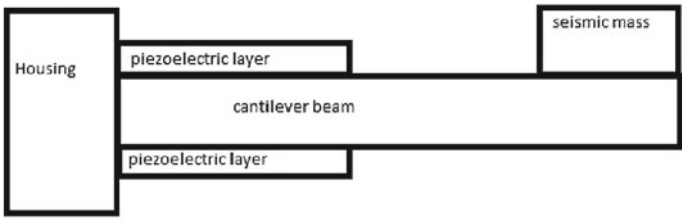


Fig. 1 Structure of a piezoelectric cantilever beam

Table 2 The parameters and their value of piezo stack (PZT-5H) [8, 9]

Symbol	Description	Values
C_b	Elastic constant for piezo component	63 Gpa [9]
d	Piezoelectric charge coefficient	320×10^{-12} C/N [9]
k	Coupling coefficient	0.43 [9]
ξ	Damping ratio	0.0541 [9]
$\varepsilon/\varepsilon_0$	Relative permittivity	3400 [9]
C	Capacitance	7.568nF [9]
l_m	Length of mass	17 mm [8]
h_m	Height of mass	7.7 mm [8]
w_m	Width of mass	3.6 mm [8]
l	Length of beam	11 mm [8]
l_e	Length of electrode	11 mm [8]
w	Width of beam	3.2 mm [8]
t_p	Thickness of piezo layer	0.28 mm [8]
t_c	Thickness of center shim	0.1 mm [8]

The cantilever beam as shown in Fig. 1 is used for the amplification of displacement of seismic mass; moreover, it enhances the mechanical stress applied to the piezoelectric material for generating high output power. Since the scavenged power depends on the properties of detected vibrations, i.e., maximum input acceleration and resonance frequency, the proposed SIMSCAPE model includes an ideal force sensor that is responsible for applying mechanical stress to the piezo stack such that it produces charges across its end and produces electrical energy. Across load resistance, power is calculated theoretically and practically via simulation of proposed model as well.

An expression for maximum RMS power recovered from the piezoelectric transducer have been expressed in [3] as follows:

$$P = \frac{1}{2\omega^2} \frac{RC_b^2 \left(\frac{2C_p d_{31} t_c}{ka\varepsilon}\right)^2 A_{\text{in}}^2}{\left(4\xi + k_{31}^4\right)\left(RC_b\omega_n^2\right)^2 + 4\xi k_{31}^2 (RC_b\omega_n) + 4\xi^2} \tag{4}$$

where A_{in} is the input acceleration, C_b and C_p are the capacitance of piezoelectric plate, ξ is the Young's modulus of the piezoelectric plate, k is the damping ratio, t_c is the beam thickness, R is the resistance, and ω and ω_n are the vibration frequency and the resonance frequency of the beam. Since parameters depend on the static and dynamic force, the coefficients related to the beam dimensions are given in Table 2. The parameters and their value of piezo stack (PZT-5H) as listed in Table 2.



Fig. 2 Used vehicle engine

3 Vibration in a Vehicle

There are different points in a vehicle where a certain amount of vibration is produced which is shown in Table 1. In this section, engine is selected as an area of experiment because the vehicle engine is considered as the major source of vibration generator. This paper measures vibration at different rotational velocity (RPM).

3.1 Engine as a Vibration Generator

This paper experimentally analyzes the vibration produced by a Tata Indica car engine as shown in Fig. 2 whose maximum power is 55.23bhp @5200 rpm and has 1193 cc engine with inbuilt tachometer used to analyze the vibration produced at different rotational velocity (rpm) as well as obtains the maximum amount of acceleration produced at different frequencies.

3.2 Measurement Equipment

In this measurement, the SKF Microlog Analyzer CMXA 75 wireless vibration monitoring system is used. It has high performance, triaxial accelerometer, and an FFT analyzer having high-speed data processor model which effectively acquires dynamic vibrations. This measurement equipment is comprised of Marvell 806 MHz PXA 320 processor for calculating fast real-time rate and equivalently displays the results. CMAC 4370-K triaxial accelerometer is used for measuring the acceleration



Fig. 3 Measurement setup

with sensitivity (±5%) 100 mVg, and the FFT analyzer is used for data analysis. Thus, the vibration measurement setup is as shown in Fig. 3.

3.3 Detected Vibration

The vibrations observed at different rpm and maximum received acceleration are 6.25 g at 55 Hz frequency as shown in Fig. 4. This maximum acceleration is achieved at 3250 rpm up to 100 Hz; moreover, the experiment goes up to 2500 Hz for data validation.

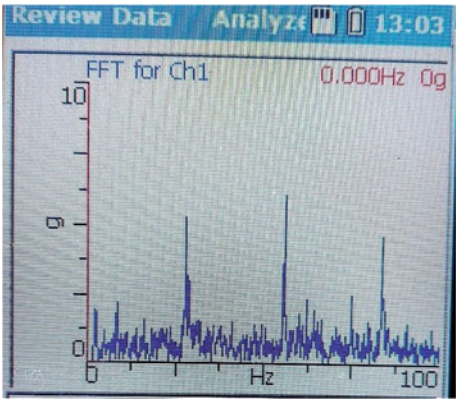


Fig. 4 Experimental vibration result

4 Piezoelectric RMS Power SIMSCAPE Model

MATLAB-based SIMSCAPE software allows a single environment for simulating multi-domain physical system. SIMSCAPE components represent physical elements that correspond to the physical connections in the real system. This tool allows simulation of multi-physical model of a piezoelectric transducer which acts as a generator without any electrical circuit for recovering maximum power generation at a maximum acceleration to power vehicle wireless sensors in Vehicular Ad Hoc Network. This simulation model represents the analogy of mechanical setup for scavenging vibration. The piezo stack represents the electrical and force characteristics by using the constitutive equations developed by Smith et al. [7]. Piezo stack works according to Eqs. (1) and (2) and the parameters used in piezo stack are listed in Table 2. In this environment, sine wave is used as an input of a vibration source and its parameters (amplitude and frequency) are set by Table 3 which is obtained by a practical experiment. The SIMSCAPE model is shown in Fig. 5.

5 Power Recovered by Transducer

The power recovered by the transducer has been analyzed accurately and maximum rms power is obtained by varying the load resistance as shown in Fig. 6. A maximum power 56.72 μw is obtained at a load resistance of 225 Ω. This simulated result plot is shown in Fig. 6.

Table 3 Maximum acceleration at different frequencies

Maximum acceleration (g)	Frequency (Hz)
6.25	55
5.5	27.5

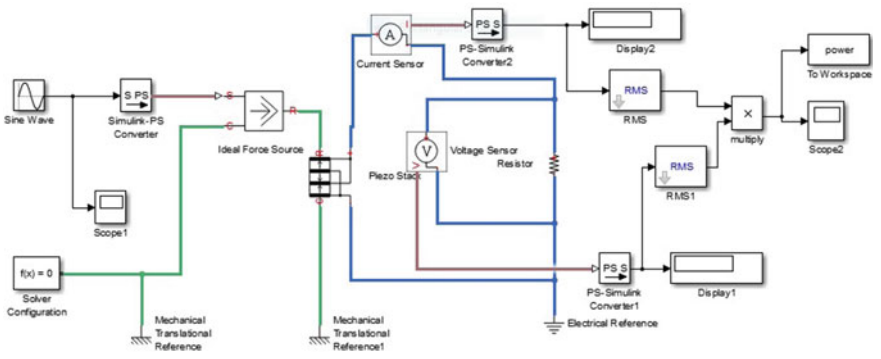


Fig. 5 Piezoelectric transducer SIMSCAPE model

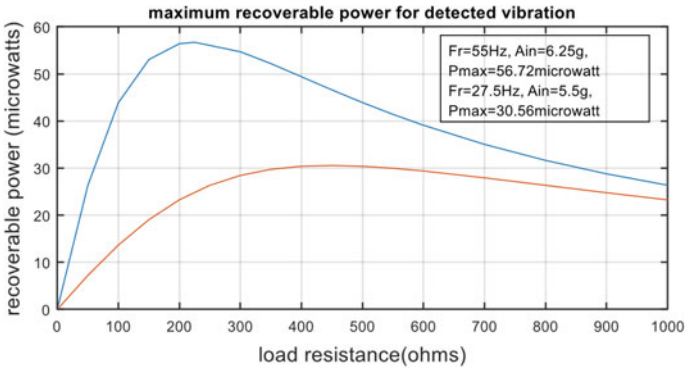


Fig. 6 Maximum scavenged power for detected vibrations

6 Conclusion

Scavenging ambient energy is an ingenious way for generating another power supply solution which is applicable to power vehicle sensor. In the present work, engine-induced vibrations were investigated with the use of piezoelectric technology in which piezoelectric material has been used for converting ambient vibration into electrical energy. Lead zirconate titanite (PZT) material is used due to its real-time applications. The proposed model is designed in MATLAB SIMSCAPE model for simulation purposes and for calculating the maximum power recovered from the real-time vibrations present in a vehicle. This paper estimates the maximum power of 56.72 μw at the maximum acceleration of 6.25 g. This recovered power is used for powering the sensor node to make the vehicle autonomous, to enhance vehicle safety, and to reduce crashes in VANET. Moreover, this will further use power sensors that are used for communication between vehicles, between vehicle and infrastructure and in vehicle safety communication which uses different protocols for efficient communication in Vehicular Ad Hoc Networks, and is useful in powering integrated networks designed for roadside and vehicle sensors to make vehicles more intelligent such that it reduces the risk of human lives which is very precious.

References

1. R. Kaur, T.P. Singh, V. Khajuria, Security issues in vehicular ad-hoc network (VANET), in *IEEE International Conference on Trends in Electronics and Informatics* (ICOEI 2018)
2. Z. Qingyuan, G. Mingjie, H. Yuanqin, Vibration energy harvesting in automobiles to power wireless sensors, in *International Conference on Information and Automation*, Shenyang, 6–8 June 2012, pp. 349–354.
3. A. Mouapi, N. Hakem, G.Y. Delisle, Autonomous wireless sensors network based on piezoelectric energy harvesting. Open J. Antennas Propag. **4**(03), 138–157 (2016)

4. S.H. Mohamad, M.F. Thalas et al., A potential study of piezoelectric energy harvesting in car vibration. ARPN J. Eng. Appl. Sci. **10**(19) (2015)
5. A. Mouapi, N. Hakem, Vibrational-powered vehicle's mesh wireless sensor network: performance evaluation, in *2018 IEEE* (2018)
6. A. Mouapi, N. Hakem, G.Y. Delisle, N. Kandil, A novel piezoelectric micro-generator to power wireless sensors networks in vehicles, in *IEEE International Conference on Environment and Electrical Engineering*, Rome, 10–13 June 2015, pp. 1089–1092.
7. J.G. Smits, W. Choi, The constituent equations of piezoelectric heterogeneous bimorphs. IEEE Trans. Ultrason. Ferroelectr. Freq. Control **38**(3), 256–270 (1991)
8. S. Roundy, K.W. Paul, A piezoelectric vibration based generator for wireless electronics. Smart Mater. Struct. **13**, 11311142 (2004). http://dx.doi.org/10.1088/09641726/13/5/018
9. S.P. Beeby, M.J. Tudor, N.M. White, Energy harvesting vibration sources for microsystems application. Meas. Sci. Technol. **17**, R175–R195 (2006). http://dx.doi.org/10.1088/0957-0233/17/12/r01

Cuckoo Search Algorithm and Ant Lion Optimizer for Optimal Allocation of TCSC and Voltage Stability Constrained Optimal Power Flow



Sheila Mahapatra, Nitin Malik and A. N. Jha

1 Introduction

1.1 Motivation

A deregulated power network operating close to their security limits has imminent effect on transmission price and overall operating cost. The economics of power and the increased load demand have led to concerns regarding secured optimal power flow (OPF). The OPF solution involves incorporating equality and inequality constraints to load flow problem in order to minimize the cost of operation [1, 2]. Planning and real-time operation of the power network is, therefore, critical and is hugely dependent on the secured OPF integrated with voltage stability. The evolution of the competitive power market has resulted in lower network losses and generation costs. The scenario of open market results in power network practically being operated in proximity to their stability margins for minimizing operational cost which intensifies the problem of voltage instability.

The motivation of the present study is to present a novel hybrid computational technique based on Cuckoo search (CS) and Ant Lion optimization (ALO) algorithm for secured OPF coupled with voltage stability as many research works reported in the literature consider these two vital power system operational issues independently which may adversely affect system security when operated close to stability margins. In this article, the hybrid technique has been utilized for optimal sizing and location

S. Mahapatra
Alliance College of Engineering and Design, Alliance University, Bangalore 562106, India

N. Malik (✉)
The NorthCap University, Gurgaon 122017, India
e-mail: nitinmalik77@gmail.com

A. N. Jha
Indian Institute of Technology Delhi, New Delhi 110016, India

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_92

of thyristor-controlled series compensator (TCSC) controller and ensuring consistent voltage profile of the transmission system.

1.2 Literature Review

The inclusion of stability constraints into the conventional OPF model has been presented in the literature [3–5]. Voltage instability has thus evolved as a dominant research area in power system studies after numerous occurrences of voltage instability across the world [6, 7]. In a deregulated power industry where demand for services becomes preeminent, it is imperative to explore the application of flexible a.c. transmission system (FACTS) devices [8]. FACTS controllers have the capability to enhance the reactive power profile in an electrical network and can, therefore, act as a restorative measure for voltage instability and voltage collapse. The installation of FACTS devices has become indispensable as they provide rapid and authentic control over three crucial parameters of AC transmission network—voltage, line impedance, and phase angle. TCSC is a series compensation device which has immense application potential in precisely regulating the transmission line power flow, diminishing sub-synchronous resonance, and damping inter-area power oscillations and improving transient and voltage stability [9]. The TCSC implementation for voltage stability constrained optimal power flow solution and reduction of power loss is carried out in this work.

The increased load demand has resulted in zooming anxiety with respect to the safe functioning of the power network. The contingency scenario necessitates a call for the remedial measure to avoid the voltage and the power limit infringement. The variation in generation schedule triggers line flow variations resulting in overload conditions. Numerous stochastic algorithms are being implemented to provide solutions to complex constrained OPF optimization and hybrid techniques based on population-based algorithms that have been developed and implemented in diverse power system operations. A particle swarm optimization (PSO) inspired method is proposed by Anitha et al. for constrained optimal power flow solution on IEEE 30-node test system [10]. Differential Evolution algorithm has been presented by Basu as a preferred approach for solving OPF by including TCSC and thyristor-Controlled Phase Shifter [11]. The author has worked on DE to implement FACTS devices to obtain OPF solution in IEEE 30-node and IEEE 57-node systems [12]. Voltage stability constrained OPF has been analyzed on IEEE 30-node, IEEE 57-node, and IEEE 118-node systems, respectively, using the PSO technique by Dutta and Sinha [13]. A modified genetic algorithm as a computational method to resolve OPF issue and the effectiveness of the proposed method is tested on an IEEE 30-node system and an IEEE reliability test system-96 which is a 73-node, 120-branch system [14].

1.3 Organization

The research work presented in the sections is described as follows: Sect. 2 provides the detail about the formulation of the problem statement, framing the objective function, and the system constraints. It also gives an overview of TCSC modeling. Section 3 offers a detailed view of the proposed techniques and algorithms applied to the problem followed by the result analysis and discussion in Sect. 6. The conclusion of the research work is described in Sect. 7.

2 Problem Formulation

In this paper, the optimal TCSC placement to attain secured OPF solution integrated with preserving the voltage stability is formulated as a multi-objective optimization problem with fuel cost, power flow index (PFI), voltage deviation, active power loss, and TCSC cost being considered as objectives that are required to be minimized.

2.1 Objective Function

The fuel cost is given by the quadratic function of active power generated as in Eq. (1).

$$\text{Minimize Fuel cost, } F_c = \sum (a_i + b_i P_{\text{gi}} + c_i P_{\text{gi}}^2)\$/\text{hr} \tag{1}$$

where F_c is aggregate fuel cost, P_{gi} is active power generated at the ith bus, and a_i, b_i, and c_i are fuel-cost coefficients and are subject to the following equality and inequality constraints:

(1) Real and reactive power generation and load power balance

$$P_{\text{inj}} = P_{\text{g i}} - P_{\text{Li}} \tag{2}$$

$$Q_{\text{inj}} = Q_{\text{gi}} - Q_{\text{L i}} \tag{3}$$

where P_{inj} and P_{Li} are the injected active power and the active load power at the ith node, respectively.

(2) Lower and upper bounds on the active power generated and on the voltage magnitude at all the nodes as in Eqs. (4) and (5)

$$P_{\text{g,i}}^{\min} \leq P_{\text{g,i}} \leq P_{\text{g,i}}^{\max} \tag{4}$$

$$V_i^{\min} \leq V_i \leq V_i^{\max} \tag{5}$$

where V_i is the voltage magnitude at node i of the power network.

Power Flow Index
A new power flow index (PFI) is proposed to optimally place TCSC:

$$PFI = \left(\frac{S_{\text{opf}}}{S_{\max}} \left(\frac{S_{\text{rL}} - S_{\text{rn}}}{S_{\text{rL}} - S_{\text{opf}}} \right) \right) \tag{6}$$

where S_{rn} and S_{rL} are the apparent power for of the rth line at normal and overloaded conditions, respectively. S_{opf} refers to the power flow during the optimal condition. $S_{\max}$ indicates the maximum value of power in the transmission line at overloaded condition.

Node Voltage Deviation
Voltage deviation (V_D) is minimized so that the stable voltage profile is achieved along with secured OPF.

$$Minimize\, V_D = \left(\sum_{i=1}^{NL} \left| V_i - V_i^j \right| \right) \tag{7}$$

Node voltage is one of the most critical indicators of security and service quality for a power network. Improved voltage profile can be realized in a real-time system by reducing the load node voltage deviations with respect to the desired voltage, which is usually set to 1 p.u.

Cost of TCSC
The genre and site of FACTS controller applied in a power network has a great influence on the objective function and so TCSC cost optimization is viewed as an important objective function to be minimized in the problem formulation. The cost for TCSC controller is given in terms of its operating range/coefficients (MVAr) in Eq. (8) [15].

$$Minimize\, C_{\text{TCSC}} = 0.0015s^2 - 0.7130s + 153.75 \tag{8}$$

2.2 Model of TCSC Device

The TCSC model considered for the study is shown in Fig. 1 [16]. The TCSC is a combination of the successive arrangement of capacitors placed parallel to the thyristor-controlled reactor. In this model, the reactance of transmission line is made controllable by integrating TCSC in series with the line such that impedance value is

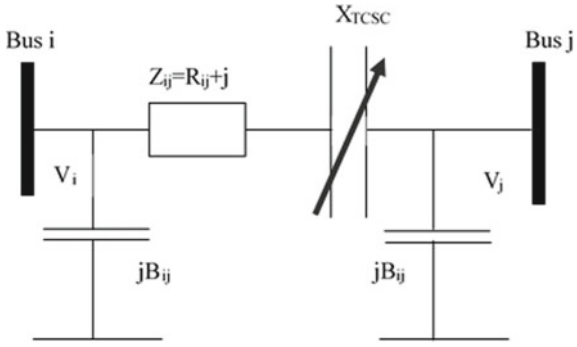


Fig. 1 Static model of TCSC [16]

adapted to provide suitable series compensation. The TCSC provides series compensation resulting in modified bus admittance matrix between ith and jth bus. The new value of net reactance X_{ijnew} with TCSC and degree of series compensation acquired by installing TCSC is given in Eqs. (9) and (10), respectively.

$$X_{\text{ijnew}} = (1 - k)\text{X}_{\text{ij}} \tag{9}$$

$$\text{where } k = \frac{X_{\text{TCSC}}}{X_{\text{ij}}} \tag{10}$$

Here, k represents the degree of series compensation and X_{TCSC} and X_{ij} are the reactance of TCSC and the line, respectively. The injection capacity of TCSC and its location will modify the bus admittance matrix and is updated for secured OPF under overloaded conditions. As a result, the apparent power flowing through the lines gets modified depending on the loading conditions.

The bounds on TCSC reactance is given in Eq. (11). A compensation value ranging up to 70% of line reactance is typically selected to overcome the issue of series resonance in the system and it also curtails overcompensation.

$$-0.7X_{\text{ij}} \leq X_{\text{TCSC}} \leq 0.2X_{\text{ij}} \tag{11}$$

3 Cuckoo Search Algorithm

Cuckoo search algorithm [17] is a metaheuristic algorithm which was first developed by Yang and Deb in 2009. The algorithm derives its inspiration from the strange behavior of cuckoo to lay their eggs in the nest of other birds. The steps involved in cuckoo search algorithm are [18]

i. Cuckoo bird would lay only one egg at a randomly selected nest where each egg in the nest would bear resemblance to a solution.

ii. The nest which holds the best quality eggs will survive to give rise to the next generation.
iii. The search process is designed with the number of nests being predetermined and the host bird operating with a probability of $P = [0, 1]$ to discover the alien egg and throw it or abandon the nest to build a new one.
iv. Fitness is evaluated and inferior nests are abandoned based on probability Pa and a new solution is generated which leads to the establishment of a new nest. In the final phase, greedy search algorithm is applied to retain the best solution.

3.1 Utilization of Levy Flight

A levy flight is considered which enhances the local and global searching in CS algorithm as given by Eq. (12)

$$\delta_i^{t+1} = \delta_i^t + \alpha \oplus Levy(\lambda) \tag{12}$$

The terms in δ_i^{t+1} represent a new location, δ_i^t represents the current location, and Levy (λ) is the transition probability. The notation ⊕ represents entry-wise multiplications. A levy flight is performed where the randomly selected egg moves to a new location if the objective function has an improved value as random search pattern is initiated. This enables the CS algorithm to simultaneously find all optima in the search space thus elevating its search performance. The flowchart for the CS algorithm is given in Fig. 2.

4 Ant Lion Optimization

The Ant Lion Optimizer (ALO) is a nature-inspired, metaheuristic algorithm which imitates the hunting pattern of ant lion larvae and has been proposed by Mirjalili [19]. It is a stochastic optimization technique for attaining global optima in a search space. It has the advantage of circumventing local optima and has the benefit of problem independency. The ant lions are predators who dig traps and wait for the prey or ants to slide into the traps where the trap size is directly related to the hunger level and size of moon which it has developed instinctively [20]. The ALO algorithm models and captures the interactivity between the ant lion and ants in the search space. The algorithmic steps for ALO are mentioned below.

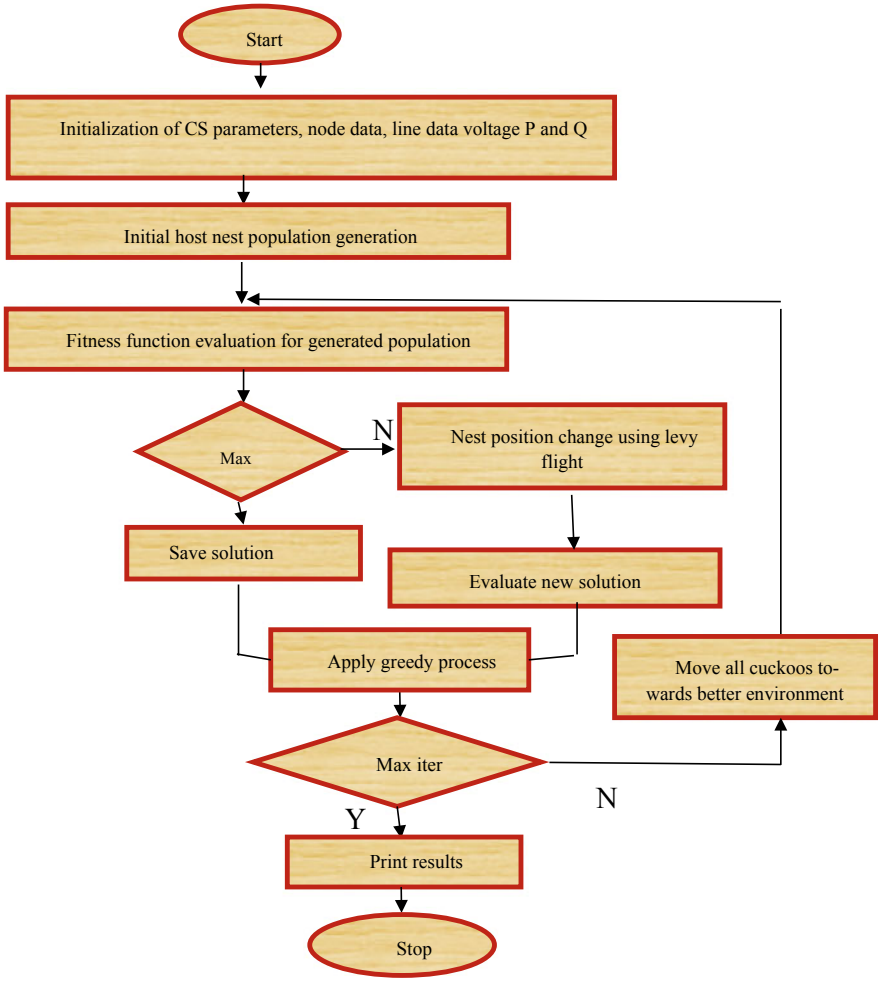


Fig. 2 Flowchart of Cuckoo search algorithm

4.1 Initialization—Random Movement of Ants

The stochastic movement of ant is a random walk and ant lions hunt them by setting traps and evolve as fitter ones. The random positions of the ants are saved in the matrix Y_{ant} which resembles the particles in the search space.

$$Y_{\text{ant}} = \begin{pmatrix} Y_{1,1} & Y_{1,2} & \dots & Y_{1,n} \\ Y_{2,1} & Y_{2,2} & \dots & Y_{2,d} \\ \vdots & \vdots & \vdots & \vdots \\ Y_{n,1} & Y_{n,2} & \dots & Y_{n,d} \end{pmatrix} \tag{13}$$

where $Y_{\mathrm{n,d}}$ is the value of the nth variable (dimension) of dth ant; n is the population size of ants. Cumulative sum of random movement of ants is given in Eq. (14).

$$Y(t) = [0, sum(2r(t_1) - 1), sum(2r(t_2) - 1), \ldots, sum(2r(t_n) - 1)] \tag{14}$$

where t is random walk step/iteration number.

4.2 Fitness Evaluation and Updating Process

The evaluation of the fitness function of the ants is done during optimization and the fitness value is stored in matrix N_B in terms of objective function f. The position of ant lion in search space is stored in matrix N_{BP}. The matrix N_{BL} provides the fitness function of ant lion. To restrain randomness of the ant within the search space, min-max normalization is implemented and position of the ants is updated as given in Eq. (18).

$$N_B = \begin{pmatrix} f\left(\left|B_{1,1},\ B_{1,2}, \ldots, B_{1,d}\right|\right) \\ f\left(\left|B_{2,1},\ B_{2,2}, \ldots, B_{21,d}\right|\right) \\ \vdots \\ f\left(\left|B_{n,1},\ B_{n,2}, \ldots, B_{n,d}\right|\right) \end{pmatrix} \tag{15}$$

$$N_{\mathrm{BP}} = \begin{pmatrix} BL_{1,1} & BL_{1,2} & \ldots & BL_{1,n} \\ BL_{2,1} & BL_{2,2} & \ldots & BL_{2,d} \\ \vdots & \vdots & \vdots & \vdots \\ BL_{n,1} & BL_{n,2} & \ldots & BL_{n,d} \end{pmatrix} \tag{16}$$

$$N_{\mathrm{BL}} = \begin{pmatrix} f\left(\left|BL_{1,1},\ BL_{1,2}, \ldots, BL_{1,d}\right|\right) \\ f\left(\left|BL_{2,1},\ BL_{2,2}, \ldots, BL_{2,d}\right|\right) \\ \vdots \\ f\left(\left|BL_{n,1},\ BL_{n,2}, \ldots, BL_{n1,d}\right|\right) \end{pmatrix} \tag{17}$$

$$Y_m^t = \frac{\left(Y_m^t - a_m\right)\left(d_m - c_m^t\right)}{\left(d_m^t - a_m\right)} + c_t \tag{18}$$

where a_m and d_m are minimum and maximum of random walk of ants, respectively. c_m^t and d_m^t represent minimum and maximum mth variable at tth iteration.

4.3 Trap Building by Ant Lion

The construction of pits for trapping of ants by ant lion is modeled mathematically as

$$c_m^t = Ant - lion_n^t + c^t \tag{19}$$

$$d_m^t = Ant - lion_n^t + d^t \tag{20}$$

The ants slide into the pits and the fittest ant lion will be selected among the rest by roulette wheel method. The updating of position is done once an ant falls into the trap.

4.4 Prey Catching and Process of Trap Reconstruction

The ant is trapped and caught by an ant lion when it descends into the trap and goes to the bottom of the pit. It has to relocate to a new position for catching a new prey and its position is updated. To obtain the best solution in each iteration, elitism is applied in every stage [21].

5 Solution Methodology

The parameters such as fuel cost, voltage deviation, power loss, and the cost function of TCSC are minimized to boost the dynamic stability of the electrical network when the power line is subjected to overloading conditions. The power flow under base-case and overloaded condition are determined using the NR algorithm. The ALO algorithm is further applied to optimize the objective function and relieve the system during overloading thus retaining network security by improving the voltage profile. The ALO algorithm ensures voltage stability with secured OPF solution. The proposed methodology is implemented on IEEE 30-node network.

5.1 Determination of Optimal TCSC Location

The CS algorithm is applied to the current problem implemented to determine the line with maximum PFI value which dictates the TCSC controller optimal location in the test system and implements the optimization to generate the best solution based on the reduction of the fuel cost. The computational steps for determining the ideal TCSC location are as follows:

Step 1: Initialization of node data, line data, voltage, and real and reactive power.
Step 2: Setting the limiting condition for control variables such as voltage and reactive power, number of TCSC devices, and the number of iterations and search agents.
Step 3: Generating population matrix using a D-dimensional parameter vector which is defined by the maximum and minimum limits as follows:

$$\delta_{\min} = \left\{\delta_1^{\min}, \delta_2^{\min}, \ldots . \delta_D^{\min}\right\} \tag{21}$$

$$\delta_{\max} = \left\{\delta_1^{\max}, \delta_2^{\max}, \ldots \ldots . \delta_D^{\max}\right\} \tag{22}$$

Here, $\delta_{\min}$ and $\delta_{\max}$ represent the limits on objective function applied in the problem such as the equality and inequality constraints applied to the system.
Step 4: The inequality constraints are checked for with respect to search agent's position in population matrix. Proceed to Step 5 if the condition is satisfied. Else, go to Step 3 to regenerate the population matrix.
Step 5: The objective function is evaluated by performing load flow analysis by N–R method for base-case uncompensated system.
Step 6: Line data and node data are updated with the new population of search agents.
Step 7: The Y-bus matrix is modified due to the series compensation introduced by TCSC installation in the test system.
Step 8: The overloading on transmission lines is analyzed.
Step 9: The PFI values of all the lines are generated and the line corresponding to maximum PFI is identified as the weakest node and considered as the optimal location for TCSC.
Step 10: If a limit violation exists, then the cycle is repeated by increasing the TCSC compensation until the best solution is attained or maximum iterations are reached.

5.2 Determination of Optimal TCSC Size

The ALO algorithm is applied to the current problem by evaluating fitness function of the ant lions in Eq. (19) and it signifies the parameters required to be minimized to achieve multi-objective optimization of the secured optimal power flow problem while preserving voltage stability at the network nodes. The algorithm is further applied to optimize the objective function and relieve the system during overloading, thus retaining network security by improving the voltage profile. The optimization is terminated when maximum iteration is reached after attaining the best solution by minimizing the objective function based on the reduction of the fuel cost with the highest PFI value and minimizing power loss, voltage deviation at nodes, and the cost of TCSC. The ALO algorithm ensures voltage stability with a secured OPF solution.

$$F_i = \min \begin{pmatrix} PFI \\ V_D \\ P_L \\ C_{\text{TCSC}} \end{pmatrix} \tag{23}$$

6 Numerical Results and Discussion

This section describes the improvement in voltage profile and secured an optimal power flow solution for IEEE 30-node transmission network which consists of 41 branches, six generators, four transformers, and 21 load buses. The line, load, and generator data have been taken from [22, 23]. The fuel cost coefficients and the generator limits for the test system are referred from [24, 25] and are tabulated in Table 2.

The security assessment is evaluated by increasing the demand at the load buses. The system performance for secured power flow is determined and the results are verified by comparing them with the other techniques already existing in the literature. The proposed techniques are simulated using MATLAB to validate their performance. The computational methods employed are subjected to effectively maintain the voltage levels at load buses within the maximum and minimum limit of 1.06 and 0.98 p.u., respectively, and slack bus voltage of 1.06 p.u.

The combined application of CS–ALO algorithms is implemented in two phases to render optimal solutions for the power network. The initial power flow conditions of voltage, real and reactive power flow, and the power losses are solved for base-case load of uncompensated system by using the N–R method. Then, the introduction of fault at the system node renders instability to the system. The CS algorithm generates PFI values of all the lines and the line corresponding to maximum PFI is identified as the weakest node subject to limit violation constraint for ensuring secured OPF solution. This identifies the TCSC location.

The optimal rating of TCSC is determined using the ALO algorithm and TCSC cost is optimized by using Eq. (8). The power flow security of the network is analyzed by the algorithm during the line outage condition which reduces the load power limits. The restoration is eventually achieved by TCSC optimal location thus rebuilding secured network flow and holding on to stable voltage profile. Table 1 provides details about the various parameters required for both the CS–ALO algorithms when applied to the current problem (Table 2).

The metaheuristic optimization tool of CS–ALO is applied to provide the OPF solution and voltage profile on the IEEE 30-node test system. The overloading condition of load buses is instrumental to obtain the security assessment and optimized location of TCSC by the proposed approach is found to be on transmission lines 19, 26, and 30. The TCSC placement in these optimal locations enhances the power flow in the congested lines thus restoring secured operation with the dual benefit

Table 1 Implementation parameters for CS and ALO algorithms

Algorithm	Description	Values
CS	Maximum iterations	100
	Rate of finding solutions	3/2
	Highest (ub) and Lowest bound (lb)	1, 0.1
	Number of Nests	20
ALO	Highest (ub) and Lowest bound (lb)	100, −100
	Number of population size (N)	10
	Dimension	1

Table 2 Active power limits and fuel cost constants of IEEE 30-node system

Generator bus no.	P_{min} (MW)	P_{max} (MW)	a_i	b_i	c_i
1	50	200	0	2	0.0038
2	20	80	0	1.75	0.0175
5	15	50	0	1	0.0625
8	10	35	0	3.25	0.0083
11	10	30	0	3	0.0250
13	12	40	0	3	0.0250

of retaining voltage levels at the load buses well within the prescribed limits. The results for TCSC evaluation for its optimized position on each of the transmission lines are provided in Table 3.

The optimal TCSC placement between node 10 and node 17 yields the best results with respect to voltage deviation, power loss, and TCSC cost resulting in the enhancement of system security margin. The active power generations of all the six generators of the IEEE 30-node test system is tabulated below.

Table 4 depicts that sufficient active power is generated by optimal TCSC placement so as to relieve the system from overloading. The results are also verified with other hybrid techniques of Fuzzy-gravitational search algorithm (GSA), Radial basis function neural network (RBFNN)-GSA and Improved GSA–Firefly (FA) algorithm [26] and is presented in Table 5.

Table 3 TCSC allocation based on PFI in IEEE 30-node network

Branch no.	From node	To node	TCSC rating (MVAr)	Reactance pu	Voltage deviation pu	Power loss (MW)	TCSC Cost ($/MVAr)
19	12	16	0.174821	0.0845	0.0473	6.3822	149.9435
26	10	17	0.184650	0.0845	0.0023	2.3822	130.3561
30	15	23	0.175948	−0.202	0.2113	10.5427	148.625

Table 4 Active power generation with and without TCSC in IEEE-30 node network

Generator	Base-case active power (MW)	Active power after the introduction of TCSC (MW)			
		Fuzzy-GSA [26]	RBFNN-GSA [26]	IGSA-FA [26]	CS–ALO
G_1	111.5234	111.523	142.102	130.019	127.734
G_2	58.41801	58.418	43.5944	75.0611	56.2042
G_6	32.88061	32.8806	30.3378	19.0955	33.6475
G_{13}	25.46275	25.4628	17.913	21.096	20.9767
G_{22}	22.9112	22.9112	15.921	21.096	20.9767
G_{27}	34.62022	34.6202	33.6361	17.6868	22.8917

Table 5 Comparison of results with other methods

Parameters	Base-case loading (uncompensated)	Fuzzy-GSA [26]	RBFNN-GSA [26]	IGSA-FA [26]	CS–ALO
Active power generation (MW)	295.70	285.8162	283.5041	284.0543	282.4307
Power loss (MW)	12.004	7.5797	3.7917	2.6351	2.3822
Fuel cost ($/hr)	828.3393	809.2	800.317	795.675	764.2472
TCSC cost ($/MVAr)		148.2325	142.2168	138.4178	130.3561
Voltage deviation (p.u.)		0.5634	0.1241	0.0684	0.0473

To validate the effectiveness of the proposed approach in OPF solution using the TCSC controller, the comparison is done with other existing methods for a compensated system and the results for the real power loss and generator fuel cost are obtained and compared to [26] in Table 5. The result analysis from Table 5 candidly projects that the CS–ALO algorithm has generated superlative results with respect to active power loss reduction, fuel cost, TCSC cost, and voltage deviation. The generated power under base-case condition is 285.8162 MW which is raised to 295.70 MW for loaded and uncompensated systems. With the optimal TCSC placement, the real power losses are reduced thus propelling improved generator power profile and total generation is in close proximity to the base-case value with the implementation of the CS–ALO algorithm. This is also evident from the minimum real power losses attained by the proposed method. The total reduction in fuel cost and TCSC cost using CS–ALO algorithm compared to IGSA-FA is 3.95% and 5.824%, respectively. The OPF result for TCSC compensated system is also compared with other stochastic algorithms as reported in the literature. The proposed technique implementation yields a fuel cost reduction by 7.73% and real power loss is reduced by 1.1103 MW

compared to GA [27]. A similar comparison with [28] for TCSC compensated IEEE 30-bus system shows a reduction of generator fuel cost by 4.75% and real power loss reduced by 7.3656 MW.

Table 6 presents the statistical analysis of the cases when TCSC is optimally placed in the 19th, 26th, and 30th transmission lines, respectively. Figure 3 depicts the power loss across the transmission lines for the test system. The power loss is increased under overloaded condition which is substantially reduced by the optimal TCSC placement with CS–ALO implementation. Figures 4 and 5 give a comparative account of all the techniques for voltage profile at node and power loss reduction.

Figure 6 shows the comparison of convergence of various algorithms and establishes that CS–ALO algorithm can be successfully implemented for multi-objective optimization problem.

Table 6 Statistical analysis for voltage deviation, power loss, and TCSC cost

TCSC location	Parameters	Voltage deviation (pu)	Power loss (MW)	TCSC cost ($/MVAr)
19	Mean	0.1627	8.0464	155.1888
	Std. Dev.	0.0739	0.8674	1.4887
	Min	0.0473	6.3822	149.9435
	Max	0.2983	9.4852	155.1888
26	Mean	0.2482	4.3782	135.2096
	Std. Dev.	0.0773	1.1623	2.8338
	Min	0	2.3822	130.3561
	Max	0.2482	6.4596	139.9977
30	Mean	0.357	11.7027	151.5222
	Std. Dev.	0.0808	0.6656	1.9682
	Min	0.2113	10.5427	148.625
	Max	0.4961	12.7885	154.8982

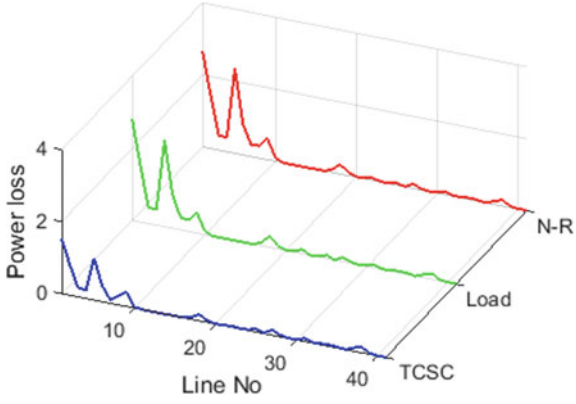


Fig. 3 Power loss for base-case, overloaded, and TCSC compensated condition

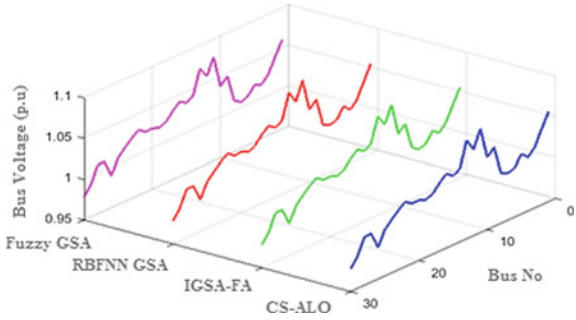


Fig. 4 Comparison analysis for voltage profile using different algorithms

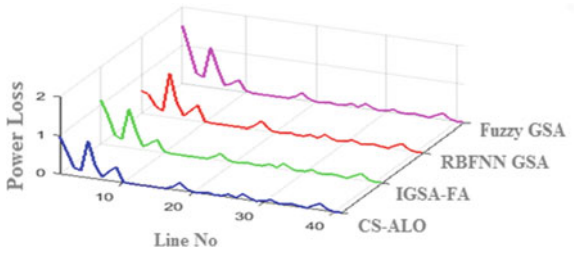


Fig. 5 Comparative analysis of power loss with different algorithms

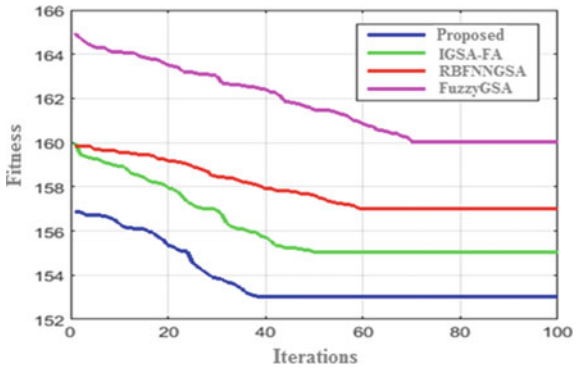


Fig. 6 Comparative analysis of the convergence of various algorithms for IEEE 30-node system

7 Conclusion

The research work furnished in this article focuses on determining a voltage stability constrained optimal power flow solution even under severely loaded environment and the selected algorithm of CS–ALO generates superior outcomes with TCSC optimal placement thus providing reduced cost of generation, less cost for TCSC, minimized power losses, and reduced node voltage deviation. The methodology has been implemented and tested on the standard IEEE benchmark 30-node test system. The power flow security is suitably analyzed, and the proposed methodology has been compared with other techniques such as GA, Firefly, GSA, and other algorithms as

already published in the literature. The statistical analysis signifies the robustness of the proposed methodology. The outage of transmission line directly impacts the reduction in load power limits which are restored eventually by optimized TCSC setting to re-establish secured power flow of transmission lines and retain stable voltage profile.

References

1. M. Huneault, F.D. Galiana, A survey of the optimal power flow literature. IEEE Trans. Power Syst. **6**(2), 762–770 (1991)
2. J.A. Momoh, M.E. EL-Hawary, R. Adapa, A review of selected optimal power flow literature to 1993, Part II: Newton, linear programming and interior point methods. IEEE Trans. Power Syst. **14**(1), 104–111(1999)
3. F. Dong, B.H. Chowdhury, M. Crow, L. Acar, Cause and effects of voltage collapse-case studies with dynamic simulations. IEEE Power Eng. Soc. Gen. Meet. **2**, 1806–1812 (2004)
4. J. Carpentier, Toward a secure and optimal automatic operation of power systems, in *IEEE PICA Conference Proceedings* (Montreal, Canada, 1987), pp. 2–37
5. I.A. Momoh, R.J. Koessler et al., Challenges to optimal power flow. IEEE Trans. Power Syst. **12**(1), 444–455 (1997)
6. T.V. Custem, C.D. Vournas, *Voltage Stability of the Electric Power Systems* (Kluwer Academic, Norwell, 1998)
7. A. Berizzi, The Italian 2003 blackout, in *IEEE Power Engineering Society General Meeting* (Denver, CO, 2004), pp 1673–1679
8. Y.H. Song, X.F. Wang, *Operation of Market Oriented Power System*. Springer (2003). ISBN: 1-85233-670-6
9. F.D. Galiana, K. Almeida, M. Toussaint, J. Griffin, D. Atanackovic, Assessment and control of the impact of FACTS devices on power system performance. IEEE Trans Power Syst **11**(4), 1931–1936 (1996)
10. M. Anitha, S. Subramanian, R. Gnanadass, FDR PSO-based transient stability constrained optimal power flow solution for deregulated power industry. Electr. Power Compon. Syst. **5**(11), 1219–1232 (2007)
11. M. Basu, Optimal power flow with FACTS devices using differential evolution. Electr. Power Energy Syst. **3**(2), 150–156 (2008)
12. M. Basu, Multi-objective optimal power flow with FACTS devices. Energy Convers. Manage. **52**(2), 903–910 (2011)
13. P. Dutta, A.K. Sinha, Voltage stability constrained multi-objective optimal power flow using particle swarm optimization, in *First International conference on Industrial and Information Systems (ICIISs)*, Sri Lanka, 8–11 Aug 2006, pp. 161–166
14. A.G. Bakirtzis, P.N. Biskas, C.E. Zoumas, V. Petridis, Optimal power flow by enhanced genetic algorithm. IEEE Trans. Power Syst. **17**(2), 229–230 (2002)
15. K. Habur, D. Oleary, FACTS—Flexible AC transmission systems, for cost effective and reliable transmission of electrical energy (2008). http://www.siemenstd.com/
16. S.N. Singh, A.K. David, A new approach for placement of FACTS devices in open power markets. IEEE Power Eng. Rev. **21**(9), 5–7 (2001)
17. P. Sekhar, S. Mohanty, An enhanced cuckoo search algorithm based contingency constrained economic load dispatch for security enhancement. Electr. Power Energy Syst. **75**, 303–310 (2016)
18. S.M. Abd-Elazim, E.S. Ali, Optimal location of STATCOM in multi machine power system for increasing loadability by Cuckoo Search algorithm. Electr. Power Energy Syst. **80**, 240–251 (2016)

19. S. Mirjalili, The ant lion optimizer. Adv. Eng. Softw. **83**, 80–98 (2015)
20. M. Raju, L.C. Saikia, N. Sinha, Automatic generation control of a multi-area system using ant lion optimizer algorithm based PID plus second order derivative control-ler. Int. J. Electr. Power Energy Syst. **80**, 52–63 (2016)
21. H.M. Dubey, M. Pandit, B.K. Panigrahi, Ant lion optimization for short-term wind inte-grated hydrothermal power generation scheduling. Int. J. Electr. Power Energy Syst. **83**, 158–174 (2016)
22. S. Hadi, *Power System Analysis* (Tata McGraw-Hill, New Delhi, 1999)
23. O. Alsac, B. Scott, Optimal load flow with steady-state security. IEEE Trans. Power Appar. Syst. **93**(3), 745–751 (1974)
24. E.I. De Oliveira, J.M. De Lima, K.C. Almeida, Allocation of FACTS devices in hydro-thermal systems. IEEE Trans. Power Syst. **15**, 276–282 (2000)
25. IEEE-30 bus transmission system data is refereed from: http://www.ee.washington.edu/research/pstca/pf30/ieee30cdf.txt
26. S. Mahapatra, N. Malik, A.N. Jha, B.K. Panigrahi, Voltage Stability Enhancement by IGSA-FA hybrid technique implementation for optimal location of TCSC. J. Eng. Sci. Technol. **12**(9), 2360–2373 (2017)
27. G.V. Lakshmi, K. Amaresh, Optimal power flow with TCSC using genetic algorithm, in *2012 IEEE International Conference on Power Electronics, Drives and Energy Systems (PEDES)*, pp. 1–6 (2012)
28. N.P. Padhy, M.A. Abdel Moamen, A generalized Newton's optimal power flow modelling with facts devices. Int. J. Model. Simul. **28**(3), 229–238 (2008)

A New Approach to Smart Knock Detection Based Security System for Door Lock



Amit Chaurasia, Umesh Kumar Dwivedi, Amita Chaurasia and Shubham Kumar Jain

1 Introduction

What is smart? Smart devices are those that can make their own decisions. Traditionally, we have a key and lock type door lock system: the mechanical bolt and the knob to lock and a metal key to unlock the cylinder which can be easily found anywhere [1]. This key and lock type door lock system is the most common door lock system currently used in the world. We face many problems with this key and lock type door lock system say if multiple persons want to unlock the door at different times, it might cause inconvenience to them [2]. Similarly, if the keys are lost, then the door cannot be opened and if it is found by someone else, then it will impair your security. So in this case, we have only two options whether we have to call the locksmith to replace the lock or buy a new one. Also, it is very easy to make a duplicate key of the respective lock. So there are various locking systems available today; choosing one out of them for your application is an important aspect. Let us consider an example: you are the CEO of one of the companies and you want to open a lock of a room in which all the important files are there and you are the only person who can open it. Today, there are door locks such as biometrics, fingerprint scanner, face recognition or numerical codes to open or close the door, but they are highly expensive and need

A. Chaurasia (✉) · U. K. Dwivedi · A. Chaurasia
Amity University, Jaipur, Rajasthan, India
e-mail: achaurasia2583@gmail.com

U. K. Dwivedi
e-mail: umeshkudwivedi@gmail.com

A. Chaurasia
e-mail: amita1610@gmail.com

S. K. Jain
ASET, Amity University, Jaipur, Rajasthan, India
e-mail: skjshubhamkumarjain@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_93

high maintenance costs. Keeping in mind all these difficulties and mishaps, we make a very simple system to unlock or lock a door. In this project, we came up with the idea of developing an electronic door lock system not costly enough but more importantly, it serves the purpose without compromising the security. The solution that we came up with is a smart knock detecting door lock. In this, a specific pattern can be assigned where it can even be music [3]. If anybody wants to open the door at their convenience, he or she should knows the pattern with certain accuracy. The main principle is sensing specific knocks and performing the required action. This is not a detection-based security system for door lock which is made by using Arduino and its IDE (integrated development environment) software. This smart lock can replace the traditionally used key-lock system open any lock without a key. As a potential to make our life, it is more convenient and fast, and thus increases our standard of living [4].

2 Implementation

The project is made using a custom Arduino board and to program the microcontroller on the board, we have used Arduino's IDE software which is free software. This project is tested on a miniature door replica. The device is mounted on top of the existing door lock on the inside of the replica. The custom board is tested on a breadboard and then finalized on a DOT PCB for ease of installation. A piezo sensor is mounted on the same side of the door to detect the vibrations created when someone knocks the door from outside. A momentary push button is provided to record the new knock and save it to our board. One needs to push the button and knock the pattern in a 500 ms gap as we have set this much delay between two knocks in our code. This delay determines the input as 0 or 1. While we knock, the time at each knock is observed by Arduino and put into an array. Assuming we knocked 5 times, so we have four time periods. Arduino then checks the time period between the first and the second knock and if the difference is less than 500 ms, then it will be 0, else it will be 1 and this will be saved in a variable. Now, it checks the time period between subsequent knocks and finally we get our output as 0 and 1. A servo motor is mounted to lock and unlock the door if open signal is received from the microcontroller. LEDs are used to show the lock status to the user and also to indicate the recording mode on/off status and this whole package is mounted inside a box. The code also called sketch for the system was written in an Arduino integrated development environment. The code includes a pattern-detecting algorithm to unlock the door lock. The hardware detects knocks, record knocks and the software compares knock sequences, detects the state of the system, determines whether the door should be unlocked or locked, and sends signal to unlock and lock the door [5]. Many input/output operations become much easier because the Arduino programming environment comes with a software library called "Wiring" from the original wiring project. Arduino is an open-source single-board microcontroller intended to make interactive objects. It is a tool for making computers that can sense and control more of the physical world as compared to our

desktop computers. The system requires 9 V power supply and can be used with a multi-bolt locking provided by motorized lock installed inside a metal door as seen in the main entrance door, fire doors, emergency doors, etc., in public buildings, offices, and institutes. A relay will be required to control the motorized lock. GSM module is used to notify the owner in the event of multiple mismatches of knocking sequence thus indicating forced entry.

3 Hardware and Software Requirements

3.1 Hardware

Arduino: It is an open-source prototyping platform based on easy-to-use hardware and software. It has the capability to read inputs in the forms of a finger on a button, a sensor input or a message. This input turns into an output of a motor, turning on an LED, publishing the outputs, etc. By sending a set of instructions to the microcontroller on the board, you can tell your board what to do. The reason to choose Arduino is its simplicity to learn and interface different electronic components or modules and they are not expensive. Taking advantage of Arduino being open-source, we have decided to build our own Arduino in an attempt to further reduce the cost of smart door locks. The objective is to find the fine line between cost cutting and reliability without compromising security.

2N2222 bipolar transistor: In this circuit, the NPN bipolar transistor is used. The biasing of this transistor is reverse biasing. And therefore, when the base pin is grounded and forward biased when a signal is given to the base pin. Transistor 2N2222 has a gain of 110–800; this value gives the amplification capacity of the transistor and is beneficial for amplification and switching purposes. The device is designed for a medium range of power applications and is housed in the TO-92 package.

1N4007 rectifier diode: According to the definition, the diode allows the current flow through one direction from the anode to cathode. The cathode terminal can be identified by using a gray bar on the diode. For the 1N4007 diode, the maximum current carrying capacity is 1 A and can peak up to 30 A. Hence, we can use this in circuits that are designed for less than 1 A. The negligible reverse current is flow, i.e., 5 μA. The power dissipation of this diode is 3 W. Here, we are using this as a protection device as it can be used to prevent reverse polarity problem.

Piezoelectric sensor: The piezoelectric sensor consists of a diaphragm that is made up of polymer polyvinylidene difluoride (PVDF) and works as a sensing and transducing element. The pressure is directly converted into a voltage upon the application of pressure to the diaphragm.

LM 7805 voltage regulator IC: It is a member of 78xx series of the fixed voltage regulators ICs that maintain the output voltage at a constant value where xx indicates

the output voltage it supplies. LM 7805 IC is used to prevent heat and provide regulated power supply of 5 V.

ATmega328 microcontroller: It is a high-performance microchip 28-pin dual in-line package Alf and Vegard's RISC based 8-bit microcontroller and has 32 kb of In-System Programming (ISP) flash with read while write capabilities. Its operating voltage is 1.8–5.5 V. It is an advanced microcontroller as it can control instructions in one clock cycle approaching 1 MIPS (million instructions per second) per MHz, thus controlling power consumption and processing speed [6].

Resistor: A resistor is an important and passive electrical component that provides resistance in a circuit. The task of the resistors is to reduce the current flow and to provide low voltage levels within the loop.

LED: Basically, it is a normal p–n diode, which emits the light when activated, and a semiconductor light source. Electrons are able to recombine with holes when a suitable voltage is applied. It releases energy in the form of photons which is also known as the electroluminescence effect. The bandgap of energy in the semiconductor determines the color of the light.

SG90 digital servo motor: A servo motor encompasses a regular motor along with a sensor for position feedback which in turn enables a motor to precisely control the position, velocity, and acceleration, unlike the dc motor which turns continuously when power is supplied. They are quite useful as you can instruct them how far to move. The servo motor has three wires—ground, power, and a third wire through which those commands are carried out. SG90 micro-servo's operating voltage is 4.8 V and has a stall torque of 1.8 kg/cm.

Capacitor: It is a two-terminal passive electronic component generally used to locally store energy, filter out voltage spikes, and complex signal filtering.

GSM module: Global system for mobile communication (GSM) is a mobile communication device which can be either a mobile phone or a modem device and can be used to make a computer or any other processor communicate over a network. We will be using SIM300 which is a tri-band GSM modem that can operate in 900, 1800, and 1900 MHz bands with operating voltage between 3.4 and 4.5 V. It provides a serial TTL interface for direct and easy interface to microcontrollers and features low power consumption of 0.25 A on normal operations and around 1 A during transmission.

3.2 Software

The instructions are written in the Arduino programming language (based on Wiring) on the Arduino software (IDE) and uploaded to the microcontroller using the cable provided. Even its software is easy to use for beginners, yet flexible enough for advanced users. It is also a free, open-source and cross-platform. We have also used Fritzing to prototype the system and Eagle to design the circuit layout.

4 Circuit Diagram and Layout

For this project, the abovementioned dataflow diagram gives the flow of information in the system and makes it easy to understand Fig. (3).

The basic knowledge of microcontrollers and soldering is required to make the custom Arduino. First, the circuit is made on a breadboard which is also useful for testing the system before finalizing it on DOT PCB (Figs. 1, 2, and 4).

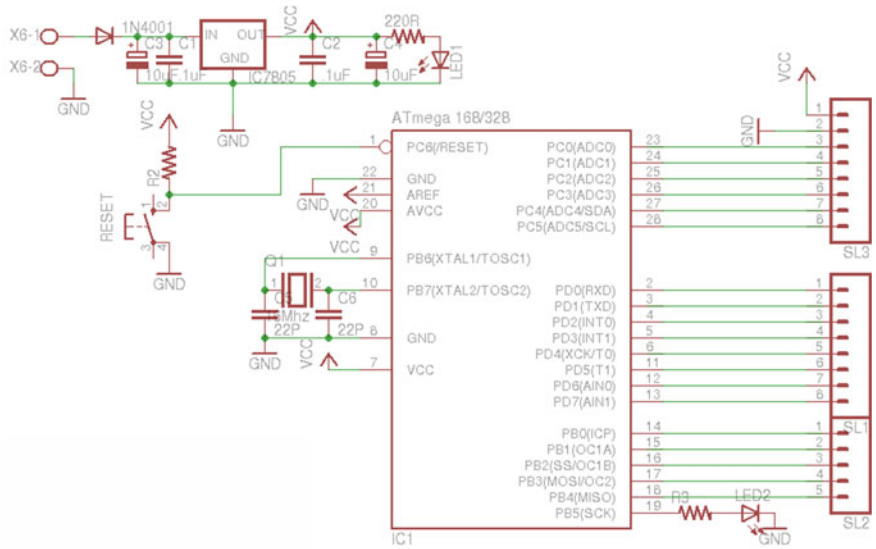


Fig. 1 Circuit diagram for custom Arduino board

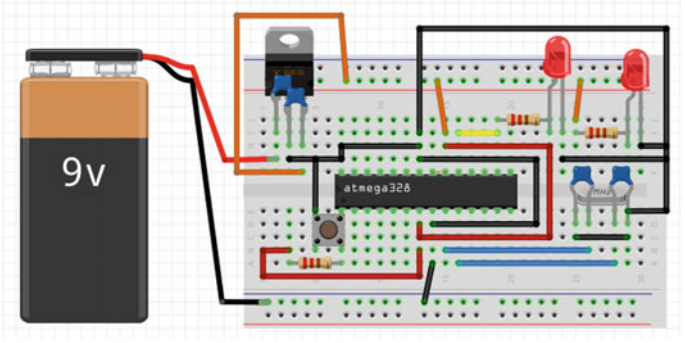


Fig. 2 Breadboard connections of the custom Arduino

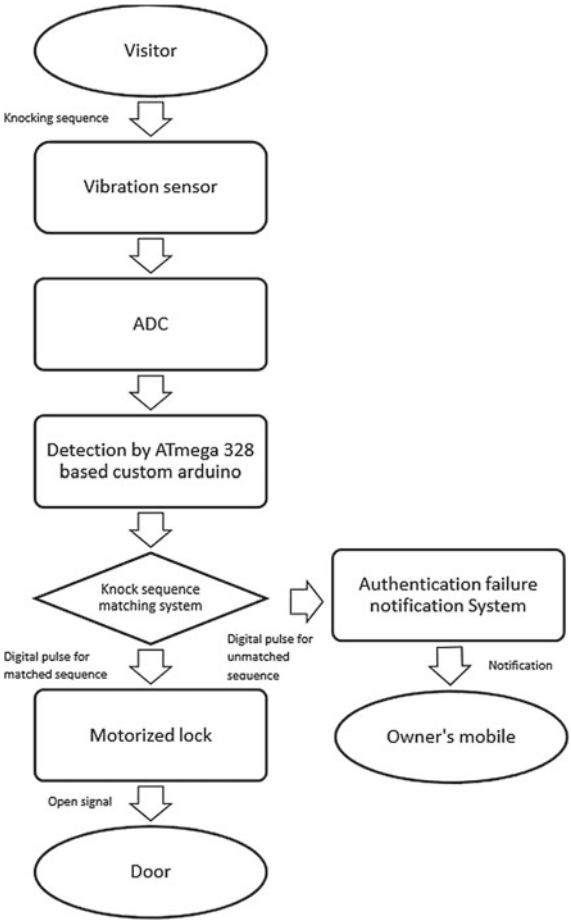


Fig. 3 Data flow diagram

5 Conclusion

Technology has made its way into stepping up security for our homes with nifty gadgets—one example is smart door locks. In a nutshell, smart door locks are small devices that either stick onto or replace your normal deadbolt door lock. Smart door locks are based on existing wireless technologies, including Wi-Fi, cellular, and bluetooth. These enhance the overall security; let us say, of your home by having smart and convenient features such as remote locking and forced entry detection, which then lets you know if any door in your home happens to gets compromised. Thanks to the existing technologies, many households are already taking advantage of stepped-up security and convenience brought by smart door locks. These gadgets are also relatively cost-effective, with prices hovering between INR 6000 and INR

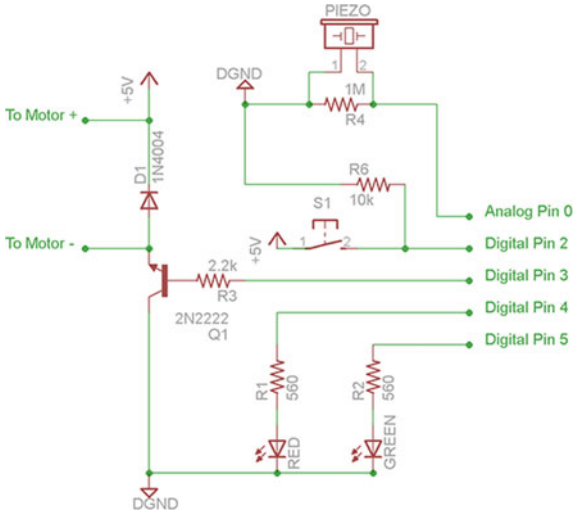


Fig. 4 Knock detecting and actuator circuit

12,000. As technology advances, smart door locks might see more security features in the near future, such as biometric recognition. But right now, smart door locks prove that they represent the future of secured households.

6 Future Outcome

As expected, the door gets locked or unlocked based on the input vibrations made by a user when trying to unlock the door by knocking in a specific pattern. If the knock sequence matches with the registered sequence, the door gets unlocked and the user can open the door, else when the user fails to knock the correct knock sequence, the door remains locked and a message is sent to the owner notifying him about the intrusion.

References

1. G. Jewani, Review on-a knock based security system. Int. J. Recent Innov. Trends Comput. Commun. **3**(2) (2015). ISSN: 2321-8169
2. G. Jamous, E. Saad, A. Kassem, S. El Murr, M. Geagea, A smart lock system using Wi-Fi security, in *3rd International Conference on Advances in Computational Tools for Engineering Applications (ACTEA)* (2016). ISBN: 978-1-4673-8523-7
3. Project article: Secret knock detecting door lock, Instructables by Autodesk, Incorporation, US, MIT Media Lab

4. K. Upadhyay, A.K. Yadav, P. Gandhi, A review of internet of things from Indian perspective, in *Engineering Vibration, Communication and Information Processing*. Lecture Notes in Electrical Engineering, vol. 478 (Springer, Singapore, 2019) pp. 621–632
5. S. Mondal, B. Nandi, R. Biswas, D. Das, R. Mutt, Knock to unlock. Int. J. Emerg. Res. Manag. Technol. **4**(9) (2015). ISSN: 2278-9359
6. R.B. Pandhare, N.D. Chhabile, V.U. Gaikwad, M.B. Bawaskar, R.A. Kapse, Home automation and security using arduino, bluetooth and GSM technology. Int. J. Res. Advent Technol. (IJRAT), (E-ISSN: 2321-9637), Special Issue National Conference "CONVERGENCE 2017", 9 April 2017

Effective Control Strategy in Microgrid



Ankur Maheshwari, Yog Raj Sood, Aashish Goyal, Mukesh Singh and Sumit Sharma

1 Introduction

A microgrid (MG) can be viewed as an independent power system operating as a single controllable entity that comprises a cluster of loads, distributed generation (DG), energy storage devices [1], and other devices such as electric vehicles that have the capability to consume and produce electrical energy. It is a small source of energy that comprises a wide range of sustainable power sources, for example, wind, solar, Micro hydro, geothermal, tidal, and so on [2]. Two modes of operation of a microgrid are grid-connected mode (nonautonomous mode) and islanded mode (autonomous mode). In grid-connected mode, there is an advantage of power trading between the main grid and a microgrid [3]. However, if there is any disturbance such as a fault in the main grid or if it is scheduled, microgrid needs to be disconnected from the main grid and operate in the autonomous mode [4].

Inverters are a very essential part of the microgrid. Inverters are required to interface the distributed generators such as solar and wind generators with the microgrid. Also, system performance is improved by the parallel operation of the inverters in a microgrid. It also enhances the reliability [5]; in case of fault occurrence in any inverter, power is still delivered to the load by the rest of the units. For reliable and efficient operation of the distributed energy resources, robust design of voltage source inverter (VSI) controller is necessary [6]. There are several control techniques that can be utilized for inverter parallel operations in a microgrid. Peer-to-peer and master–slave control techniques are generally used and widely studied for the controlling of a microgrid inverter. In the master–slave control strategy, one or more distributed generators within the microgrid act as a master and rest of the DGs as slaves. However, it has certain disadvantages in master–slave control itself [7]. In order to effectively operate the microgrid in both autonomous and nonautonomous modes by satisfying the voltage/frequency [8] and load demand, droop control based on current and voltage dual closed-loop control and PQ control strategy is described in this paper. Section-wise sequence of this paper is as follows. In Sect. 2, the proposed microgrid

A. Maheshwari (✉) · Y. R. Sood · A. Goyal · M. Singh · S. Sharma
National Institute of Technology Hamirpur, Hamirpur, HP, India
e-mail: ee.ankur.maheshwari@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_94

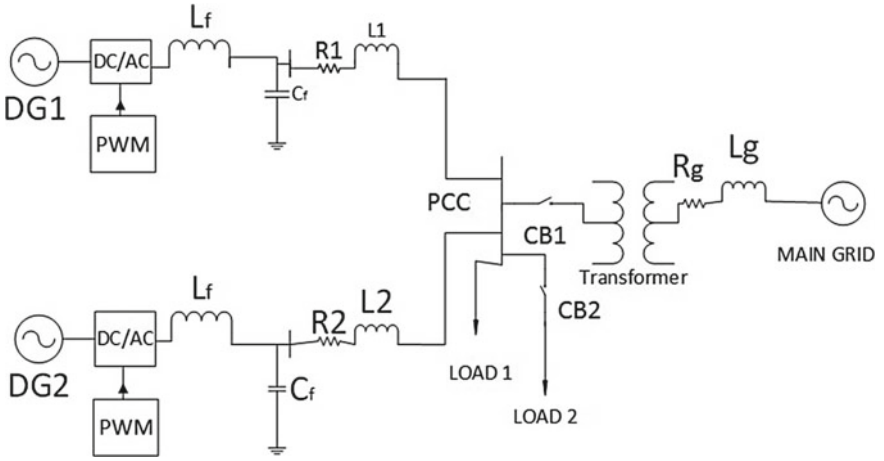


Fig. 1 Architecture of microgrid

architecture is presented. In Sect. 3, PQ control theory under grid-connected mode is explained. In Sect. 4, droop control strategy during islanded mode is presented. In Sect. 5, the MATLAB/Simulink simulation results are presented which verify the proposed control theory. Section 6 concludes this paper.

2 Architecture of Microgrid

Figure 1 shows the structure of the microgrid considered in this paper to check the adequacy of the controllers. It comprises two distributed generators (DG1 and DG2) connected to voltage source inverter with pulse width modulation (PWM).

The DGs can be PV, fuel cells, energy storage devices, etc. Inverter output power is fed to an LC filter in order to filter out the harmonics. The main grid is connected to the microgrid at the point of common coupling (PCC) via a static switch. Circuit breaker, CB1, operates seamlessly to transfer the microgrid from the grid-connected mode to islanded mode in case of any fault or power quality issue in the main grid. Circuit breaker, CB2, is used to disconnect and reconnect the load 2. For simplifying the research objective, the distributed generators at the inverter DC side are replaced by a constant DC source [7].

3 PQ Control of Microgrid

In the nonautonomous mode of a microgrid, by PQ control, the real and reactive power output of the distributed generators can be controlled according to the reference provided by the grid operator [1]. In this mode, voltage and frequency support is

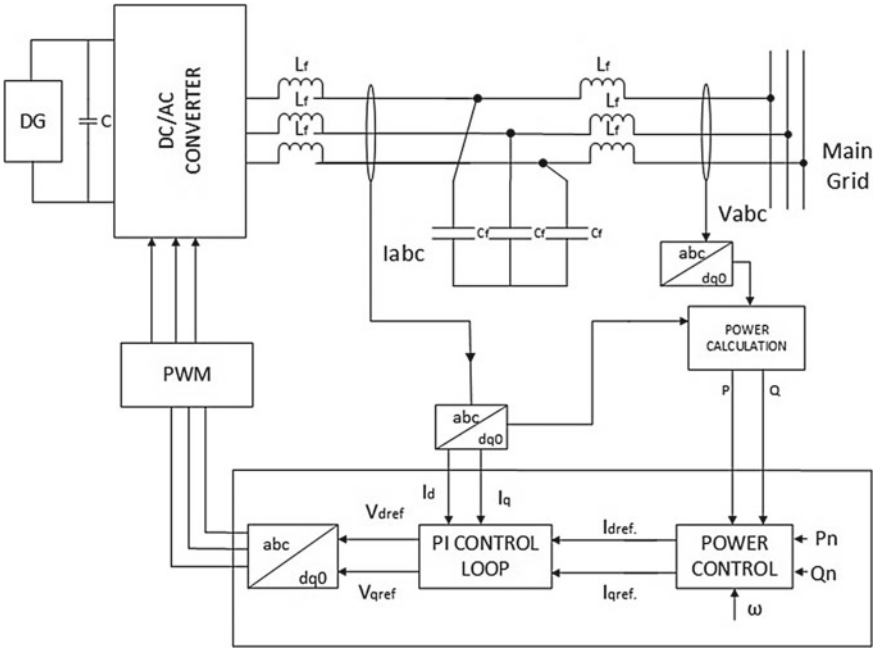


Fig. 2 Schematic of PQ control

provided by the main grid and the distributed generators just inject or absorb power without any active participation in voltage and frequency regulation [7]. Avoiding the direct participation of the distributed generators in feeder voltage regulation eliminates the adverse impact on the power system.

3.1 PQ Controller Design

Schematic of a PQ controller design is shown in Fig. 2. Due to better control provided by proportional plus integral (PI) controllers in the dq rotating reference frame, the voltage and current are collected and converted into a dq rotating reference frame with the help of Park's transformations. The current in the dq frame is proportional to the active and reactive power [9]. Therefore, by controlling the d-axis current, active power can be controlled and by controlling the q-axis current, reactive power can be controlled.

4 Droop Control in Autonomous Mode

When microgrid have several distributed generators, the sharing of the power among them is made legitimately with the assistance of droop control strategy. The transition of microgrid from grid connected mode to autonomous mode and vice versa also becomes smooth with the help of droop control. The active and reactive power sharing among the distributed generators in droop control strategy is determined by frequency and voltage control of microgrid. Active and reactive power of a transmission line is given by Eqs. (1) and (2).

$$P = \frac{E_1^2}{Z}\cos(\theta) - \frac{E_1 E_2}{Z}\cos(\theta + \delta) \tag{1}$$

$$Q = \frac{E_1^2}{Z}\sin(\theta) - \frac{E_1 E_2}{Z}\sin(\theta + \delta) \tag{2}$$

where θ and δ are the impedance and power angle, respectively. The inductance (L) is much higher than the resistance (R) in the practical overhead transmission line so resistance can be neglected. Taking this into account, the above equations can be transformed into

$$P = \frac{E_1 E_2}{X}\sin(\delta) \tag{3}$$

$$Q = \frac{E_1^2}{X} - \frac{E_1 E_2}{X}\cos(\delta) \tag{4}$$

The power angle δ is also small in these transmission lines so it can be assumed that $\cos(\delta) = 1$ and $\sin(\delta) = \delta$; Eqs. (3) and (4) can be rewritten as

$$\delta \cong \frac{XP}{E_1 E_2} \tag{5}$$

$$E_1 - E_2 \cong \frac{XQ}{E_1} \tag{6}$$

It is observed from Eqs. (5) and (6), the power angle δ is proportional to the active power P while the voltage (E1 E2) is proportional to the reactive power Q. Power angle δ is constrained by the generator's torque which implies that control of active power P is acknowledged by frequency control and the voltage control is realized by the reactive power control. It can be expressed by the following equations.

$$f - f_n = m(P - P_n) \tag{7}$$

$$V - V_n = n(Q - Q_n) \tag{8}$$

where P_n and Q_n are the reference points for active and reactive power, V_n and f_n are base voltage and frequency, m is the frequency coefficient and n is the voltage coefficient, f and V are the new operating points of frequency and voltage, respectively. These reactive power/voltage (Q/V) and active power/frequency (P/Q) droops for the control of inverters are called conventional droops, which are shown in Fig. 3 [10, 11].

4.1 Droop Controller Design

The schematic of a droop controller is shown in Fig. 4. Primarily, the three-phase output voltage and current of the converter are collected by the Park transform block and the magnitude of the reference voltage, angular frequency is calculated by the power calculation and droop control block. The voltage synthesis block controls the voltage amplitude and angular frequency to get the droop control voltage [11].

At last, the voltage generated by droop control is controlled by the voltage outer loop and a feedback signal is controlled by the current inner loop. Finally, the dual closed-loop control generates a PWM signal through the PWM module to control the switching of the inverter. Figure 5 shows the details of voltage synthesis and droop control.

5 Simulation Result

So as to confirm the correctness and viability of the proposed control methodology, in MATLAB/Simulink a microgrid simulation model is presented in a simulation platform, which consists of two distributed generators (DG1 and DG2) and both controllable and noncontrollable loads. In Fig. 1, the proposed model is shown. The droop control of both the distributed generators and output impedance parameter setting of both inverters are set uniquely. The simulation results in this paper represent the effectiveness of PQ and droop control based on current and voltage dual closed

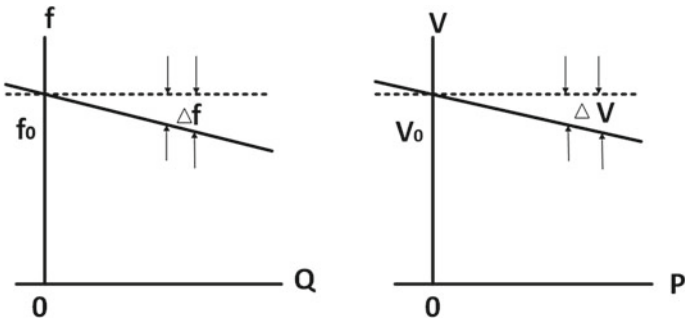


Fig. 3 Conventional droop characteristics curves

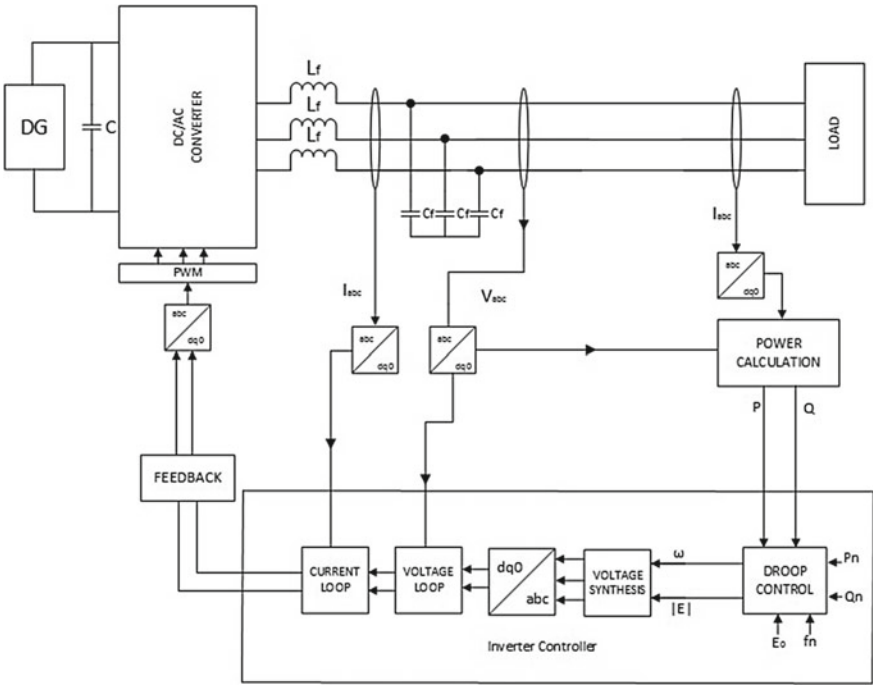


Fig. 4 Schematics of droop control

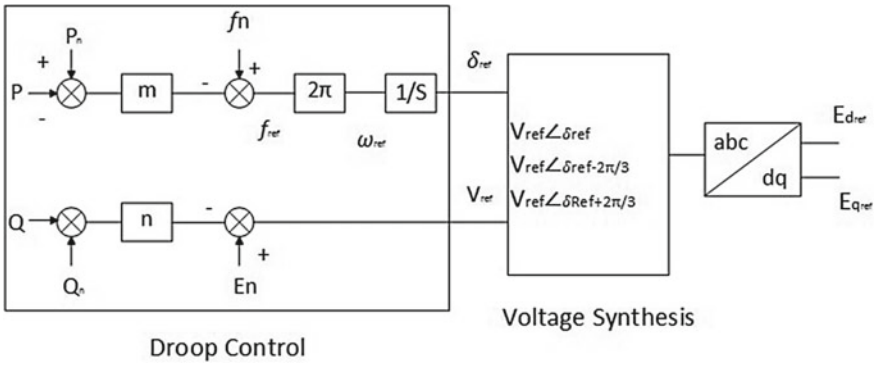


Fig. 5 Details of voltage synthesis and droop control

loop control of inverters along with the two distributed generators. The results are broken down under two cases.

Case 1. Initially, the microgrid is operating by PQ control when it is operating under grid-connected mode. The microgrid is scheduled to work in islanded mode at time $t = 0.4$ s and again in grid-connected mode at time $t = 0.6$ s. The results under this condition are shown in Figs. 6, 7, and 8. The time of the simulation is 1.5 s. Figure 6 shows the sharing of the active power of both the distributed generators under grid-connected and isolated modes of the microgrid. Because the parameters of both the converters associated with the distributed generators are unique, the change in the active power of DG1 and DG2 are the same. Figure 7 shows that the reactive power support provided by the microgrid to the main grid becomes zero during the period when the MG is operated in the isolated mode. Figure 8 interprets that the frequency of both the distributed generators rises when the main grid is disconnected at 0.4 s and again decreases when it is reconnected at 0.6 s. The frequency retains in 50–50.5 Hz range which meets the necessities of the above control attributes. Analysis of Figs. 6 and 7 shows that the active power output of both the distributed generators decreases when the main grid is disconnected at 0.4 s and increases when the main grid is reconnected to the microgrid.

In the meantime, Figs. 6, 7, and 8 show inverter real and reactive power output of both distributed generators assuring the balance between the generation and load demand. It also shows that the power can be balanced rapidly during the transition of microgrid from grid-connected mode to isolated mode and vice versa and the distribution of power is rapidly realized.

Case 2. Initially, only a fixed load is connected and the microgrid is coupled to the main grid. The microgrid is scheduled to operate in the isolated mode after 0.4 s permanently and at the same time, another load is connected to the isolated microgrid. The simulation time is 1 s.

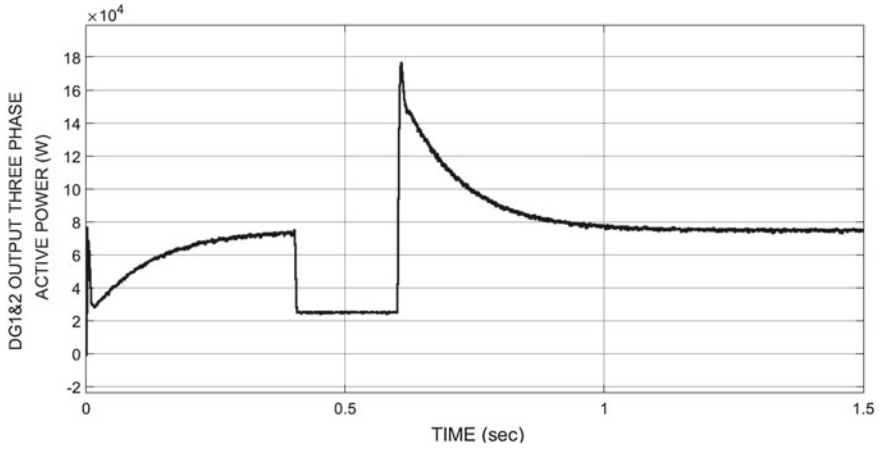


Fig. 6 Active power output of DG1 and DG2

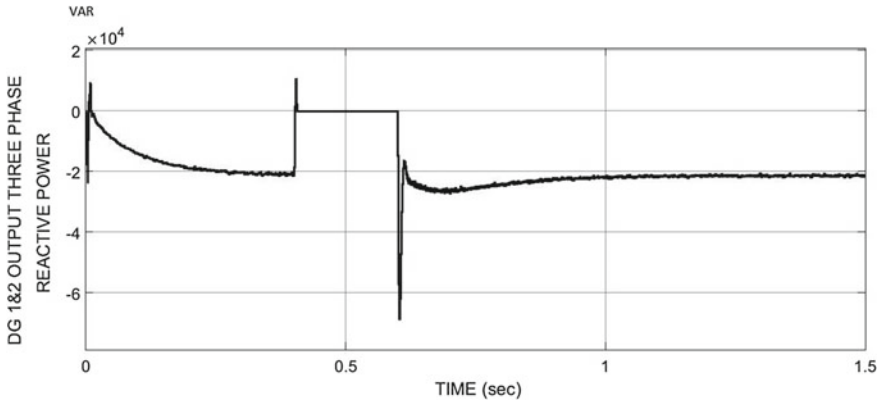


Fig. 7 Reactive power output of DG1 and DG2

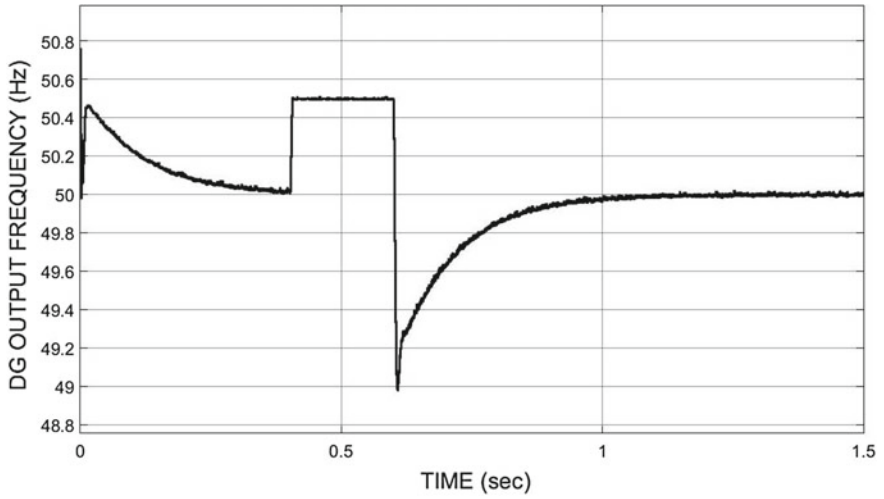


Fig. 8 Output frequency of DG1 and DG2

Figures 9 and 10 show the variations in active and reactive power of DG1 and DG2, respectively. Figure 11 shows that the frequency is maintained at 50 Hz. These cases show that the PQ and droop control methods work effectively; voltage and frequency are well-maintained during load changes.

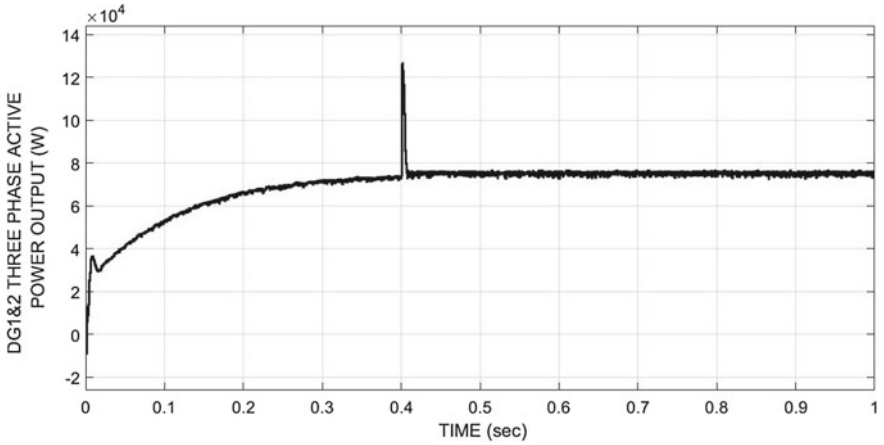


Fig. 9 Active power output of DG1 and DG2

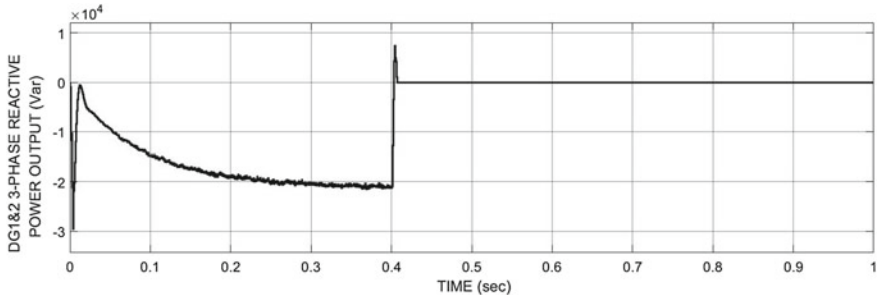


Fig. 10 Reactive power output of DG1 and DG2

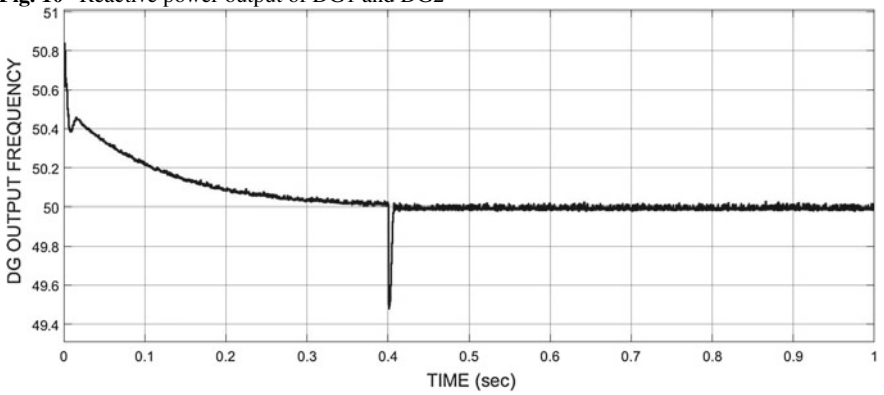


Fig. 11 Output frequency of DG1 and DG2

6 Conclusion

The PQ and droop controller was introduced in this paper for the controlling of distributed generators interfaced voltage source inverters. Smooth switching between both isolated and grid-connected mode was ensured by droop controller. Finally. the simulations results verified that the proposed controller effectively tracks the real and reactive power changes, shares the power among distributed generators according to their droop characteristics, and maintains the voltage and frequency stability in both operating modes of the microgrid.

References

1. X. Zhou, T. Guo, Y. Ma, An overview on operation and control of microgrid, in *International Conference on Manufacturing Construction and Energy Engineering (MCEE)* (2016), pp. 223–229. https://doi.org/10.12783/dtetr/mcee2016/6412
2. M.S. Mahmoud, *Microgrid* (Butterworth-Heinemann, Oxford, 2017)
3. M. Anwar, M.I. Marei, A.A. El-Sattar, Generalized droop-based control for an islanded microgrid, in *2017 12th International Conference on Computer Engineering and Systems (ICCES)* (IEEE Press, Egypt, 2017), pp. 717–722. https://doi.org/10.1109/ICCES.2017.8275399
4. A. Banerji, D. Sen, A.K. Bera, D. Ray, D. Paul, A. Bhakat, S.K. Biswas, Microgrid: a review, in *2013 IEEE Global Humanitarian Technology Conference: South Asia Satellite (GHTC-SAS)* (IEEE Press, Trivandrum, India, 2013), pp. 27–35. https://doi.org/10.1109/GHTC-SAS.2013.6629883
5. T.F. Wu, Y.E. Wu, H.M. Hsieh, Y.K. Chen, Current weighting distribution control strategy for multi-inverter systems to achieve current sharing. IEEE Trans. Power Electron. **22**(1), 160–168 (2007). https://doi.org/10.1109/TPEL.2006.886622
6. K. Abo-Al-Ez, A. Elaiw, X. Xia, A dual-loop model predictive voltage control/sliding-mode current control for voltage source inverter operation in smart microgrids. Electr. Power Compon. Syst. **42**(3–4), 348–360 (2014). https://doi.org/10.1080/15325008.2013.862319
7. F. Li, R. Li, F. Zhou, *Microgrid Technology and Engineering Application*, 1st edn (Joe Hayton, United States, 2014)
8. W. Bai, K. Lee, Distributed generation system control strategies in microgrid operation. IFAC Proc. Vol. **47**(3), 11938–11943 (2014). Springer, Heidelberg. https://doi.org/10.3182/20140824-6-ZA-1003.02116
9. M. Chamana, S.B. Bayne, Modeling, control and power management of inverter interfaced sources in a microgrid, in *2011 IEEE 33rd International Telecommunications Energy Conference (INTELEC)* (IEEE Press, Amsterdam, 2011), pp. 1–7. https://doi.org/10.1109/INTLEC.2011.6099766
10. C. Natesan, S. Ajithan, S. Mani, P. Kandasamy, Applicability of droop regulation technique in microgrid—a survey. Eng. J. **18**(3), 23–36 (2014). https://doi.org/10.4186/ej.2014.18.3.23
11. P. Ye, J. He, G. Wang, S. Li, F. Sun, Y. Han, T. Zhang, An improved droop control strategy for parallel inverters in microgrid, in *2017 IEEE Conference on Energy Internet and Energy System Integration (EI2)* (IEEE Press, China, 2017), pp. 1–5. https://doi.org/10.1109/EI2.2017.8245514

An Efficient Trust-Based Approach to Load Balanced Routing Enhanced by Virtual Machines in Vehicular Environment



Rakhi and G. L. Pahuja

1 Introduction

A Vehicular Ad Hoc Network (VANET) comprises vehicles as wireless nodes communicating with each other without a fixed infrastructure. VANET is formed either by transfer of information between the two vehicles that is Vehicle to Vehicle (V2V) communication or between vehicle to the roadside equipment that is Vehicle to Infrastructure (V2I) communication. The major applications of VANETs on roads is to transfer both critical and noncritical messages such as safety and warning instructions to avoid possible travel accidents, also messages related to entertainment and multimedia to provide comfort to the fellow travelers [1]. However, with these vast applications of advanced communication technologies, there come the challenges faced by these vehicular networks. VANETs work in unpredictable and unreliable conditions, as these are extremely dynamic in nature due to the high mobile nature of the vehicular nodes. Thus, the major burden on the network is to fulfill the service-related parameters like energy consumption, Packet Delivery Ratio(PDR), Bit Error Rate(BER), and total transmission time [2, 3]. With the increasing number of vehicles on road and growing communication needs, the VANET network is becoming extremely complex. VANET is integrated with the networking routers deployed along the roadside equipment to deliver the required information of the traffic and monitor the dangerous conditions on the road. The dissemination of messages through VANET requires communication between wireless nodes within a secured reliable architecture. The three modes of vehicular communication are Vehicle to Vehicle

Rakhi (✉) · G. L. Pahuja
Electrical Engineering Department, National Institute of Technology, Kurukshetra, Kurukshetra 136119, Haryana, India
e-mail: rakhikamra@gmail.com

G. L. Pahuja
e-mail: pahuja.gl@nitkkr.ac.in

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_95

(V2V) communication, Vehicle to Infrastructure (V2I) communication, and Infrastructure to Vehicle (I2V) communication. Figure 1 displays a general architecture of the VANET [4]. On-Board Unit (OBU) is the unit placed on the vehicle responsible for setting up the short-range communication network with the sensor nodes and other network equipment placed in the fixed infrastructure along the roadside called the RoadSide Unit (RSU) such as poles, traffic light signals, and other buildings [5].

The major constraints faced by VANETs in successful data dissemination are highly dynamic vehicular mobility, dedicated short-range communication, and stringent timing (minimum delay) [6]. Since the performance of the VANETs depends on the lifetime of the sensor nodes, therefore, the prime challenge becomes to manage the energy resources of the sensor nodes. One of the efficient methods to extend the lifetime is to balance the load over the nodes or processors employed in the network. The load-balancing assignment includes optimal utilization of all resources of the network. The challenging task is to attend maximum number of vehicles with the required information in terms of available resource and minimum time. Thus, the problem is summed up as to design an efficient, reliable, secure, and trustworthy routing path with minimum delay and energy consumed. This paper aims at providing solution to the problem through investigating various Quality of Service (QoS) parameters resulting in a reliable network load balancing and routing. A comparative performance analysis is done on the basis of an optimized load balancing mobility model for VANETs available in the literature [7].

The contributions of this work are as follows:

- A system model of a network is set upin vehicular environment.
- Developed a route discovery mechanism on broadcasting technique overcoming the delay constraint.
- A framework for a trust model is proposed to optimally utilize the resources.
- Introduction of Virtual Machines (VM) to balance the load on the nodes efficiently.
- In conclusion, a validation of the proposed algorithm using various network performance measures.

The rest of the paper is organized as follows. Section 2 presents related works. Section 3 describes the network model. The proposed load balancing model using VMs is discussed in Sect. 4. Section 5 validates the performance evaluation, and lastly Sect. 6 concludes the paper.

2 Related Works

A number of research works have been studied comparing the performance of various routing protocols using different network models, in diverse traffic environment with a number of performance metrics. Routing in VANETs is quite a complex task due to highly mobile nature of the vehicular nodes. VANETs employ two types of communications, i.e., Vehicle to Vehicle (V2V) and Vehicle to Infrastructure (V2I) [5]. Routing protocols in ad hoc network for VANET applications are categorized

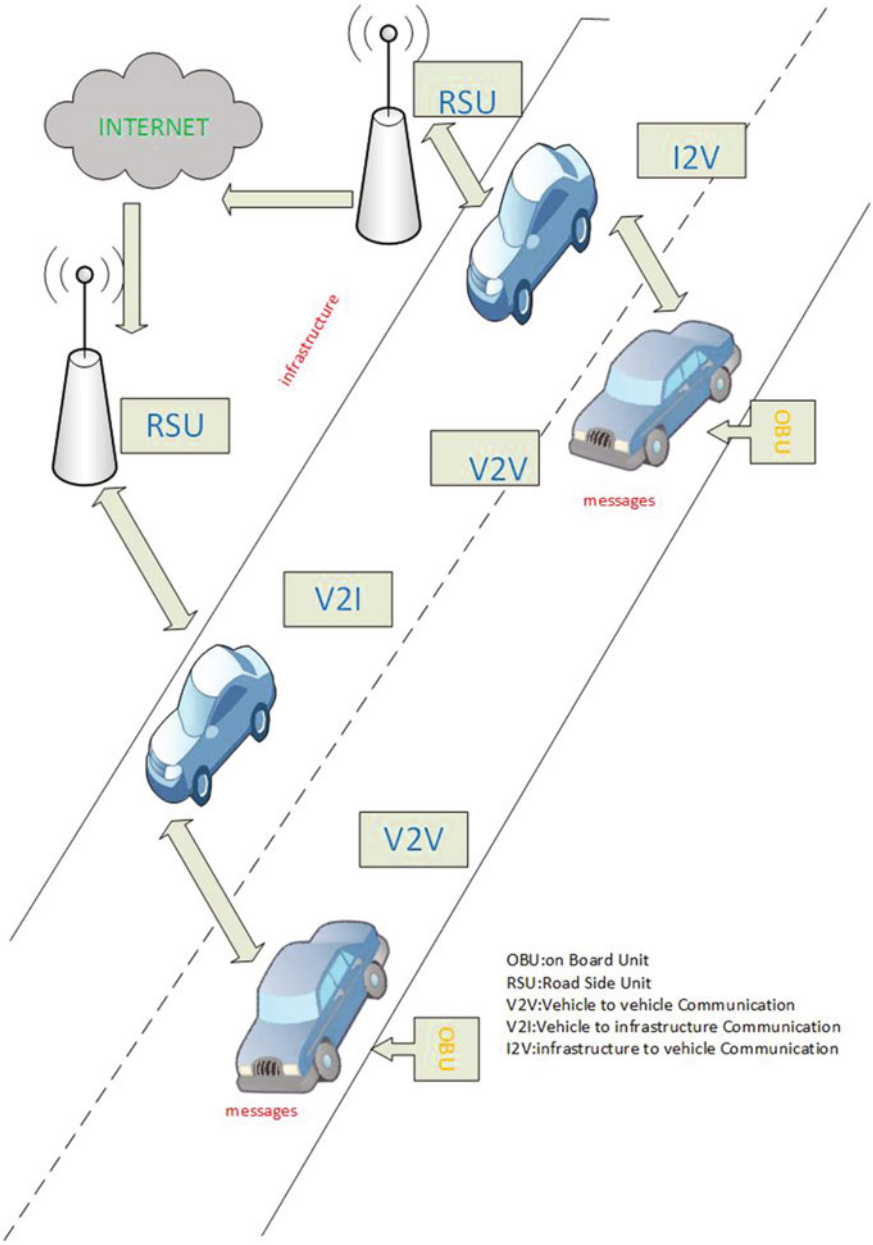


Fig. 1 Architecture of VANET

in topology and location-based protocol [8]. Kumar et al. in their paper proposed a new clustering and routing algorithm based on learning agents [9]. Hussain and Sharma in their paper [10] analyzed the two routing protocols (Distance—Effect Routing Algorithm for Mobility (DREAM) and Location Aided Routing (LAR) in city as well as highway scenario and evaluated different performance metrics like PDR, routing overhead, delay, throughput. In the paper [11], the authors proposed a new approach for cluster-based routing algorithm, i.e., message dissemination in vehicular networks based on interest of the receiver. This new algorithm suggests for reliable data transmission and the vehicles must be intelligent enough to decide the best solution for the environment changes like traffic jams, accidents, and deployment of varying nodes.

The dynamics in topology due to movement of nodes in VANET leads to overloading of nodes/access points (AP) s. To increase the lifetime of the network, dynamic load balancing and reliable data transfer among various nodes/access points(AP)s are the two main requisites. Wu et al. [12] in their paper presented a fast handoff scheme, i.e., Quality Scan for VANET which decreases the latency considering the loading states of nearby access points simultaneously. This scheme improves the QoS and efficiency in VANET. Optimal routing in VANETs ensures searching the path with the highest throughput and data transmission rate in variable constraints as road barriers and behaviors of drivers. An important issue which has not been significantly investigated by researchers is the problem of load balancing in VANETs. Appropriate distribution of traffic among all available connected paths of the network guarantees the increase in throughput and reduced congestion [13]. Effective resource utilization is possible by proper resource allocation and balancing of load on RSUs. Proper placement of RSU decides the requirement of energy in sending of the ground data uploaded to the RSUs. Various power saving schemes for appropriate placement of RSU suggest many scheduling algorithms that are proposed in [14–16]. Agarwal et al. [16] have proposed a load balanced routing algorithm in a one-dimensional scenario in VANET. The algorithm compares the performance metrics with the Nearest Neighbor Routing (NNR) scheme in their work. But this research study is supposed to work only single dimensional traffic on road. Perdana et al. [7] have compared the performance of VANET in a realistic mobility model with and without load balancing scheme. Works such as [17] have suggested an optimized way for cooperative load balancing in VANETs by shifting the load from one RSU to another RSUs Furthermore, many approaches have focused on trust issues in VANETs for secure routing mechanisms. [18, 19].

To comprehend, there is a lack of an appropriate reliable trustworthy energy-efficient solution to the problem of load balancing in VANET environment. The contribution of this research is to employ Virtual Machines (VMs) as an approach to reduce the load on the Vehicular network. Virtualization of the resources hosted on the mobile vehicular nodes provides an ample opportunity to harness the whole lot of applications through VANETs. This is the task of Infrastructure Provider (ISP) to provide a collection of virtual machines as a pool of resources.

This paper propounds a reliable solution to the traffic management by adding up the virtual gears into the network and optimizes the congestion problem by using a

trust queue updated with the broadcast concept permissible to remove the undesired nodes from the list.

3 Network Model

The proposed reliable solution is divided into the following categories:

1. Setting up the network in vehicular environment.
2. Development of optimized route discovery mechanism.
3. Development of trust framework to enhance the reliability of the network.
4. Introduction of Virtual Machines (VM) as a solution to load balancing of the nodes efficiently.

3.1 Setting up the Network in Vehicular Environment

All the vehicular nodes are deployed randomly with random parameter values of delay time, battery capacity, message transmission time, and state of the node, i.e., active state or sleep state. In real-time setup, time it is obvious that a few of the vehicular nodes will be in inactive or sleep state possibly due to battery limit or due to surrounding conditions.

For the data transmission, one vehicular node would be the source and the other would be the destination among all the randomly deployed nodes forming the network which leads to a lot of uncertainty in the network. There has to be a certain path for the successful transmission of the data packets from the source node to the destination node.

3.2 Development of Optimized Route Discovery Mechanism

In the state of uncertainty, there arises a condition where the data has to be communicated for the very first time that is no vehicle has been previously involved in the path. The VANET routing becomes more challenging as to discover a path that can meet the limited requirements of mobility, bandwidth constrained, traffic load, and security. In the route discovery mechanism for first-time communication, an algorithm is proposed here using the broadcasting technique [20]. Herein, the source vehicular node will discover its shortest path if no previous communication path is available. Figure 2 shows the scenario of first communication where vehicular node S acting as a source node wishes to send a data packet to a destination vehicular node D, but the memory cache of the source node S is void, i.e., it does not already know a route. “X” denotes unidentified vehicle thus “S” cannot directly communicate with

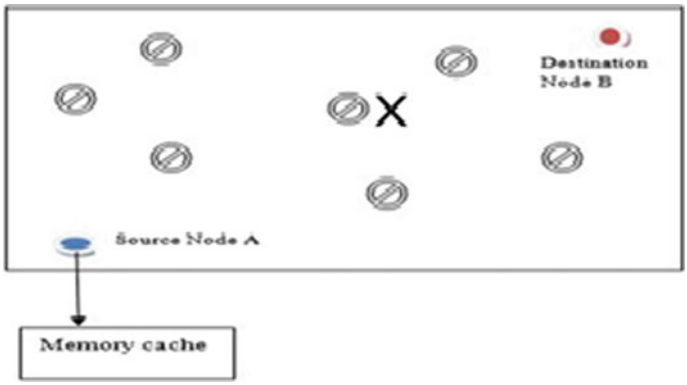


Fig. 2 Communication from source node "A" to destination node "B" (broadcasting technique)

unidentified vehicle. Thus, as per the algorithm, vehicle "S" will broadcast a route request, a data packet with its own location address and the sink/destination address within the network.

If the source node A receives acknowledgement within the permissible time frame, it confirms sending the data and a route is developed, otherwise, another route determination technique is employed which includes calculation of the coverage set of the source node. The coverage set is calculated using the Eq. 1, where d is the distance covered by the transmitting nodes of any vehicle. Also, Δx is the difference between the position coordinates on "x" axis, and Δy is the difference between the position coordinates on "y" axis.

$$d = \sqrt{(\Delta x^2 + \Delta y^2)} \tag{1}$$

The next step is to select any vehicle in the coverage set and save the selected vehicle in the cache memory of the source node and follow the steps till the destination node is encountered.

3.3 Development of the Trust Model

The very dynamic nature of the VANETs unlocks many challenges in security and reliable data communication between nodes. The noncentric and dynamic nature of the VANET make it tough for an established, long, and secure relationship among vehicular nodes. Security and authentication of the data is the prime concern in VANET, as data may also carry life-critical information. Hence, trust management is very necessary and crucial in VANETs as compared to MANETs. The major concern is to check and select which vehicle or message is trustworthy or not. These search processes for each data transmission from source node to destination node

that consumes a lot of time and add delay in the system. For secure, reliable, and timely transmission of data, a trust model is proposed in this paper along with the broadcasting technique. The key points in the trust model are as follows:

a. If a vehicle "V" delivers the data successfully between source "S" and destination "D", the proposed algorithm place the vehicle "V" in the trust list for the destined path from vehicular node "S" and vehicular node "D". The trusted "V" will remain in the trust list provided that it does not increase the delay more than permitted for the transfer of data packet.
b. The monitoring system would continue to check the position of the vehicle. For the vehicle "V" to persist in the trust list, it should still stay in the coverage set of the source node "S".
c. The trusted vehicles are those which are consuming delay less than the specified delay time or that is within the permissible time frame.

4 Introduction to Virtual Machines (VM) for Solution to Load Balancing Problem

To optimize the usage of resources and lessen energy consumption, this paper focuses on employing vehicles as VM hosts. Vehicular nodes can act as effective mobile resource creating a data center [21]. The idea here is to improve the load balancing among the nodes in the network by using Virtual Machines (VMs). To achieve this optimized solution, a well-known energy-efficient approach for VM allocation by Anton et al. [22] has been studied and simulated in the vehicular environment. The Modified Best Fit Decreasing algorithms take account of the introduction of VMs on vehicles and decide the most power-efficient nodes for the hosts to the chosen VMs. The Minimization of Migrations (MM) algorithm sorts the list of VMs in the declining order of resource consumption. This MM policy chooses the least number of VMs required to migrate from a host.

5 Performance Analysis

The performance of the proposed method for the solution of load balancing problem in VANETs using broadcasting technique with virtual machines is simulated in a vehicular environment. In this study, vehicular mobile scenario is set up and simulations are conducted using MATLAB. The scenario generated is using a grid topology of 1000 m $\times$ 1000 m. The homogeneous vehicular nodes with constant velocity are uniformly distributed in the network area. The network is initially designed and then it is tested around three times with three different sets of number of vehicles to 50,100, and 150. The simulation parameters for setting up the network are described in Table 1.

Table 1 Network model parameters

Simulation parameters	Simulation parameter value
Network Size Length Width	1000×1000 m^2 1000 m 1000 m
No. of vehicles	50, 100, 150
Coverage (range)	20% of width
Coverage parameter	Distance between the vehicles
Delay model Simulation time Message size	Random delay model 80 s 512 Kbytes packet size

The various performance measures evaluated are throughput, delay time, Packet Delivery Ratio (PDR), and energy consumption.

The throughput of the network has been analyzed five times with different sets of nodes and every time the throughput has come out to be efficient enough as depicted in Fig. 3. This is because of the optimized network, which reduces the overhead of the network due to which the nodes are capable of transferring the data as much as possible.

As Fig. 4 indicates the proposed algorithm preserves the energy in the network, this reduced amount of energy consumption is due to the routing in the saved trustworthy nodes and thus trust management leads to an energy-efficient reliable route determination solution. Moreover, sharing of the load among the working virtual

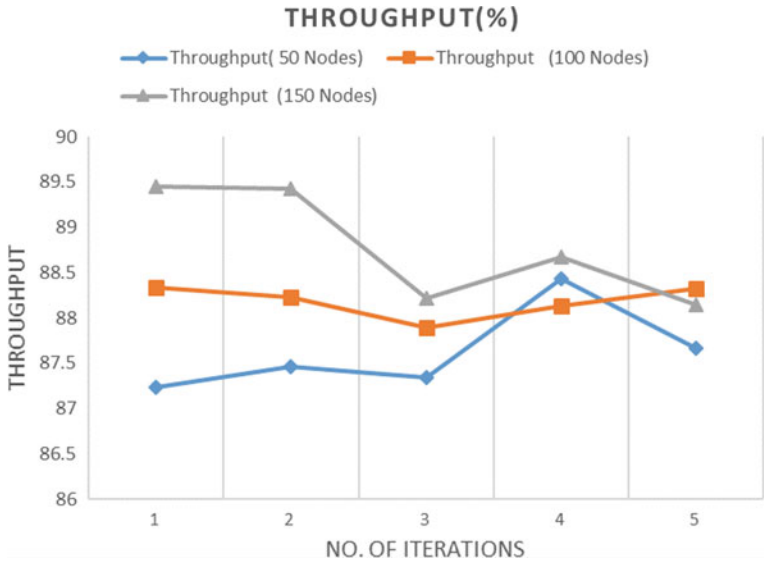


Fig. 3 Network throughput with 50,100, and 150 number of nodes

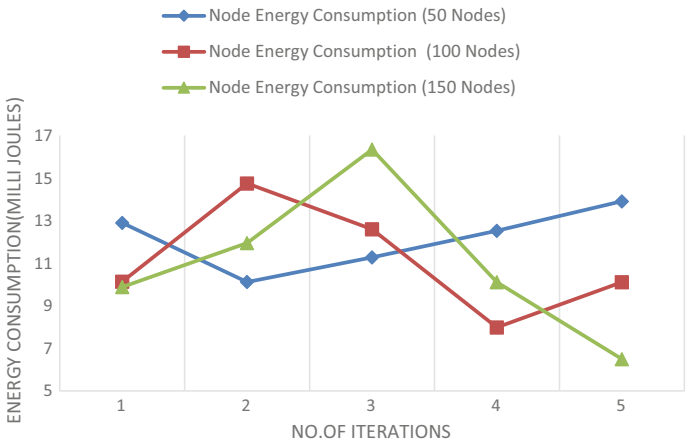


Fig. 4 Consumption of energy with 50,100, and 150 number of nodes

machines on the vehicular hosts reduce the resource consumption by each vehicular nodes. On an average 11 mJ of energy has been consumed even if the node count has been increased.

The end to end delay is the total time taken from the packet transmission from the source to the arrival of the packet at the destination. In Fig. 5 for 50,100, and 150 number of nodes, average delay is about 2.54 ms. It is noticeable that the delay time taken is considerably low with the implementation of the trust model. Also, with the increase in number of vehicular nodes in the network, end to end delay decreases.

For validation of the proposed algorithm, the average throughput is compared with the results of the load balanced mobility model using stock path algorithm proposed by Perdana et al. [7]. Figure 6 compares the network throughput for about 100 number of nodes at an average speed of 100 km/h. It clearly indicates that the

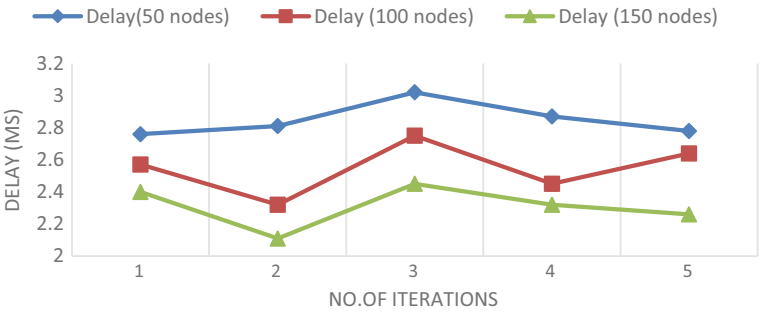


Fig. 5 Transmission delay with 50,100, and 150 nodes

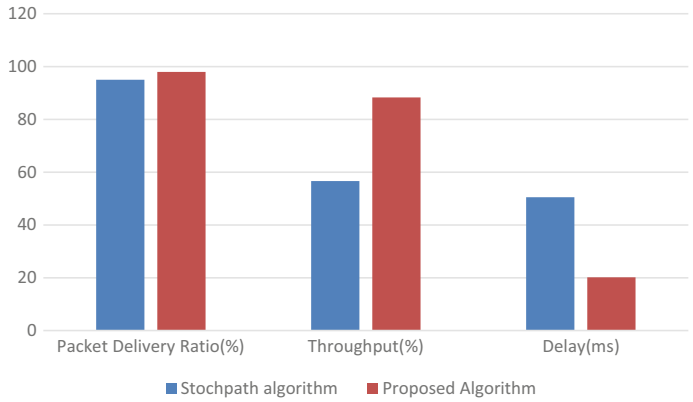


Fig. 6 Comparison of performance measures

proposed algorithm is better than the load balancing scheme "Stochpath algorithm" in [7]. The selected and reliable trustworthy node for the route determination will result in the least packet drops and increased throughput. The decrease in delay in the proposed algorithm network is widely distributed and the available virtual machines on the vehicular nodes reduce the load on each vehicle and can make the transmission faster and reliable. As implied in Fig. 6, the packet delivery ratio is comparable.

6 Conclusion

In the fast-rising vehicular movement in global scenario, the emerging communication infrastructure consists of Vehicular Ad Hoc Networks (VANET). For acquisition of reliable data and effective performance of the VANET applications, a trust model is proposed for route determination by saving considerable time in reliable node discovery process. The proposed research solution envisages the broadcasting technique in the dynamic environment of the VANETs to optimize the available resources. The incorporation of the vehicles as possible VM hosts allows the network with ample virtual resources without computing about the actual physical resource available. Such potential VMs on vehicles can share the load on physical nodes and can handle hot spots in the network. The simulation studies show that the proposed algorithm has given efficient results on various performance metrics like delay, energy, throughput, and packet delivery ratio and has been validated by a standard load balancing mobility model for VANETs. The network performance has been evaluated and compared using various Quality of Service (QoS) parameters, namely, energy, delay, and throughput.

References

1. E.C. Eze, S.J. Zhang, E.J. Liu, J.C. Eze, Advances in vehicular ad-hoc networks (VANETs): challenges and road-map for future development. Int. J. Autom. Comput. **13**, 1–18 (2016)
2. F. Cunha et al., Data communication in VANETs: protocols, applications and challenges. Ad Hoc Netw. **44**, 90–103 (2016)
3. V. Saritha, P.V. Krishna, S. Misra, M.S. Obaidat, Learning automata-based channel reservation scheme to enhance QoS in vehicular adhoc networks, in *2016 IEEE Global Communications Conference, GLOBECOM 2016 - Proceedings* (Institute of Electrical and Electronics Engineers Inc., 2016)
4. Rakhi, G. Pahuja, Component importance measures based risk and reliability analysis of vehicular ad hoc networks. Int. J. Comput. Netw. Inf. Secur. **10**, 38–45 (2018)
5. S. Dharmaraja, R. Vinayak, K.S. Trivedi, Reliability and survivability of vehicular ad hoc networks: an analytical approach. Reliab. Eng. Syst. Safety **153**, 28–38 (2016)
6. B.S. Madhusudan, Study of the effect of velocity on end-to-end delay for V2V communication in ITs. Int. J. Next-Gener. Netw. **4**, 19–26 (2013)
7. D. Perdana, Febryan, F. Dewanta, Performance evaluation of vehicle load balancing scheme on IEEE 802.11p standard. Int. J. Simul.: Syst. Sci. Technol. **17**, 30.1–30.7 (2016)
8. A. Husain, S. C. Sharma, Comparative analysis of location and zone based routing in VANET with IEEE802.11p in city scenario, in *Conference Proceeding - 2015 International Conference on Advances in Computer Engineering and Applications, ICACEA 2015* (Institute of Electrical and Electronics Engineers Inc., 2015), pp. 294–299
9. N. Kumar, N. Chilamkurti, J.H. Park, ALCA: agent learning-based clustering algorithm in vehicular ad hoc networks. Pers. Ubiquit. Comput. **17**, 1683–1692 (2013)
10. A. Husain, S.C. Sharma, Simulated analysis of location and distance based routing in VANET with IEEE802.11p. Procedia Comput. Sci. **57**, 323–331 (Elsevier B.V., 2015)
11. S. Harrabi, I. Ben Jaafar, K. Ghedira, Message dissemination in vehicular networks on the basis of agent technology. Wireless Pers. Commun. **96**, 6129–6146 (2017)
12. T.Y. Wu, M.S. Obaidat, H.L. Chan, QualityScan scheme for load balancing efficiency in vehicular ad hoc networks (VANETs). J. Syst. Softw. **104**, 60–68 (2015)
13. H.T. Hashemi, S. Khorsandi, Load balanced VANET routing in city environments, in *IEEE Vehicular Technology Conference* (2012)
14. S.I. Sou, A power-saving model for roadside unit deployment in vehicular networks. IEEE Commun. Lett. **14**, 623–625 (2010)
15. C.H. Ou, A roadside unit-based localization scheme for vehicular ad hoc networks. Int. J. Commun. Syst. **27**, 135–150 (2014)
16. S. Agarwal, A. Das, N. Das, An efficient approach for load balancing in vehicular ad-hoc networks, in *2016 IEEE International Conference on Advanced Networks and Telecommunications Systems, ANTS 2016* (Institute of Electrical and Electronics Engineers Inc., 2017)
17. G.G.M.N. Ali, M.A. Rahman, P.H.J. Chong, S.K. Samantha, On efficient data dissemination using network coding in multi-RSU vehicular ad hoc networks, in *IEEE Vehicular Technology Conference* (Institute of Electrical and Electronics Engineers Inc., July 2016)
18. S.S. Manvi, S. Tangade, A survey on authentication schemes in VANETs for secured communication. Veh. Commun. **9**, 19–30 (2017)
19. X. Yao, X. Zhang, H. Ning, P. Li, Using trust model to ensure reliable data acquisition in VANETs. Ad Hoc Netw. **55**, 107–118 (2017)
20. Rakhi, G.L. Pahuja, A reliable solution to load balancing with trust based authentication enhanced by virtual machines. Int. J. Inf. Technol. Comput. Sci. **9**, 64–71 (2017)
21. T.K. Refaat, B. Kantarci, H.T. Mouftah, Virtual machine migration and management for vehicular clouds. Veh. Commun. **4**, 47–56 (2016)
22. A. Beloglazov, J. Abawajy, R. Buyya, Energy-aware resource allocation heuristics for efficient management of data centers for cloud computing. Future Gener. Comput. Syst. **28**, 755–768 (2012)

Parameter Optimization of a Modified PID Controller Using Symbiotic Organisms Search for Magnetic Levitation Plant



D. S. Acharya and S. K. Mishra

1 Introduction

Proportional-Integral-Derivative (PID) controller is the most commonly used controller in most of the applications. More than 90% of the closed loops in industries employ PID controller [1–3]. The simple structure and ease of implementation are the main reason for the popularity of PID controller. However, the performance of the PID controller depends on proper choice of parameters. The process of finding the most suitable values of the controller parameters, for optimum performance, is called *tuning*. Several methods have been proposed for fast and proper tuning of the controller. Industries, mostly, rely on Ziegler-Nichols [4, 5], Cohen Coon, [6] etc., to tune the controller parameters.

Several tuning methods and algorithms have been devised in the recent past. Effort has been focused on formulating simple, fast, and exact methods for tuning the controller parameters. In many cases, the time domain and frequency domain specifications such as settling time, overshoot, gain margin, phase margin, slope of phase curve are employed to find the best set of the controller parameters. Some of the recent advancements may be found in [7–9]. The Internal Model Control (IMC)-based PID design helps in reduction of the tunable parameters [10, 11].

Evolutionary algorithms have been extensively applied to the tuning of PID controllers since very long. Algorithms, similar to Genetic Algorithm (GA) [12, 13], Particle Swarm Optimization (PSO) [14, 15], Artificial Bee Colony(ABC) [16], Symbiotic Organisms Search(SOS) [17], Grey Wolf Optimizer (GWO) [18, 19], etc., have been repeatedly applied for the purpose.

This work proposes a novel structure for PID controller. The proposed controller is then applied to stabilize second-order unstable plant. The plant chosen for this work

D. S. Acharya (✉)
Department of EEE, Birla Institute of Technology,
Mesra, Off-Campus, Deoghar, India
e-mail: acharya.deep@gmail.com

S. K. Mishra
Department of EEE, Birla Institute of Technology, Mesra, Ranchi, India
e-mail: sudhansu.nit@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_96

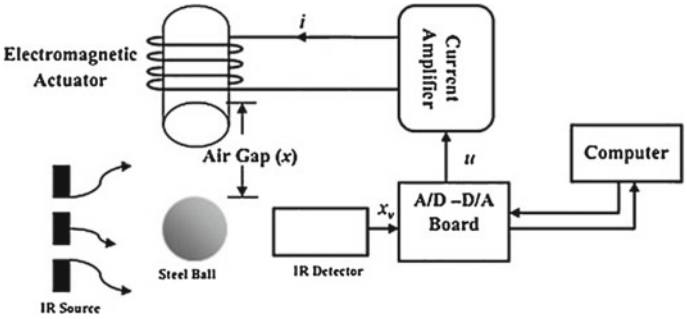


Fig. 1 Block diagram of magnetic levitation plant (Model 33-210)

is the magnetic levitation plant (Model No. 33-210) [20] of the Feedback Instruments Pvt. Ltd. The proposed controller has a modified series structure. The modification in the structure results in the reduction in the number of tunable parameters of the controller. This helps in simpler and faster tuning of the controller parameters. The major contributions of this work may be enumerated as follows:

1. Propose a novel structure for PID controller.
2. Propose the application of SOS algorithm for tuning the proposed controller.
3. Present a comparative analysis.

The remaining part of the paper is structured as follows: Sect. 2 describes the mathematical model of the magnetic levitation plant. Section 3 presents the proposed structure. Section 4 illustrates the SOS algorithm. Section 5 discusses the simulation results obtained therefrom. Finally, the conclusion is presented in Sect. 6.

2 Magnetic Levitation Plant

The Magnetic Levitation (maglev) plant considered in this work is the Model No. 33-210 of Feedback Instruments Pvt. Ltd. [20]. Magnetic levitation is the process of suspending an object in air using electromagnet forces, without any support. The block diagram of the system is shown in Fig. 1. The metal ball is suspended, in air, by the force produced by the electromagnet, excited by current i. The position of the ball (distance x from the electromagnet), is sensed by the IR detector. The control input u (voltage signal) is changed, depending on the distance x. The dynamic of the system may be expressed by

$$m\ddot{x} = mg - F_e \tag{1}$$

where m is the mass of the metal ball, x is the distance of the ball from the electromagnet, F_e is the electromagnetic force, and g is the acceleration due to gravity. F_e is expressed as [21, 22]:

$$F_e = k\frac{i^2}{x^2} \tag{2}$$

where k is a constant depending on system parameters. So, (1) may be expressed as [21, 22]:

$$m\ddot{x} = mg - k\frac{i^2}{x^2} \tag{3}$$

It is obvious, from (3), that maglev is a nonlinear system. The plant dynamic (3) needs to be linearized for analysis purpose. The equilibrium point (i_0, x_0) as provided by the manufacturer is (0.8 A, 0.009 m). By putting $\ddot{x} = 0$, the constant k is found to be $mg\frac{i_0^2}{x_0^2}$. The system (3) is linearized, with the help of Taylor's series as [21, 22]:

$$\Delta\ddot{x} = \frac{\partial f(i,x)}{\partial i}\bigg|_{i_0,x_0} \Delta i + \frac{\partial f(i,x)}{\partial x}\bigg|_{i_0,x_0} \Delta x \tag{4}$$

where Δx and Δi are infinitesimal perturbations in position and current, respectively, with respect to the equilibrium point x_0 and i_0. Applying Laplace transform to (4), the following transfer function is obtained:

$$\frac{\Delta x}{\Delta i} = \frac{-k_i}{s^2 - k_x} \tag{5}$$

where $k_i = \frac{2g}{i_0}$ and $k_x = \frac{2g}{x_0}$. Since x and i are proportional to the voltage level of sensor output x_v and input level of control voltage u, the transfer function of (5) is expressed as

$$\frac{\Delta x}{\Delta i} = \frac{-k_1 k_2 k_i}{s^2 - k_x} \tag{6}$$

where k_1 is the coil gain and k_2 is the sensor gain. The parameters of the maglev plant of Feedback Instruments Ltd. (Model No. 33-210) are referred from [20]. Substituting the values, (6) may be expressed as [21, 22]:

$$G(s) = \frac{\Delta x_v}{\Delta u} = \frac{-3518.85}{s^2 - 2180} \tag{7}$$

The system (7) is open loop unstable. One pole is located to the right side of the s-plane. There is a need to design a proper controller to stabilize this plant and achieve satisfactory response.

3 The Proposed PID Controller

This work proposes a PID controller to stabilize the open loop unstable maglev plant described in Sect. 2. The modified PID has a series structure. The block diagram of the proposed structure is shown in Fig. 2. The transfer function of the proposed PID is expressed as follows:

$$G_{PID}(s) = K_c s(1 + \frac{1}{T_i s})^2 \tag{8}$$

where K_c is the controller gain and T_i is the integral time. The correlation between gains of the proposed controller and the conventional PID controller (parallel structure) $G(s) = K_p + K_i \frac{1}{s} + K_d s$ may be expressed as follows:

$$\begin{aligned} K_p &= \frac{2K_c}{T_i} \\ K_i &= \frac{K_c}{T_i^2} \\ K_d &= K_c \end{aligned} \tag{9}$$

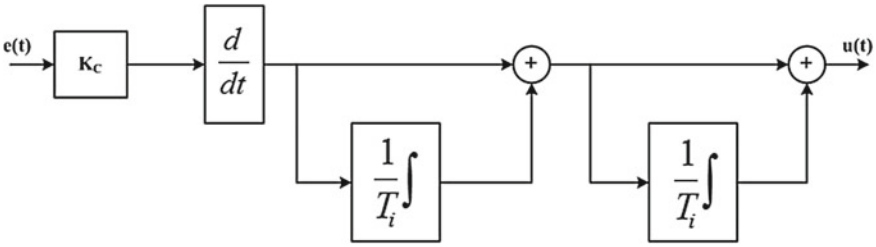


Fig. 2 Block diagram of the proposed structure of PID controller

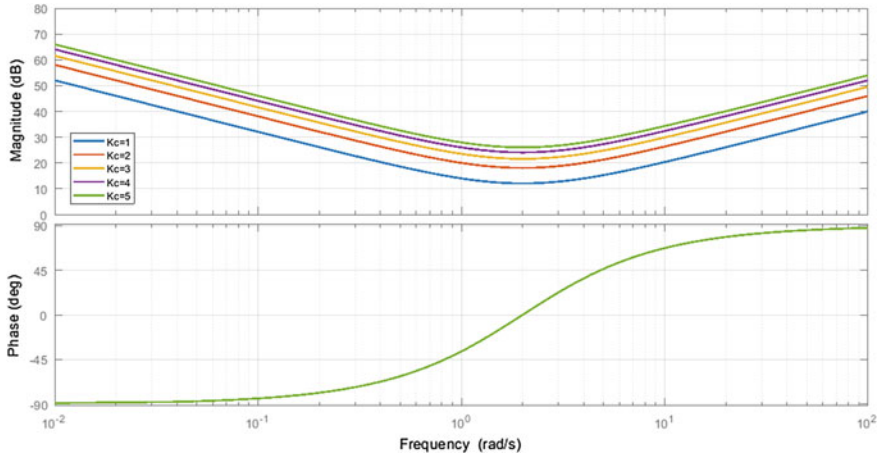


Fig. 3 Effect of variation in K_c

The advantage of (8) as a PID controller may be enumerated as:

1. Number of tunable parameters are reduced from three to two.
2. The double integral aids in faster elimination of steady-state error.
3. The zero has the advantage of increased relative stability.

The effect of changing the two parameters K_c and T_i is shown in Figs. 3 and 4, respectively. It is observed that K_c affects only the magnitude plot, whereas T_i affects both magnitude and phase plots.

Looking closely at Fig. 4, it can be inferred that variation in T_i has no effect in the high frequency range of the magnitude plot. But, as per Fig. 3, variation in K_c

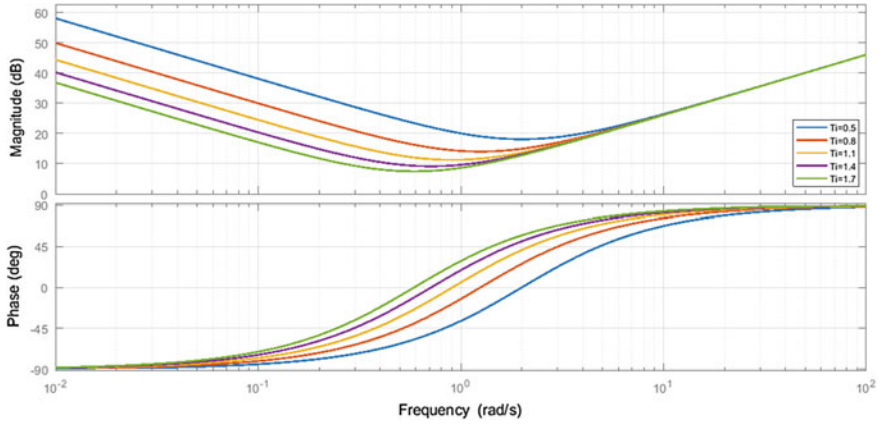


Fig. 4 Effect of variation in T_i

affects the magnitude in the entire range of frequency. Also, on observing the phase plot it is inferred that changes in T_i affect the phase plot in the mid-frequency range. It can thus, be stated that with proper combination of K_c and T_i, we can modify the entire range of the frequency response.

4 Symbiotic Organisms Search

The Symbiotic Organisms Search (SOS) is a metaheuristic optimization algorithm proposed by Prayogo and Cheng [23]. The algorithm mimics the symbiotic relations existing among various organisms to mutually survive in the ecosystem.

In the context of problem-solving, an *organism*, in SOS, signifies a *candidate solution* of the optimization problem. The algorithm works on three operators, namely, *mutualism, commensalism*, and *parasitism*. The first two operators help in exploration of the search space. The third operator is employed to eliminate the poor solutions, if any and reduce the chances of getting trapped in a local optima. A simple outline of the algorithm is presented below [23].

- Initialization
- Start
  - mutualism
  - commensalism
  - parasitism
- Repeat until termination criteria is fulfilled.

4.1 Mutualism

This operator mimics the relation between two organisms wherein both the organisms receive benefits from each other; thereby surviving in the ecosystem with mutual sharing. A relation between an Oxpecker and Rhinoceros is an example of mutualism. The bird feeds on the insects on the rhino's body; in turn, cleaning the animal. They both receive benefit from each other.

Since both organisms receive benefits, two new organisms are produced as per (10)–(12) [23]. In such a relation, however, the factor of benefit may vary depending on the organisms. Thus, a *benefit factor* is involved in such a relation.

$$v_{in} = v_i + rand(0, 1) \times (v_{best} - mv \times bf_1) \tag{10}$$

$$v_{jn} = v_j + rand(0, 1) \times (v_{best} - mv \times bf_2) \tag{11}$$

$$mv = \frac{v_i + v_j}{2} \tag{12}$$

where v_i and v_j are the two organisms participating in mutualism, v_{best} is the organism having best fitness in the ecosystem, mv is the mutual vector, rand(−1,1) is a uniformly distributed random number between 0 and 1. bf_1 and bf_2 are benefit factors which may be either 1 or 2 signifying half and full benefit, respectively.

4.2 Commensalism

This operator mimics the relation, between two organisms, in which one of the organisms receives benefits, from the other, without affecting the other. The relation between Shark and Remora fish is an example of such a relation. The remora attaches itself to the shark and feed on its leftovers. So, here, only the remora gets all the benefits, while the shark remains unaffected. Since only one organism receives the benefits, only one new organism is produced as per (13) [23]

$$v_{in} = v_i + rand(-1, 1) \times (v_{best} - v_j) \tag{13}$$

where v_i and v_j are the two organisms participating in commensalism, rand(−1,1) is a uniformly distributed random number between -1 and 1, v_{best} is the organism having best fitness in the ecosystem.

4.3 Parasitism

This operator mimics the relation between two organisms, in which one of the organisms receives benefits at the cost by harming the other organism. This operation may best be understood by taking an example of virus. A virus lives in the host receiving all the benefits from it, eventually killing the host.

In SOS, a *parasite vector* is created by modifying a random dimension of the i-th organism [23]. The fitness of the parasite vector is evaluated and compared to that of another organism (v_j) chosen randomly from the ecosystem. The parasite replaces v_j if its fitness is better.

5 Simulation and Results

The simulation block diagram is shown in Fig. 5. The transfer function shown in the figure is that of the magnetic levitation system. The transfer of the controller used in the simulation is that shown in (8). The closed loop characteristic equation of the system in Fig. 5 may be expressed as follows:

$$\delta(s) = T_i^2 s^3 + bK_c T_i^2 s^2 + (2bK_c T_i + T_i^2 a_0)s + bK_c \tag{14}$$

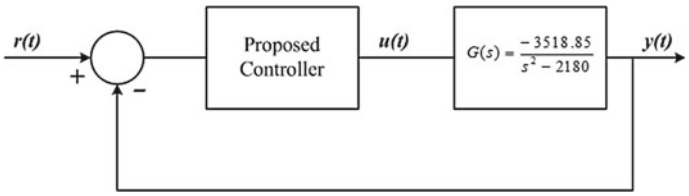


Fig. 5 Simulation block diagram

where $b = -3518.85$ and $a_0 = -2180$. For the closed loop system to be stable, the coefficients in (14) should be positive. So we get $T_i^2 > 0$, $bK_cT_i^2 > 0$, $(2bK_cT_i + T_i^2a_0) > 0$ and $bK_c > 0$. From the inequalities, we get $T_i > 0$ and $K_c < -\frac{T_ia_0}{2b}$.

$$\begin{aligned} T_i &> 0 \\ K_c &< -\frac{T_ia_0}{2b} \end{aligned} \tag{15}$$

Note that the DC gain of the maglev plant (7) is negative. It is thus, obvious the controller gains should be negative, so that the closed loop system be stable. For $T_i = 0.5$, $K_c < -0.155$. The range of K_c for a value of T_i is validated from the system responses shown in Fig. 6. It is seen that the system becomes unstable for $K_c = -0.1$ (violation of (15)) and remains stable for $K_c = -0.18$ (in compliance with (15)).

The effect of variation in K_c and T_i on the step response of the system is shown in Fig. 7. It has been observed that an increase in T_i reduces the settling time and rise time of the system and vice-versa. But, it increases the overshoot in the response. Decreasing the value of K_c introduces oscillations in the system response and also increases the overshoot in the response and vice-versa.

The PID controller has been tuned using the SOS algorithm. The population size and maximum number of iteration were set at 50 and 500, respectively. bf_1=1 and $bf_2 = 2$. The objective function employed for the purpose is

$$J(K_c, T_i) = t_s + M_p$$

where t_s is the settling time and M_p is the maximum overshoot. The best result of the SOS algorithm is given in Table 1. The table also presents a comparison of the proposed controller with conventional PID (parallel structure). The values of K_p, K_i and K_d has been calculated using (9). It is clear from the table that the Integral Square Error (ISE), Integral Absolute Error (IAE), settling time (t_s), and maximum overshoot (M_p) are lowest when the proposed controller is used. Figure 8 compares the system response of both the controllers, listed in Table 1. Usage of the proposed controller results 72% and 80% reduction in settling time and maximum overshoot, respectively. It has also been verified that the closed loop system is stable for the

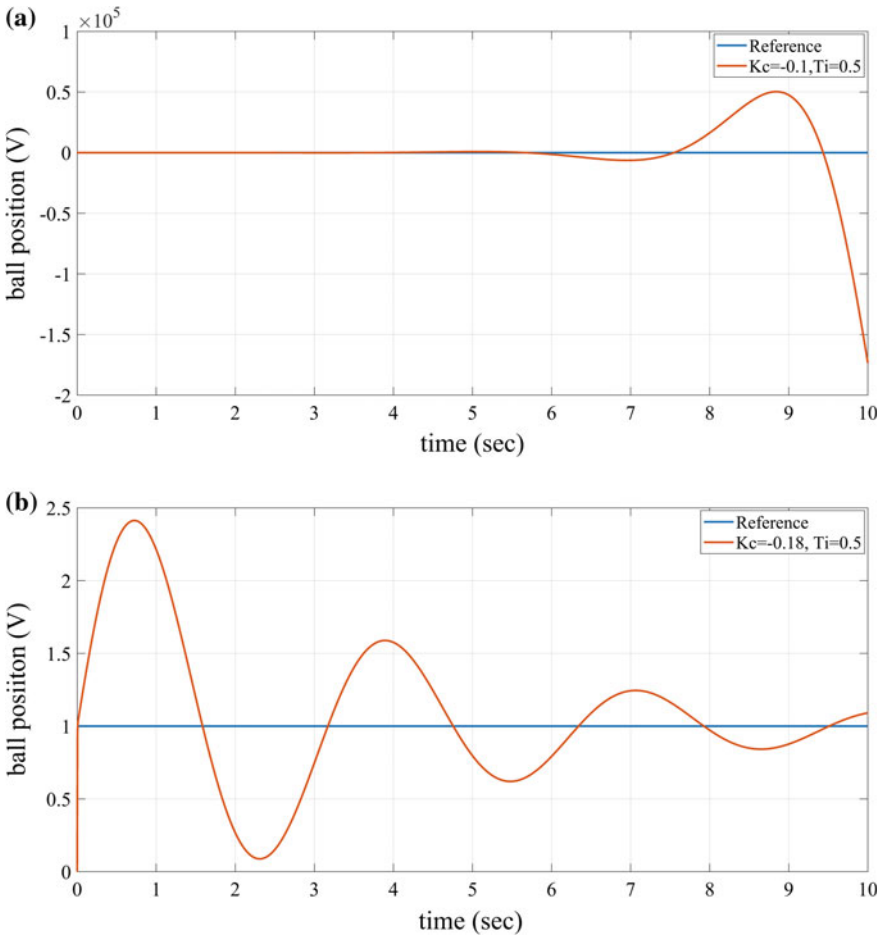


Fig. 6 Response of the maglev system with **a** unstable PID, **b** stable PID

values of K_c and T_i given in Table 1. The roots of the closed loop characteristic equation are -178.9, $-22 \pm j4.6$, located in the left side of the s-plane.

Figures 9 and 10 present the stability results using the Bode plot and Nichols chart of the system, respectively. The gain margin is $-$**29.7dB** and phase margin is **88.6**°. It should be noted that **although the gain margin is negative, the system is closed loop stable**. This is also confirmed by the results shown in Fig. 8.

A close look at the bode plot also reveals that the phase plot is nearly flat, in the vicinity of the gain crossover frequency, 1830 rad/s. This is the *iso-damping* property, which makes the system robust to gain variations. In other words, the proposed controller makes the system robust variations and/or uncertainty in system gain.

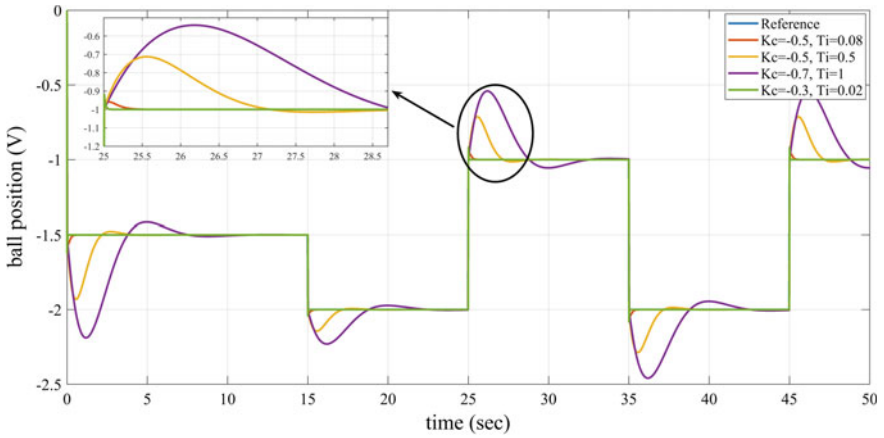


Fig. 7 System response for different K_C and T_i

Table 1 Result of tuning by SOS algorithm

Controller	K_p	K_i	K_d	K_c	T_i	t_s	M_p	IAE	ISE
Proposed	–	–	–	−0.521	0.045	0.067	2.83	0.00037	0.0039
Conventional	−23.15	−257.28	−0.5210	–	–	0.24	15.75	0.0352	0.0119

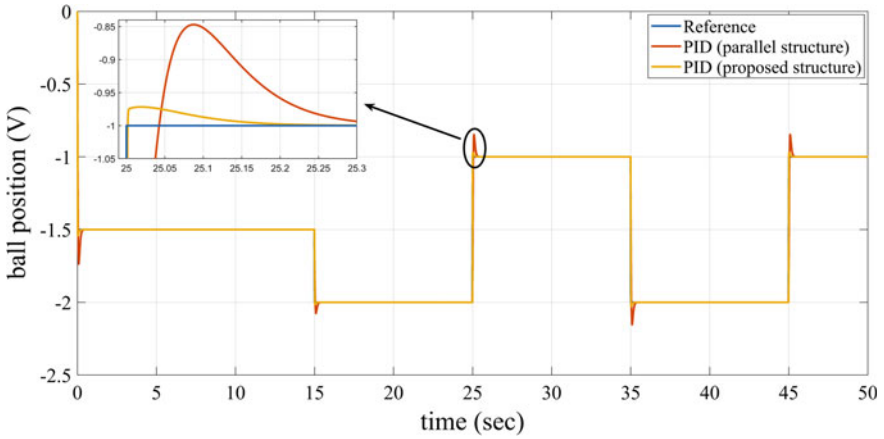


Fig. 8 Comparison of system response

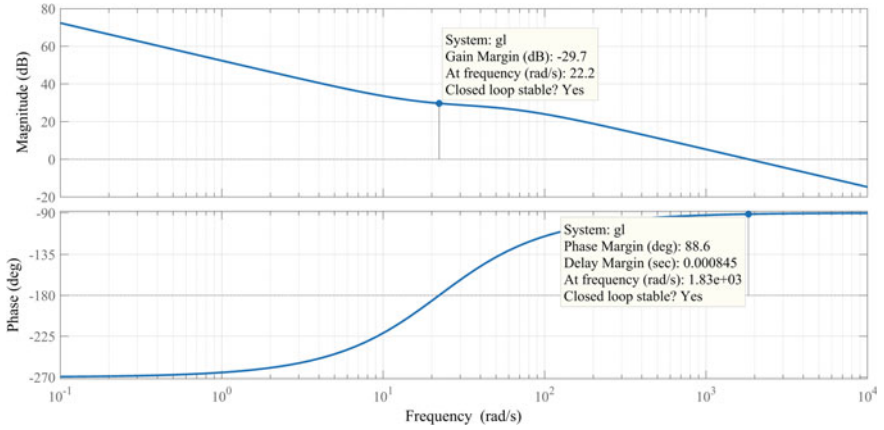


Fig. 9 Bode plot of the loop transfer function

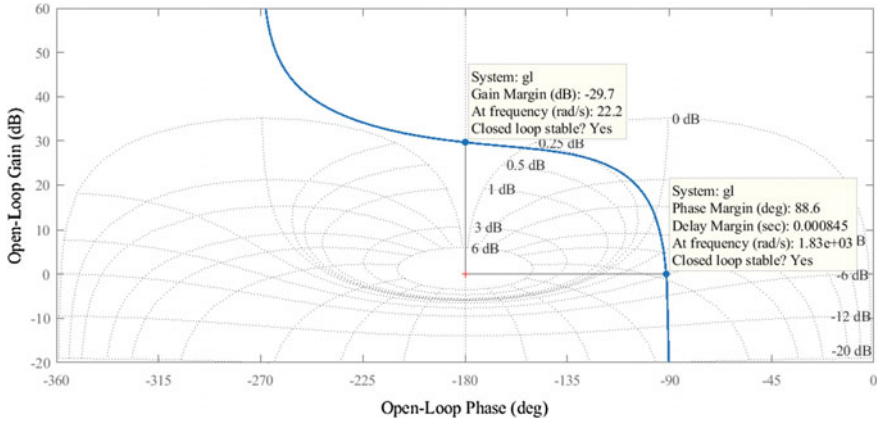


Fig. 10 Nichols chart of the loop transfer function

6 Conclusion

This work proposes a PID controller, which is a modified version of the series structure. The proposed controller has been applied to stabilize a second-order unstable plant. The plant chosen for this purpose is the magnetic levitation system. Advantage of using the proposed controller is the reduced number of tunable parameters and faster system response. This work also presents the application of SOS algorithm for tuning the controller. With proper tuning, the proposed controller gives stable and satisfactory response. Frequency domain analysis reveals that the system exhibits iso-damping property, thereby making the system robust to variations/uncertainty in system gain. Finally, the simplicity in tuning process and superior response, makes the controller suitable to be applied in other control/automation fields.

References

1. K.H. Ang, G. Chong, Y. Li, PID control system analysis, design, and technology. IEEE Trans. Control Syst. Technol. **13**(4), 559–576 (2005)
2. K.J. Åström, T. Hägglund, The future of PID control. Control Eng. Pract. **9**(11), 1163–1175 (2001)
3. H. Takatsu, T. Itoh, M. Araki, Future needs for the control theory in industries-report and topics of the control technology survey in Japanese industry. J. Process Control **8**(5–6), 369–374 (1998)
4. K.J. Åström, T. Hägglund, Revisiting the Ziegler-Nichols step response method for PID control. J. Process Control **14**(6), 635–650 (2004)
5. A.M. De Paor, M. O'Malley, Controllers of Ziegler-Nichols type for unstable process with time delay. Int. J. Control **49**(4), 1273–1284 (1989)
6. W.K. Ho, O. Gan, E.B. Tay, E. Ang, Performance and gain and phase margins of well-known PID tuning formulas. IEEE Trans. Control Syst. Technol. **4**(4), 473–477 (1996)
7. W. Hu, G. Xiao, X. Li, An analytical method for PID controller tuning with specified gain and phase margins for integral plus time delay processes. ISA Trans. **50**(2), 268–276 (2011)
8. C. Lorenzini, A.S. Bazanella, L.F.A. Pereira, G.R.G. da Silva, The generalized forced oscillation method for tuning PID controllers. ISA Trans. (2018)
9. L.J. da Silva Moreira, G.A. Júnior, P.R. Barros, Time and frequency domain data-driven PID iterative tuning. IFAC-PapersOnLine **51**(15), 1056–1061 (2018)
10. S. Saxena, Y.V. Hote, Internal model control based PID tuning using first-order filter. Int. J. Control Autom. Syst. **15**(1), 149–159 (2017)
11. B. Verma, P.K. Padhy, Indirect IMC-PID controller design. IET Control Theory Appl. **13**(2), 297–305 (2018)
12. F. Cao, PID controller optimized by genetic algorithm for direct-drive servo system. Neural Comput. Appl. 1–8 (2018)
13. T. Samakwong, W. Assawinchaichote, PID controller design for electro-hydraulic servo valve system with genetic algorithm. Procedia Comput. Sci. **86**, 91–94 (2016)
14. H. Freire, P.M. Oliveira, E.S. Pires, From single to many-objective PID controller design using particle swarm optimization. Int. J. Control Autom. Syst. **15**(2), 918–932 (2017)
15. K. Nimisha, R. Senthilkumar, Optimal tuning of PID controller for switched reluctance motor speed control using particle swarm optimization, in *2018 International Conference on Control, Power, Communication and Computing Technologies (ICCPCCT)* (IEEE, 2018), pp. 487–491
16. A. Rajasekhar, R.K. Jatoth, A. Abraham, Design of intelligent PID/PIλDμ speed controller for chopper fed DC motor drive using opposition based artificial bee colony algorithm. Eng. Appl. Artif. Intell. **29**, 13–32 (2014)
17. E. Çelik, R. Durgut, Performance enhancement of automatic voltage regulator by modified cost function and symbiotic organisms search algorithm. Eng. Sci. Technol. Int. J. **21**(5), 1104–1111 (2018)
18. M. Ghanamijaber, A hybrid fuzzy-PID controller based on gray wolf optimization algorithm in power system. Evol. Syst. 1–12 (2018)
19. S. Yadav, S. Verma, S. Nagar, Optimized PID controller for magnetic levitation system. IFAC-PapersOnLine **49**(1), 778–782 (2016)
20. M. Levitation, *Control Experiments Feedback Instruments Limited* (2011)
21. A. Ghosh, T.R. Krishnan, P. Tejaswy, A. Mandal, J.K. Pradhan, S. Ranasingh, Design and implementation of a 2-dof PID compensation for magnetic levitation systems. ISA Trans. **53**(4), 1216–1222 (2014)
22. S.K. Swain, D. Sain, S.K. Mishra, S. Ghosh, Real time implementation of fractional order PID controllers for a magnetic levitation plant. AEU-Int. J. Electron. Commun. **78**, 141–156 (2017)
23. M.Y. Cheng, D. Prayogo, Symbiotic organisms search: a new metaheuristic optimization algorithm. Comput. Struct. **139**, 98–112 (2014)

Review of Literature—Analysis and Detection of Stress Using Facial Images



Ayusha Harbola and Ram Avtar Jaswal

1 Introduction

1.1 Stress

Stress refers to how the body and mind react to everyday problems, challenges, and surroundings. It occurs due to inconsistency and incongruity between circumstantial demand and individual's ability. It consists of many interdependent and interconnected parameters (comprehensive, discernment, mental and physiological) that comprises person's reaction to varying internal and/or external parameters and psychological demands. "Dr. Hans Selye", described stress as "wear and tear of life." Stress is either negative or positive. It is positive when the situation provides a chance to achieve something and to stimulate oneself for high performance and results, whereas works as a negative force when a person deals problems socially, physically, and emotionally. In recent times, one cannot think of life without stress on educational as well as organizational fields.

Evolution and growth of Stress.

- Dr. Hans Selye (1920)—The term stress was borrowed from physics. In his study, he demonstrated stress as a nonspecific circumstance wide variety constituting numerous harmful negotiators.
- John Mason (1968)—measured stress hormone amount in a person undergoing various conditions that may be stressful according to him. He showed that

A. Harbola (✉) · R. A. Jaswal
Department of Electrical Engineering UIET, Kurukshetra University, Thanesar 136119, Haryana, India
e-mail: ayeshaharbola83@gmail.com

R. A. Jaswal
e-mail: ramavtar.jaswal@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_97

Table 1 Stress categorization

Stress category	Stimuli
Interpersonal	Quarrel with friends and parents, crisis in family or conflicts
Intrapersonal	Public speech, financial constraints, personal health issues
Academic	Workload, meeting deadlines, poor performances, fear of failure, competition with peers

Table 2 Causes of stress

Biologic stressor	Sickness or laceration
Natural or habitat stressor	Inadequacy, noise, and foulness
Comprehensive or perceptive stressor	Situation handling capacity
Personal performance or deportment	Lacking physical work
Life circumstances stressor	Casualty, legal dissolution, and situation with squint

stress hormone liberation is made feasible through neuroendocrine axis named hypothalamic-pituitary-adrenal (HPA) axis triggering.

- Walter (1984)—He was the first person to demonstrate the body reaction toward stress.
- Tucker (1996)—After encountering a stressful situation, hormonal level governs the mind-set, energy level, and depression of the person (Tables 1 and 2).

Behavioral outcomes Stress damages a person internally, but there are external indications too. Some of the commonly seen behavioral effects are

- Fights and arguments on relatively unimportant issues.
- More dependence on others as well as communication gap.
- Less attention toward family.

Cognitive outcomes Stress and functioning of mind are correlated. For subconscious work, such as heed, learning, solving problems, and artistry, moderate stress levels are considered optimal. When stress is at lower level, one fails to pay attention (may show all the signs of apathy, dullness, lack of interest, torpor, stagnation), and high distortions of perceptions may occur at high levels. Distortions are

- More consciousness paid toward obstructive strand of life and work.
- Continuous worry and anxiety, resulting in poor concentration.
- Problems in recovering/ recall from memory.
- Narrowing span of attention.

1.2 Objective of the Study

A deep study on the research areas which have been carried out in the field of stress and detection using facial images recognition.

1.3 Methodology

The study is thoroughly based on secondary data. In this regard, various books were reviewed and few online journals were also assessed in this direction.

2 Literature Review

On observing the present lifestyles, various questions regarding stress and arguments have been cropped up in the researcher's minds. Having deep knowledge of the factors leading to stress and angles related to it is very crucial. The present paper aims at the glimpse of the research work, so far done in stress and its analysis using facial frames and examining the stress level.

Foregoing works on stress investigation concentrate on gathering the study of physiological information and recognizing the association between stress and several physiological traits with selecting features and optimization techniques. Most previous work on stress detection is based on digital signal processing, taking into consideration Galvanic skin response, blood volume, pupil dilation, and skin temperature. And, the other work on this issue is based on several physiological gestures or cues and optical traits such as eye-opening or closure and motion of head to observe and examine level of stress in any individual while he is working or in daily life routine. Hans Selye in 1936 borrowed stress from field of engineering and talked about stress as being a nonspecific phenomenon representing wide variety of noxious agents. In 1968, John Mason, in his study stress hormone levels in an individual was subjected to different states that would be stressful in his thought. He showed that stress hormone liberation is made feasible through neuroendocrine axis named hypothalamic-pituitary-adrenal (HPA) axis triggering. In 1975, Selye and Mason combinedly provided path to further researchers which led to promote studies in stress which established the study regarding determinants of stressors and their responses that proved to be specific, potentially predictable and measurable.

In September 2015, Manolis Tsiknakis elaborated a combined data of facial cues for stress and anxiety identification. In his study, he discussed the method of stress and anxiety evaluation using facial signals consisting of mouth movement, head rotation, heart rate, and rate of blinking and eye pupil movements. This paper employs methods that aim each facial region one by one and were able to recognize prompt signs of stress and anxiety. 73% of accuracy was achieved. Another work in November 2015

was demonstrated, in which stress and anxiety were examined separately and also illustrated the use of baseline personnel and various methods using facial expressions were used.

In 2017, authors Kan Hong and Guodong Liu explored a noninvasive method to access stress in real time. Investigation of correlation among the thermal imprints and established stress markers was done in this study with 20 subjects. K-means method was used to perform cluster (for achieving ROI signal). A FLIR thermal IR imager SC7600 with noise equivalent temperature difference of 17mK working in usable spectral ranges was utilized. FFT was used to get or achieve a pulse signal. Correlation comes out to be 96%.

In 2017, author G. Giannakakos developed a structure for detecting and analyzing stress/anxiety emotional conditions via video-recorded facial expressions. An establishment of an experimental rule was done to persuading systematic variables in pretentious states (neutral, relaxed, and stressed/anxious) through a diversity of intramural and extramural stressors. The study focuses on deliberate and semi-deliberate facial signs importantly for denoting the emotion rendering more rationally. Features under analysis were eye-related, mouth movement activity, head movement factors, and heart rate calculation by a camera. To select the most robust features, the feature selection method was employed followed by classification schemes differentiating stress/anxiety and neutral states taking reference a relaxed state in each investigating phase. The results showed that the facial signs, acquired from an eye movement, mouth activity, head movements, and heart movement achieve nice accuracy level and are relevant as discriminative denotations of stress and anxiety.

In 2013 authors Nisha Raichur, Nidhi Lonakadi, and Priyanka Mural explained the main causes of stress and evolved a stress detection layout showing inspection of the facial signs or expression. A nonintrusive layout is taken in the study. The front view of the subject's face is captured by a camera with the required pixel while working in front of the computer. The video taken is classified into three parts of equal length and set of an equal number of image frames are extracted from each section correspondingly and are analyzed. Scanning of an image is done for eyebrow coordinates by calculating the displacement of an eyebrow from its position. If the person is found stressed in the consecutive sections of the time interval, the stress detection decision is made for a person working in front of the system. Theano, which is a python framework aimed at improving both the execution time and development time.

Since past years, studies are made to correlate face features with stress and emotions. Suwa et al., in 1978 introduced facial expression classification system. Image processing, extracting features, detecting face, and classification are the main building block for detection and classification of facial expression system. Contractions of official muscles, resulting in temporal deformation of facial cues like eyes, nose, lips, and texture of skin, eyebrows movement, sometimes by revelation by face wrinkles and eye bulges, generate facial signs analyzing as well as deformations. Numerous facial cues detecting methods from consistent images are taken in literature. Ekman and Friesen evolved Facial Action Coding System (FACS) for explaining the articulations like stress, revulsion, neutral, grief, and afflicted. Padgent, Hara and Kobayashi,

Zhang and Zhan used neural network method for facial eloquence detection. In this analysis, he operated classification of the face looks into three sections with image processing feature extraction procedure and multilayered artificial neural network (ANN) as classification.

In June 2014 authors Samin Iftikhar, Rabia Youns, Noshaba Nasir, and Kashif Zafar presented a system depending on multilayered artificial neural network for the apprehension and analysis of frontal facial models. Preprocessing, face discovery, stresses extraction, and emotion disclosure are the stages via which mechanism is carried out. The emotion experimented is astonished, felicitous, and neutral. The entire average precision obtained of the class is 86.1% and algorithm used was backpropagation. This paper differentiates the stress level and emotions. Average correctness of the distribution is 86.1% and the misclassification or failure rate comes out to be 13.9%.

In 2010, authors Raheja, Jagdish Lal, and Umesh Kumar described architecture for human face expression detection. They studied individual facial eloquence from the taken picture using backpropagation network. The study comprises five stages: image acquiring, face detection, image preprocessing, neural network and recognition. Enhancing image, edge detection, thinning, and tokenization come under image processing. The image gets converted to neural network input after the thinning process in the feature extraction method. Backpropagation algorithm is used to implement neural network. Gaurav B. Vasani et al. in 2013 studied the face appearance detection using principal component analysis (PCA). In this paper, the images of four training models and six testing models are taken for the study. The identical approach was reiterated by reducing the number of testing images and raising the number of training images from a specific collection of expression. The train images create a low dimensional face space. On the training image set, component analysis (PCA) is applied and principal components are taken to reduce the eigenspace. The input test images were classified by using Euclidian distance. The technique was tested by taking facial expressions from 30 subjects with also the limited quantity of training images the apprehension rate was higher which verified that it is easy, fast, and operates greater in a stifled position.

In 2012, facial expressions of depressed person were studied. Having the volunteer's feedback in thought, attempts will be done to enhance the document model and present the speech model computationally effective. The project's main aim implies to motivate people and prevent them from taking drastic steps such as suicide. This further purpose at distinguishing oneself who is experiencing indigenous abuse or any additional kind of difficulty that people are not suitable for sharing. A feature that is in being promoted in this design is to accommodate periodic status summaries of the user's mental health to the Doctor, whose details would be presented by the user.

M. Pediaditis and G. Giannakakis, in their study employed various techniques that target specific facial area personally. The features have been shortlisted and the execution is estimated on the basis of a dataset of 23 cases with an overall precision of 73%. This paper illuminates a passageway of stress and anxiety evaluation employing facial signs consisting of the mouth action, head movement, heart pace, and blink

rate and eye gestures. Section 2 explains the schemes used for the extraction of the traits associated with these symbols. The ROI disclosure is based on the Viola–Jones detection algorithm, as performed in OpenCV 2.4.8, which utilizes a rejection cascade of hoisted classifiers. A motion filter is used that reduces irregular trajectories and fractionable points. The mean, median, and standard deviation of X distance are traits obtained from the filtered point trajectories. Data mining software Weka v3.7.12 was employed for labeling experiments and ANN classifier was practiced.

In 2013, Surbhi and Vishal Arora proposed facial appearance detection from human facial image by employing a neural network. In this, the image database contains numerous expressions of various poses. An input image is read from the database and expression face is localized. The suggested system can automatically identify human facial features expertly than any additional method for ranking of facial appearances that is, hence, experimentally proved. The recognition rate and backpropagation algorithm accuracy is 100% on the Japanese female facial expressions database (JAFFE) which has 23 images of the identical person.

Mira Chandra Kirana, S. T., M. T., and Yogsal Ramadan Putra in their study used principal component analysis method for feature extraction to recognize the traits of every facial picture data. To recognize the stressed facial picture and the normal picture they used surveys on depression anxiety stress scale (DASS). At the concluding stage, matching of the obtained normal face and stressed face was made. The outcomes of the studied graph of the image information on the stressed face has a percentage amount above normal image amounted to 50.38% and the percentage under the average image amounted to 49.63%, while the data of natural face have a percentage over ordinary image to 43.81% and the percentage image under the average to 56.19%. Depending on the investigation of the model of the facial model data, the recognized stress, irregular pattern of about 60% was computed, while an average identifiable face picture has an unusual model percentage of 80%. In 2016 Camille Daudelin-Peltier1, Hélène Forget in their work studied the impact of social stress on the identification of facial expressions in a healthful young man. Members were tested by a (TSST-G) regulated psychosocial laboratory stressor as well as the manual control condition.

Ira Cohen, Thomas S. Huang, and Ashutosh Garg in their work focused on automated facial eloquence identification from live video input utilizing temporal signs. Schemes for template matching employed dynamic programming techniques and hidden Markov models (HMM). This study gave a new layout of HMMs by exploiting existing ways or methods. Both the segmentation and identification of the facial appearances were made automatically employing a multilevel HMM structure while enhancing the discrimination ability among the different classes. In this work, person-dependent and person-independent identification of appearances were investigated. The main drawbacks in all of the tasks done on emotion or stress identification from facial cues videos face the absence of a benchmark database to examine various algorithms. Further, Kanade et al. created a database that proved to be a valuable mechanism for testing certain algorithms. They further provided a beneficial continuation of their task by giving the idea of constructing a real-time method that

contains a quick and reliable face tracking algorithm coupled with the multilevel HMM formation by which a better interaction can be achieved.

In 2013 Parul Sood, Sushri Priyadarshini, Palok Aich in their study did a questionnaire and profiled molecular-based assessment of PS and calculated the effectiveness in a random group of people. Utmost other considerations on psychological stress are also addressed with the population which is projected with similarity in age, gender, ethnicity, including commercial growth while presumably, these studies undergo significant limitations that the inequality among individuals in a group may not be separated out with certainty. They grouped score arrangement for both methodologies in three as stressed (S), non-stressed (NS), and borderline (BL). This grouping underwent discriminant grouping to confirm the classification by grouping metabolite data independently. In 2009, Jeffery F. Cohn, Tomas Simon Kruez, Ying Yang in their study did clinical analysis of higher stress and depression with automatically graded facial performances in patients experiencing medication for depression. FACS coding, active appearance modeling, and pitch wrenching were used to extract and investigate facial stress and vocal eloquence. SVM classifiers for FACS and AAM were incorporated during the study. Face plus voice both displayed average cooperative validity with stress and depression. Precision was about 88% for hand operated FACS and 79% during AAM modeling. Vocal prosody certainty was about 79%.

Vinay Bettadapura in his paper discussed facial parameterization applying FACS action units (AUs), MPEG-4 facial animation parameters (FAPs) and the modern face detection, tracking plus trait extraction methods. He further studied emotions, feelings, eloquence, and facial traits and discussed some six prototypic appearances and the current investigations on expression classifiers. Pardas and Bonafonte studied facial features for detecting face reactions. Their work reveals that surprise, delight, and hatred have pretty high identification rates (of about 100%, 93.4%, and 97.3%, respectively) because of the fact that they include clear movement of the mouth including the eyebrows. They also justified that the mouth carries higher information than that of the eyebrows. The experiments they accompanied with simply mouth being noticeable granted a recognition efficiency of about 78%, whereas tests which were carried out by the eyebrows only provided a recognition accuracy of only 50%. Bourel et al. Kotsia et al.'s worked on the consequence of occlusions on facial expression recognizers and noted that this occlusion of the mouth decreases the output result over more than 50%. This quantity ideally matches the outcomes of Pardas and Bonafonte. Kotsia et al. further confirmed that occlusions at the left half or the right half of every face do not alter the overall execution. This is so because the facial appearances adjust symmetric at the perpendicular plane that partitions the face in left and right halves.

Kanade–Lucas–Tomasi tracker was among the leading approaches manifested in the 1990s to identify and trace features. Lucas and Kanade proposed the fundamental structure of the method and then were further developed by Tomasi and Kanade. In 1998, Kanade including Schneider man developed a strong face detection algorithm employing statistical techniques, which has been practiced extensively after its recommendation. Viola and Jones in 2004 developed a technique practicing AdaBoost

learning algorithm that proved to be quite fast and could identify frontal view faces quickly. With the use of novel methods that computed the facial features very quickly, they were able to achieve excellent results and separate the background from the face within no seconds. Few other detections and tracking techniques were recommended by Sung and Poggio in 1998, Rowley et al. during 1998 and Roth et al. near 2000 which proved effective during detection.

During 2000, Bourel et al., an advanced version of the Kanade–Lucas–Tomasi tracker was introduced, which was termed as the EKLT tracker (Enhanced KLT tracker). They practiced the arrangements and visual properties of the face to practice the KLT tracker in an extended form which presented the KLT tracker strong toward transient occlusions and approved recovery from the lack of features produced by tracking drifts. Another face trackers are the ones based on Kalman Filters, extended Kalman Filters, Mean-Shifting, and mostly Particle Filtering. Tian et al. in 2001 in his study recognizes facial features and expressions in a real-time system using head motion and used facial feature tracker to reduce processing time. He practiced Optical Flow, the Gabor Wavelets, Multistate Models, and Canny edge detection methods for feature extraction. Cohn-Kanade and Ekman–Hager facial action database were taken in his study. For the uppermost face, 50 specimen sequences of 14 subjects and for lower face 63 specimen sequences of 32 subjects were exerted.

Bourel et al. in 2001 in his study experimented with facial signs during occlusions and proposed utility of modular classifiers rather than uniform classifiers. He did distribution sectionally and then joined the classifier results. Cohn-Kanade database was implemented for the study. Feature extraction was done by local spatial-temporal vectors obtained from the EKLT tracker. He took 30 subjects, 25 sequences for $ expressions for the study. In 2002, Pardas and Bonafonte studied automatic extraction of MPEG-4FAPs for extracting facial features. An advanced active contour algorithm and action evaluation were used for extracting MPEG-4FAPs. The overall efficiency of about 84% was calculated.

In 2003, Cohen et al. in his study worked on the real-time system and expression classification from the video for further detecting stress level and emotions and used semi-supervised training to operate with labeled data including a large quantity of unlabeled data. A vector of extracted motion units using PBVD tracker was used for feature extraction. 53 subjects including the Cohn-Kanade database was used for the study purpose.

In 2003, Bartlett et al. studied a fully automatic system and successfully deployed real-time identification at a large level of precision on a pet humanoid and CU Animator. Cohn-Kanade database was used including 313 sequences from 90 subjects for the study. SVM classifier was used with AdaBoost and for feature extraction, Gabor wavelets were implemented and accuracy reached was 84.8%.

In 2003, Michel and Kaliouby studied real-time system without any preprocessing. He used 10 examples for training and 15 examples for testing purpose. Cohn-Kanade database was used followed by SVM Classifier and vector of trait displacements (Euclidean distance separating neutral and peak) for feature extraction. In 2004, Pantic and Rothkrantz recognize facial expressions in frontal and profile view for further calculating the stress level and expression. In his paper, he suggested a system

to do automated AU coding in profile pictures. He used the MMI database including 25 subjects. A rule-based classifier with frontal and profile facial points was used for extracting required facial features with an accuracy of about 86% was reached.

Buciu and Pitas in 2004 studied PCA and also performs its comparison with local non-negative matrix factorization (LNMF) and non-negative Matrix factorization (NMF). He concluded that LNMF exceeded both PCA and NMF, while NMF returns the worst performance among all. He further compared CSM and MCC classifiers and concluded that CSM classifier implies more substantial results than MCC plus provides better identification of facial features and expressions. Cohn-Kanade and JAFFE database were compared which resulted in the conclusion that Cohn-Kanade LNMF including MCC provided the largest accuracy of 81.4% moreover JAFFE for all three methods only gave 55–68%.

In 2006, Zheng et al. in his study employed KCCA to identify facial appearances and improved singularity difficulty of the Gram matrix by KCCA algorithm. JAFFE dataset was used for the study and the correlation that is learned is practiced to evaluate semantic appearance vector which is later applied for classification. 34 position points transformed into a marked Graph (LG) applying the Gabor wavelet transform was employed for feature extraction.

In 2006, Anderson and McOwen studied a fully automated and multistage system that was authorized to operate efficiently within cluttered views. Motion equating is done to consolidate the data that is served to the classifiers SVM and MLP. CMU-Pittsburg AU coded including a nonrepresentative database was taken for the study. For feature extraction motion marks taken by tracking applying spatial ratio template tracker and making optical flow on every face using a multichannel gradient model (MCGM) was used by the researchers.

In 2006, Pantic and Patras studied automated segmentation of input video into facial appearances and automatic identification of AUs from profile images. The study was about 86.6% efficient on 96 test profile sequences. 1500 specimens of both inactive and profile views were examined and rule-based classifier was used. The tracking of 15 facial points using particle filtering was used for extracting features generated mid-level parameters.

In 2007, Sebe et al. recognized spontaneous facial expressions for studying several parameters involved with facial cues. He built a genuine database where subjects are displaying their normal facial emotions and assessed numerous machine learning algorithms. He created a spontaneous images database and also used the Cohn-Kanade database. MUs generated from the PVBD tracker is used for extracting required features. SVM and decision tree classifiers and voting algorithms and for improving results bagging and boosting were used. Cohn-Kanade showed accuracy from 72.46 to 93.06% and the created database accuracy was from 86.77 to 95.575%. In 2007, Kotisa and Pitas in their study identify both the fundamental facial expressions and the set of preferred AUs with a pretty high recognition speed. Cohn-Kanade database is utilized for the research and the study used geometric displacement regarding Candide nodes for feature extraction. SVM was used to recognize facial expression. The study concluded with 99.7% efficient during facial

appearance identification and 95.1% during facial expression recognition based on AU detection.

In 2007, Wang and Yin examined the robustness toward the falsification of the identified face area and the various concentrations of facial appearances. 53 subjects and 4 images per subject were taken for every facial expression from the Cohn-Kanade database also from MMI 5 different subjects, 6 images per subject during each facial cue was taken. Topographic context (TC) expression descriptors for extracting required features and QDC, LDA, SVC, and NB classifiers are used. Person-dependent test results using MMI are as follows: by QDC: 92.78%, by LDA: 93.33%, by NB: 85.56%, using Cohn-Kanade: by QDC: 82.52%, by LDA: 87.27%, by NB: 93.29%. Person-independent tests using Cohn-Kanade are as follows: by QDC: 81.96%, by LDA: 82.68%, by NB: 76.12%, by SVC: 77.68%.

In 2008, Domaika and Davoine proposed a structure for simultaneous face tracking including expression identification and used several video sequences for frames where subjects were permitted to represent any facial expression in any method for any continuation. Using online appearance models, first head pose is determined and then expressions are identified applying a stochastic strategy. Candide face model is employed to track traits or features. Kotsia et al. in 2008 developed an arrangement to identify expressions in spite of occlusions and examines the impact of occlusion on each six prototypic facial appearances. Gabor features, DNMF algorithm including geometric displacement vectors are obtained applying Candide tracker. Subsequent results were achieved—Applying JAFFE: by Gabor: 88.1%, by DNMF: 85.2% Applying Cohn-Kanade: by Gabor: 91.6%, by DNMF: 86.7%, by SVM: 91.4%.

3 Conclusion

This literature review and the studies have given an overall acquaintance of the stress and the various factors related to it. In nutshell it can be stated that currently, the stress has got its deep roots and clearly can be seen in the person's face. The further research will surely be helpful to unleash the various techniques to analyze and detect stress as to enrich the harmony and stability among human beings so that it does not cause depression which further leads to severe issues by detecting it in time.

References

1. M. Pediaditis, G. Giannakakis, F. Chiarugi, D. Manousos, A. Pampouchidou, E. Christinaki, G. Iatraki, E. Kazantzakis, P.G. Simos, K. Marias, M. Tsiknakis, Extraction of facial features as indicators of stress and anxiety
2. K. Hong, G. Liu, Facial thermal imaging analysis for stress detection (2017). N. Raichur, N. Lonakadi, P. Mural, Detection of stress using image processing and machine learning techniques (2013)

3. S. Iftikhar, R. Younas, N. Nasir, K. Zafar, Detection and classification of facial expressions using artificial neural network (2014)
4. J.L. Raheja, U. Kumar, Human facial expressions detection from detected in captured image using back propagation network (2010)
5. G.B. Vasani, Human state recognition using facial expression detection (2013)
6. V. Vanitha, P. Krishnan, Real time stress detection system based on EEG signals (2016)
7. M.Z. Poh, D.J. McDuff, R.W. Picard, Non-contact, automated cardiac pulse measurements using video imaging and blind source separation. Opt. Express **18**(10), 10762–10774 (2010)
8. E. Christinaki, G. Giannakakis, F. Chiarugi, M. Pediaditis, G. Iatraki, D. Manousos, K. Marias, M. Tsiknakis, Comparison of blind source separation algorithms for optical heart rate monitoring, in *Proc*eedings *of the EAI 4th International Conference on Wireless Mobile Communication and Healthcare (Mobihealth)* (2014), pp. 339–342, 3–5 Nov 2014
9. N.S. Thejaswi, S. Sengupta, Lip localization and Viseme recognition from video sequences, in *National Communications Conference (NCC)*, Mumbai, India (2008)
10. A. Pampouchidou, M. Pediaditis, F. Chiarugi, K. Marias, F. Meriaudeau, F. Yang, M. Tsiknakis, Mouth activity recognition based on template-matching and eigen-features, in *International Conference on Image Analysis and Processing* (2015)
11. J. Shi, C. Tomasi, Good features to track, in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'94)* (1994), pp. 593–600
12. C. Daudelin-Peltier, The effect of acute social stress on the recognition of facial expression of emotions (2017)
13. V. Bettadapura, Face expression recognition and analysis: the state of the art (2012)
14. J.F. Cohn, T.S. Kruez, Y. Yang, Detecting depression from facial actions and vocal prosody (2009)
15. P.R. Ekman, *What the Face Reveals: Basic and Applied Studies Spontaneous Expression Using Action Coding System (FACS)* (Oxford University Press, New York, 1997)
16. I. Kotsia, I. Buciu, I. Pitas, An analysis of facial expression recognition under partial facial image occlusion. Image Vis. Comput. **26**(7), 1052–1067 (2008)
17. MPEG Video and SNHC, Text of ISO/IEC FDIS 14 496-3: audio, in *Atlantic City MPEG Mtg.*, Doc. ISO/MPEG N2503 (1998)
18. M. Pardas, A. Bonafonte, Facial animation parameters extraction and expression recognition using Hidden Markov Models. Sig. Process. Image Commun. **17**, 675–688 (2002)
19. M.A. Sayette, J.F. Cohn, J.M. Wertz, M.A. Perrottand, D.J. Parrott, A psychometric evaluation of the facial action coding system for assessing spontaneous expression. J. Nonverbal Behav. **25**(3), 167–185 (2001)
20. M.C. Kirana, Y.R. Putra, Comparison of facial feature extraction on stress and normal using principal component analysis method
21. I. Cohen, A. Garg, T.S. Huang, Emotion recognition from facial expressions using multilevel HMM
22. V. Bettadapura, Face expression recognition and analysis: the state of the art
23. Y. Tian, T. Kanade, J. Cohn, Recognizing action units for facial expression analysis. IEEE Trans. Pattern Anal. Mach. Intell. **23**(2), 97–115 (2001)
24. I. Cohen, N. Sebe, F. Cozman, M. Cirelo, T. Huang, Learning Bayesian network classifiers for facial expression recognition using both labeled and unlabeled data, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR),* vol. 1 (2003), pp. I-595–I-604
25. I. Cohen, N. Sebe, A. Garg, L.S. Chen, T.S. Huang, Facial expression recognition from video sequences: temporal and static modeling. Comput. Vis. Image Underst. **91**, 160–187 (2003)
26. M.S. Bartlett, G. Littlewort, I. Fasel, R. Movellan, Real time face detection and facial expression recognition: development and application to human computer interaction, in *Proceedings of the CVPR Workshop on Computer Vision and Pattern Recognition for Human-Computer Interaction,* vol. 5 (2003)
27. P. Michel, R. Kaliouby, Real time facial expression recognition in video using support vector machines, in *Proceedings of the 5th International Conference on Multimodal Interfaces,* Vancouver, BC, Canada (2003), pp. 258–264

28. M. Pantic, J.M. Rothkrantz, Facial action recognition for facial expression analysis from static face images. IEEE Trans. Syst. Man Arid Cybern. Part B **34**(3), 1449–1461 (2004)
29. I. Buciu, I. Pitas, Application of non-negative and local non negative matrix factorization to facial expression recognition, in *Proceedings of ICPR*, Cambridge, U.K., (2004), pp. 288–291, 23–26 Aug 2004
30. W. Zheng, X. Zhou, C. Zou, L. Zhao, Facial expression recognition using kernel canonical correlation analysis (KCCA). IEEE Trans. Neural Netw. **17**(1), 233–238 (2006)
31. K. Anderson, P.W. McOwan, A real-time automated system for recognition of human facial expressions. IEEE Trans. Syst. Man Cybern. Part B **36**(1), 96–105 (2006)
32. M. Pantic, I. Patras, Dynamics of facial expression: recognition of facial actions and their temporal segments form face profile image sequences. IEEE Trans. Syst. Man Cybern. Part B **36**(2), 433–449 (2006)
33. N. Sebe, M.S. Lew, Y. Sun, I. Cohen, T. Gevers, T.S. Huang, Authentic facial expression analysis. Image Vis. Comput. **25**, 1856–1863 (2007)
34. I. Kotsia, I. Pitas, Facial expression recognition in image sequences using geometric deformation features and support vector machines. IEEE Trans. Image Process. **16**(1), 172–187 (2007)
35. J. Wang, L. Yin, Static topographic modeling for facial expression recognition and analysis. Comput. Vis. Image Underst. **108**, 19–34 (2007)
36. F. Dornaika, F. Davoine, Simultaneous facial action tracking and expression recognition in the presence of head motion. Int. J. Comput. Vis. **76**(3), 257–281 (2008)
37. P. Viola, M.J. Jones, Robust real-time object detection. Int. J. Comput. Vis. **57**(2), 137–154 (2004)
38. H. Schneiderman, T. Kanade, Probabilistic modeling of local appearance and spatial relationships for object recognition, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (1998), pp. 45–51
39. B.D. Lucas, T. Kanade, An iterative image registration technique with an application to stereo vision, in *Proceedings of the 7th International Joint Conference on Artificial Intelligence (IJCAI)* (1981), pp. 674–679
40. C. Tomasi, T. Kanade, Detection and tracking of point features, in *Carnegie Mellon University Technical Report CMU-CS-91-132* (1991)
41. N. Sebe, I. Cohen, A. Garg, T.S. Huang, *Machine Learning in Computer Vision,* 1st edn. (Springer, New York, 2005)
42. http://www.icg.isy.liu.se/candide/main.html

A Survey on Text Detection from Document Images



M. Ravikumar and G. Shivakumar

1 Introduction

The problem of text information extraction needs to be defined more precisely before processing further. Text Information Extraction (TIE) system receives on input still image and sequence of images. The images can be in gray scale or color, compressed, or uncompressed and the text in images may or may not move. The problem of text extraction can be divided into the following steps. They are: (1) detection (2) localization (3) tracking (4) extraction and enhancement (5) recognition (OCR).

The presence of text data in images and videos containing information that is useful for clearing automatically, indexing, and structuring an image. Nevertheless, because of differences changes the text size and style, orientation and alignment, and low image contrast and complex background to the problem of automation text extraction extremely challenging although a comprehensive survey of related problem such as face detection, analysis, documents, and indexing images and videos can be found. The problem of extracting the information is not well explored.

2 Related Work

In this section, we discuss the related work on text extraction/detection from both printed as well as handwritten document images.

Text detection and extraction can be achieved using the feature extraction methods, namely, the Haar and Discrete Wavelet Transform (DWT) is proposed [1]. Various text regions such as horizontal edge, vertical edge, and diagonal edges from compound images and complex background images (collection of graphs, tables, charts,

M. Ravikumar · G. Shivakumar (✉)
Department of Computer Science, Kuvempu University, Shimoga 577451, Karnataka, India
e-mail: g.shivakumarclk@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_98

etc., and also text) are removed by using Field Programmable Gate Array [FPGA] technique. Text and non-text and skewed regions are also removed using mathematical morphological operations. Comparative analysis of various text extraction methods is discussed. The average time taken for M-band is 286 s, Cluster center approach is 158 s, Fisher classifier is 139 s, Haar and DWT is 96 s.

Using different text to extraction methods like region-based method, edge method, texture method, morphological-method, text from an image is extracted [2]. After the detailed survey given on comparison and performance evaluation and different text extraction methods, it is found that region and text-based methods give poor result compared with the morphological and edge-based methods.

By using edge-based and K-means clustering algorithm, text extracted from live captured image is with diversified background [3]. K-means clustering algorithm is performed on dataset, which partitions into group according to some distinct distance measure. Non-text region from an image is removed using morphological operations. After the experimentation the overall precision rate and recall rate result was compared with the edge-based algorithm and connected component-based algorithm is 47.4–75.09% and 50.10–73.42% is obtained.

A two-dimensional wavelet transform used for text extraction is proposed [4]. For the purpose of classification of the images into text region, simple background region, and complex background region, k-means clustering algorithm is used. After the classification is performed, clustering is done using morphological operation. Experimentation is carried out on 100 diverse gray-scale pictures, which contain content data with distinctive languages, textual styles, and sizes. The Correct Rate (DR) gives of 94.5, 91.8, 88.6, and 86.3% and False Alarm Rate (FAR) gives of 13.6, 8.5, 5.1, and 4.8% data were obtained.

Text detection and extraction from natural scene images, which are captured through mobile camera and digital devices is proposed [5]. The proposed algorithm also tackles the complications involved in scene images like uneven illumination and reflection, poor lighting conditions and complex background analysis. Sharp transitions are detected using a revised Prewitt edge detection algorithm. The image is segmented into several regions. Each region can be regarded as an object. Finally, it is considered as abnormal objects (area too large or too small, width is far longer than height, etc).

Text regions are detected by 8-connected objects in natural scene images using which region-based extraction method is proposed [6]. Image can be detected using the median filter which is used to reduce the noise present in the image. In order to improve the precision of edge detection methods (Sobel, Prewitt, Laplacian Gaussian and canny method), experimentation is carried out on ICDAR-2014 dataset containing 509 English images, in which 258 images are taken for training set and 251 images are taken for testing set. To improve the performance of Stroke Width Transform method (SWT), the modified method is conducted. The Precision (p), Recall (r), and f-measure are dropped from 0.65, 0.71, 0.64 to 053, 0.55, 0.63, respectively. The computational time is also decreased from 4.2 to 1.7 s.

In this paper [7], the authors have proposed a Gabor function-based multichannel directional filtering approach that is used for separation of text and non-text regions

from the images (containing graphs, natural images, and other kinds of sketches drawn with lines) is proposed. Experimentation is carried out on images of 1000 words (Tamil, Hindi, Odiya, English). The documents considered for experimentation are bilingual, where English is a common script. Using linear discriminant function, script is identified for document containing Hindi and English which gives the highest accuracy of reported 99.56.

Image segmentation and text extraction from natural scene images are proposed [8]. Using Otsu method where Geometric Properties are used text localization and extraction are performed by using connected component algorithm and Run Length Smoothing Algorithm (RLSA) approaches. The proposed algorithm gives a reliable OCR results.

Using text-specific properties, text image deblurring is proposed for text-specific image deconvolution approaches [9], where the proposed algorithm not only estimates a more accurate blur kernel but also restores sharper texts. Experimentation is carried out on both blurred and also deblurred images, the Peak Signal Noise Ratio (PSNR) value is found to be 15.66 for blurred and 28.52 for deblurred images. Experimental results show that the proposed method generates higher quality results on deblurring text images.

Text extraction from natural images with a particular font size in ASCII text (English language) is proposed [10], which accepts an input image with complex background. The proposed algorithm is limited to only binary images. Text localization can be performed by using circumscribing the character of the text one after the other. Segmenting done by generating the distinct windows for each and every character in the image.

In this paper, an author proposed the Gabor filter that is used for extraction of text in images and video frames, where text regions are localized using morphological operation and heuristic filtering process [11]. After detection of text region, Gabor filters are used for identification of text within complex images. The experimentation is conducted in large dataset and the proposed method performs efficiently when compared with the other existing techniques. The result is investigated on recall and precision rate and analyzed for both proposed and existing method, it gives 99.11% for recall and 94.67% precision rate and computational average time taken is 5.28 s/images.

Extraction of text from a degraded document image is proposed [12], which consists of three stages such as preprocessing, text-area detection, and post-processing. The proposed method is textual extensively on DIBCO dataset, the performance of evaluation is analyzed based on F-measure (FM), pseudo F-measure (psFM), Peak Signal-to-Noise Ratio (PSNR), Distance Reciprocal Distortion (DRD), pseudo-Recall (Rps), and pseudo-precision (Pps). To test suitability the proposed approach, dataset of Guajarati degraded documents is created and the algorithm performs equally well on Guajarati document images. The main limitation of the proposed method is highly dependent on parameters that are in turn dependent on the set of images used.

The authors have proposed [13] a method to extract text from images with complex background which can be achieved using an edge-based text extraction algorithm

based on the fact that edges are reliable feature of text regardless of font sizes, styles, color/intensity, layout, orientation, etc. Experimental results show that this method is very effective and efficient in localizing and extracting text-based features.

Using the proposed method, separating texts from a textured background with similar color to texts is performed [14]. Experimentation is carried out with their own dataset containing 300 image blocks in which several challenges like manually generated images by adding text on top of relatively complicated background. From the experimentation with other methods, proposed method achieves more accurate results, i.e., precision of 95%, recall of 92.5%, and $F1$ score of 93.7%, and the proposed algorithm is robust to the initialized value of variables.

In this survey paper, different issues like text detection, segmentation, and recognition natural scene images are discussed [15]. Comparison of different text detection methods based on the Maximally Stable Extremal Regions (MSER) is highlighted followed by advantages and disadvantages. From the survey, it is observed that detecting and recognizing text from natural scene images is more difficult task than all other existing methods. Even though there are many algorithms, no single unified approach can fit for all applications.

For text extraction in complex natural scene images, different methods based on color and gray information are proposed [16]. The proposed method works even if the document containing skew and perspective of candidate text regions. The method is tested in 128 natural scene images that are captured in various places such as in schools, hospitals, subway stations, and streets. The dataset is classified into two categories, simple and complex from the experimentation, color-based method gives better results than gray-based method for complex images but it has more false detections. The gray-based method has better performance for simple images. The combination of both the methods gives better results than that of each method.

To detect text from natural scene images and also extracts them with any orientation, an algorithm is proposed [17]. For exact binary regions MSER detector is used, since it is robust lighting conditions. To remove non-test pixels, stroke width detector and several heuristics are used. To evaluate the performance of proposed approach, experimentation is carried out on two different datasets, i.e., ICDAR and KAIST dataset. The proposed method gives a performance gain of 8% in terms of Precision Rate (PR), 2% in terms of Recall Rate (RR), and 13% in terms of *f*-measure.

In this paper, the authors have proposed a new algorithm for the extraction of text from an image [18]. In addition to that, problem with unconstrained complex background in the scene has also been addressed. Experimentation is carried out on grabbed 270 image samples captured through a camera from the ICDAR dataset as well as collected from still images. A detection rate text is almost 93%, whereas rate of false positive detection of 4% is obtained.

In this paper, the authors have proposed a morphological technique for text extraction from images, where image is insensitive to noise, skew, and orientation [19]. Text extraction algorithm is providing an OCR system that recognizes the contained information and also, reduces the number of false text candidate regions.

Segmentation and classification from text and non-text using document images are proposed [20]. Heuristic rules have been used in segmenting and classifying

the text and non-text blocks like technical documents into table, graph, and figure. For classification, Backpropagation Neural Network (BPNN) and Support Vector Machine (SVM) are used the results it is observed that more features would produce better classification result. The results for block classification using BPNN and SVM based on different zones are as follows, i.e., for 9 zones BPNN gives 67.5% and SVM gives 81.3% similarly 36 zones BPNN gives 78.8 and 92.5%.

Using LBP base feature, separation of text and non-text from handwritten document images is proposed [21]. Texture-based feature like Gray Level Co-Occurrence Matrix (GLCM) is proposed for classifying the segmented regions. In this paper, a detailed analysis of how accurately features are extracted by different variants of Local Binary Pattern (LBP) operator is given a database of 104 handwritten engineering lab copies and class notes collected from an engineering college are used for experimentation purposes. For classification of text and non-text, Naïve Bayes (NB), Multilayer Perceptron (MLP), K-Nearest Neighbor (K-NN), Random Forest (RF), and Support Vector Machine (SVM) classifiers are used. It is observed that Rotation Invariant Uniform Local Binary Pattern (RIULBP) performed better than the remaining feature extraction methods.

Text and non-text classification of connected components in document images are proposed [22]. In which Multilayer Perceptron (MLP) and Convolution Neural Network (CNN) classifiers are used. The proposed method is compared with other method and achieves an accuracy of 98.68%. Different 19 geometrical features are given to the classifier. Experimentation is carried out on ICDAR-2009 dataset containing 55 images, in which 28 images are selected for training and 25 images for testing.

Based on color enhanced Contrasting External Region (CER) and Neural Network, an approach is proposed [23] for text detection from natural scene images to evaluate the performance of proposed method experimentation is carried out on ICDAR-2013 computation dataset. Performances of four different classifiers are analyzed and the proposed method is compared with ICDAR-2013 and ICDAR-2011 benchmark dataset, the proposed method gives highest recall rate of 92.14%, precision of 94.03%, and 93.08% of *F*-score and 92.22% Recall, 91.13%, 91.63% *F*-score, respectively.

A method is proposed for text detection from natural scene image consisting of two steps: connected components (CCs) extraction and non-text filtering [24]. To accomplish this, a multiscale adaptive color clustering strategy combining color histogram analysis with K-means++ algorithm is used to filter non-text and Text Covariant Descriptor (TCD) with histogram oriented gradients are considered. Experimentation is conducted on two publicly available datasets, i.e., ICDAR 2013 and ICDAR 2011, where 233 images are training and 255 are testing images and 299 images are training and 255 for testing, respectively.

The authors proposed [25] a robust Uyghur text localization method in complex background images which provide a CPU-GPU various parallelization system. A two-stage component classification system is used to filter out non-text components

and a component connected graph algorithm is used to constructe text lines. Experimentation is conducted on UICBI400 dataset; the proposed algorithm achieves best performance, i.e., achieves 12.5 speed up.

Text localization in video also from scene images is achieved using Kirsch Directional Masks [26]. The performance of the method is done with respect to *f*-measure (F), precision (P), and recall (R) and the method gives better results in terms of Detection Rate (DR) and Misdetection Rate (MDR). Experimentation is carried out on benchmark dataset horizontal-1, horizontal-2, and Hua dataset. The proposed method performs better as horizontal-1 dataset.

Based on corner points from document images, whatever might be their resolution, component/edges, level of degradation, a method is proposed for text extraction in document images [27]. To accomplish this, a simple approach based on FAST key points, the proposed method is evaluated on different languages Telugu, Arabic, and French which are handwritten printed, skewed documents, and images with different resolutions and noises. The proposed method gives an average of Precision and Recall rate of 95% (Table 1).

3 Challenges

From the literature review, we found that many challenging issues are still existing in text extraction from images: some of them are as follows:

(1) Extraction of text from blurred images.
(2) Text extraction from handwritten document images.
(3) Text extraction from multi-scripts.
(4) Text extraction from documents containing multiple skews.
(5) Text extraction from documents containing logos.
(6) Text extraction from partially visible documents.

Table 1 We have given the detailed summary of text extraction, segmentation, and classification methods

Citations	Preprocessing	Feature extraction methods	Segmentation	Classifiers	Dataset	Performance evaluation
[1]	Mathematical morphological operations	Edge detection and connected component	HAAR-DWT & FPGA (Field programmable gate array)	M-band and fisher	33—document images with ref IEEE	M band: 286 s, Cluster center approach: 158 s, Fisher classifier: 139 s, HAAR & DWT 96 s
[2]	Gabor filter	Region based	Bottom-up approach for text lines		Handwritten text	Recall rate: 92% Precision rate:
			Clustering, top-down successive splitting, bottom-up region growing		Colored book and journal covers	Recall rate: 83.3%. Precision rate: 83.5%
		Edge based	Automated system for text extraction		Scene text and superimposed text images	Recall rate-92% Precision rate-85%
			Multiscale strategy, clustering		Complex printed document images	Recall rate 96.6%. Precision rate-91.8%
		Texture based	Multiscale texture-based method using local energy analysis		Hybrid Chinese/English text detection in images	Recall rate-93.5%, Precision rate-1.8%
			Hough transform technique combined with an extremity segments neighborhood analysis		Detect text in video images	Recall rate-94%. Precision rate-2.48%
		Morphological-based	A measure of accumulated gradients and morphological post-processing		Artificial text in images and video	Recall rate-93.5%. Precision rate-85.4%
			Novel set of morphological operations and an x-projection techniques		Cluttered images	Recall rate-96.3%. Precision rate-99.4%

(continued)

Table 1 (continued)

Citations	Preprocessing	Feature extraction methods	Segmentation	Classifiers	Dataset	Performance evaluation
[3]	Morphological operations	Edge based algorithm + connected component algorithm	K-means clustering algorithm		150-different types of images	Precision rate-47.4–75.09% Recall rate-50.10–73.42%
[4]	Mathematical morphological operations + erosion & dilation	Discrete wavelet transform (DWT) + texture feature extraction	K-means clustering algorithm		100 different gray-scale images	The correct rate (DR)-94.5%, 91.8%, 88.6%, 86.3% and false alarm rate (FAR)-13.6, 8.5, 5.1 and 4.8%
[5]	Gray scale transformation + smoothing + contrast enhancement	Prewitt edge detection algorithm + Canny edge detector	Canny edge detector + ANN		10-natural scene images captured by mobile	
[6]	Image edge detection + Sobel + Prewitt + Laplacian of Gaussian	8-connected objects detection algorithm + SWT			ICDAR-2014	Precision-0.65, Recall-0.71, *f*-measure-0.64 Time (s)-4.2
[7]	Gabor function-based multichannel direction filtering	Gabor filter	Texture-based segmentation algorithm		Natural images and text portions given a block of text area	Hindi-99.56 Tamil-96.02 Odiya-97.1
[8]	Re-sampling + geometrical properties	CBDAR	Otsu + RLSA	MLP + SVM		
[9]	Deblurring methods + kernel estimation	SWT	Text binarization with thresholding based technique + FFT		Latent image	PSNR value blurred images-15.66 and deblurred images-28.22
[10]	Gray-scale thresholding	Line removal + discontinuity + dot removal	8-connected pixel connectivity + NN matrices		ASCII text (English)	
[11]	Gabor filter	Gabor filter + morphological + heuristic filtering	4-neighbor connected component		32 camera images (1578 words)	Recall rate-99.11% Precision rate-94.67%

(continued)

Table 1 (continued)

Citations	Preprocessing	Feature extraction methods	Segmentation	Classifiers	Dataset	Performance evaluation
[12]	Contrast edge detection using rough set theory	PCA + connected component + Block blob			DIBCO	FM-96.88-92.29 psFM-97.65-94.29 PSNR-22.66-18.54 DRD-0.9.2-2.66 Rps-0-98.09
[13]	TIE	Edge-based text extraction algorithm + Otsu method + connected component				
[14]	Text extraction algorithm	Edge-based text extraction algorithm	Text segmentation + k-means + sparse decomposition		Standard dataset-300 images	Precision-95% Recall-92.5% *F*1-score-93.7%
[15]	Text detection	MSER + SIFT + connected component	MRF + CRF	HOG + SVM	ICDAR 2011	
[16]	Geometrical clustering + color reduction + noise reduction + canny edge detection	Connected component + RLS			120 natural scene images	
[17]	SWT + canny edge detector	MSER + connected component			ICDAR and KAIST	Performance gain-8% Precision rate (PR)-2% *f*-measure (RR)-13%
[18]	TIE	Color feature + spatial distribution	Color chain segmentation		ICDAR 2003	
[19]	Block based local thresholding	Morphological gradient + Nonlinear filter + CCs			CD cover image	
[20]		Dilation morphological operation	Block segmentation	BPNN + SVM	240 images	BPNN-9 zones-65.5% & 81.3% SVM-36 zones 78.8% & 92.5%

(continued)

Table 1 (continued)

Citations	Preprocessing	Feature extraction methods	Segmentation	Classifiers	Dataset	Performance evaluation
[21]	Gray-scale image + BBs	Connected component + GLCM + LBP		NB + MLP + SMO + K-NN + RF	104 handwritten images	NB-77.92 MLP-90.78 SMO-88.62 KNN-90.20 RF-91.96
[22]	Geometric layout analysis + Otsu algorithm	Connected component		MLP + CNN	ICDAR-2009	**Non-text**: recall-98.68, precision-98.15 **Text**: recall-98.88 precision-99.30 Accuracy: 98.68
[23]	Color space conversion + geometrical properties	Connected component + CER		Neural network (NN)	ICDAR-2011 ICDAR-2013	Recall-92.14%, 92.32% & 92.22%, precision-94.03%-92.12% & 91.13%, *F*-score-93.08%, 92.22% & 91.65%
[24]	MSER + SWT	Multi-scale adaptive color clustering + CCs + TCD + HOG		SVM	ICDAR 2013 ICDAR 2011	Recall-0.70, precision-0.84, and *F*-measure-0.76 detection rate-0.90, proposal number-1226
[25]	Heterogeneous parallelization scheme	MSER + CCA + HOG		SVM	UICB I400	12.5 speedup
[26]	Canny, Sobel, Prewitt	CCA + KDM		SVM and ANN	Hua dataset	DR-89.38, FPR-13.08, MDR-10.52
[27]	Document image analysis (DIA)	component/edges extraction	FAST (features from accelerated segment test)		HDLAC2011	Recall—87.45% Precision—89.70%

4 Conclusion

In this review article, we have given a detailed approach on text detection from document images which include printed, handwritten, and also multilingual document images. The paper is mainly focused on the researchers who are planning to start their research on domain of text detection from document images.

References

1. A. Kumar, P. Rastogi, P. Srivastava, Design and FPGA implementation of DWT image text extraction technique. Procedia Comput. Sci. **57**, 1015–1025 (2015)
2. K.R. Soumya, T.V. Ancy, A. Chacko, Text extraction from images: a survey. Int. J. Adv. Comput. Sci. Technol. **3**(2), 100–104 (2014)
3. A. Singh, A.S. Bhide, P. Singh, Text extraction from live captured image with diversified background using edge based & K-means clustering. Int. J. Innov. Eng. Technol. (IJIET) **3**(11), 11–17 (2014)
4. X.-W. Zhang, X.-B. Zheng, Z.-J. Weng, Text extraction algorithm under background image using wavelet transforms, in *International Conference on Wavelet Analysis and Pattern Recognition* (2008), pp. 200–204
5. J. Yuan, Y. Zhang, K.K. Tan, T.H. Lee, Text extraction from images captured via mobile and digital devices, in *International Conference on Advanced Intelligent Mechatronics* (2009), pp. 566–571
6. Z. Huang, J. Leng, Text extraction in natural scenes using region-based method. J. Digit. Inf. Manage. **12**(4), 246–254 (2014)
7. P.B. Pati, S.S. Raju, N. Pati, A.G. Ramakrishnan, Gabor filters for document analysis in Indian bilingual documents, in *ICISIP* (2004), pp. 123–126
8. D.M. Nor, R. Omar, M. Zarar, M. Jenu, J.-M. Ogier, Image segmentation and text extraction: application to the extraction of textual information in scene images, in *ISASM* (2011), pp. 01–08
9. H. Cho, J. Wang, S. Lee1, Text image deblurring using text-specific properties, in *ECCV, Part V* (2012), pp. 524–537
10. P. Agrawal, R. Varma, Text extraction from images. IJCSET **2**(4), 1083–1087 (2012)
11. A. Kumar, An efficient approach for text extraction in images and video frames using Gabor filter. Int. J. Comput. Electr. Eng. **6**(4), 02–07 (2014)
12. R. Patel, S.K. Mitra, Extracting text from degraded document image, in *IEEE (NCVPRIPG)* (2015), pp. 01–04
13. B.S. Mamatha, B.P. Chaithra, Extraction of text from images. IJECS **3**(8), 7583–7587 (2014)
14. S. Minaee, Y. Wang, Text extraction from texture images using masked signal decomposition, in *IEEE (Global SIP)* (2017), pp. 01–05
15. U.B. Karanje, R. Dagade, Survey on text detection, segmentation and recognition from a natural scene images. Int. J. Comput. Appl. **108**(13), 39–43 (2014)
16. H.-R. Byun, M.-C. Roh, K.-C. Kim, Y.-W. Choi, S.-W. Lee, Scene text extraction in complex images. DAS, LNCS **2423**, 329–340 (2002)
17. N.-M. Chidiac, P. Damien, C. Yaacoub, A robust algorithm for text extraction from images, in *IEEE: (TSP)* (2016), pp. 493–497
18. S.P. Chowdhury, S. Dhar, A.K. Das, B. Chanda, K. McMenemy, Robust extraction of text from camera images, in *ICDAR* (2009), pp. 1280–1284
19. Y.M.Y. Hasan, L.J. Karam, Morphological text extraction from images. IEEE Trans. Image Process. **9**(11), 1978–1983 (2000)
20. Z. Ibrahim, D. Isa, R. Rajkumar, Text and non-text segmentation and classification from document images, in *IEEE/CSSE* (2008), pp. 01–04

21. S. Ghosh, D. Lahiri, S. Bhowmik, E. Kavallieratou, R. Sarkar, Text/non-text separation from handwritten document images using LBP based features: an empirical study, in *MDPI* (2018), pp. 01–15
22. F.D. Julca-Aguilar, A.L.L.M. Maia, N.S.T. Hirata, Text/non-text classification of connected components in document images, in *SIBGRAPI* (2017), pp. 01–06
23. L. Sun, Q. Huo, W. Jia, K. Chen, A robust approach for text detection from natural scene images. Pattern Recogn. **48**, 2906–2920 (2015)
24. H. Wu, B. Zou, Y. Zhao, Z. Chen, C. Zhu, J. Guo, Natural scene text detection by multi-scale adaptive color clustering and non-text filtering. Neuro Computing **214**, 1011–1025 (2016)
25. Y. Song, J. Chen, H. Xie, Z. Chen, X. Gao, X. Chen, Robust and parallel Uyghur text localization in complex background images. Mach. Vis. Appl. **28**, 755–769 (2017)
26. B.H. Shekar, M.L. Smitha, Text localization in video/scene images using kirsch directional masks, in *IEEE (ICACCI)* (2015), pp. 1436–1440
27. V. Yadav, N. Ragot, Text extraction in document images: highlight on using corner points, in *IAPR-Document Analysis Systems* (2016), pp. 281–286

Object Recognition Using SBMHF Features



M. Ravikumar, S. Sampathkumar, M. C. Prashanth and B. J. Shivaprasad

1 Introduction

Recognition of objects from an image is fundamental and also a challenging issue in the domain of computer vision from the olden days. Object recognition is not only for the primary research purpose even for automated the computer vision system, it had been gradually increased the productive capacity of organizations. The application of object recognition has used in different fields like biological sciences, medical applications, biometrics, later by knowing the importance it was successfully inherited for classification of the agricultural process followed by made easy for suppression of capsules and tablets in the pharmaceutical industry, automation of assembly and industrial inspection process of the electromechanical industry. In the domain of machine and deep learning, object recognition is a key output.

The paper is discussed as follows: in the next section, we have discussed the detailed work done by the researchers in recognition of the object. After Sect. 2, the next Sect. 3 provides a description of the proposed method. The experimentation part is discussed in Sect. 4. Finally, the last part of the paper Sect. 5 will conclude as which is the best among the performed experimentation.

The remaining part of the paper is discussed as follows: in Sect. 2 we discuss the detailed work done by researchers on object recognition. Section 3 gives a description

M. Ravikumar · S. Sampathkumar (✉) · M. C. Prashanth · B. J. Shivaprasad
Department of Computer Science, Kuvempu University, Shimoga 577451, Karnataka, India
e-mail: sampath1447@gmail.com

M. Ravikumar
e-mail: ravi2142@yahoo.co.in

M. C. Prashanth
e-mail: prashanth.m.c87@gmail.com

B. J. Shivaprasad
e-mail: shivaprasad1607@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_99

of the proposed method. The experimentation part is discussed in Sect. 4. Finally, conclusion is given in Sect. 5.

2 Related Work

Here, the detailed work which relates the object recognition has been discussed. An algorithm to match the features in an image which are corresponding to the physical point from an object seen in two arbitrary viewpoints is discussed [1]. The proposed method consists of three different steps; they are the detection of scale-space features, calculation of affine invariant and matching. Experimentation is concluded on two small images of a box having different viewpoints. The images are taken by a static camera by rotating the object by 15°.

For object recognition, an efficient binary descriptor FAST algorithm based on BRIEF and ORB is proposed [2]. The experimental result shows how ORB magnitudes faster than SIFT. The efficiency of the proposed method is tested on different real-world applications including object detection and patch tracking on a Smartphone.

For image matching which helps in object recognition some different approaches like SIFT, SURF, and ORB are proposed [3]. In this experimentation a 45° of the rotation image is matched and experimentation SIFT achieves 65% of matching rate. The SIFT approach has observed the best performance in most of the cases. The ORB concentrated the central features in objects at the image.

For precisely localizing objects of various classes and for classification, DNNs are used for the problem of object detection as proposed [4]. Simple bounding box interference is used to extract detections from masks. For experimentation, approximately 20 classes of 5000 test images are taken. The results of the experiment computational cost at training time. No need to train a network per object type and mask type.

Finding an object to localize, to categorize by extracting the features, and to get the appearance information various techniques are used in images and videos [5]. OD is classified into five major categories and other approaches like shape-based detection and Steiner tree based are also summarized.

Different methods and techniques are presented for detecting and recognizing an object that is proposed [6]. To enhance efficiency and robustness different super-resolution based method is used. For object detection, it is showed that shadow c-means approach will give better results for occlusion patterns.

Various methods for object recognition and a technique for multiple object detection in an image is proposed [7]. On experimentation, global features and shape-based method yields good results and are effective as compared with local features. The proposed technique helps in easy to access images. They help in different fields of application.

Object detection and recognition in images are proposed [8]. For object tracking along with detection, Easy net model is unified and implemented to find an object

in the image taken from a single camera, background subtraction approach has been used. The model is compared with another detection system like RCNN.

Object detection is performed using a top-down and bottom-up approach for segmentation [9]. A modified context shape feature is designed which is a robust object deformation and background clutter. To prune out false positives a method is proposed based on pruning. The result demonstrates that the proposed detection algorithm gives high precision and recall rates, yields good results, and is efficient as compared with local features. The proposed technique helps in easy access to images. They help in different fields of application.

To recognize an image, object segmentation and detection in the photos having a white background the deep learning method is proposed [10]. The recognition model is trained on the Google network to check whether a photo is having the background white or not. The implemented algorithms in Real time with k80 Telsa and Caffe from NVidia. On experimentation, the accuracy of recognition is 96% and detection is 94%.

Some of the works related to object recognition can also be found [11–15]. In the next section, the proposed method is discussed.

3 Proposed Methodology

The proposed method for object recognition from complex scene images is addressed and given in the form of a block diagram in Fig. 1.

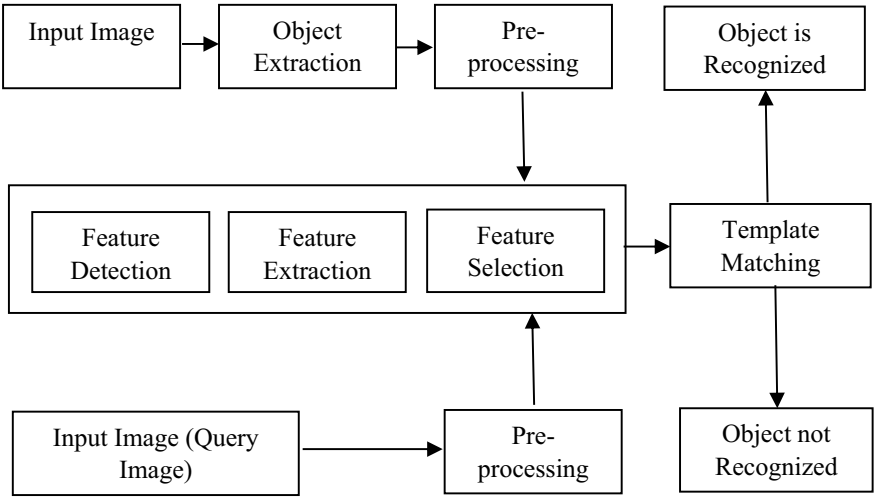


Fig. 1 Block diagram of the proposed method

Initially, an input image containing different objects is taken, from which we extracted the objects of our interest (in this case, key is our object) and preprocessing can be performed using a median filter. Afterward we detect features like blobs, edges, and corners from the image. once the features are detected, feature selection can be done by applying Surf, MinEigen, Fast and Harris features, from this we select only the strong feature points in order to improve the performance and all feature points are then compared with query objects, if the query object is same as the objects which are extracted from the input image, then the feature points are compared, if matching takes place then the object is recognized. If the matching doesn't take place the object is not recognized. While doing experimentation, intentionally we have captured an image without key. To check the correctness of the proposed algorithm in that image, the key result is not found.

4 Results and Discussion

In this work, we carried out experimentation in order to prove efficacy of the proposed method. For this purpose, experiment is carried out on our dataset containing different objects spread across an image with different orientations also occluded and overlapping the objects. From this dataset our focus is on to segment the object (key). Here, different features like Surf, Harris, MinEigen, Fast, and Brisk are used. The results are given in Tables 1, 2, 3, 4, 5, 6 and all the tables (except Brisk) is normalized to only twelve pair points because each feature will give the random number of features according to their functionality, as it leads too much of space we have reduced to twelve rows to display. In the tables image seven which not contain the object i.e. key.

In this approach, 118 features are obtained wherein 30 strong features are considered for the purpose of comparison.

In this case, 120 features are generated in which 11 strong features are taken for comparison purposes.

Here image three and image six will not give any features because of the specific property of brisk.

Here 130 features are obtained out of that only 12 strong features are selected which are used for comparison.

The table shows 12 strong features from 60 features and is listed in the table. The Fast algorithm detects the corners of an object in the image.

After the experimentation, it is observed that the MinEigen feature method performs better than the remaining approaches because the MinEigen detects more features even under the rotation, scaling changes, and introduction of noise.

Table 1 SURF features

Image 1		Image 2		Image 3		Image 4		Image 5		Image 6		Image 7	
Key	General	Key	General	Key	General	Key	General	Key	General	Key	General	Key	General
1	734	2	336	3	172	7	166	1	817	1	13	1	203
5	1064	3	16	7	817	9	132	2	896	3	122	4	111
7	263	5	20	11	1484	11	708	3	60	4	77	5	770
8	228	9	326	12	1746	13	260	4	960	7	599	6	153
11	152	18	486	15	6165	18	710	10	261	8	102	7	90
12	550	24	516	25	307	19	2157	12	337	10	140	9	230
15	1151	26	1230	26	716	23	1110	13	515	11	393	13	188
17	738	35	1856	29	869	34	3762	14	288	12	393	18	735
19	320	50	2350	30	842	38	1308	15	742	15	224	20	876
23	1790	56	3084	33	784	45	2796	16	682	16	500	21	749
26	683	71	2425	38	920	72	5054	17	333	24	715	24	1619
27	1233	88	5137	39	1337	77	2654	18	1000	38	1060	26	4092

Table 2 Harris features

Image 1		Image 2		Image 3		Image 4		Image 5		Image 6		Image 7	
Key	General	Key	General	Key	General	Key	General	Key	General	Key	General	Key	General
13	2739	94	6274	10	3069	13	2739	1	5531	305	3544	7	2140
79	2502	145	6080	173	2621	79	2502	3	5526	392	3477	34	1990
86	2487	146	6079	209	2525	86	2487	4	5525			36	1989
94	2465	147	6078			94	2465	9	5496			38	1982
98	2455	153	6055			98	2455	11	5469			40	1975
102	2449	172	5950			102	2449	16	5446			49	1949
108	2435					108	2435	20	5419			56	1933
109	2432					109	2432	21	5418			57	1905
114	2396					114	2396	25	5400			60	1888
133	2360					133	2360	26	5399			61	1889
134	2352					134	2352	31	5386			66	1877
								38	5356			70	1868

Table 3 BRISK features

Image 1		Image 2		Image 3		Image 4		Image 5		Image 6		Image 7	
Key	General	Key	General	Key	General	Key	General	Key	General	Key	General	General	Key
29	835	3	388			4	182	26	600			19	458
								30	575			23	449
												33	399

Table 4 MinEigen features

Image 1		Image 2		Image 3		Image 4		Image 5		Image 6		Image 7	
Key	General	Key	General	Key	General	Key	General	Key	General	Key	General	Key	General
1	87,960	12	94,828	51	83,131	227	62,761	6	91,590	130	154,290	1	118,247
8	87,850	104	93,919	441	80,273	228	62,700	9	91,635	137	154,232	4	118,191
10	87,795	105	93,918	1061	76,154	229	62,657	17	91,481	164	154,129	8	118,163
19	87,693	498	91,126	1210	75,179	238	62,583	35	91,183	179	154,062	17	118,048
54	87,283	555	90,686	1223	75,098	248	62,655	65	90,819	207	153,856	18	118,047
66	87,049	666	89,835	1244	74,943	278	62,402	72	90,719	213	153,857	34	117,984
80	86,996	706	89,569	1388	73,829	291	62,286	76	90,664	221	153,854	53	117,868
91	86,826	934	87,793	1430	73,564	317	62,144	83	91,313	245	153,663	54	117,896
105	86,610	962	87,536	1530	72,935	327	62,086	94	90,439	258	153,597	70	117,780
126	86,311	1224	85,486	1708	71,644	441	61,456	100	90,385	278	153,465	118	117,405
151	86,030	1256	85,242	1900	69,960	599	60,291	101	90,384	348	153,083	143	117,110
160	86,029	1258	85,241	1947	69,615	656	60,065	112	90,334	371	152,977	145	117,133

Table 5 FAST features

Image 1		Image 2		Image 3		Image 4		Image 5		Image 6		Image 7	
Key	General	Key	General	Key	General	Key	General	Key	General	Key	General	Key	General
8	3239	1	934	17	1608	7	342	3	1719			2	1581
16	3230	6	764			10	340	19	1677			5	1569
18	3229	7	763					24	1622			9	449
21	3221	8	760					25	1621			16	1281
24	3220	10	756					29	1618			23	1182
30	3194	13	691					31	1617			24	1180
38	3145							43	1584			29	1142
39	3144							45	1581			31	1138
67	3007							52	1499			42	1066
68	3006							55	1476			47	1044
71	2911							59	1462			48	1043
75	2925							60	1461			50	1041

Table 6 The overall analysis of the above tables

SURF		HARRIS		MINEIGEN		FAST		BRISK	
Obtained features	Strong features	Obtained features	Strong features	Obtained features	Strong features	Obtained features	Strong features	Obtained features	Strong features
118	30	120	11	130	13	60	25	10	2

5 Conclusion

In this work, instead of using feature extraction methods to obtain features, directly five features (SBMHS) are considered. Experimentation is carried out on our own dataset and from the result it is observed that the MINEIGEN feature outperforms well when compared with the remaining approaches.

References

1. A. Baumberg, Reliable feature matching across widely separated views, in *Proceedings IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2000* (2000) pp. 01–08
2. E. Rublee, V. Rabaud, K. Konolige, G. Bradski, Orb: an efficient alternative to sift or surf, in *International Conference on Computer Vision* (2011)
3. E. Karami, S. Prasad, M. Shehata, Image matching using sift, surf, brief and orb: performance comparison for distorted images, in *Newfoundland Electrical and Computer Engineering Conference*, St. Johns, Canada (2015)
4. C. Szegedy, A. Toshev, D. Erhan, Deep neural networks for object detection, in *Neural Information Processing Systems* (2013)
5. K.U. Sharma, N.V. Thakur, A review and an approach for object detection in images. Int. J. Comput. Vis. Robot. **7**, 196 (2017)
6. D. Patel, P.K. Gautama, A review paper on object detection for improve the classification accuracy and robustness using different techniques. Int. J. Comput. Appl. **112**(11), 05–07 (2015)
7. K. Khurana, R. Awasthi, Techniques for object recognition in images and multi-object detection. Int. J. Adv. Res. Comput. Eng. Technol. (IJARCET) **2**, 1383–1388 (2013)
8. S. Kumar, A. Balyan, M. Chawla, Object detection and recognition in images. Int. J. Eng. Dev. Res. (IJEDR) **5**(4), 1029–1034 (2017)
9. L. Wang, J. Shi, G. Song, I.-F. Shen, Object detection combining recognition and segmentation, in *Asian Conference on Computer Vision*, *Accv 2007* (2007), pp. 189–199
10. X. Ning, W. Zhu, S. Chen, Recognition, object detection and segmentation of white background photos based on deep learning, in *32nd Youth Academic Annual Conference of Chinese Association of Automation* (2017), pp. 182–187
11. P.F. Felzenszwalb, R.B. Girshick, D. Mcallester, D. Ramanan, Object detection with discriminatively trained part-based models. IEEE Trans. Pattern Anal. Mach. Intell. **32**(9), 1627–1645 (2010)
12. J. Tyckowski, C. Kyrtsos, T. Davies, C. Hopson, F. Breynaert, P. Bonduel, Object detection by pattern recognition. U.S. Patent Number 6,154,149 (2000), pp. 01–07
13. Z. Sun, G. Bebis, R. Miller, Object detection using feature subset selection. Pattern Recogn. 2165–2176 (2004)
14. K. Mikolajczyk, B. Leibe, B. Schiele, Local features for object class recognition, in *Tenth IEEE International Conference on Computer Vision*, vol. 1 (2005)
15. D.M. Ramík, C. Sabourin, R. Moreno, K. Madani, A machine learning based intelligent vision system for autonomous object detection and recognition. Appl. Intell. **2**, 358–375 (2014)
16. C. Leng, H. Zhang, B. Li, G. Cai, Z. Pei, L. He, local feature descriptor for image matching: a survey. IEEE. Transl. Content Min. **7**, 6424–6434 (2019)
17. H.-C. Shih, H.-Y. Wang, A robust object verification algorithm using aligned chamfer history image, in *Multimedia Tools and Applications* (Springer Nature, 2019)

DWT Based Compression Algorithm on Acne Face Images



Garima Nain, Ashish Gupta and Rekha Gupta

1 Introduction

The medical images acquired during the examination, along with the other information like personal data of the patient, sort of medical examination, diagnostic reports, etc., are systematized in medical files. Various kind of digital medical image sequences is present that includes computed tomography (CT) images, magnetic resonance imaging (MRI), ultrasound images, X-ray, and capsule endoscopy (CE) images. Many compression standards for medical image, such as Joint Photographic Experts Group (JPEG) [1] and JPEG–LS [2] build on Discrete Cosine Transform (DCT) [3] and JPEG2000 [4] based on wavelet transform [5, 6] are projected for various medical images. Various compression methods concerning a type of image such as the MRI images [7–9], CT images [9], coronary angiographic images, [10] and computed radiography images [11] are available in literature. A literature survey is also done to unearth the performance assessment of various compression techniques on the medical skin images and it is found that the paramount method would be using DWT which is the lossless image compression technique in medical imaging [12–14]. In fact, the survey shows that medical images may possibly not accomplish superior compression ratios in order to uphold the required grade of quality at the image repossession to the treatment center. Moreover, if less compression is done then it provides less amount of storage space and less bandwidth economy. But here a special type of image is intended. These include the acne face images which are the

G. Nain · A. Gupta (✉) · R. Gupta
Department of Electronics Engineering, MITS Gwalior, Gwalior, Madhya Pradesh, India
e-mail: agashishietk@gmail.com

G. Nain
e-mail: garimanain96@gmail.com

R. Gupta
e-mail: rekha652003@yahoo.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_100

records stored by the dermatologist about the tribulations on the face. Acne appears on the skin as blackheads or whiteheads, pustules (bumps containing pus), cysts (deep pimples, boils), red bumps also known as pimples, etc. Therefore, the finest technique is tossed that can be applied to such type of acne face images.

From the surveillance of such images, it was found that the entire information in the image is obligatory to the dermatologist and thus the proposed algorithm is based on it. Three iterations of DWT [15] are applied to the image followed by the arithmetic coding [16]. The subsequent analysis was performed on a set of five images and the following conclusions were made that even if the entire image is compressed greatly the uncompressed patch provides the complete obligatory information for diagnosis. Mean Square Error (MSE) and Peak signal to noise ratio (PSNR) is enhanced in contrast to applying DWT [17] to the images exclusive of the uncompressed patches.

2 Two-Dimensional DWT and Arithmetic Coding

This section gives a brief description of image compression using 2-D wavelet transform and arithmetic coding.

2.1 Two-Dimensional DWT

Wavele-based image compression provides extensive precision in image quality at higher compression ratios chiefly due to one of the properties of the wavelet transform such as enhanced energy squeezing [18].

A two-dimensional scaling function, $\phi(x, y)$ and three, 2-D wavelets function $\psi^H(x, y)$, $\psi^V(x, y)$ and $\psi^D(x, y)$ are essential in the wavelet transform technique. Each is the product of its corresponding wavelet function $\psi(\mathrm{x})$ and a one-dimensional scaling function $\phi(\mathrm{x})$ [19].

$$\phi(x, y) = \phi(x)\phi(y) \tag{1}$$

$$\psi^H(x, y) = \psi(x)\phi(y) \tag{2}$$

$$\psi^V(x, y) = \phi(x)\psi(y) \tag{3}$$

$$\psi^D(x, y) = \psi(x)\psi(y) \tag{4}$$

where ψ^H evaluates changes along rows, i.e., vertical edges and ψ^D corresponds to disparity along diagonals and ψ^V responds to disparity along columns, i.e., horizontal

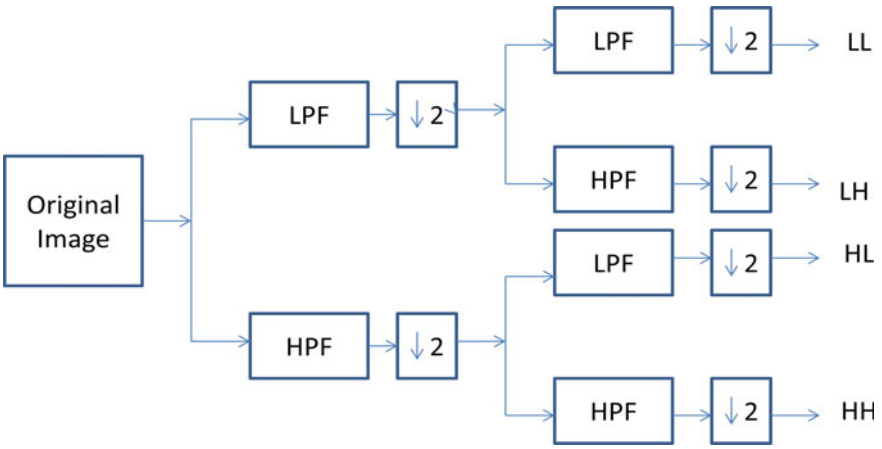


Fig. 1 The analysis filter bank of 2-D DWT

edges. At each of the iteration level of DWT, the rows and columns of the specified image are low-pass and high-pass filtered [19].

Then values of the two obtained images secured at the output of the two filters, i.e., a low and a high-pass filter are decimated by a factor of 2. Next, the filtration of columns is done using given set of filters. The new obtained columns of filtered images are also down-converted by a factor of 2 as shown in Fig. 1.

Figure 2 shows how the image decomposition takes place when DWT is applied to it. The figure shows image decomposition at each iteration level reducing image size significantly. The LH, HL, and HH components contain high-frequency information which is ignored at the image retrieval.

The Haar' wavelet is used in our DWT algorithm and its mother wavelet function can be described as

$$\psi(t) = \begin{cases} 1 & 0 \le t < \frac{1}{2}, \\ -1 & \frac{1}{2} \le t < 1, \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

Its scaling function can be described as

$$\varphi(t) = \begin{cases} 1 & 0 \le t < 1, \\ 0 & \text{otherwise.} \end{cases} \tag{6}$$

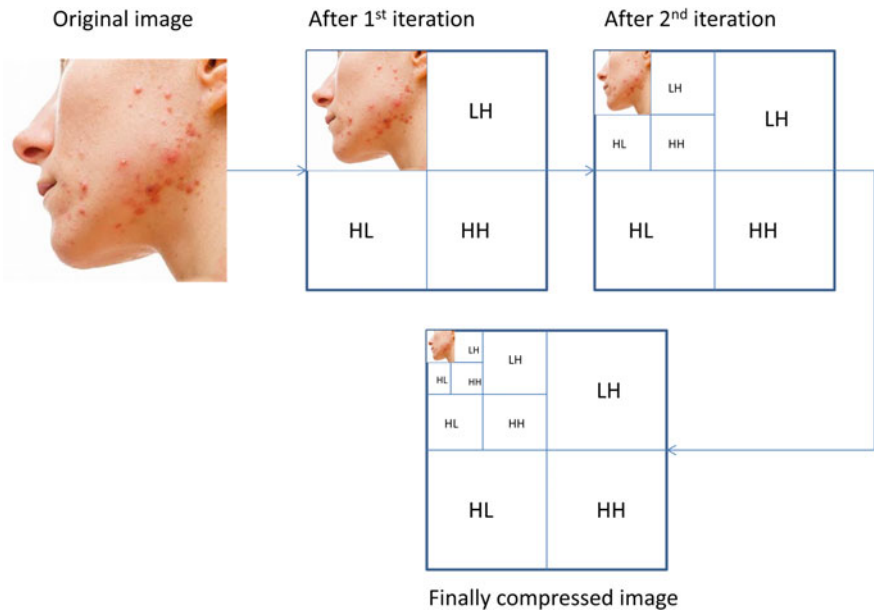


Fig. 2 Three level image decomposition using 2-D DWT

2.2 Arithmetic Coding

After applying DWT, the next step is to encode data for transmission and further compression for which we used arithmetic coding [16, 20]. Arithmetic coding is an outline of an entropy encoding algorithm applied in lossless data or image compression. In this encoding scheme a source ensemble is constituted in an interval between 0 and 1 on the real number line. Arithmetic coding assumes an unambiguous probabilistic model of the input source. It is a defined word format in which the probabilities of the resource messages are used to sequentially narrow the period used to represent the ensemble. The method proceeds with feeding the unordered list of DWT image matrices to the arithmetic coding and then generating the raw data.

3 Proposed Algorithm

As explained in previous sections, medical image compression suffers from the problem of information loss on retrieval when highly compressed using DWT. For that reason, in the present section, a new algorithm is proposed, which gives better image compression without loss of necessary information.

The algorithm is proposed in the following key steps

(i) First divide the image into $N \times N$ parts depending upon the compression requirements.
(ii) Then find the R to G ratio of the image.
(iii) Observe the ratio values in the patches where the acne is the most or part which contains the most information.
(iv) Also, use histogram of that patch to specify its ratio's upper threshold and lower threshold.
(v) Verify the appropriate thresholds from the count values of a number of pixels falling in given threshold range for the required part which should be maximum.
(vi) Keep one of $N \times N$ parts uncompressed and apply compression on all other parts.

The specified acne face image is alienated into various parts as 2, 4, 9, etc. The authors made 3×3 parts of the images going first from top to bottom and then going left to right. Then the pixel values of RGB matrices of the images were observed. Then the ratios of the R to G matrix, G to B matrix and R to B matrix were intended. It is found in hand that maximum number of pixels of ratios of the R to G matrices is in the range 1.4–1.6. In this range the patch with maximum acne is detected and transmitted uncompressed and the whole image is passed through the three iterations of the DWT. At the receiving end the patch is superimposed on its actual location at the time of image repossession.

4 Result and Discussion

In the current section, the helpfulness of our proposed algorithm is revealed on different acne images. Figure 3 shows the various test images [21, 22] and their respective retrieved images using proposed algorithm. The patches seen in original images with most grimy parts or acne are clearly seen in their respective retrieved images as they are uncompressed and selected automatically by the proposed algorithm in the range 1.4–1.6.

Figure 3b is a retrieved image of Fig. 3a in which 9 parts are made from 1 to 9 going from top to bottom and then left to right. In this it is patch 8 that is uncompressed, also from the given Table 1 it can be seen that the count bar for the various blocks of the images has maximum counts of R to G pixels ratio in this 8th patch in the range 1.4–1.6 of the ratios. The respective histogram shows that most of the ratios in the patch 8 fall in the same range.

Further, Table 2 shows the size of the image 1, with its corresponding average compression ratio which is 20.271. This image takes very less storage space after compression, PSNR is in the acceptable range, i.e., 28.58 dB and MSE is 0.0014.

Similarly in Fig. 3d, it is the patch 9 that is uncompressed, in Fig. 3f, it is the patch 2 that is uncompressed, in Fig. 3h it is the patch 8 that is uncompressed and in Fig. 3j it is the patch 4. From the illustration as well from the results it is validated. Table 1 shows the value of R to G ratio counts in all of the 9 parts of an image where

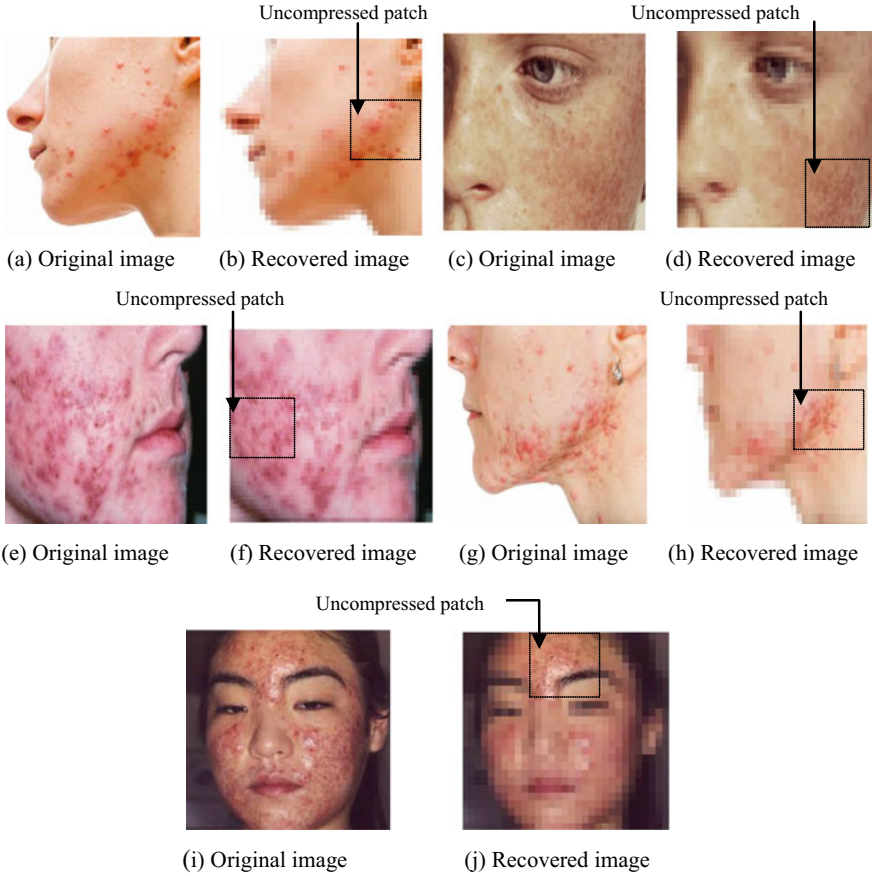


Fig. 3 Various test images [21, 22] and their respective retrieved images

on x-axis it is the patch/part number from patch 1 to 9 and on the *y*-axis it is the number of ratios in the range of 1.4–1.6.

It can be seen from Table 1 that the number of counts is maximum in that patch where the image encompasses most of the spots or pimples. Also respective histogram shows the range of ratio in the selected patch.

Table 2 shows the respective values of the image sizes and their average compression ratios that are 18.77. Also, the PSNR and MSE are of the order of 26.56 dB and 0.0023, respectively.

Table 1 Below are tabulated the image number, their ratio counts in the range of 1.4–1.6 and their respective histograms of uncompressed image patch

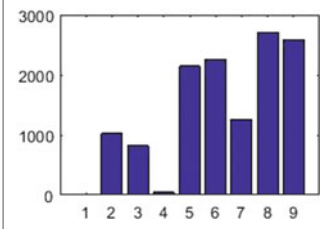
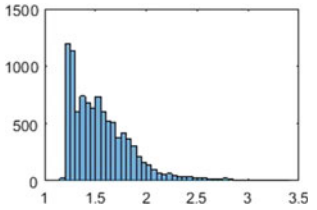
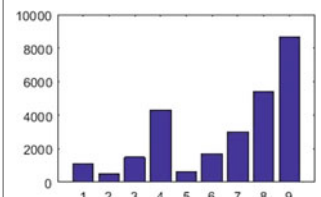
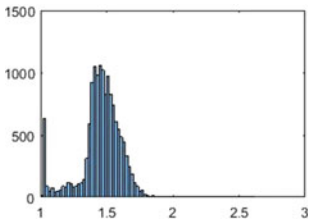
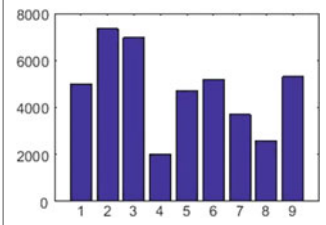
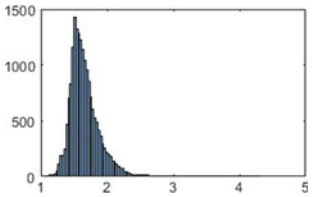
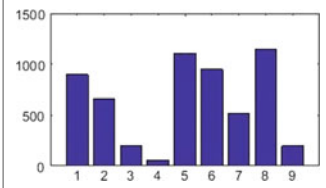
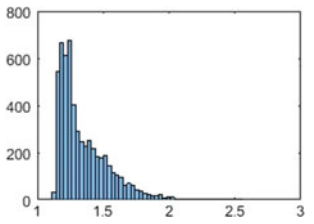
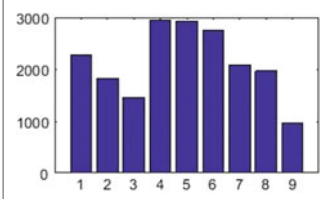
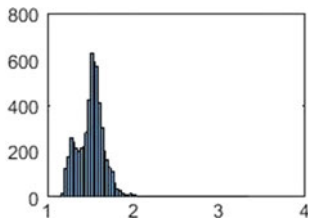
Figure 3	Count Bar in between 1.4 and 1.6	Histograms of uncompressed blocks
(a)		
(c)		
(e)		
(g)		
(i)		

Table 2 Below are the size (in bytes), compression ratio, PSNR (in dB) and respective MSE of the test images

Image	Size (bytes)	Average compression ratio	PSNR (dB)	MSE
(a)	300 × 300	20.271	28.58	0.0014
(c)	308 × 428	17.718	25.44	0.0029
(e)	335 × 493	18.012	24.18	0.0038
(g)	195 × 258	19.900	27.56	0.0018
(i)	210 × 243	17.965	27.04	0.0020

5 Conclusion

The method proposed in segment 3 and the conclusion obtained in segment 4 introduces a novel method for such images. The proposed method is uncomplicated and at the same time, it helps in considering only the region of interest area and applying a very high compression on the rest part of the image. The uncompressed patch is sufficient for a doctor to generate essential analysis and diagnosis without getting affected by the whole image quality. Better PSNR is observed when compared to the compression applied to an image without the superposition of an uncompressed patch. Although the compression ratios achieved are low but this is not an issue when medical imaging is considered as no information loss can be tolerated and a minute (small) error would lead to misdiagnosis. As the image is divided into more number of parts, better compression is achieved giving higher bandwidth economy.

References

1. Y. Zeng, F. Huang, H.M. Liao, Compression and protection of JPEG images, in *18th IEEE International Conference on Image Processing,* Brussels (2011), pp. 2733–2736
2. S. Miaou, F. Ke, S. Chen, A lossless compression method for medical image sequences using JPEG-LS and interframe coding. IEEE Trans. Inf. Technol. Biomed. **13**(5), 818–821 (2009)
3. Y.-Y. Chen, Medical image compression using DCT-based sub band decomposition and modified SPIHT data organization. Int. J. Med. Inform. **76**, 717–725 (2007)
4. S. Sebastian, M.A.P. Manimekalai, Color image compression using JPEG2000 with adaptive color space transforms, in *International Conference on Electronics and Communication Systems (ICECS)* (2014), pp. 1–5
5. L. Shen, R.M. Rangayyan, A Segmentation-based lossless image coding method for high-resolution medical image compression. IEEE Trans. Med. Image **16**(3), 301–307 (1997)
6. A. Nashat, N.M.H. Hassan, Image compression based upon wavelet transform and a statistical threshold, in *International Conference on Optoelectronics and Image Processing (ICOIP)* (2016), pp. 20–24
7. M. Midtvik, I. Hovig, Reversible compression of MR images. IEEE Trans. Med. Imaging **18**(9), 795–800 (1999)
8. R. Srikanth, A.G. Ramakrishnan, Contextual encoding in uniform and adaptive mesh based lossless compression of MR images. IEEE Trans. Med. Imaging **24**, 1199–1206 (2005)

9. J. Taquet, C. Labit, Hierarchical oriented predictions for resolution scalable lossless and near-lossless compression of CT and MRI biomedical images. IEEE Trans. Image Process. **21**(5), 2641–2652 (2012)
10. A. Munteanu, J. Cornelis, P. Cristea, Wavelet-based lossless compression of coronary angiographic images. IEEE Trans. Med. Imaging **18**(3), 272–281 (1999)
11. S. Wong, L. Zaremba, D. Gooden, H.K. Huang, Radiologic image compression-a review, in *Proceedings of the IEEE*, vol. 83, no. 2 (1995), pp. 194–219
12. S. Grgic, M. Grgic, B. Zovko-Cihlar, Performance analysis of image compression using wavelets. IEEE Trans. Ind. Electron. **48**, 682–695 (2001)
13. D. Ravichandran, M.G. Ahamad, M.R.A. Dhivakar, Performance analysis of three-dimensional medical image compression based on discrete wavelet transforms, in *22nd International Conference on Virtual System & Multimedia (VSMM)*, Kuala Lumpur (2016), pp. 1–8
14. A. Baviskar, S. Ashtekar, A. Chintawar, J. Baviskar, A. Mulla, Performance analysis of sub-band replacement DWT based image compression technique, in *Annual IEEE India Conference (INDICON)* (2014), pp. 1–6
15. D. Marpe, G. Blattermann, J. Ricke, P. Maass, A two-layered wavelet-based algorithm for efficient lossless and lossy image compression. IEEE Trans. Circuits Syst. Video Technol. **10**(7), 1094–1102 (2000)
16. H.Y. El-Arsh, Y.Z. Mohasseb, A new light-weight JPEG2000 encryption technique based on arithmetic coding, in *MILCOM 2013–2013 IEEE Military Communications Conference*, San Diego, CA (2013), pp. 1844–1849
17. S.-G. Miaou, S.-T. Chen, Automatic quality control for wavelet-based compression of volumetric medical images using distortion-constrained adaptive vector quantization. IEEE Trans. Med. Imaging **23**(11), 1417–1429 (2004)
18. M. Antonini, M. Barlaud, P. Mathieu, I. Daubechies, Image coding using wavelettransform. IEEE Trans. Image Process. **1**, 205–220 (1992)
19. R.C. Gonzalez, R.E. Wood, *Digital Image Processing Using MATLAB*, 2nd edn. (McGraw Hill Companies, Reading, 2011), pp. 331–439
20. T. Koya, S. Chandran, K. Vijayalakshmi, Analysis of application of arithmetic coding on DCT and DCT-DWT hybrid transforms of images for compression, in *International Conference on Networks & Advances in Computational Technologies (NetACT)* (2017), pp. 288–293
21. United States National Library of Medicine. http://www.medlineplus.gov. Accessed 5 Jan 2019
22. http://www.medicinenet.com. Accessed 7 Jan 2019

Segmentation of Blood Vessels from Retinal Fundus Images Using Bird Swarm Algorithm and River Formation Dynamics Algorithm



Jyotika Pruthi, Shaveta Arora and Kavita Khanna

1 Introduction

The computer-assisted segmentation of blood vessels from retina plays a very important role in the diagnosis of various eye diseases [1]. Segmentation of blood vessels from retina is helpful for the detection of diseases like hypertension [2], glaucoma [3], obesity [4], diabetic retinopathy [5], and narrowing of arteriolar [6], etc. The retinal vessels are comprised of arteries and veins and form the connected tree-like branch structure. All the vessels might appear similar but they at times vary from image to image in terms of shape, size, and intensity of gray level. Blood vessel extraction from fundus images becomes challenging in the presence of noise, variations in intensity or underlying poor contrast.

1.1 Related Work

Many researchers have developed different algorithms for retinal blood vessel segmentation. In 1999, algorithms were developed for rapid and accurate extraction of blood vessels that dealt even with the images having unconnected vasculature [7]. The major drawback of this approach is the non-consideration of low contrast images. Particularly, to deal with the low contrast images, another vessel tracking algorithm was presented for the automated detection of diabetic retinopathy by Lalonde et al.

J. Pruthi · K. Khanna
Department of Computer Science and Engineering, The NorthCap University, Sector 23A, Gurgaon 122017, Haryana, India

S. Arora (✉)
Department of Electrical, Electronics and Communication Engineering, The NorthCap University, Sector 23A, Gurgaon 122017, Haryana, India
e-mail: shavetaarora@ncuindia.edu

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_101

[8]. This algorithm not only produced accurate results but also handled the broken edges.

Furthermore, in 2004, a novel automatic system was developed to extract the retinal vessels and was the first system that could identify the bifurcations and crossings of vessels [9]. In 2011, a supervised model utilizing the capabilities of neural network was presented by Diego et al. for vessel segmentation [10]. This approach outperformed many segmentation techniques existing in the literature as it was highly robust, simple, and fast in implementation. As the metaheuristic algorithms started becoming popular, ant colony optimization (ACO) algorithm was also applied for retinal vessel detection [11]. ACO provided output having high visual quality of the vessels that are identified. In the same trend, flower pollination search algorithm was also utilized for retinal vessel segmentation that further used the concept of clustering for better detection [12]. In 2016, vessel segmentation problem was remolded in terms of cross-modality data transformation [13]. Deep neural networks were employed in this model and approach required no artificially designed features. In 2017, flower pollination search algorithm was also improved by combining it with pattern search algorithm [14]. This hybrid approach was capable enough to achieve multi-objective retinal blood vessels localization.

In 2018, many metaheuristic algorithms were used for vessel segmentation like Particle swarm optimization (PSO), artificial bee colony (ABC) [15], and BAT algorithm [16], etc. PSO was applied to handle the problem of vessel detection in case of images having variation in diameters [17]. It even improved multiscale line detection approach. Although many techniques have been developed in the literature for retinal vessel detection, there is a scope for bringing improvement in this domain. Few parameters can be considered like the variations in luminance, contrast, and the gray level values of the background to achieve better results.

In this work, we utilize the capabilities of bird swarm algorithm (BSA) [18] along with river formation dynamics (RFD) algorithm [19] to optimize the segmentation of blood vessels. The hybrid approach developed with BSA and RFD proves to be robust even in case of noisy images. In this approach, we have used Kapur's entropy thresholding and OTSU thresholding function simultaneously in order to take the spatial information under consideration and select the class separating vessels and background in the best possible manner.

2 BSA-RFD Vessel Extraction Approach

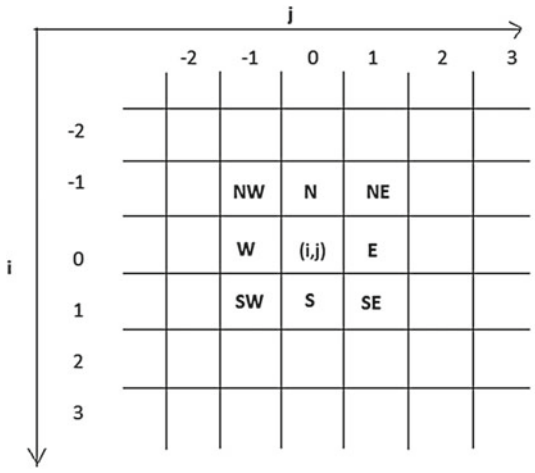
BSA is an optimization algorithm that imitates the behavior of the birds in swarms [18]. The algorithm models the foraging behavior, vigilance behavior, and flight behavior of the birds. The hybrid approach of BSA-RFD for vessel extraction is given in the following sections.

2.1 Exploration Region

The exploration region is the area in which the birds fly while searching for the food. The retinal fundus image on which the birds explore for the food is 2D in nature.

2.2 Foraging and Vigilance

Suppose N virtual birds occupy positions on the image in a random manner and are denoted by $p_b^t (b \in [1, \ldots, N])$ at the time t and $p^t = \left(p_1^t, p_2^t, \ldots, p_b^t\right)$ denotes the solution for threshold-based segmentation for b thresholds. All these birds prefer to fly together in flocks and search for the food. As they forage, they continuously keep an eye on their surroundings in order to protect themselves from threat. Therefore, after certain intervals they keep their heads high and try to look for threats around them. When a bird detects any predator, the alarm calls are sent to all the other birds flying in a swarm and accordingly the direction in which they all should fly is decided. This is known as vigilance. In this outlook, every bird searches for the edge pixel in the retinal fundus image. The intensity variations existing between the pixels represent the edges. The next location to fly is selected by the birds using the concept of 4-connectivity or 8-connectivity neighborhood. For example, if 8-connectivity neighborhood is considered there would be 8 possible directions as shown in Fig. 1.



Position	Direction
(i-1,j-1)	NW
(i-1,j)	N
(i-1,j+1)	NE
(i,j-1)	W
(i,j+1)	E
(i+1,j-1)	SW
(i+1,j)	S
(i+1,j+1)	SE

Fig. 1 Center pixel (i, j) surrounded by 8 neighborhood pixels

As per the bird swarm theory, when birds keep vigilance each bird competes with one another to move towards the center of the swarm. All the birds that have higher reservoir would have higher chance to get near to the middle of the swarm in comparison to the ones with lower reservoir. Therefore, each bird is not able to get towards the middle of the swarm in a direct manner. Hence, the next direction is found for the bird using Eq. 1 that comes from Bird Swarm algorithm.

$$x_{i,j}^{t+1} = x_{i,j}^{t} + W\left(\text{mean}_j - x_{i,j}^{t}\right) * \text{rand}(0, 1) + Z\left(p_{k,j} - x_{i,j}^{t}\right) * \text{rand}(-1, 1) \tag{1}$$

where

$$W = c * \exp\left(\frac{p\text{Fit}_i}{\text{sumFit} + \varepsilon} * N\right), \tag{2}$$

$$Z = d * \exp\left(\left(\frac{p\text{Fit}_i - p\text{Fit}_k}{|p\text{Fit}_k - p\text{Fit}_i| + \varepsilon}\right)\frac{N * p\text{Fit}_k}{\text{sumFit} + \varepsilon}\right) \tag{3}$$

where c and d denote two constants with positive values lying in $[0, 2]$. $p\text{Fit}_i$ shows the fitness value of ith birds and sumFit is the summation of the best fitness value of swarm. ε is the smallest constant that helps to avoid zero division error. mean_j denotes the jth element of average position of entire swarm. rand$(0, 1)$ denotes independent uniformly distributed numbers in $(0, 1)$. k is a positive integer and is chosen in a random manner between 1 and N. x_i^t at time step t represents the bird's position who search for the food.

As the next pixel is encountered by the bird, it is checked whether it is noisy or not. The image histogram technique is used to obtain noise boundaries. The pixels which are free of noise are picked up as the edge pixels and the ones with noise are removed using fuzzy impulse noise detection and reduction method [20].

The birds keep updating their best previous experience along with the best experience of swarm about the edge pixel using Eq. 4.

$$x_{i,j}^{t+1} = x_{i,j}^{t} + \left(r_{i,j} - x_{i,j}^{t}\right) * A * \text{rand}(0, 1) + \left(s_j - x_{i,j}^{t}\right) * B * \text{rand}(0, 1) \tag{4}$$

where $j \in [1, \ldots, N]$, rand$(0, 1)$ represents numbers in $(0, 1)$ which are uniformly distributed. A and B are two numbers that hold positive values and are known as cognitive and social accelerated coefficients, respectively. $r_{i,j}$ is the best previous position of ith bird and s_j denotes the best previous position of the swarm.

2.3 Selecting Optimal Path Using RFD

As the birds fly and forage for the food, it becomes very important to follow the optimal path. In our approach, the path that the birds follow while searching for the food has been optimized using another optimization algorithm known as river formation dynamics algorithm [19]. This algorithm mimics the behavior of riverbed formation.

The drops (birds) that are positioned at some seed pixel is influenced by gravitational forces due to which they move towards the center of the earth (center of the swarm). Therefore, these drops (birds) get distributed throughout the fundus image, move on edges and explore the search space to obtain the best solution. It is achieved through the process of erosion and soil sedimentation, associated with the changes in altitude associated with each pixel on the image. Drops (birds) when move throughout the search space; modify the altitude values of the pixels that lie along their path. RFD algorithm is a variant of the ant colony optimization algorithm based on gradient. The probability of choosing the next edge pixel is dependent on the gradient value.

Drops (birds) move (fly) till they reach the goal (edge pixel) or traverse maximal number of prescribed pixels in an image. This maximum number of pixels is considered to be the total number of pixels on an image. The probability $P_{bi}(a, b)$ that a bird bi residing on node a would select the next node b is given as per Eq. 5.

$$P_{bi}(a,b) = \begin{cases} \frac{\text{gradient}(a,b)}{\text{total}}, & \text{for } b \in N_k(a) \\ \frac{\omega/|\text{gradient}(a,b)|}{\text{total}}, & \text{for } b \in B_k(a) \\ \frac{\delta}{\text{total}}, & \text{for } b \in E_k(a) \end{cases} \tag{5}$$

where

$$\text{gradient}(a,b) = \frac{\text{altitude}(a) - \text{altitude}(b)}{\text{distance}(a,b)} \tag{6}$$

$$\text{total} = \left(\sum_{m \in N_k(a)} \text{gradient}(a,m) \right) + \left(\sum_{m \in B_k(a)} \frac{\omega}{|\text{gradient}(a,m)|} \right) + \left(\sum_{m \in E_k(a)} \delta \right) \tag{7}$$

$N_k(a)$ denotes the set of neighboring pixels with gradient of positive value (pixel a) has a higher value of altitude than pixel b, $B_k(a)$ represents a set of neighboring pixels with gradient having negative value (altitude of pixel b is greater) and $E_k(a)$ denotes the neighboring pixels with flat gradient. ω and δ are the coefficients that have a small certain fixed value.

When the birds pass through all the pixels, an erosion process gets implemented on all traveled paths. The amount of erosion for each pair of pixels a and b depends upon the number of birds B, number of all pixels on the image P, and an erosion coefficient E.

$$\forall a, b \in \text{Path}_b, \quad \text{altitude}(a) := \text{altitude}(a) - \frac{E}{(P-1).B}.\text{gradient}(a, b) \tag{8}$$

where Path_k is a path traveled by bird b.

The original RFD algorithm has been modified by Redlarski et al. so that the transition probability for the next pixel to be selected is based on an exponential function [21]. The modified probability shown in Eq. 9 has two coefficients: pBase denotes the base for the exponent and α denotes the convergence tuning coefficient. The e_j denotes the Euclidean distance from pixel j to the goal.

$$P_b(a, b) = \frac{\text{gradientparam}_{i,j}}{(d_b)^{\alpha}}/\text{total}, \tag{9}$$

$$\text{total} = \sum_{m \in N_k(a) \cup U_k(a) \cup F_k(a)} \frac{\text{gradientparam}_{i,m}}{(d_m)^{\alpha}} \tag{10}$$

where

$$\text{gradientparam}_{i,j} = p\text{Base}^{\text{gradient}(a,b)} \tag{11}$$

2.4 Selecting the Best Threshold

The unhealthy birds that do not have much contribution in obtaining the optimal solution are separated from the swarm. A threshold value thresh is considered. As the next pixel is accessed by the bird, the intensity value of that pixel is compared to the threshold value. In case the intensity of the pixel is identified to be less than the threshold value thresh, the bird is treated as an unhealthy one, is removed from the swarm and replaced by healthy birds. When all the iterations get over and birds lying on the optimal path show the intensity value to be larger than the threshold value, the final vessel map is obtained.

3 Results and Discussions

The proposed approach discussed in the above section has been tested and analyzed using the publicly available dataset namely STARE (Structured Analysis of the Retina) which consists of 20 images out of which 10 reflect the symptoms of eye diseases and rest do not show any signs. The retinal fundus images are obtained using high fundus camera keeping the illumination and field of view to be constant. The results have been obtained using bird swarm algorithm and river formation dynamics algorithm and have been compared with Flower Pollination Search algorithm [12], Ant Colony optimization algorithm [11] and Matched filter [22] against the ground truth (manual segmentation by two experts) with respect to accuracy, sensitivity, and specificity. The retinal fundus images from STARE database and the results obtained by the proposed approach are shown in Fig. 2.

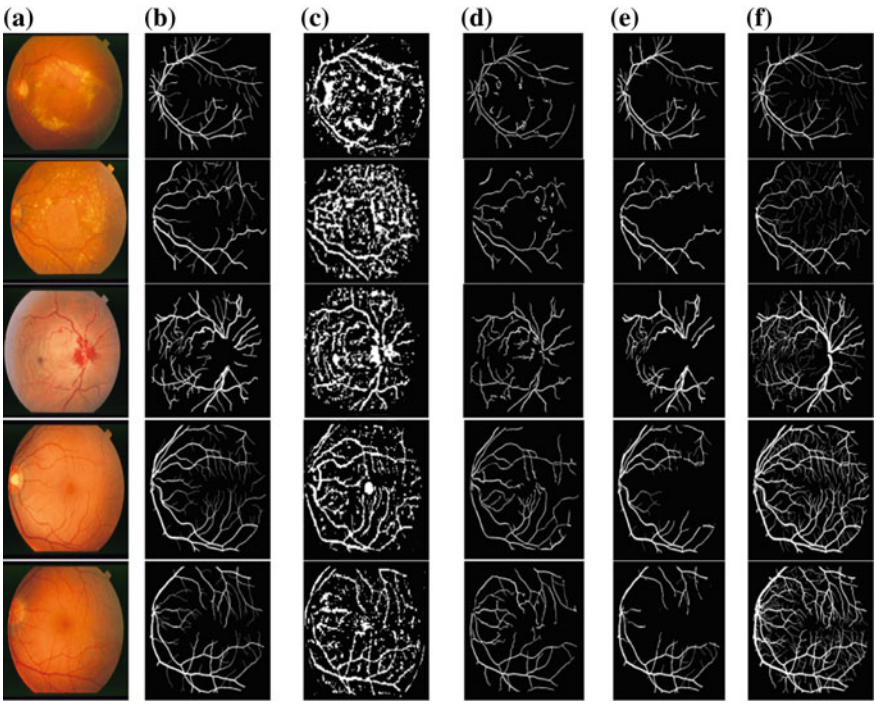


Fig. 2 Retinal vessel detection in few images from database. **a** Unprocessed fundus images. **b** Ground truth. **c** Ant colony optimization algorithm. **d** Matched filter technique. **e** Flower pollination algorithm. **f** Proposed Approach using BSA and RFD

It can be observed in Fig. 2 that the results produced by BSA and RFD are visually better as compared to the rest of the techniques. Ant colony optimization is producing lot of information and the vessels detected are very thick. Matched filter technique is producing thin vessels but many of the blood vessels are missing and only few features have been detected clearly. In case of flower pollination, the results are quite near to proposed approach in terms of visual quality but few thin blood vessels are missing. In case of proposed approach all the blood vessels have been detected and are thin in nature. The parameters namely sensitivity, specificity, and accuracy have been evaluated and the values have been found for 20 images as displayed in Tables 1, 2 and 3, respectively. It can be observed that all the parameters have shown the highest value for the proposed approach. Figure 3 shows the flowchart summarizing the implementation of the proposed approach.

Table 1 Sensitivity with respect to different approaches

Image number	Ant colony optimization [11]	Matched filter [22]	Flower pollination algorithm [12]	Proposed approach
1	88.90	71.32	89.67	93.45
2	87.76	72.21	87.63	92.45
3	87.52	71.45	88.52	91.27
4	88.31	74.28	87.29	93.25
5	85.29	72.18	86.24	91.82
6	88.23	74.28	87.38	93.27
7	84.83	71.26	88.82	93.85
8	86.33	74.22	89.25	91.36
9	87.93	73.24	89.76	92.54
10	88.67	71.22	88.37	92.53
11	85.53	72.54	89.53	92.72
12	84.12	75.82	88.72	94.52
13	82.71	77.16	87.42	90.81
14	87.23	73.27	89.32	94.22
15	82.81	74.27	89.81	92.54
16	84.34	72.21	88.22	92.37
17	87.33	75.29	88.28	94.28
18	81.82	74.27	89.84	92.87
19	85.34	72.22	86.25	92.36
20	83.93	76.24	88.96	94.54
Average	85.94	73.44	88.46	92.85

Table 2 Specificity with respect to different approaches

Image number	Ant colony optimization [11]	Matched filter [22]	Flower pollination algorithm [12]	Proposed approach
1	88.82	86.34	89.82	92.34
2	87.46	85.24	87.63	93.65
3	88.22	86.46	88.62	92.77
4	87.11	84.18	89.19	92.25
5	86.92	85.58	88.24	93.52
6	88.79	84.98	89.38	92.17
7	87.53	84.23	88.62	93.55
8	88.23	85.82	89.55	92.26
9	86.43	86.44	88.46	93.54
10	88.57	85.22	87.27	93.43
11	84.93	83.84	88.43	91.32
12	87.72	84.82	87.42	93.42
13	88.77	86.16	88.42	91.51
14	87.13	85.27	89.52	93.62
15	88.85	83.27	87.81	94.54
16	87.24	85.21	86.52	94.37
17	86.33	84.29	88.88	93.18
18	87.82	83.27	89.74	93.87
19	88.24	84.22	87.75	92.36
20	87.95	85.24	88.66	91.54
Average	87.65	85.00	88.49	92.96

4 Conclusion

An automated system for the extraction of retinal blood vessels has been proposed using the hybrid approach of bird swarm algorithm and river formation dynamics algorithm. The imitation of behavior of birds and formation of river beds has been combined together and the path taken by the birds in BSA has been optimized using RFD. The results obtained using the proposed approach has been compared with ant colony optimization algorithm, matched filter, and flower pollination algorithm. The analysis of the outputs has been done in terms of sensitivity, specificity, and accuracy and it is observed that the results obtained through proposed approach are better and promising with respect to all the mentioned parameters. As a future direction, the proposed hybrid approach can even be employed for automated diagnosis of other eye-related diseases.

Table 3 Accuracy with respect to different approaches

Image number	Ant colony optimization [11]	Matched filter [22]	Flower pollination algorithm [12]	Proposed approach
1	90.32	89.56	93.67	95.23
2	91.38	89.32	93.25	94.27
3	90.62	88.63	91.33	95.33
4	89.55	89.62	92.45	95.22
5	88.46	89.79	93.55	95.34
6	90.27	88.14	92.26	95.21
7	89.43	89.28	93.54	95.28
8	91.42	88.52	93.43	95.38
9	89.42	88.35	92.52	95.22
10	89.62	89.46	92.52	92.33
11	90.81	88.27	93.81	95.76
12	88.52	87.43	93.22	95.24
13	90.52	88.42	91.54	95.35
14	91.17	86.27	92.37	95.29
15	88.55	89.43	93.28	94.29
16	90.26	88.42	91.87	95.35
17	90.54	90.42	92.36	95.10
18	89.43	89.52	93.54	95.28
19	90.32	88.81	92.53	95.38
20	91.42	87.52	93.72	95.32
Average	90.10	88.75	92.83	95.05

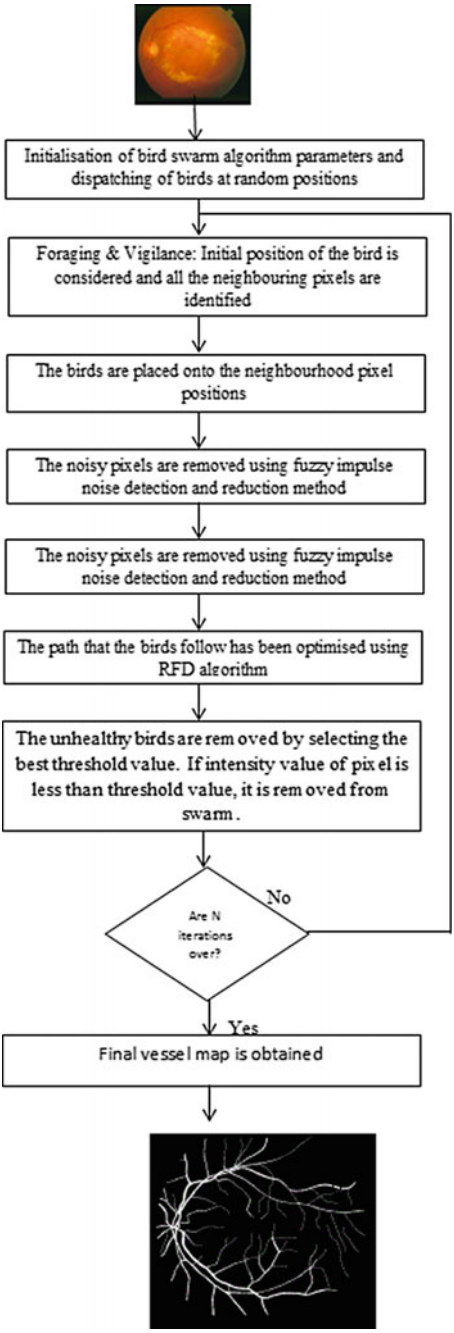


Fig. 3 Flow-chart showing the summary of implementation of the proposed algorithm

References

1. M.M. Fraz, P. Remagnino, A. Hoppe, B. Uyyanonvara, A.R. Rudnicka, C.G. Owen, S.A. Barman, Blood vessel segmentation methodologies in retinal images—a survey. Comput. Methods Programs Biomed. **108**, 407–433 (2012). https://doi.org/10.1016/j.cmpb.2012.03.009
2. M. Forracchia, E. Grisan, A. Ruggeri, Extraction and quantitative description of vessel features in hypertensive retinopathy fundus images, in *Abstracts of 2nd International Workshop on Computer Assisted Fundus Image Analysis* (2011)
3. P. Mitchell, H. Leung, J.J. Wang, E. Rochtchina, A.J. Lee, T.Y. Wong, R. Klein, Retinal vessel diameter and open-angle glaucoma: the Blue Mountains Eye Study. Ophthalmology **112**, 245–250 (2005). https://doi.org/10.1016/j.ophtha.2004.08.015
4. J.J. Wang, B. Taylor, T.Y. Wong, B. Chua, E. Rochtchina, R. Klein, P. Mitchell, Retinal vessel diameters and obesity: a population-based study in older persons. Obesity **14**, 206–214 (2006). https://doi.org/10.1038/oby.2006.27
5. K. Goatman, A. Charnley, L. Webster, S. Nussey, Assessment of automated disease detection in diabetic retinopathy screening using two-field photography. PLoS ONE **6**, e27524 (2011). https://doi.org/10.1371/journal.pone.0027524
6. E. Grisan, A. Ruggeri, A divide et impera strategy for automatic classification of retinal vessels into arteries and veins, in *Proceedings of the 25th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (IEEE Cat. No.03CH37439)*. IEEE (2003), pp. 890–893
7. A. Can, H. Shen, J.N. Turner, H.L. Tanenbaum, B. Roysam, Rapid automated tracing and feature extraction from retinal fundus images using direct exploratory algorithms. IEEE Trans. Inf. Technol. Biomed. **3**, 1–14 (1999)
8. M. Lalonde, L. Gagnon, M.-C. Boucher, Non-recursive paired tracking for vessel extraction from retinal images, in *Proceedings of the Conference Vision Interface* (2000), pp. 61–68
9. E. Grisan, A. Pesce, A. Giani, M. Foracchia, A. Ruggeri, A new tracking system for the robust extraction of retinal vessel structure. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. **3**, 1620–1623 (2004). https://doi.org/10.1109/IEMBS.2004.1403491
10. D. Marín, A. Aquino, M.E. Gegúndez-Arias, J.M. Bravo, A new supervised method for blood vessel segmentation in retinal images by using gray-level and moment invariants-based features. IEEE Trans. Med. Imaging **30**, 146–158 (2011). https://doi.org/10.1109/TMI.2010.2064333
11. G. Kavitha, S. Ramakrishnan, Detection of blood vessels in human retinal images using Ant Colony Optimisation method. Int. J. Biomed. Eng. Technol. **5**, 360 (2011). https://doi.org/10.1504/IJBET.2011.039926
12. E. Emary, H.M. Zawbaa, A.E. Hassanien, M.F. Tolba, V. Snášel, Retinal vessel segmentation based on flower pollination search algorithm, in *Advances in Intelligent Systems and Computing*, ed. by P. Kömer, A. Abraham, V. Snášel (Springer International Publishing, Cham, 2014), pp. 93–100
13. Q. Li, B. Feng, L. Xie, P. Liang, H. Zhang, T. Wang, A cross-modality learning approach for vessel segmentation in retinal images. IEEE Trans. Med. Imaging **35**, 109–118 (2016). https://doi.org/10.1109/TMI.2015.2457891
14. E. Emary, H.M. Zawbaa, A.E. Hassanien, B. Parv, Multi-objective retinal vessel localization using flower pollination search algorithm with pattern search. Adv. Data Anal. Classif. **11**, 611–627 (2017). https://doi.org/10.1007/s11634-016-0257-7
15. B. Khomri, A. Christodoulidis, L. Djerou, M.C. Babahenini, F. Cheriet, Retinal blood vessel segmentation using the elite-guided multi-objective artificial bee colony algorithm. IET Image Process (2018), pp. 2–12 https://doi.org/10.1049/iet-ipr.2018.5425
16. V. Sathananthavathi, G. Indumathi, BAT algorithm inspired retinal blood vessel segmentation. IET Image Process **12**, 2075–2083 (2018). https://doi.org/10.1049/iet-ipr.2017.1266
17. B. Khomri, A. Christodoulidis, L. Djerou, M.C. Babahenini, F. Cheriet, Particle swarm optimization method for small retinal vessels detection on multiresolution fundus images. J. Biomed. Opt. **23**, 1 (2018). https://doi.org/10.1117/1.jbo.23.5.056004

18. X.-B. Meng, X.Z. Gao, L. Lu, Y. Liu, H. Zhang, A new bio-inspired optimisation algorithm: Bird Swarm Algorithm. J. Exp. Theor. Artif. Intell. **28**, 673–687 (2015). https://doi.org/10.1080/0952813X.2015.1042530
19. P. Rabanal, I. Rodríguez, F. Rubio, Using river formation dynamics to design heuristic algorithms, *Unconventional Computation* (Springer, Berlin, Heidelberg, 2007), pp. 163–177
20. S. Schulte, M. Nachtegael, V. De Witte, D. Van der Weken, E.E. Kerre, A fuzzy impulse noise detection and reduction method. IEEE Trans. Image Process. **15**, 1153–1162 (2006). https://doi.org/10.1109/TIP.2005.864179
21. G. Redlarski, M. Dabkowski, A. Palkowski, Generating optimal paths in dynamic environments using River Formation Dynamics algorithm. J. Comput. Sci. **20**, 8–16 (2017). https://doi.org/10.1016/j.jocs.2017.03.002
22. A. Hoover, V. Kouznetsova, M. Goldbaum, Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response. IEEE Trans. Med. Imaging **19**, 203–210 (2000). https://doi.org/10.1109/42.845178

Image Processing for UAV Using Deep Convolutional Encoder–Decoder Networks with Symmetric Skip Connections on a System on Chip (SoC)



Abhiraj Hinge, Pranav Garg and Neena Goveas

1 Introduction

Many UAV based applications being developed require analysis of images which have been captured by its onboard cameras [1]. This analysis typically involves identification of the presence and location of an object from a given input set. An analysis using a ground-based computer would involve (1) Compression and transmission of the image (2) Analysis on the ground-based computer and (3) retransmission back to the UAV. This is not feasible when dealing with analysis which is continuous in nature, for example, a tracking application [2, 3]. Another reason for such an approach failing could be the poor quality of the transmission link. This becomes worse in case of long distance applications where connections may not be available. The only feasible solution is to perform the analysis onboard the UAV [4, 5]. Here we propose an onboard UAV image processing application using a Deep Convolutional Encoder–Decoder Networks with Symmetric Skip Connections running on a SoC with a GPU. In our proposed setup we pretrain our neural network on a computer with a GPU. The trained neural network can then be executed on a SoC with GPU. The neural network has been tested with a Nvidia Jetson TX2 embedded computer and is found to accurately and efficiently perform the task of image segmentation on the SpaceNet [6] dataset. Thus, it is feasible to use the proposed setup for image analysis tasks onboard a UAV. The rest of this paper is organized as follows: In

A. Hinge · P. Garg · N. Goveas (✉)
Department of Computer Science & Information Systems, BITS Pilani, Goa Campus, Sancoale 403726, Goa, India
e-mail: neena@goa.bits-pilani.ac.in

A. Hinge
e-mail: f20160053@goa.bits-pilani.ac.in

P. Garg
e-mail: f20150295@goa.bitspilani.ac.in

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_102

Sect. 2, the past work on image processing has been described. It also describes the minimum criteria to be satisfied for any methods proposed for aerial image analysis. In Sect. 3 a brief description of deep neural networks is given. The network that has used, Deep Convolutional Encoder–Decoder Networks with Symmetric Skip Connections, is explained in Sect. 4. The results from the implementation and evaluation are presented in Sect. 5. Section 6 presents the conclusions.

2 Image Capturing and Analysis in UAV

Many UAV applications require the analysis of aerial images captured by onboard cameras for navigation and obstacle avoidance, among others. The images captured by onboard cameras have several issues that need to be solved before their usage. These include poor environmental conditions, nonstandard angles for capture, vibrations due to UAV operations, etc.

2.1 UAV Applications Using Images

Uses of images acquired by UAVs include autonomous navigation, locating and tracking of objects, segmentation and comparison with stored images. Tracking applications rely on a continuous stream of images being taken and locating the object being tracked in sequential images. This movement of the tracked object also needs to be translated to inputs to the trajectory of the UAV motion. Autonomous navigation is mostly done using satellite navigation systems as the main position source. In locations where the radio signal from the global satellite navigation systems (GNSS) is not present, UAV positioning is not possible. Closed spaces, areas in the shadow of larger structures cannot be explored without an alternate positioning mechanism. One of the proposed approaches is autonomous navigation using vision-based techniques [3, 7]. Other applications include survey of locations, identifying and positioning based on landmarks, etc. These and other future applications all need effective image analysis being performed within a short time-span. In addition, if the UAV has to fly long distances in remote areas, it needs autonomous functionality. This is difficult to achieve if any image analysis based processing with transmission to a ground-based machine is required.

2.2 Issues with UAV Images

Images that UAV applications deal with are aerial views of the surroundings. These are typically low resolution images due to the hardware constraints. Any object detection technique developed for these images will have to work towards detecting

objects in a size invariant manner. In addition, the angle of the image, occlusion, light and shadows all play a role in complicating this process. Traditional image processing techniques have been used for object detection and collision avoidance [8]. These techniques work well if high accuracy Inertial Navigation Systems (INS) is present to provide an appropriate level of image stabilization. Without preprocessing and modifications, traditional image processing techniques which may have worked successfully in non-aerial applications may not work in UAV. In addition these preprocessing stages take time to execute making it difficult to use for applications requiring continuous image analysis.

3 Deep Neural Networks Based Image Analysis

Deep learning allows computational models that are composed of multiple processing layers to learn representations of data with multiple levels of abstraction [5] and has been successfully used in state of the art object recognition and object detection. A convolutional neural network (CNN) is a computational model that works by passing images through a series of kernels which serve as filters and effectively perform a pixel-wise analysis of the image to segment and classify parts of the image. Since the introduction of GPU optimization for training a neural network, deep learning has provided solutions for a wide range of image processing problems. Recently, UAV equipped with NVIDIA TX2 card and a customized convolutional neural network (CNN) has been shown to be effective for fast onboard vision processing capable of navigation [9].

4 Deep Convolutional Encoder–Decoder Networks with Symmetric Skip Connections

Development of image segmentation applications is now at an interesting stage due to several challenges being proposed and datasets being available for researchers. We use the network proposed by Mao et al., which has been successfully used for image restoration [10].

The network used is an end to end convolutional network which functions as an autoencoder. This network as shown in Fig. 1 consists of four convolutional layers which serve to compress the image, following them, the network has four de-convolutional layers which decode the output from its compressed state. The network then adds the original input image to the output of these eight layers and runs it through two convolutional layers with the intention of passing on data to aid the reconstruction.

Description of the layers is as follows:

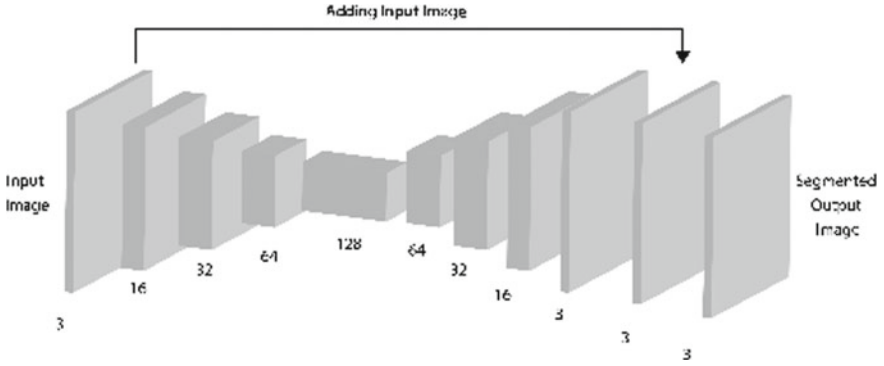


Fig. 1 Visual representation of the network

1. First Convolutional Layer (inchannels = 3, outchannels = 16, kernelsize = 4, stride = 1, padding = 0)
2. Max Pooling Layer 1 (kernelsize = 2, stride = 2)
3. Second Convolutional Layer (inchannels = 16, outchannels = 32, kernelsize = 5, stride = 1, padding = 0)
4. Max Pooling Layer 2 (kernelsize = 2, stride = 2)
5. Third Convolutional Layer (inchannels = 32, outchannels = 64, kernelsize = 3, stride = 1, padding = 0)
6. Max Pooling Layer 3 (kernelsize = 2, stride = 2)
7. Fourth Convolutional Layer (inchannels = 64, outchannels = 128, kernelsize = 2, stride = 1, padding = 0)
8. First De-Convolutional Layer (inchannels = 128, outchannels = 64, kernelsize = 2, stride = 1, padding = 0)
9. Max UnPooling Layer 1 (kernelsize = 2, stride = 2)
10. Second De-Convolutional Layer (inchannels = 64, outchannels = 32, kernelsize = 3, stride = 1, padding = 0)
11. Max UnPooling Layer 2 (kernelsize = 2, stride = 2)
12. Third De-Convolutional Layer (inchannels = 32, outchannels = 16, kernelsize = 5, stride = 1, padding = 0)
13. Max UnPooling Layer 3 (kernelsize = 2, stride = 2)
14. Fourth De-Convolutional Layer (inchannels = 16, outchannels = 3, kernelsize = 4, stride = 1, padding = 0)
15. Adding the input to the output obtained from Fourth De-Convolutional Layer
16. End Convolutional Layer 1 (inchannels = 3, outchannels = 3, kernelsize = 5, padding = 2)
17. End Convolutional Layer 2 (inchannels = 3, outchannels = 3, kernelsize = 3, padding = 1).

5 Results

The network has been trained on a multimode Tesla K-20 m GPU machine. The model was trained for six epochs and took 300 min. The time for running the trained network over 3850 images was 4.81 min.

The Vegas section of the SpaceNet dataset has been used to train the model [6]. This consists of 30 cm imagery collected from WorldView-3 Satellite. It contains footprints for buildings which have been used. The footprints were plotted onto the satellite images using QGIS [11] and these images were used as labelled images to calculate the loss and train the network. The processed dataset used had 3465 images for training. The aim is to train the network to perform instance segmentation of rooftops of buildings using satellite images. Figure 2 shows a sample image from the Vegas section, Fig. 3 shows a labelled training image from the dataset. Output from our network is shown in Fig. 4.

We find that the trained network gave MSE (Mean Squared Error) loss on the unseen test dataset as 660. The PSNR (Peak Signal-to-Noise Ratio) obtained on test dataset was 20. The SSIM score obtained over the test dataset is 0.7435. We also benchmarked the trained model on a Nvidia Jetson TX2 and found that it evaluated results for 3850 satellite images in 13.2 min. The results obtained for a complex task shows that the SoC can accurately perform the task of image segmentation in less than 5 s per image.

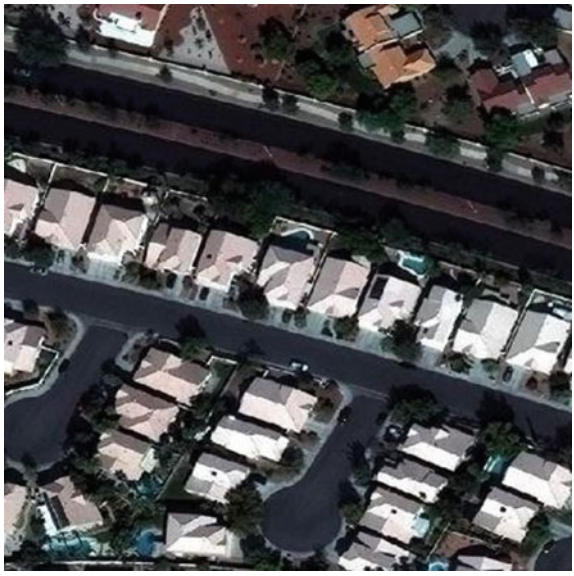


Fig. 2 Training image



Fig. 3 Ground truth



Fig. 4 Trained network output

6 Conclusions

In this paper, we propose a computationally efficient approach to solve the problem of onboard image analysis for UAV and successfully test it on the SpaceNet Dataset. We show that a pretrained Deep Convolutional Encoder–Decoder Network with Symmetric Skip Connections can perform the image analysis task of segmentation. This network can be used onboard a UAV with a GPU and thus be used for onboard image analysis. In our study we have used a NVIDIA Jetson TX2 as the GPU. The task of image segmentation is complex and the time per image of a few seconds makes it possible to design applications like navigation using the results of onboard image analysis. This enhanced analysis capability will result in design of UAVs with autonomous capabilities as a result of image capture and analysis performed onboard.

Acknowledgements Authors acknowledge the GPU grant of a Titan XP by NVIDIA for this work.

References

1. L. Merino, F. Caballero, J.R. Martinez-de Dios, I. Maza, A. Ollero, An unmanned aircraft system for automatic forest fire monitoring and measurement. J. Intell. Robot. Syst. **65**(1–4), 533–548 (2012)
2. Z.X. Liu, C. Yuan, Y.M. Zhang, J. Luo, A learning-based fault tolerant tracking control of an unmanned quadrotor helicopter. J. Intell. Robot. Syst. **84**(1), 145–162 (2016)
3. P. Burdziakowski, M. Przyborski, A. Janowski, J. Szulwic, A vision-based unmanned aerial vehicle navigation method, in *1st International Conference on Innovative Research and Maritime Applications of Space Technology* (2015)
4. C. Szegedy et al., Going deeper with convolutions, in *CVPR* (2015)
5. A. Krizhevsky, I. Sutskever, G. Hinton, ImageNet classification with deep convolutional neural networks. Proc. Adv. Neural Inf. Process. Syst. **25**, 1090–1098 (2012)
6. https://spacenetchallenge.github.io/AOI_Lists/AOI_2_Vegas.html. Accessed 5 Feb 2019
7. A. Stateczny, Artificial neural networks for comparative navigation, in *Artificial Intelligence and Soft Computing*, ed. by L. Rutkowski, J. Siekmann, R. Tadeusiewicz, et al. *ICAISC 2004. LNAI, 3070* (2004), pp. 1187–1192
8. D. Bratanov, L. Mejias, J.J. Ford, A vision-based sense-and-avoid system tested on a ScanEagle UAV, in *International Conference on Unmanned Aircraft Systems (ICUAS)*, 13–16 June 2017 (2017), pp. 1134–1142
9. P. Burdziakowski, M. Przyborski, J. Szulwic, A vision-based unmanned aerial vehicle navigation method, in *IRMAST* (2015)
10. http://papers.nips.cc/paper/6172-image-restoration-using-very-deep-convolutional-encoder-decoder-networks-with-symmetric-skip-connections. Accessed 5 Feb 2019
11. QGIS Development Team, *QGIS Geographic Information System*. Open Source Geospatial Foundation Project. http://qgis.osgeo.org. Accessed 5 Feb 2019
12. https://www.cv-foundation.org/openaccess/content_cvpr_2015/app/2B_011.pdf. Accessed 5 Feb 2019

A Comparison of Different Filtering Strategies Used in Attribute Profiles for Hyperspectral Image Classification



Arundhati Das, Kaushal Bhardwaj and Swarnajyoti Patra

1 Introduction

The hyperspectral image (HSI) acquisition tools are constantly improving the acquisition quality both in terms of spectral and spatial resolution. HSI is capable of providing accurate identification of objects in an acquired scene because of its hundreds of spectral bands. Each band of the acquired scene covering heterogeneous objects is one gray-level image with high spatial resolution [9]. The incorporation of spatial or contextual knowledge in the objects appearing on the acquired scene can be successfully achieved with a set of operators provided by mathematical morphology (MM) [1, 5, 7]. The morphological profiles (MPs) defined in the MM framework, are suggested as an effective tool for fusing the spatial information with the spectral data of hyperspectral images [1, 8]. The MPs are constructed by concatenating the filtered images obtained by the operators namely, opening and closing by reconstruction on gray-level images. An extended MP (EMP) is obtained by concatenating MPs constructed on few component images extracted by any feature reduction technique (e.g., principal component analysis (PCA)) computed on an HSI. Although MPs incorporate important information on the geometry of the structures, they can only model the information of a particular size due to the fixed size of structuring elements (SEs) used while filtering the image [5, 7]. The other limitation is that the image cannot be filtered based on other geometrical or gray-level properties of the objects. These limitations of MPs are overcome by using the morphological attribute profiles (APs) [5]. APs provide incorporation of spatial information at multiple levels by concatenating a series of filtered images obtained using morphological attribute filters (AFs). The connected filters AFs incorporate spatial information by processing connected

A. Das (✉) · K. Bhardwaj · S. Patra
Department of CSE, Tezpur University, Tezpur 784028, India
e-mail: arundha@tezu.ernet.in

K. Bhardwaj
e-mail: kauscsp@tezu.ernet.in

S. Patra
e-mail: swpatra@tezu.ernet.in
URL: http://agnigarh.tezu.ernet.in/~swpatra/index.html

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_103

components of a gray-level image. They can be efficiently implemented with the help of a tree structure [10]. The attribute filtering operations can be completed through a three-step process. The first step is the creation of a tree structure (max-tree and its dual min-tree). Max-tree (min-tree) possesses the connected components with the lowest (highest) gray-level at the root and the connected components with the highest (lowest) gray-level at the leaves. In the second step, some attribute values are calculated on the tree nodes and depending on a predefined criterion filtering is done where filtering strategies play a vital role. The filtering strategy decides whether the nodes satisfying the criterion are to be merged into their background or not. It also decides what is to be done with the children of the nodes. The third step is the restitution of the processed tree back to a gray-level image.

In this paper, an empirical study to compare the four different filtering strategies, namely *subtractive*, *max*, *min*, and *direct*, used for attribute filtering operations are presented. The effectiveness of these strategies are compared in respect of overall accuracy, kappa accuracy, standard deviation of overall accuracy obtained by an SVM classifier on five attributes of diverse nature. The experimental results obtained on two real HSI data sets show that the strategies *subtractive*, *min*, and *direct* perform robustly for all the five considered attributes.

The paper is ordered as follows: Sect. 2 recalls the construction of APs and explains the four filtering strategies used in the AF. Section 3 gives a brief description of the hyperspectral data sets considered in our experiments. Section 4 demonstrates the experimental results. Finally, Sect. 5 draws the conclusion.

2 Attribute Filtering Strategies and Attribute Profiles

Attribute filters perform filtering by merging the connected components (iso-level flats zones) in a gray-level image. So, the filtered images preserve the geometrical structures of various objects in the original image. Additionally, AFs provide the flexibility to incorporate spatial information according to any characteristics or property of the connected components as an attribute (eg., size, shape, gray-level values, orientation, etc.) [6]. Furthermore, an efficient tree representation (max-tree/min-tree) can evaluate an AF by representing the image in a hierarchical manner.

2.1 Attribute Filtering Strategies

As mentioned, the AFs follow a three-step procedure where the different filtering strategies are adopted in the filtering step. Figure 1 demonstrates the steps of an AF with the help of a synthetic gray-level image. Firstly, the original gray-level image is represented as a max-tree. For a detail construction of max-tree from an image, the readers are referred to [2, 10]. Next, the filtering of the max-tree is done considering a predefined criterion (for example $Attribute(N_j^i) > \lambda$). Here, the attribute value of a node N_j^i is evaluated and compared against a threshold value λ. In this case, N_j^i

having attribute value less than λ are to be merged to their background depending on the filtering strategy being used. A filtering strategy is a rule adopted during the merging of the nodes not satisfying a criterion that decides whether the node has to be merged or not as well as what is to be done with their children. A node is preserved if its attribute value is greater than the threshold value else it is merged into its immediate background. During the merging of the nodes to their parents, a filtering strategy or rule is adopted. The filtering strategies are distinguished as a pruning strategy or non-pruning strategy. In the pruning strategy, a node is merged or preserved along with its children. In the non-pruning strategy, the children of a merged node are preserved and linked to the parent of the merged node. The *max* and *min* are pruning strategies whereas *direct* and *subtractive* are non-pruning strategies [5, 11]. The filtering strategies are described in Table 1.

After processing the max-tree according to the filtering criterion adopting one of the filtering strategies, a filtered max-tree is obtained as shown in Fig. 1. Finally, the restitution of the filtered gray-level image is carried out from the filtered max-tree. AF, when performed using a max-tree, is termed as thinning and when it is performed using the min-tree (dual of a max-tree) it is termed as thickening.

2.2 Attribute Profiles

In case of remote sensing images the objects present in the scene generally possess heterogeneous characteristics in terms of size, shape etc. Thus, incorporation of

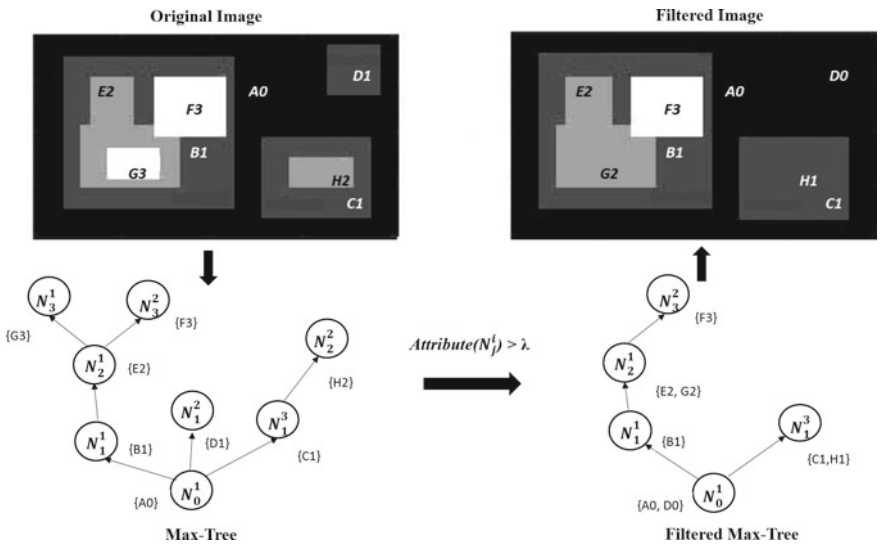


Fig. 1 The attribute filtering operation applied on a gray-level image

Table 1 A brief description of different filtering strategies

Strategy	Description
Subtractive	A node is removed when it does not satisfy the criterion and the level of all its child nodes are lowered by the difference of the gray-level between the removed node and its parent
Max	A node is removed only when neither it nor any of its children satisfies the criterion
Min	The node that does not satisfy the criterion is merged to its parent node along with all it's children
Direct	A node is removed if it does not satisfy the criterion and its descendants are linked to the parent of the removed node

the spatial information is required to perform at multiple levels. By performing multiple thickening and thinning operations with a sequence of threshold values (λs), incorporation of spatial information at multiple levels can be achieved for a remote sensing image. An AP is concatenation of the original gray-level image with filtered images obtained from such multiple thinning and thickening operations. An attribute profile $AP(I)$ for a grayscale image I is defined as:

$$AP(I) = \{\phi_{\lambda_n}(I), \phi_{\lambda_{n-1}}(I), \ldots, \phi_{\lambda_1}(I), I, \gamma_{\lambda_1}(I), \gamma_{\lambda_2}(I), \ldots, \gamma_{\lambda_n}(I)\} \quad (1)$$

Here, the AP stores the original image I with its *thickening* transforms ϕ_{λ_i} and *thinning* transforms γ_{λ_i} where λ_i is the ith threshold parameter value used during filtering operation. The threshold parameter value λ_i varies from $i = 1, 2, \ldots, n$ resulting in n *thickening* and n *thinning* transforms.

In case of hyperspectral data, the concept of APs is extended by Dalla Mura et al. and termed as extended APs (EAPs) [4]. Similar to an EMP, the EAP is obtained by concatenating APs constructed on a few component images extracted by the feature reduction method PCA computed on an HSI. For an HSI H taking into account the first L PCs, the EAP is formulated as:

$$EAP(H) = \{AP(PC_1), AP(PC_2), \ldots, AP(PC_L)\} \quad (2)$$

3 Hyperspectral Data Sets

The effectiveness of different filtering strategies on various attributes is assessed by conducting experiments on two hyperspectral image data sets.[1] The first data set is the University of Pavia captured by ROSIS sensors while the second data set is the Kennedy Space Center (KSC) captured by AVIRIS sensors. In the following, a brief description of the data sets are given.

[1] Accessible at: http://www.ehu.eus/ccwintco/index.php?title=Hyperspectral_Remote_Sensing_Scenes.

Table 2 The classes and corresponding number of available labeled samples for The University of Pavia data set

Class no.	Class name	#Labeled samples	Class no.	Class name	#Labeled samples
1	Asphalt	6631	6	Bare Soil	5029
2	Meadows	18649	7	Bitumen	1330
3	Gravel	2099	8	Self-Blocking Bricks	3682
4	Trees	3064	9	Shadows	947
5	Metal Sheets	1345		**Total**	**42776**

ROSIS University of Pavia: The first data set represents a hyperspectral image of size 610×340 pixels captured over a portion of North Italian urban area of Pavia including The University of Pavia. The scene has a total of 103 number of spectral bands available after discarding 12 noisy bands and has a spatial resolution of 1.3 m. The coverage range of spectral bands is from 0.43 to 0.86 μm. Figure 2a shows a three-band color image of the University of Pavia along with its ground truth. The ground truth reference of the acquired scene is pointing out nine thematic classes whose details are listed in Table 2.

AVIRIS Kennedy Space Center: The next data set captured a scene of the Kennedy Space Center situated in Florida having a spectral coverage range 400–2500 nm. Each captured hyperspectral image has a size of 512×614 pixels with 18 m spatial resolution and total there are 224 spectral bands out of which 176 spectral bands are available for use. Figure 2b shows a three-band color image of the KSC data and its ground truth. Table 3 shows the thirteen classes for which ground truth is available.

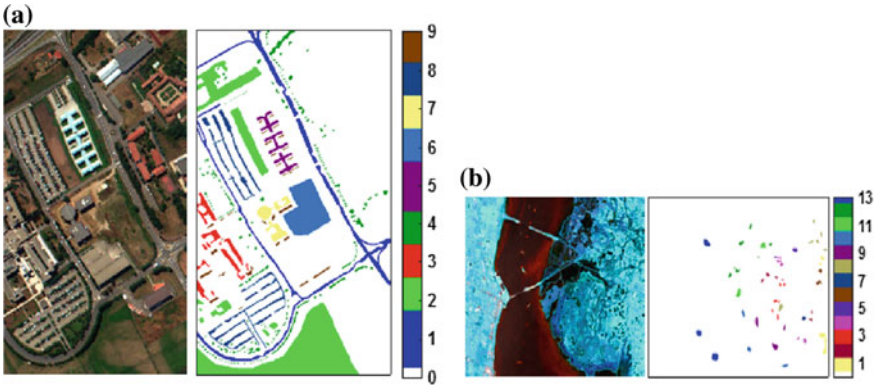


Fig. 2 Three-band color image of **a** the University of Pavia and **b** the KSC data sets along with their ground truth

Table 3 The classes and corresponding number of available labeled samples for KSC data set

Class no.	Class name	#Labeled samples	Class no.	Class name	#Labeled samples
1	Scrub	761	8	Cabbage palm hammock	256
2	Willow swamp	241	9	Cabbage palm/oak hammock	251
3	Slash pine	161	10	Oak/broadleaf hammock	229
4	Hardwood swamp	105	11	Graminoid marsh	431
5	Spartina marsh	520	12	Cattaial marsh	377
6	Salt marsh	419	13	Mud flats	462
7	Water	908		**Total**	**5221**

4 Experimental Results

Experimental setups: The effectiveness of the different filtering strategies are compared by considering attributes of diverse nature in the experiments performed on the aforementioned data sets. The spectral-spatial profiles are constructed on the first five PCs in the reduced space of the HSI data sets obtained after applying PCA. The spectral-spatial profiles constructed for five different attributes are termed as EAP_a (area), EAP_d (diagonal of bounding box), EAP_{ach} (area of convex hull), EAP_{com} (complexity) and EAP_e (entropy). In our experiments, to model the spatial information in maximum possible ways nine filtering threshold values (λs) are considered in the construction of the EAPs. Thus each EAP consists of filtered images from nine *thinning* and nine *thickeing* operations done on five PCs. The λs used for constructing the EAPs are mentioned in Table 4.

For all the considered filtering strategies the constructed EAPs are classified using a *one-Vs-all* multi-class SVM classifier. The kernel chosen for the SVM classifier is radial basis function (RBF) whose parameters $\{\sigma, \mathrm{C}\}$ are derived by using grid search with fivefold cross-validation. SVM is trained with randomly selected 30% labeled samples for each class. The LIBSVM library [3] is used in the implementation of SVM. The experiments are carried out using 64-bit Matlab (R2015a). The classification results after running ten times with ten sets of randomly selected samples for both the data sets are reported using the indices overall accuracy ($\overline{OA}$), average kappa accuracy (kappa), and standard deviation (std) of $\overline{OA}$.

Table 4 The threshold values (λs) used for constructing the EAPs

Attributes	Threshol values
Area	[50 250 450 650 850 1050 1250 1450 1650]
Diagonal of bounding box	[9 16 25 36 49 64 81 100 121]
Area of convex hull	[100 200 300 600 900 1200 1500 2000 2500]
Complexity	[.1 .2 .3 .4 .5 .6 .7 .8 .9]
Entropy	[2 2.5 3 3.5 4 4.5 5 5.5 6]

4.1 Results on University of Pavia Data Set

The first experiment is conducted on the University of Pavia data set. In this experiment the spectral-spatial profiles EAP_a, EAP_d, EAP_{ach}, EAP_{com} and EAP_e are constructed considering the four different filtering strategies *subtractive*, *max*, *min*, and *direct*. The classification results for the profiles are reported in Table 5. One can see from the table that the spectral-spatial profiles (EAP_a, EAP_d, EAP_{ach}, EAP_{com} and EAP_e) provide better classification results than those provided by only the spectral features. This shows the potentiality of considering spectral-spatial information over spectral values alone. Further, one can observe from the table that the accuracies for the strategies *subtractive*, *min* and *direct* are almost similar for all the five considered attributes. Whereas the strategy *max* performs well only with the attributes area and complexity. In case of the rest of the attributes (i.e., diagonal of bounding box, area of convex hull, and entropy) the *subtractive*, *min*, and *direct* have performed considerably better than the strategy *max*.

4.2 Results on KSC Data Set

The next experiments are carried out with the KSC data set and the results are shown in Table 6. For this data set also, one can see that the spectral-spatial profiles (EAP_a, EAP_d, EAP_{ach}, EAP_{com} and EAP_e) provide significantly better classification results than those provided by only the spectral features. As observed in the previous data set for this data set as well, the filtering strategies namely *subtractive*, *min* and *direct* have performed considerably better through all the different attributes whereas, the strategy *max* has performed well only in case of area and complexity. This may be due to the reason that *max* does not merge the required nodes (which were supposed to be merged) simply because their at least one child is not merged. This may lead to poor modeling of the spatial information in case of the attributes diagonal of bounding box, area of convex hull and entropy.

Thus, it can be concluded that the effects of the filtering strategies namely *subtractive*, *min*, and *direct* are similar irrespective of the attribute used during filtering while the effectiveness of the strategy *max* differs according to the attribute considered.

Table 5 Average class-wise accuracy, average overall accuracy ($\overline{OA}$), related standard deviation (std), and average kappa accuracy (kappa) obtained using SVM classifier on ten randomly selected data runs by considering spectral features and EAPs of different attributes obtained using four filtering strategies (University of Pavia data set). The best results for each attribute are highlighted in boldface

Class	Spectral	EAP_a				EAP_d				EAP_{ach}				EAP_{com}				EAP_e			
No.		sub	max	min	direct	sub	max	min	direct	sub	max	min	direct	sub	max	min	direct	sub	max	min	direct
1	95.88	99.56	99.58	99.58	**99.67**	**99.73**	95.90	99.67	99.63	99.61	96.11	**99.68**	**99.68**	99.85	**99.94**	99.81	99.87	**99.54**	95.72	**99.54**	99.50
2	98.75	**99.78**	99.75	99.75	99.74	**99.80**	98.18	**99.80**	99.79	**99.80**	98.31	99.76	99.76	99.97	**99.98**	99.97	99.97	**99.88**	98.37	99.86	99.85
3	85.28	**97.66**	97.33	97.33	97.46	97.15	87.65	**97.23**	97.20	**97.53**	85.39	97.44	97.44	99.47	**99.80**	99.52	99.61	95.81	88.92	**97.65**	95.57
4	97.35	99.16	99.09	99.09	**99.33**	**99.15**	94.53	**99.15**	99.10	**99.14**	92.13	98.99	98.99	99.59	99.61	99.40	**99.63**	**99.20**	93.30	99.13	99.11
5	99.61	99.73	**99.84**	**99.84**	99.80	99.87	99.72	99.84	**99.91**	99.72	99.66	**99.86**	**99.86**	99.88	**99.90**	99.83	**99.90**	99.85	99.54	99.76	**99.90**
6	93.67	98.63	98.60	98.60	**98.71**	98.55	93.61	98.54	**98.63**	98.52	92.03	**98.62**	**98.62**	99.94	99.95	99.88	**99.96**	99.13	95.94	**99.56**	99.24
7	90.11	**97.93**	97.83	97.83	97.67	**99.87**	95.17	99.79	99.83	**98.29**	94.99	98.07	98.07	99.94	**99.96**	99.88	99.92	99.79	96.46	**99.90**	99.89
8	91.65	98.60	**98.68**	**98.68**	98.66	**98.63**	93.30	98.59	98.54	98.56	93.29	**98.63**	**98.63**	99.60	**99.73**	99.51	99.62	97.29	91.97	**98.15**	97.32
9	99.74	99.73	**99.77**	**99.77**	99.76	**99.96**	98.63	**99.96**	99.89	99.77	97.87	**99.84**	**99.84**	99.96	99.94	**99.98**	99.94	99.87	97.93	**99.91**	99.86
$\overline{OA}$	96.12	99.30	99.28	99.28	**99.31**	**99.37**	96.06	99.36	99.35	**99.30**	95.65	**99.30**	**99.30**	99.86	**99.91**	99.83	99.88	99.27	96.26	**99.46**	99.25
kappa	0.948	**0.991**	0.990	0.990	**0.991**	**0.992**	0.948	**0.992**	0.991	**0.991**	0.940	**0.991**	**0.991**	0.998	**0.999**	0.998	0.998	**0.990**	0.951	**0.990**	**0.990**
std	0.041	**0.026**	0.048	0.048	0.029	0.033	0.079	0.056	**0.029**	**0.046**	0.152	0.053	0.053	0.018	**0.015**	0.019	0.022	0.036	0.077	**0.028**	0.042

Table 6 Average class-wise accuracy, average overall accuracy ($\overline{OA}$), related standard deviation (std), and average kappa accuracy (kappa) obtained using SVM classifier on ten randomly selected data runs by considering spectral features and EAPs of different attributes obtained using four filtering strategies (KSC data set). The best results for each attribute are highlighted in boldface

Class	Spectral	EAP_a				EAP_d				EAP_{ach}				EAP_{com}				EAP_e			
No.		sub	max	min	direct	sub	max	min	direct	sub	max	min	direct	sub	max	min	direct	sub	max	min	direct
1	96.89	99.26	99.65	**99.82**	99.38	**99.49**	97.49	99.34	99.40	**99.50**	97.69	99.44	99.25	99.68	99.61	**99.79**	99.76	**99.54**	94.70	99.36	99.29
2	93.25	97.45	**97.49**	97.37	97.24	**97.37**	77.08	96.21	97.00	96.75	79.42	96.83	**97.20**	96.21	**98.15**	96.91	98.07	96.38	81.73	96.17	**96.42**
3	95.16	97.89	97.73	98.16	**98.20**	**98.36**	86.95	98.28	98.16	**98.52**	82.93	98.01	**98.52**	**98.32**	**98.32**	97.70	98.20	97.89	77.58	**98.36**	97.89
4	81.87	94.33	94.40	**94.96**	94.05	94.92	25.32	**96.51**	96.31	**95.32**	12.90	94.25	95.00	94.25	**95.52**	93.93	94.33	94.05	60.12	94.52	**94.84**
5	82.73	**94.66**	92.98	93.23	93.04	93.35	64.10	92.98	**93.48**	**93.60**	56.46	93.35	93.42	91.61	**94.53**	92.86	91.43	**93.04**	84.04	92.86	92.05
6	81.44	**96.51**	93.76	93.36	96.24	**96.86**	26.33	96.24	96.11	94.93	25.72	94.37	**95.46**	97.16	97.07	**98.38**	97.95	95.28	44.59	**95.94**	95.63
7	94.95	97.05	**98.29**	98.10	96.10	**98.29**	56.86	97.14	97.81	95.43	18.00	**96.95**	96.76	98.48	**98.76**	96.29	96.86	97.62	67.62	**99.24**	98.67
8	96.91	**98.56**	97.89	98.35	98.14	97.68	74.85	97.75	**97.94**	97.01	72.07	**97.08**	96.91	**99.44**	99.28	99.35	97.96	**99.19**	92.09	98.31	98.75
9	98.71	99.27	99.06	99.13	**99.35**	**99.58**	90.06	99.38	99.35	99.35	84.96	99.27	**99.37**	**99.52**	99.33	99.25	99.25	**99.50**	96.13	99.33	99.27
10	97.72	97.18	96.46	96.41	**97.40**	**97.77**	91.06	97.35	97.55	**96.83**	89.80	96.44	96.76	97.15	97.23	**97.57**	97.38	94.58	85.84	94.63	**95.52**
11	98.78	99.40	**99.55**	99.45	99.45	99.55	96.75	99.14	**99.64**	99.38	98.35	**99.59**	99.38	99.14	99.31	**99.59**	99.57	**99.55**	98.64	99.50	99.40
12	95.41	97.83	**98.45**	**98.45**	98.01	98.19	85.39	97.81	**98.21**	97.93	85.47	**98.17**	97.77	98.11	97.93	**98.19**	98.03	**97.50**	86.44	97.10	96.48
13	99.62	99.48	99.74	**99.91**	99.88	99.75	99.20	99.60	**99.84**	99.64	99.51	**99.76**	99.71	99.72	99.81	99.87	**99.92**	**99.95**	99.56	99.80	99.70
$\overline{OA}$	95.50	98.26	98.15	**98.27**	**98.27**	**98.44**	83.52	98.23	98.43	**98.09**	81.17	98.04	**98.09**	98.39	**98.60**	98.51	98.41	**98.08**	88.13	97.99	97.95
kappa	0.950	**0.981**	0.979	**0.981**	**0.981**	**0.983**	0.816	0.980	**0.983**	**0.979**	0.789	0.978	**0.979**	0.982	**0.984**	0.983	0.982	**0.979**	0.868	0.978	0.977
std	0.349	0.260	0.209	**0.144**	0.180	**0.163**	0.668	0.200	0.243	**0.198**	0.352	0.298	0.261	0.191	**0.120**	0.171	0.138	**0.202**	0.638	0.220	0.234

5 Conclusion

High-quality hyperspectral imagery is nowadays available for various applications. Accurate classification of HSI is tremendously improved with the help of morphological attribute profiles. In this paper, the various strategies adopted during a filtering operation in the construction of APs for accurate classification of HSI are considered for comparison. The filtering strategies considered in this paper are *subtractive*, *max*, *min*, and *direct*. The classification accuracies obtained for two real HSI data sets reveal that the strategies *subtractive*, *min*, and *direct* work robustly for both the data sets and in case of all the different attributes used while the strategy *max* performs well only in case of area and complexity. Therefore, it is suggested to use the strategies *subtractive*, *min*, and *direct* in case of any existing or new attribute for HSI classification.

References

1. J.A. Benediktsson, J.A. Palmason, J.R. Sveinsson, Classification of hyperspectral data from urban areas based on extended morphological profiles. IEEE Trans. Geosci. Remote. Sens. **43**(3), 480–491 (2005)
2. K. Bhardwaj, S. Patra, An unsupervised technique for optimal feature selection in attribute profiles for spectral-spatial classification of hyperspectral images. ISPRS J. Photogramm. Remote. Sens. **138**, 139–150 (2018)
3. C.C. Chang, C.J. Lin, LIBSVM: a library for support vector machines. ACM Trans. Intell. Syst. Technol. (TIST) **2**(3), 27 (2011)
4. M. Dalla Mura, J.A. Benediktsson, B. Waske, L. Bruzzone, Extended profiles with morphological attribute filters for the analysis of hyperspectral data. Int. J. Remote. Sens. **31**(22), 5975–5991 (2010)
5. M. Dalla Mura, J.A. Benediktsson, B. Waske, L. Bruzzone, Morphological attribute profiles for the analysis of very high resolution images. IEEE Trans. Geosci. Remote. Sens. **48**(10), 3747–3762 (2010)
6. A. Das, K. Bhardwaj, S. Patra, Morphological complexity profile for the analysis of hyperspectral images, in *2018 4th International Conference on Recent Advances in Information Technology (RAIT)*, (IEEE, 2018), pp. 1–6
7. P. Ghamisi, M. Dalla Mura, J.A. Benediktsson, A survey on spectral-spatial classification techniques based on attribute profiles. IEEE Trans. Geosci. Remote. Sens. **53**(5), 2335–2353 (2015)
8. S. Patra, K. Bhardwaj, L. Bruzzone, A spectral-spatial multicriteria active learning technique for hyperspectral image classification. IEEE J. Sel. Top. Appl. Earth Obs. Remote. Sens. **10**(12), 5213–5227 (2017)
9. S. Patra, P. Modi, L. Bruzzone, Hyperspectral band selection based on rough set. IEEE Trans. Geosci. Remote. Sens. **53**(10), 5495–5503 (2015)
10. P. Salembier, A. Oliveras, L. Garrido, Antiextensive connected operators for image and sequence processing. IEEE Trans. Image Process. **7**(4), 555–570 (1998)
11. E.R. Urbach, J.B. Roerdink, M.H. Wilkinson, Connected shape-size pattern spectra for rotation and scale-invariant classification of gray-scale images. IEEE Trans. Pattern Anal. Mach. Intell. **29**(2), 272–285 (2007)

Standard Statistical Feature Analysis of Image Features for Facial Images Using Principal Component Analysis and Its Comparative Study with Independent Component Analysis



Bulbul Agrawal, Shradha Dubey and Manish Dixit

1 Introduction

The two most important techniques have been explored here as these approaches are very useful in many fields. One of the most interesting fields, scilicet as face recognition has become a hot topic in computer vision algorithms. Similarly, dimensionality reduction, compression, signal separation, image filtering, and many other statistical approaches are becoming popular and of great use in many fields day by day. Therefore, these approaches have been discussed here along with their applications, problems, and mathematical approaches. The process of identifying the lesser number of uncorrelated variables, i.e., principal components from a huge set of data is known as Principal Component Analysis (PCA) [1]. PCA just removes correlation and not higher order dependency as it works only in second-order statistics. The technique is widely used to emphasize variation and capture strong patterns in a dataset [2]. PCA is a technique that is used in predictive models and exploratory data analysis was invented by Karl Pearson in 1901, PCA is considered as a useful statistical method and used in fields such as image compression, face recognition, neuroscience, and computer graphics. In the method of ICA, not only statistical characteristics in second-order or higher order are considered, but also basis vectors decomposed from face images obtained by ICA are more localized in distribution space than those by PCA [3]. In the computer vision, PCA is a popular method and

B. Agrawal · S. Dubey · M. Dixit (✉)
Department of Computer Science & Engineering and Information Technology, Madhav Institute of Technology and Science, Gwalior, India
e-mail: dixitmits@gmail.com

B. Agrawal
e-mail: bulbulag04@gmail.com

S. Dubey
e-mail: shradha290594@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_104

applied mainly in face recognition whereas ICA was initially originated for distinct and mixed audio signals into independent sources. In recent times, ICA has been applied to recognize the faces and expression and image analysis [4].

The literature on the subject is conflicting. Some assert that PCA outperforms ICA others claim that ICA outperforms PCA and some claim that there is no difference in their performance statistically [5]. Thus, this paper compares the PCA technique to a newer technique ICA.

The paper is organized as the second section shows the literature review, in the third and fourth section, there is a discussion about both the techniques, i.e., PCA and ICA, respectively. Further, their applications have been discussed along with their comparative study and some parameters are also calculated.

2 Literature Review

Bing Luo et al. This paper discusses the PCA and ICA and compares them in the face recognition application and gives suitable differences between them. They used PCA derived from eigenfaces and ICA derived from the linear representation of non-Gaussian data.

Diego A. Socolinsky et al. This paper presents an exhaustive performance analysis of multiple appearance-based tactics, on visible and thermal infrared imagery. Different algorithms have been compared to achieve performance levels.

Bruce A. Draper et al. This paper compares PCA and ICA. Depending upon the task statement, the ICA algorithm, PCA subspace distance metric relative performance has been measured. In the context of a baseline face recognition system, a comparison motivated by contradictory.

Zaid Abdi Alkareem Alyasseri. In this paper, two algorithms are discussed to recognize a face, namely, Eigenface and ICA. This paper shows how the error rate between the original face and the recognized face has been improved and the results are shown by the various graphs.

Gupta, Varun, et al., PCA is a certain case of factor analysis. Here is an example that shows the step-by-step process of finding the value of the principal component. PCA has a little fundamental assumption like principal components, variances, and linearity are orthogonal. Data classification and dimensionality reduction techniques used the concept of principal component analysis. This paper shows the importance of biomedical signal processing. PCA is applicable to both batch process and continuous process.

Ramsay, J.O. et al., while some type of smoothness is absorbed into the PCs then functional data of PCA is more appealing; smoothed FPCA as shown by the author of this paper. This paper shows the minimum two methods that carry out smoothed FPCA. Primary is, before applying FPCA, to make the functional data smooth and secondary is to define smoothed PCs by adding a roughness penalty term to maximize the sample covariance.

Chawla M.P.S., This paper shows the ICA and PCA procedure to remove the noise and artifacts. ECG signals have various types of noise and artifacts that may conceal critical information. In this paper, a transformation technique is used to find good accuracy, even in the existence of low- or high-frequency noise. This paper describes the importance of ICA for separating the ECG complex signals and PCA for dimensional reduction.

3 Principal Component Analysis (PCA)

PCA is the analytical technique commonly used in image processing to the data dimension reduction or data décor-relation [6, 7].

3.1 Dimensionality Problem

The main objective of dimensionality reduction techniques is to transform the data or features from a higher dimensional space to a lower dimensional space [8].

Suppose that there are N no. of features of the object and combination of these features is called pattern. So, a question arises that how many features are to be selected and which features are to be selected? Taking too many features will not help as the training sample size required to train the classifier will also increase. Greater the size of a feature vector, the requirement of training sample data will also increase. If this set of training is small relative to the no. of features, then increased number of features will degrade the classify performance. This effect is called the peaking phenomenon. So, it is also desirable to keep the feature's size as small as possible without compromising the performance of the classifier [9].

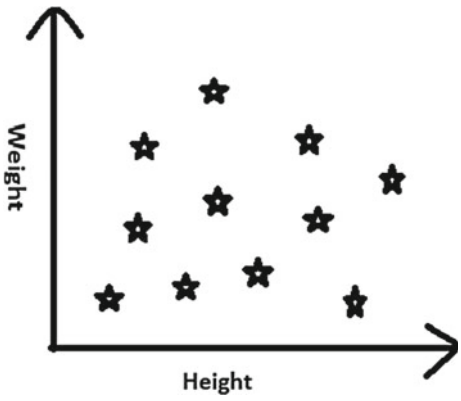


Fig. 1 shows the number of students with two features, i.e., weight and height

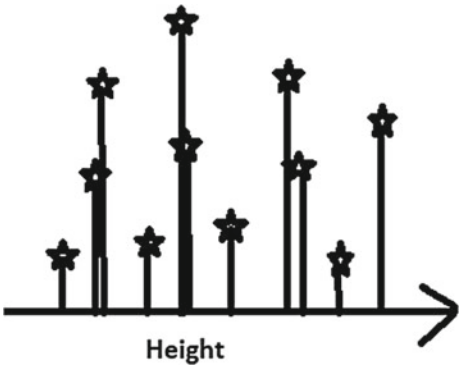


Fig. 2 Projection of the points on the single dimension

In the above Fig. 1, we take an example of students in 2D space. Now, we want to reduce this feature space into 1D. So, one of the solutions to this problem is to select one of the features from height and weight. If we project all data points on the single dimension height, then Fig. 2 shows all the data points on one dimension or can say that intuitive approach is followed to bring down the dimensionality of this 2D feature space by just omitting one of the data from the feature vector. From Fig. 2, we can see that these points are very close to each other and it is difficult to classify the students on the basis of height only. That's why we've chosen two features to separate the boys from girls.

Here, only one feature is not enough to represent all students. So, the solution to this problem is Principal Component Analysis (PCA).

3.2 PCA

The objective of the PCA technique is to obtain a lower dimensional space or PCA space, i.e., used to transform the data $X = \{x_1, x_2 \ldots x_N\}$ from a higher dimensional space to a lower dimensional space. Here, N shows the total no. of observations or samples [10]. The direction of the PCA space serves as the direction of the maximum variance of the given data points [11]. PCA space consists of K no. of principal components. Each PC has a different vitality per portion of the variance in its direction. The PCs are uncorrelated and orthonormal. The 1PC (First principal component) of the PCA space serves as the direction of maximum variance of the given data points, the 2PC (Second principal component) has the second largest variance of the samples, and so on [12].

In the Fig. 3, R^M is the higher dimensional space and R^K is the lower dimensional space, PC1 and PC2 are the primary principal component and secondary principal component, respectively.

The concept of PCA is very simple; we can understand it by an example. So, from Fig. 2, we can easily understand that there is a need to find a new dimension that can

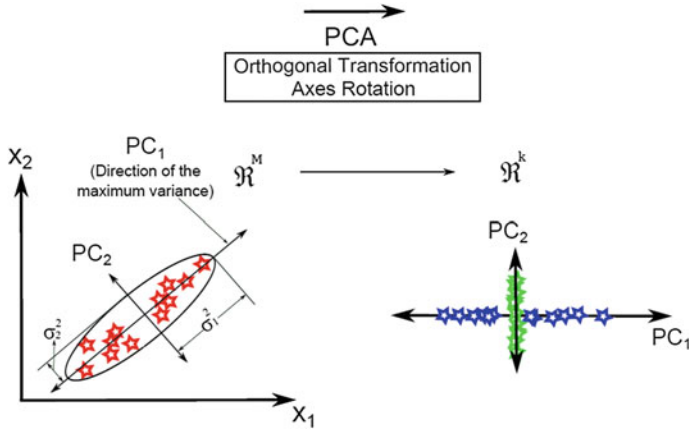


Fig. 3 Illustration of 2D projection of data points (X_1, X_2)

represent all the data points individually. Here a question arises that which line we've to select. So, it is very intuitive; if we select line Z_1, then points are projected on them (Fig. 4 represents this example). These points are very adequately separated from each other. Now, one more question arises that how to find out this line or technically this line can be called as the principal component. So, the method of finding the PCs of the PCA space is given below [13].

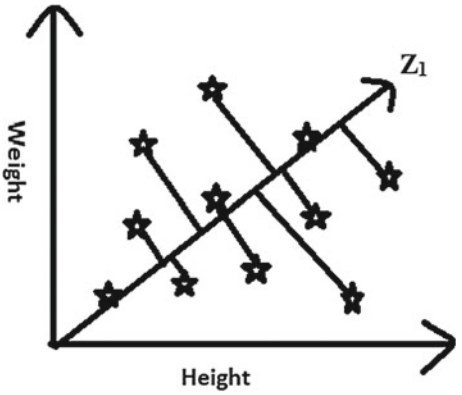


Fig. 4 Introducing a new line and projection of all the data points on them

3.3 *Process of Finding the Principal Component (PC)*

1. **Acquire some data**: To find the PC, first acquire some data points from the dataset and plot these data points on the 2D space. Block A in Fig. 5 shows this step in a very well manner.

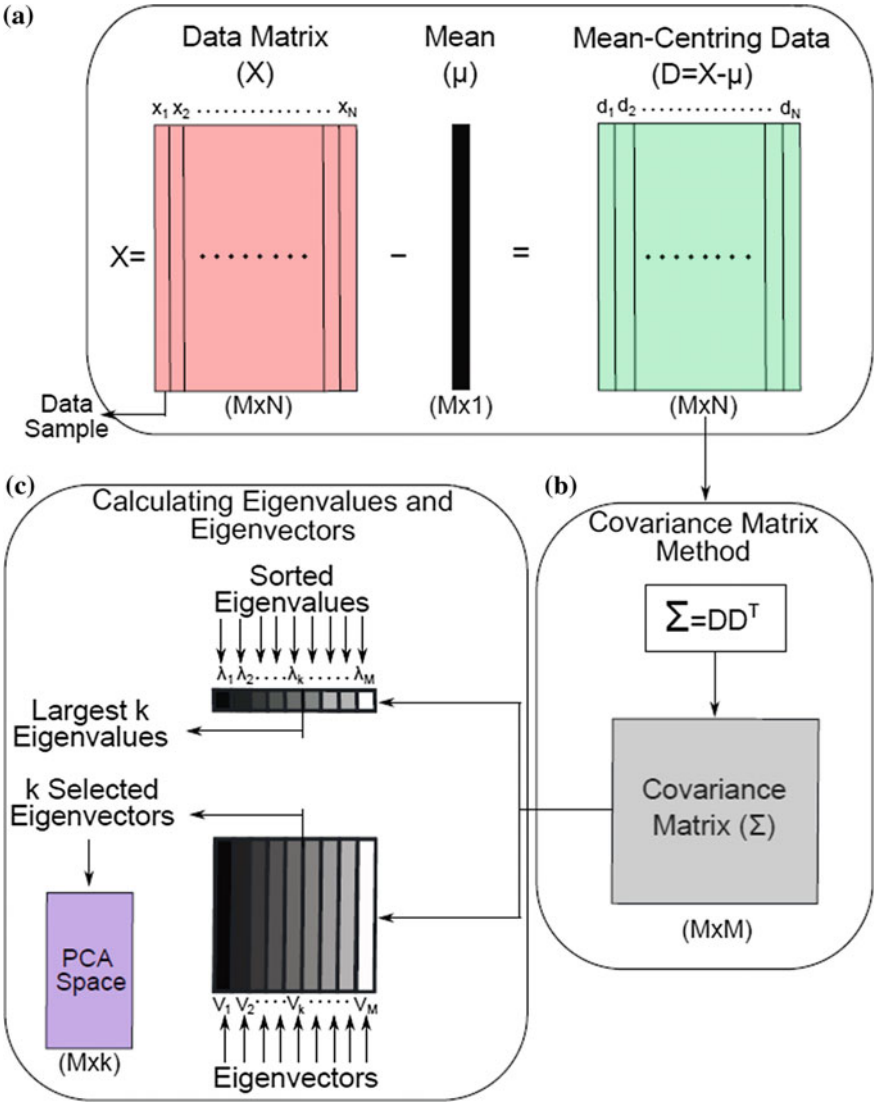


Fig. 5 Process of computing the covariance matrix in PCA

2. **Calculate the covariance matrix**: It is the main step to find the PCs of the PCA space. We know that if the no. of variables is more than one, then the covariance matrix is used. Then after acquiring the data points (data matrix) $\{X = [x_1, x_2 \ldots x_N]\}$ where N serves as the total number of samples and x_i represents the ith sample, so first, we need to calculate the mean of all sample points [6]. Value of mean (μ) is given by [14]

$$\mu = \frac{1}{N} \sum_{i=1}^{N} x_I \tag{1}$$

Now, subtract the mean μ from the sample points, given by

$$D = \{d_1, d_2 \ldots d_N\} = \sum_{i=1}^{N} x_i - \mu \tag{2}$$

Finally, a matrix of covariance is calculated as [15];

$$\sum = \frac{1}{N-1} D \times D^T \tag{3}$$

This step explains in block B of Fig. 5.

3. **Compute the Eigenvector and Eigenvalue**: The solution of the covariance matrix is given by eigenvalue λ and eigenvector V.

The formulation is $V \sum = \lambda V$, here λ and V are eigenvalues and eigenvectors, respectively. Eigenvectors are nonzero vector whereas eigenvalues are scalar values. So, each eigenvector shows the one principal component. So here eigenvector acts for the directions of the PCA and the corresponding eigenvalues act for the length, scaling factor, robustness, or the magnitude of the eigenvectors [6]. The first principal component represents the eigenvector with the highest eigenvalues and it has a maximum variance [16]. So, we've to sort the eigenvectors per eigenvalues and select those eigenvectors which have maximum eigenvalues. Eigenvalue is also known as characteristic roots. In the block, C of Fig. 5 represents the process of this step.

When PCA is applied for dimensionality reduction, normally PCs are a dump with near-zero and zero eigenvalues [17].

3.4 PCA with Singular Value Decomposition (SVD) Method

Now using this method, we calculate the PC with the singular value decomposition method (SVD). SVD is a method of matrix factorization. SVD is one of the most crucial linear algebra principles and utilized in many numerical applications such as

PCA [18]. SVD and PCA both are eigenvalues methods which are used to minimize the high-dimensional dataset into a small number of dimensional while keeping important information. Diagonalizing the data matrix into three matrices is the main goal of this method.

$$X \in R^{m \times n} = X = LSR^T = \begin{bmatrix} l_1 \ldots l_m \end{bmatrix} \begin{bmatrix} s_1 & 0 & 0 & 0 \\ 0 & s_2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & s_n \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} -r_1^T - \\ -r_2^T - \\ \vdots \\ -r_n^T - \end{bmatrix}$$

Here, $L(m \times m)$ are left singular vectors, $R(n \times n)$ shows the right singular vectors, and $S(m \times n)$ shows the singular values and it is a diagonal matrix that is organized from high to low. It means that in the upper left index of S is the highest singular value, hence $s_1 \geq s_2 \geq \cdots \geq s_n \geq 0$. The left singular matrices L and the right singular matrices R are the orthonormal bases. To determine the SVD, first S and R^T are computed by diagonalizing $X^{\mathrm{T}}X$ as follows [19]:

$$X^T X = (LSRT)^T \left(LSR^T\right) \tag{4}$$

$$X^T X = RS^T L^T LSR^T \tag{5}$$

$$X^T X = RS^2 R^T \left(\text{Here } L^T L = \mathrm{I}\right) \tag{6}$$

And left singular vector L is computed as $L = XRS^{-1}$. The columns of the R, i.e., right singular vectors show the eigenvectors of $X^T X$ or principal components of the PCA space and s_i^2; $\forall\ i = 1, 2 \ldots n$ shows their corresponding eigenvalues [18] (As shown in Block B and C of Fig. 6)

Since the eigenvalues of PC and this no. of PCs are equal to n, so the dimension of the actual data matrix must be reversed to be adaptable with the SVD method. Or we can say that in other words, before computing the SVD method, first the mean-centric matrix is transposed. We can see that in Fig. 6, block (A) represents each sample by one row. In the following steps, we can understand the Fig. 6

1. In the block A, data matrix $X = [x_1, x_{2 \ldots} x_N]$ is a given data matrix, here N shows the total no. of samples and x_i represents the ith sample of $(M \times 1)$.
2. After that compute the mean of all samples and subtract the mean also from all samples (as shown in block A)
3. Compute a matrix

$$Z = \frac{1}{\sqrt{\mathrm{N} - 1}} D^T,\ Z(N \times M). \tag{7}$$

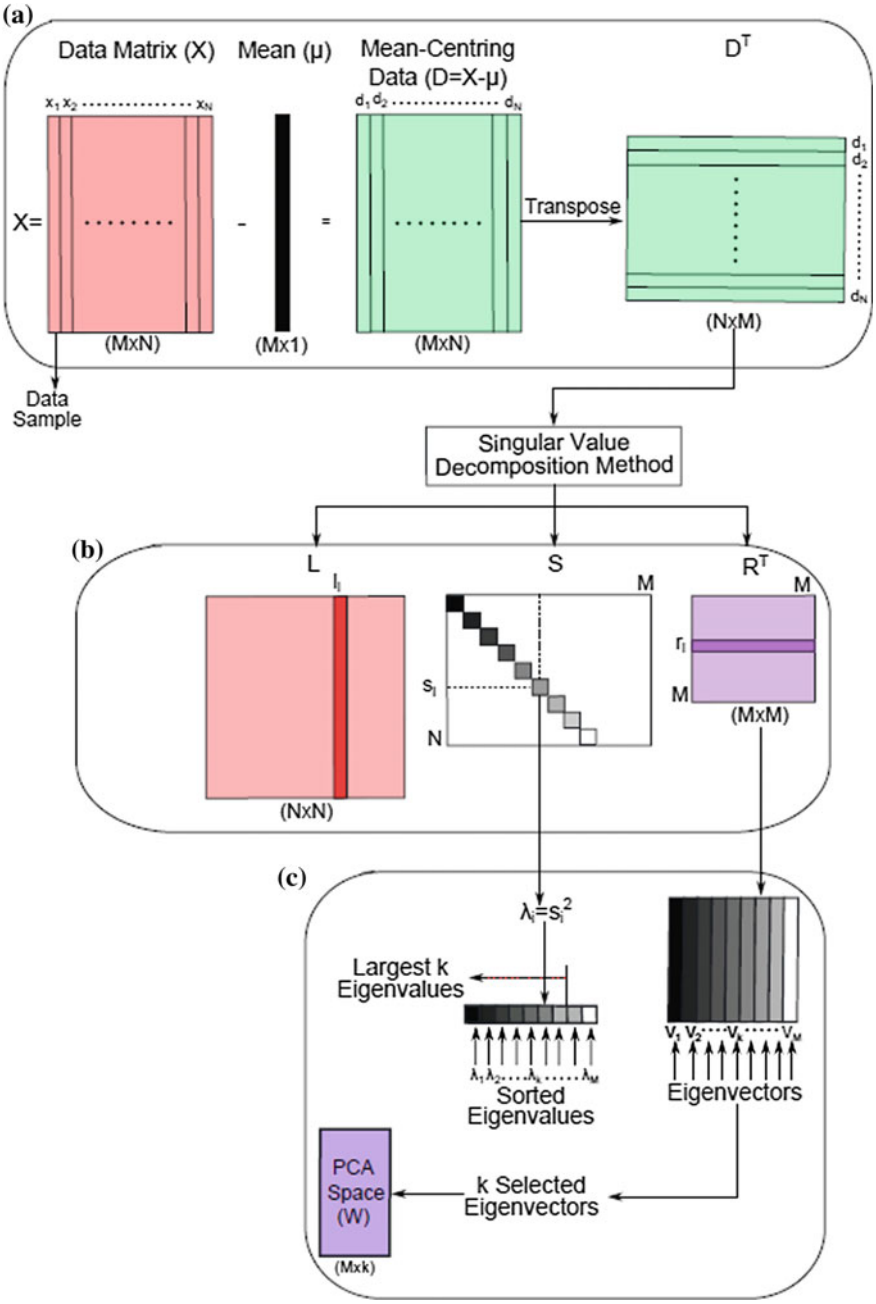


Fig. 6 Computation of the PCA space with SVD method

Table 1 Calculate the value of entropy for PCs and EVs

Description	Entropy
Initially data matrix	0.3373
First Principal Component (PC1)	0
Project the data on the PCA (Yv2)	1.4056
Second Principal Component (PC2)	1
Project the data on the PCA (Yv1)	1.5488
Data is reconstructed from the projected data on PC1 (Ev1)	1.7744
Data is reconstructed from the projected data on PC2 (Ev2)	2.5000

4. Now compute the SVD for Z matrix (as shown in block B)
5. $\lambda = \text{diag}(S^2)$ means that the diagonal elements of S represent the square root of the sorted eigenvalues and PCs are represented by the columns of R (as shown in block C).
6. To construct the PCA space, chose the eigenvectors that have the highest eigenvalues $W = \{R_1, R_2 \ldots R_k\}$.

Now, on the lower dimensional space of PCA (W), all samples are projected on them as follows:

$$Y = W^T D. \tag{8}$$

Here, the one thing which is important about the SVD and the PCA is that both are similar thing and it is better to use the SVD of the centered data matrix because SVD algorithms are numerically more stable and faster than PCA.

Here, we show the example of PCA from the data matrix, in which 1PC and 2PC reconstruct the data from the projected data on PC1 and PC2, are calculated and these entropy values are shown in Table 1.

$$X = \begin{bmatrix} 1\ 1\ 2\ 0\ 5\ 4\ 5\ 3 \\ 3\ 2\ 3\ 3\ 4\ 5\ 5\ 4 \end{bmatrix};$$

4 Independent Component Analysis (ICA)

The method for uncovering the key components or factors from the statistical data containing two or more variables or can be multidimensional statistical data is the basic definition of ICA (Independent Component Analysis). One of the properties of ICA is to look for the components which are statistically independent and non-Gaussian that differentiate it from other methods [11, 20].

The method based on the statistically independent assumption, i.e., the value of one does not go to provide any information about the other is called ICA method. The complex or huge dataset is fragmented into independent parts by using this statistical technique. Non-Gaussianity is a measure of independency [11].

Based on the central limit theorem, the Gaussianity of $x(t)$ must be larger than $s(t)$.

Now, we put b_i^T as mixing vector, $s_i(t) = b_i^T x(t)$. We want to maximize the Non-Gaussian of $\left(b_i^T x(t)\right)$. Then such b is a part of a solution.

For example, there are following two vector $b^{'}$ and b. We can say that b is better than $b^{'}$ (Fig. 7, 8, 9, 10).

In other words, we can say that the data collected from the real world is not theoriginal data; the original data comes up as a combination of mixed strings and random signals [21].

In the given example, the picture of the monkey is mixed up with the strings, in our example, a monkey has come along and mixed up our strings. So, we look for a method to disentangle them.

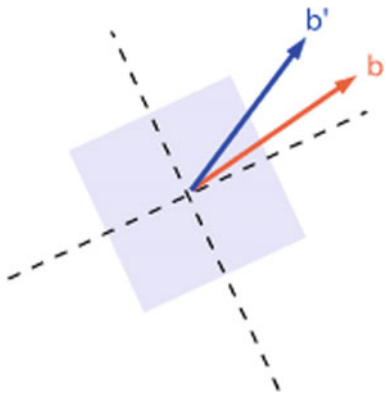


Fig. 7 Illustration of Non-Gaussianity



ORIGINAL STRINGS (ORIGINAL DATA)

UN-MIXING MATRIX, W (A^{-1})

MIXED STRINGS (OBSERVED) DATA)

Fig. 8 An e.g. of how original data is calculated from mixing and unmixing matrices

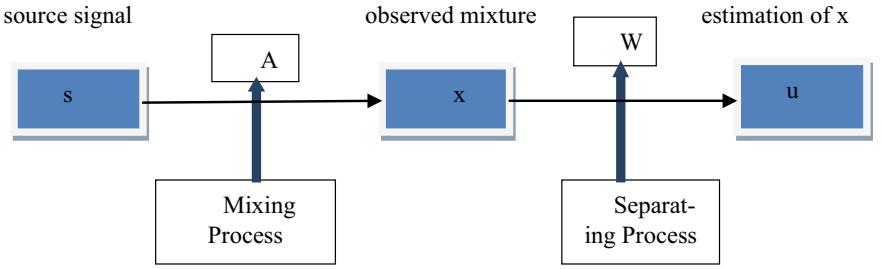


Fig. 9 Model of blind source separation

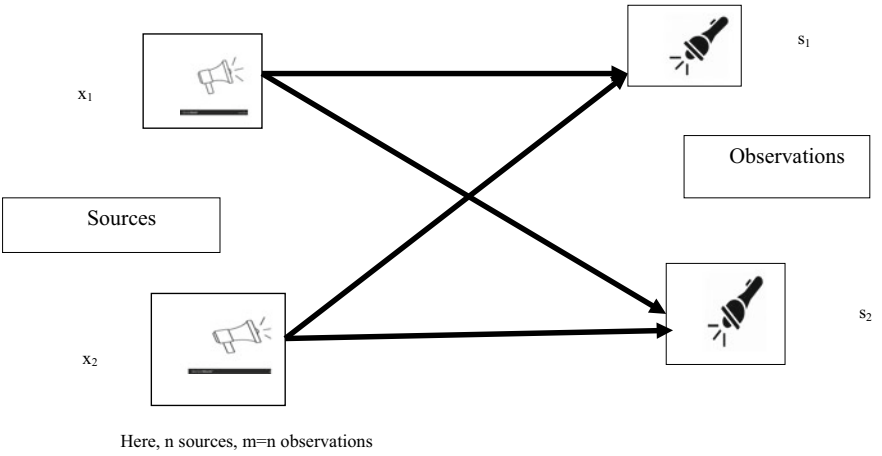


Fig. 10 Illustration of the cocktail party problem



Here,

X ➔ Observed Data
A ➔ Mixing Matrix
S ➔ Original Data

We can unmix the strings to get the original data by knowing something exceptional about each string may be a feature like color but when we work with the very large dataset and does not have any idea about the special feature, then it becomes a tedious task to unmix the strings and then the ICA comes in and assumes that the mixed data consist of independent signals [22].

Mixing matrix is the sequence of variations to the unmixed data (s); these variations are applied by the monkey in the given example. The observed data (X) is generated by multiplying vector of numbers with s [23, 24].

The equation for s must be solved to resolve this problem and reclaim original strings from mixed ones. As X is already known to us, we just calculate A^{-1} (known as unmixing matrix "W")

$$S = A - 1 * X \tag{9}$$

What is basically done is that we model the CDF of each signal's probability as the sigmoid function because it increases from 0 to 1, the derivative of the sigmoid is the density function, and then we would iteratively maximize that function until convergence to find the weights, this inverse matrix.

In a face recognition problem, mostly the information contains is of higher order relationships between the pixels of the image which is difficult to generalize with the help of PCA as it generally works with second-order relationships. So, we use ICA for such problems which can be able to reduce both second-order as well as higher order relationship and works well with such type of huge dataset [25].

ICA is closely interrelated to the blind source separation problem, where the observed signal has degenerated into a linear combination of independent signals which is not known.

The method to approximate original signals from observed signals which consist of noise and mixed signals is called Blind Source Separation (BSS) method [26].

The difficulty of BSS is standardizing as follows:

The matrix $X \in R^{m \times d}$ denotes original signals, where m is the number of original signals, and d is the dimension of one signal. We consider that the observed signals $Y \in R^{n \times d}$ is given by linear mixing system as

$$Y = AX + E \tag{10}$$

where $A \in R^{n \times m}$ is the unknown mixing matrix and $E \in R^{n \times d}$ denotes a noise. Basically, $n \geq m$.

$$Y = AX + E \tag{11}$$

To estimate A and X, the degree of freedom of BSS model is very high as there are many combinations (A, X) which satisfy $Y = AX + E$. Therefore, some constraint is required to figure out the BSS problem such as

PCA (Principal Component Analysis): Orthogonal constraint

SCA (Sparse Component Analysis): Sparsity constraint
NMF (Nonnegative Matrix Factorization): Nonnegativity constraint
ICA (Independent Component Analysis): Independency constraint and many other techniques that rely on the constraints are used to solve the BSS problem.

The cocktail party problem is also called BSS problem. In the basic cocktail party problem to make the components statically independent or as independent as possible, a linear representation of non-Gaussian data is found and various applications such as feature extraction and signal separation.

The basic "Cocktail Party" problem promotes a method in which the main aim is to make the components statistically independent or as independent possible and for this, it searches a linear representation of non-Gaussian [23]. This type of representation obtains the vital structure of the data in various applications as feature extraction and signal separation.

Here, n sources, $m = n$ observations

$$x_1(t) = a_{11}s_1(t) + a_{12}s_2(t) + a_{13}s_3(t) \tag{12}$$

$$x_2(t) = a_{21}s_1(t) + a_{22}s_2(t) + a_{23}s_3(t) \tag{13}$$

$$x_3(t) = a_{31}s_1(t) + a_{32}s_2(t) + a_{33}s_3(t) \tag{14}$$

x ➔ observed signal,
s ➔ original signal.

Let us assume $\{s_1, s_2, s_3\}$ are statistically independent of each other. ICA uses $x(t)$ to evaluate the independent components $s(t)$ [10].

$$x(t) = As(t) \tag{15}$$

$\{s_i\}$ are statistically independent of each other

$$p(s_1, s_2, \ldots s_n) = p(s_1)\, p(s_2) \cdots p(s_n) \tag{16}$$

$\{s_i\}$ follows the Non-Gaussian distribution. If $\{s_i\}$ follows the Gaussian distribution, then ICA is impossible therefore it follows the non-Gaussian.

A is a regular matrix. Therefore, the model can be rewritten as

$$s(t) = Bx(t) \tag{17}$$

where $B = A^{-1}$. It is compulsory to evaluate B so that $\{s_i\}$ is independent.

The most two important applications of ICA in face recognition are either to serve image pixels as random variables and images as observations or to treat images as random variables and pixels as observations. We call these two as ICA architecture

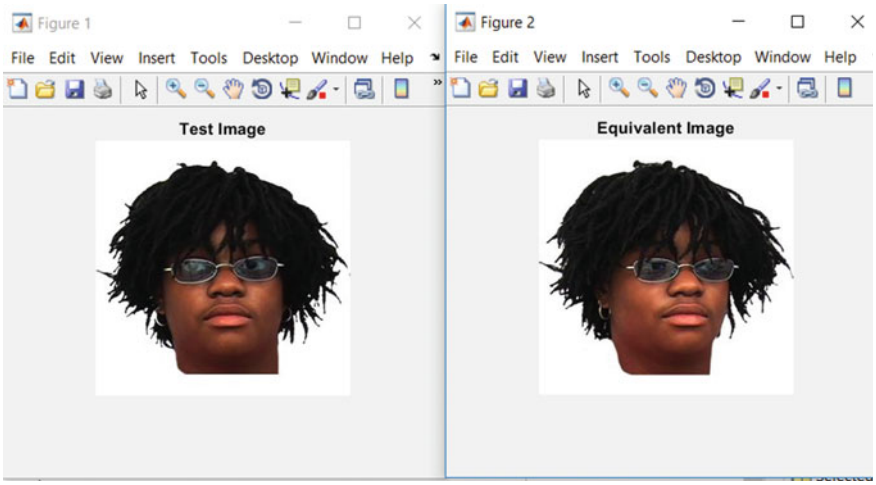


Fig. 11 Analysis of face recognition system for Africans people

I and II, respectively [5]. Here, architecture I is for holistic tasks and architecture II is for localized tasks.

ICA Ambiguities:
The two most common ambiguities that arise in ICA are the method to calculate the variances of the independent components can't be determined. Another problem is that the order of the independent components also can't be identifiable.

Here, we have implemented the face recognition system by training the images [27] on different types of datasets like Japanese, African, and Asian with the help of PCA for the different types of datasets and calculated their parameters. All the parameters of the different datasets have been calculated in the MATLAB 2018a. Figures 11, 12, 13 show the output of the face recognition system in which input is the original image and output is its equivalent image which is obtained after testing the images for different datasets based on Euclidean distance. The different datasets along with their parameters are as follows:

Parameters for America's Africans Dataset:
(Table 2).
Parameters for Japanese Dataset:
(Figure 12, Table 3).
Parameters for Asians Dataset:
(Table 4).

From Fig. 14, we can see the variance among the different types of datasets in the context of Entropy and Variance.

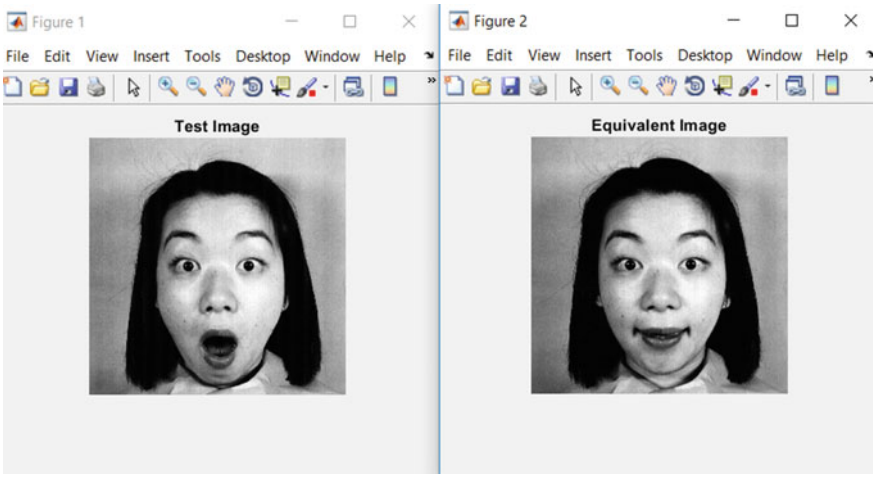


Fig. 12 Analysis of face recognition system for Japanese dataset

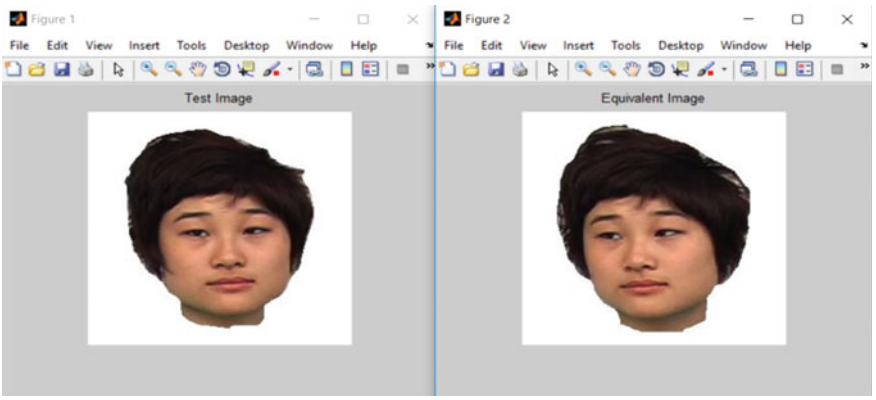


Fig. 13 Analysis of face recognition system for Asian dataset

Table 2 Calculation of different parameters for African's dataset

S.No.	Parameters	Test image	Equivalent image
1	Entropy	4.8991	4.9450
2	Standard deviation	110.4553	109.0322
3	Mean	139.2155	142.5504
4	Median	110	129
5	Variance	10.5097	10.4418
6	Mode	255	255
7	Correlation	0.8993	
8	SSIM	0.6804	

Table 3 Calculation of different parameters for Japanese dataset

S.No.	Parameters	Test image	Equivalent image
1	Entropy	7.3298	7.2780
2	Standard deviation	78.4347	76.4705
3	Mean	127.9513	121.8242
4	Median	152	142
5	Variance	8.8563	8.7447
6	Mode	2	2
7	Correlation	0.9514	
8	SSIM	0.5976	

Table 4 Calculation of different parameters for Asians dataset

S.No.	Parameters	Test Image	Equivalent Image
1	Entropy	4.6686	4.8430
2	Standard deviation	97.0983	98.4387
3	Mean	174.5845	169.2308
4	Median	255	252
5	Variance	9.8538	9.9216
6	Mode	255	255
7	Correlation	0.9075	
8	SSIM	0.7840	

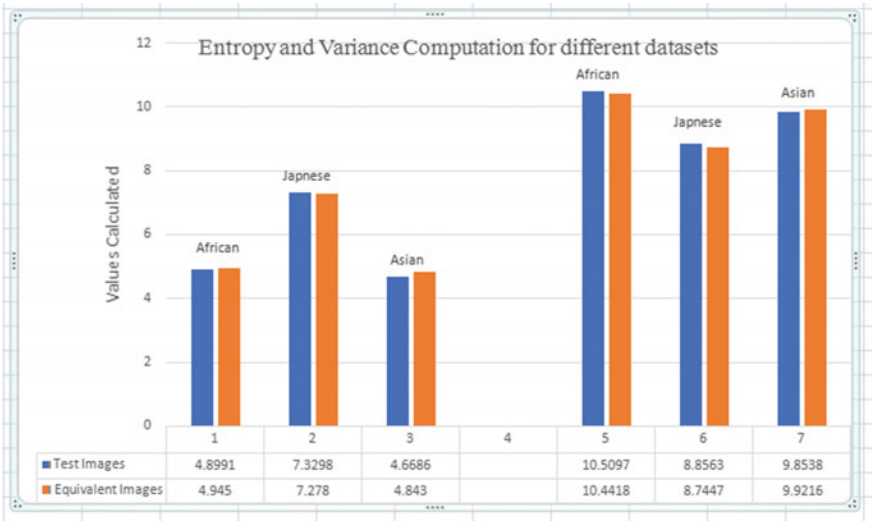


Fig. 14 Computation of Entropy and Variance for different types of datasets

Table 5 Difference between the PCA and ICA [29, 30]

S.No.	Principal Component Analysis (PCA)	Independent Component Analysis (ICA)
1	In the image database, PCA relies only on pairwise relationships between pixels	Detecting the component from a multivariate data is done by ICA
2	PCA takes details of statistical changes from second-order statistics	It can have details up to higher order statistics
3	With the help of PCA, higher order relations cannot be removed but it is useful for removing correlations	ICA, on the other hand, removes both correlations as well as higher order dependence
4	It works with the Gaussian model	ICA works with the non-Gaussian model
5	Based on their eigenvalues, some of the components in PCA are given more importance with respect to others	In ICA, all of the components are of equal importance
6	It prefers orthogonal vectors	Non-orthogonal vectors are used
7	PCA performance is based on the task statement, the subspace distance metric, and the number of subspace dimensions retained	The performance of ICA depends on the task, the algorithm used to approximate ICA, and the number of subspace dimensions retained
8	For compression purpose, PCA uses low-rank matrix factorization	It uses full rank matrix factorization to eliminate reliability on rows

5 Juxtaposition Between PCA and ICA

Data reduction, statistical data analysis, and features extraction are classical techniques of PCA whereas array processing and data analysis are techniques of ICA. The separation of sound signals is the great potential application of ICA. Table 5 discusses the comparison between PCA and ICA [28].

6 Applications of PCA and ICA

There are various applications available of PCA and ICA. So here, we'll discuss the applications of ICA and PCA in the context of several fields.

6.1 Application of PCA

Some of the areas are discussed here where PCA can be applied. PCA acts very good for dimensionality reduction that can be used to minimize the size of a dataset, i.e., from large dataset to small dataset [2, 19]. Another application is image compression [31], PCA also plays a very crucial role to compress the size of the image by detaching the redundant data. Using SVD can also compress the size of an image [18]. If image

compression is done using PCA, then it is also well known as the Karhunen–Leove (KL) transform and Hotelling. One more example of PCA is in feature extraction [5, 29, 32], there is a need to extract some useful data or process a large amount of data from big data in the today's era. So, using this feature extraction application with PCA finds a useful pattern for further analysis. PCA is also used to analyze the two-dimensional data [33], medical imaging [34], gene expression analysis [18], data classification [2], trend analysis, factor analysis, and noise reduction [35].

6.2 Applications of ICA

In this section, we will discuss some of the areas where ICA can be applied. The most standard application of ICA is the cocktail party problem which has already been discussed earlier in this paper. Some of the other applications are as follows:

Artifacts separation in MEG data:
Magnetoencephalography (MEG) is a noninvasive procedure used to measure cortical neurons with great temporal resolution and average spatial resolution. There may be a problem of taking out the useful features of neuromagnetic signals in artifacts presence while using MEG as a research tool. The amplitude of noise becomes much greater than that of the brain signals Therefore, to separate brain activity from artifacts, a new method has been proposed by authors known as ICA [10]. This method assumes that artifacts and brain activity are a physiologically separate process and this separation can be seen in the magnetic signals produced by this process.

Noise Reduction in natural images:
Depending on ICA decomposition and by filters, the images which are corrupted with additive Gaussian noise can be removed. There are different ways to remove noise from natural images like to make a transformation to spatial frequency space by DFT, do low-pass filtering, and return to the image space by IDFT but this method is not much reliable, therefore, another method came into existence in which wavelets transforms are used. For image statistics, statistically principled method sparse code shrinkage is introduced and this method is closely related to ICA.

ICA in text mining:
Independent Component Analysis (ICA) mainly had its application in signal processing. In a recent study, it has been found that if the numerical form of the text document is available, then ICA can be used to analyze text document as well. This can be useful in large databases in finding topics, grouping them appropriately. The Singular Value Decomposition (SVD) can be used to minimize the high dimensionality of data which is usually done before implementing ICA algorithms on data.

ICA for CDMA communication:
Distribution of the transmission medium among various users is the main issue in wireless communication. The use of common resources has to be used efficiently as

the number of users in the system increases. The other goal is to make every user in the system to communicate effectively [36].

Searching for hidden factors in Financial Data:
It is an appealing substitute to use ICA on financial data. Parallel time series are available in various applications like the exchange of currency rates or daily stock returns. ICA can be able to disclose some tools that are hidden. ICA can be applied in many areas: the cashflow of several stores belonging to the same retail chain, trying to find the fundamental factors common to all stores that affect the cashflow data. Thus, the cashflow effect of the factors specific to any particular store, i.e., the effect of the actions taken at the individual stores and in its local environment could be analyzed.

7 Conclusion and Future Work

In this paper, we have implemented the PCA techniques on different types of datasets namely African, Japanese, and Asian datasets and have calculated various parameters such as entropy, standard deviation, mean, median, variance, mode, correlation, and SSIM and shown the graph for SSIM, entropy, and variance for test and equivalent images, respectively. The distance between the test and equivalent images is calculated based on Euclidean distance and similarly, we can calculate these distances based on other distance measurement techniques such as city block and chessboard. We have also compared PCA and ICA based on certain specifications. In our study, we have also concluded that face-based PCA is a classical fruitful technique and PCA justifies its strength in pattern recognition and dimensionality reduction. On the other hand, in future work, ICA can be made to work efficiently on various image processing applications as it mainly works well with signal processing applications. These techniques can also be made to test for the underwater images by applying the suitable preprocessing on the dataset before giving the input to the PCA/ICA.

References

1. J.O. Ramsay, B.W. Silverman, Principal components analysis for functional data. Funct. Data Anal. 147–172 (2005)
2. S. Wold, K. Esbensen, P. Geladi, Principal component analysis. Chemometr. Intell. Lab. Syst. **2**(1–3), 37–52 (1987)
3. K. Polat, S. Güneş, An expert system approach based on principal component analysis and adaptive neuro-fuzzy inference system to diagnosis of diabetes disease. Digit. Signal Proc. **17**(4), 702–710 (2007)
4. Z. Zaplotnik, *Indepenent Component Analysis* (Apr 2014)
5. B.A. Draper et al., Recognizing faces with PCA and ICA. Comput. Vis. Image Underst. **91**(1–2), 115–137 (2003)

6. V. Gupta et al., An introduction to principal component analysis and its importance in biomedical signal processing, in *International Conference on Life Science and Technology, IPCBEE,* vol. 3 (2011)
7. H.L. Shang, A survey of functional principal component analysis. AStA Adv. Stat. Anal. **98**(2), 121–142 (2014)
8. P. Comon, C. Jutten, *Handbook of Blind Source Separation: Independent Component Analysis and Applications* (Academic press, 2010)
9. W. Zhao et al., Discriminant analysis of principal components for face recognition, in *Face Recognition* (Springer, Berlin, Heidelberg, 1998), pp. 73–85
10. A. Hyvärinen, E. Oja, Independent component analysis: algorithms and applications. Neural Netw. **13**(4–5), 411–430 (2000)
11. T.-W. Lee, M. Girolami, T.J. Sejnowski, Independent component analysis using an extended infomax algorithm for mixed subgaussian and supergaussian sources. Neural Comput. **11**(2), 417–441 (1999)
12. B.S. Everitt, G. Dunn, Principal components analysis. Appl. Multivar. Data Anal. 48–73 (2001)
13. Y. Aït-Sahalia, D. Xiu, Principal component analysis of high-frequency data. J. Am. Stat. Assoc. 1–17 (2018)
14. M.G. Borgognone, J. Bussi, G. Hough, Principal component analysis in sensory analysis: covariance or correlation matrix. Food Qual. Prefer. **12**(5–7), 323–326 (2001)
15. E. Bingham, A. Kab´an, M. Girolami, Topic identification in dynamical text by complexity pursuit. Neural Process. Lett (Submitted)
16. J. Shlens, A tutorial on principal component analysis. arXiv preprint arXiv:1404.1100 (2014)
17. X. Zheng et al., A high-performance computing toolset for relatedness and principal component analysis of SNP data. Bioinformatics **28**(24), pp. 3326–3328 (2012)
18. M.E. Wall, A. Rechtsteiner, L.M. Rocha, Singular value decomposition and principal component analysis, in *A Practical Approach to Microarray Data Analysis* (Springer, Boston, MA, 2003), pp. 91–109
19. R.E. Madsen, L.K. Hansen, O. Winther, Singular value decomposition and principal component analysis. Neural Netw. **1**, 1–5 (2004)
20. K. Raju, T. Ristaniemi, J. Karhunen, E. Oja, Suppression of bit-pulsed jammer signals in DS-CDMA array systems using independent component analysis, in *To Appear in Proceedings of the 2002 IEEE International Symposium on Circuits and Systems (ISCAS2002)* (Phoenix, Arizona, USA, 26–29 May 2002)
21. C.J. James, C.W. Hesse, Independent component analysis for biomedical signals. Physiol. Meas. **26**(1), R15 (2004)
22. A. Hyv¨arinen, Complexity pursuit: separating interesting components from timeseries. Neural Comput. **13**(4), 883–898
23. Simon HaykinZhe Chen, The cocktail party problem. Neural Comput. **17**, 1875–1902 (2005)
24. S. Goel, A. Verma, S. Goel, K. Juneja, ICA in image processing: a survey, in *IEEE International Conference on Computational Intelligence & Communication Technology* (2015)
25. T. Ristaniemi, K. Raju, J. Karhunen, Jammer mitigation in DS-CDMA array systems using independent component analysis, in *To Appear in Proceedings of the 2002 IEEE International Conference on Communications (ICC2002)* (New York City, NY, USA, 28 Apr–2 May 2002)
26. A. Hyvärinen, ErkkiOja, A fast-fixed-point algorithm for independent component analysis. Neural Comput. **9**(7), 1483–1492 (1997)
27. K.C. Chung, S.C. Kee, S.R. Kim, Face recognition using principal component analysis of Gabor filter responses, in *Proceedings of the International Workshop on Recognition, Analysis, and Tracking of Faces and Gestures in Real-Time Systems* (IEEE, 1999)
28. Z. Dabiri, S. Lang, Comparison of independent component analysis, principal component analysis, and minimum noise fraction transformation for tree species classification using APEX hyperspectral imagery. ISPRS Int. J. Geo-Inf. **7**(12), 488 (2018)
29. K. Baek, B.A. Draper, J.R. Beveridge, K. She, PCA vs. ICA: a comparison on the FERET data set, in *JCIS* (2002), pp. 824–827
30. B. Luo, Y.J. Hao, W.H. Zhang, Z.S. Liu, *Comparison of PCA and ICA in Face Recognition*

31. L.I. Smith, *A Tutorial on Principal Components Analysis* (2002)
32. L.J. Cao et al., A comparison of PCA, KPCA and ICA for dimensionality reduction in support vector machine. Neurocomputing **55**(1–2), 321–336 (2003)
33. J. Yang et al., Two-dimensional PCA: a new approach to appearance-based face representation and recognition. IEEE Trans. Pattern Anal. Mach. Intell. **26**(1), 131–137 (2004)
34. R.J. Martis, U.R. Acharya, L.C. Min, ECG beat classification using PCA, LDA, ICA and discrete wavelet transform. Biomed. Signal Process. Control **8**(5), pp. 437–448 (2013)
35. M.P.S. Chawla, PCA and ICA processing methods for removal of artifacts and noise in electrocardiograms: a survey and comparison. Appl. Soft Comput. **11**(2), 2216–2226 (2011)
36. J.M. Lee, C.K. Yoo, I.B. Lee, Statistical process monitoring with independent component analysis. J. Process Control **14**(5), pp. 467–485 (2004)
37. E. Bingham, Topic identification in dynamical text by extraction minimum complexity time components, In *Proceeding of the 3rd International Conference on Independent Component Analysis and Blind Signal Separation (ICA2001)* (San Diego, CA, USA, 9–13 Dec 2001), pp. 546–551
38. D.A. Socolinsky, A. Selinger, *A Comparative Analysis of Face Recognition Performance with Visible and Thermal Infrared Imagery*
39. R. Jayaswal, J. Jha, A hybrid approach for image retrieval using visual descriptors, in *2017 International Conference on Computing, Communication and Automation (ICCCA)* (IEEE, 2017)
40. R. Devesh, J. Jaimala, R. Jayaswal, Retrieval of monuments images through ACO optimization approach. Int. Res. J. Eng. Technol. (IRJET) **4**(7), 279–285 (2017)

Indian Currency Recognition Using Radial Basis Function and Support Vector Machine



Anand Upadhyay, Shashank Shukla and Verlyn Pinto

1 Introduction

The currencies of any country define the authenticity of that country, therefore, the currency recognition system plays a vital role because, with the help of currency recognition system, the authorities can easily make difference between fake and original currencies. Apart from the number of precaution and security mechanisms there are intruders those who make and replicate the fake currency of original currency, therefore, there is a need for very strong and well proof system for currency recognition. The currency recognition system is developed using different techniques and algorithms in past to recognize the fake and original currencies. The currency recognition system is designed to make a difference between fake and original currency, which human eye cannot make difference between fake and original currency. The currency recognition system should be recognizing the currency quickly and correctly. There are various features that are available in currency, based on that the currencies are recognized by authorities. The same features are used to recognize the currencies. There are different classifications methods are implemented for the recognition of currencies of different nation. Here the Indian currencies are used for recognition. The Indian currencies are available in two formats which are old currencies before demonetization and new currencies after demonetization. The new and old both the currencies are used for recognition. The research paper specifically uses the very important features of image called as edges of images. The edges of image are called as the high pass content of the image. The high pass content (edges) of Indian currencies image are calculated using Sobel and Prewitt filters. To make currencies recognition system more strong and accurate, the different moment's features are

A. Upadhyay · S. Shukla · V. Pinto (✉)
Department of Information Technology, Thakur College of Science & Commerce, Thakur Village, Kandivali (East), Mumbai 400101, India
e-mail: verlynpinto@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_106

calculated after the calculation of edges of Indian currencies image. The edges features are selected for recognition of Indian currencies because the size and different features available in new and old Indian currencies are different. The moment value for any kind of image is always unique and which provides statistical features. The statistical features are actually summary of entire Indian currencies image. All these features are supplied to radial basis function and support vector machine and based on that Indian currencies are recognized. The accuracy of both the techniques are calculated using confusion matrix and their comparative study is performed. There are various types of problems related to currency recognition because of very hidden and strong security features which humans cannot recognize because of human eye limitation and unawareness about all those features. A currency recognition system is developed to defeat such types of problems are related to currency recognition because for machine it is very easy to recognize and make difference. There are many places where currency recognition system is used, i.e., shops, bank counters and automated teller machines, auto vendor machines, and currency recognition systems can be used. The currency recognition system not only recognizes the currencies but also reduces human efforts and mistakes. The following paper is organized in much-formatted way which starts with introduction, literature review, theoretical study about classifiers, methodologies, and results and conclusion.

2 Literature Review

In past research work on currency recognition, there are various methods to validate the currency like optically variable ink, watermark, iridescent ink, security thread, Contour Analysis, Canny Edges, and Hough transformation algorithm [1]. Research on currency recognition has been ongoing for some time and the algorithms applied for currency recognition do not involve many artificial neural network algorithms in the early stages of research [2]. The process of validation is done by applying these feature extracting methods on an original currency image and storing them as training data set [3]. The created training dataset of the original currency image is used to compare with the training data of the tested note. The difference after comparison is calculated to find out the accuracy of the tested image. In this research paper the methods used for extracting features have been enhanced. The proposed system has additional processes like edge detection which is done using Sobel and Prewitt filters and use of statistical features for creation of training dataset [4]. The additional processes which are implemented are to make the proposed system more strong. The previous work on currency recognition has limitation with respect to the technology implemented. The technology used need to process images which involved changing the nature of an image in order to have certain set of algorithms implemented. The technology also has to implement Pattern Recognition Technique, to find an area of interest that matches patterns [5]. The image acquisition is done by scanning the image and the images are stored as JPEG. The images were initially not scanned from the same source which led to inconsistency in the values of the

images after they have been processed [6]. After which researchers ensured that all the collected images must be collected from the same source, which is only from one camera, since the previous collected image from for the research had increase or decrease the pixel value. The storing of image data had created a huge vacuum, as the few pixels can vary which used to alter the result [7]. Therefore, few researchers filled this vacuum to take clear images from only the source and not many of the stored image resources [4]. The slightly blurred image can result in a different pixel value than the other images are stored. Hence a certain amount of consistency was maintained in acquiring the currency images (Fig. 1).

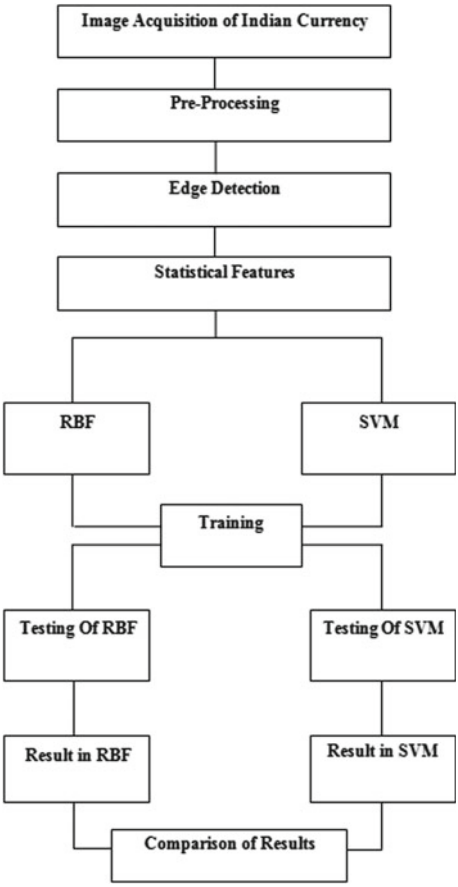


Fig. 1 Proposed architecture

3 Methodology

3.1 Image Acquisition

The collection of Indian Currency images was done using Nikon D850 high resolution. There was only one source used to collect the image to make sure all the images were of the same pixel value. The images acquired were stored in a JPEG format. No scanning of images was done. To maintain the consistency of the quality of the image the no blurred images were taken.

3.2 Image Preprocessing

The preprocessing of the image is called a beautification process of an image which is used to enhance the quality of the image and to remove the noise from the image. Here, the histogram equalization applied on an image of currencies which are acquired through the camera to improve the quality of image for proper feature extraction and to collect the detailed information from currencies.

3.3 Edge Detection

In an image, edges are high pass value which can easily be extracted by the high pass filter value. There are two very popular high pass filters are available for edge detection which is Sobel and Prewitt high pass filters. The Sobel and Prewitt filters provide the high pass information from the Indian currencies note which is further used to calculate more accurate information.

3.4 Statistical Features

The statistical feature is calculated to provide more exact and summarized value for Indian currencies recognition. The statistical feature extraction is the second level of feature extraction process which is used to prepare training and testing data sets. Here, under statistical feature different moment values of Sobel and Prewitt edges are calculated and which is used for training and testing purpose.

3.5 RBF and SVM Classifier

The python programming language is a very powerful programming language which consists of various features related to data science, data analysis, data visualization, and image processing, etc. The scikit-learn package in Python provides a very powerful tool for machine learning and data analysis. The radial basis function and support vector machine classifiers are implemented using the scikit-learn library function. The support vector machine classifiers are implemented one verse all class classifiers and radial basis function is a multiclass classifier.

3.6 Training Dataset

The statistical features which are calculated from Sobel and Prewitt edges are divided into two commas separated value CSV file. The one CSV file is used for training and another CSV file is used for the testing purpose of classifiers. The testing file is used for testing of the data. The training is a learning process which is used for learning of classifiers. The learning is the weight adjustment process of classifiers.

3.7 Testing of the Classifier

The testing file which is created at an initial level along with a training file is used for testing purpose. The testing file is used to calculate the accuracy of classifiers to do the comparative study of radial basis function and support vector machine. The confusion matrix and statistical methods which are known as the Kappa coefficient are used to calculate the accuracy of classifiers.

4 Result

The result has a comparative analysis of Radial basis function and Support vector machine. Figure 2 shows the original image and Fig. 3 shows the effect of Sobel and Prewitt filter on the original image (Table 1).

The accuracy of the confusion matrix using Radial basis function algorithm is 100 and accuracy of confusion matrix using Support Vector Machine algorithm is 84 (Fig. 4).

Kappa coefficient was 1.00 for Radial Basis Function and 0.83 for Support Vector Machine (Fig. 5).

A confusion matrix is a table which is often used to describe the performance of a classification model on a test data set known for the true values. A confusion matrix is



Fig. 2 Original image

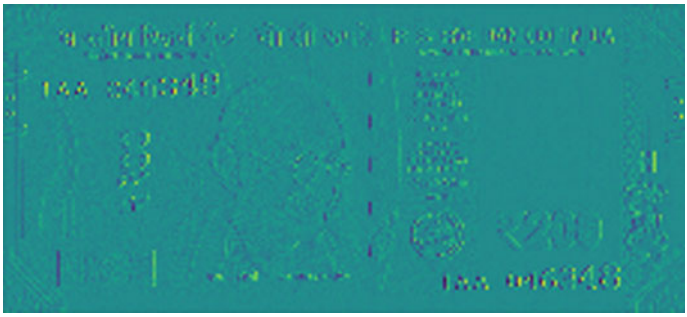


Fig. 3 After applying Sobel filter algorithm

Table 1 Kernel functions

Algorithms	Accuracy	Kappa value
RBF	100	1.00
SVM	84	0.83

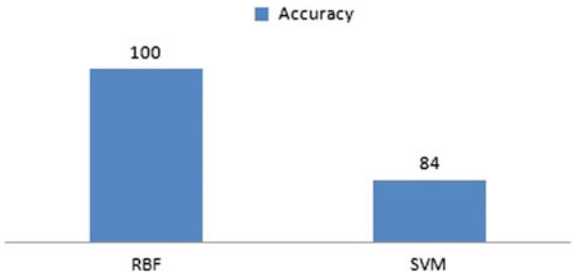


Fig. 4 Graphical representation of accuracy

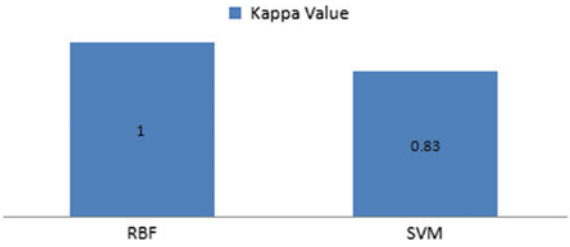


Fig. 5 Graphical representation of kappa value

a two by two table with four binary classifier results. The matrix of confusion derives from different measurements such as error rate, accuracy, specificity, sensitivity, and precision. In addition, several advanced measures are based on them, such as ROC and precision reminder. It is the tool to measure the accuracy of the given image. It consists of information in supervised learning to classify and misclassify data. Hence, the confusion matrix provides the accuracy that is effective for the detected image. Kappa coefficient is another method for the measurement of accuracy of the classifiers. This is basically to measure how well the classifier performed compared to how it will perform in a simple manner. In other words, a model will have a high Kappa score if there is a big difference between precision.

5 Conclusion

In this research work, a number of Indian currency notes have been analyzed. The research work conducted in this paper has shown the use of these steps from the use of acquisition of the image to convert it to grayscale image, Sobel, feature extraction and use of RBF and SVM Classifier have been implemented successfully. A single most important factor in achieving high recognition performance is the selection of feature extraction method. In future, inclusion of consistent pattern design of the note can help extraction using Neural Network methods with them trained on feature vectors obtained from the above system.

6 Future Enhancement

In future, we aim to apply and improve the given algorithms on notes which have been damaged. The notes which are been burnt due to natural or manmade incidents. The algorithm may focus on identifying the value of the note and certain features of the note which will help to uniquely identify whether the given note is real or fake. The future segment will involve the use of artificial neural network from the earlier stages. The future research work will implement automation of the working of the damaged note. Improving accuracy and kappa value.

Acknowledgements I want to thank all the people involved in the research for providing valuable feedback and opinion by spending valuable time there. It is not necessary to number the heading of the section Acknowledgement and the section References.

References

1. D. Alekhya, G.D. Surya Prabha, G.V. Durga Rao, Fake currency detection using image processing and other standard methods. IJRCCT **3**(1), 128–131 (2014)
2. M. Thakur, A. Kaur, Various fake currency detection techniques. Int. J. Technol. Res. Eng. **1**(11), 1309–1313 (2014)
3. D.C. Cireşan, A. Giusti, L.M. Gambardella, J. Schmidhuber, Mitosis detection in breast cancer histology images with deep neural networks, in *International Conference on Medical Image Computing and Computer-assisted Intervention* (Springer, Berlin, Heidelberg, 22 Sept 2013), pp. 411–418
4. K. Santhanam, S. Sekaran, S. Vaikundam, A.M. Kumarasamy, Counterfeit currency detection technique using image processing, polarization principle and holographic technique, in *2013 Fifth International Conference on Computational Intelligence, Modelling and Simulation (CIMSim)* (IEEE, Sept 2013), pp. 231–235
5. C.A. Witschorik, U.S. Patent No. 6,131,718, Washington, DC: U.S. Patent and Trademark Office (2000)
6. A.M. Tekalp, *Digital Video Processing* (Prentice Hall Press, 2015)
7. D.U. Mennie, B.T. Graves, U.S. Patent No. 5,806,650, Washington, DC: U.S. Patent and Trademark Office (1998)

LDA- and QDA-Based Skin Lesion Melanoma Detection



Anand Upadhyay, Arvind Chauhan and Darshan Kudtarkar

1 Introduction

Melanoma occurrence rates are rising rapidly, leading to higher death rates. Compared to other skin cancer diseases, melanoma disease is considered more dangerous. In the past three decades, a number of people have been diagnosed with this melanoma disease which has increased sharply. It causes 75% of other skin cancers which relates to deaths. The incidence rate of melanoma in India is between 0.65 and 6%. Melanoma usually looks likes colored skin where the spot is enlarged. They have different colors like brown, black, and blue. Melanoma skin disease has various stages which are stage 0, stage 1, stage 2, and stage 3. Tumors only come in stage 0 on the skin surface. In stage 1, tumors invade the skin but are not nucleated which grows at a slow mitotic rate. Stage 2 is considered intermediate melanoma and it is classified differently. Stage 3 is the most advanced stage of melanoma that affects different organs, which makes it difficult to treat [1]. Early detection of melanoma is necessary. The detection system for skin cancer saves a lot of time for the doctor and can help in diagnosis. It can also easily evaluate future skin development through dialysis at the current skin age and can present the best characteristic skin cancer project to the customer. Keep in mind that when most melanoma is found, there are no symptoms. Some are also annoying, and it can also be a late sign of injury. Melanomas are rarely irritating a bit themselves, like completely different skin cancers. They are detected by their appearance alone in a bulk of cases. It is vital to urge the early elimination of melanoma. This can usually be done due to the sensitive prognosis of "thin" melanoma results (96% cure rate). They usually develop in areas where the skin is exposed to the sun, on body parts like the back, legs, arms, and faces. Melanoma disease sign begins when the skin color changes. There are usually

A. Upadhyay (✉) · A. Chauhan · D. Kudtarkar
Department of IT, Thakur College of Science & Commerce, Thakur Village, Kandivali (East), Mumbai 400101, Maharashtra, India
e-mail: anandhari6@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_107

mixed colors (pink, red, and brown). It is most often found in women instead of men. While the most common part for women is the legs where the cancer occurs and the most common part is the back in the case of men. If it is not detected early, it can spread to other parts of the body and tissues. Skin lesion segmentation is considered due to the complexity under and over segmentation. It has a greater ability to spread to other organs, so it is responsible for the highest death rates, even with its low incidence. The most common method to detect and classify the disease is image processing. When subsequently entering an image, a series of images, or videos, output may be an image processing or a set of characters or parameters related to the images input. Images have wider future work in modern science and technology. To obtain an improved image for useful information, it is the method of performing certain image operations. Image processing plays an important role in skin cancer detection. Dermatological photographs should be used for melanoma risk analysis and assessment. Digital techniques of image processing can define an approach to skin cancer of this type. This approach provides a qualified person with reliability for diagnosis and agility. Applications for digital image processing increasingly require methods that accentuate the information contained in images for human interpretation and analysis.

2 Literature Review

In this research, we will focus on the work carried out by many people on the detection of melanoma using completely different techniques such as FCM, K-means, GLCM, and Contour Signature. Many companies have dedicated time and effort to boost the first screening method, depending on the importance of early detection [2]. Classification of skin cancer is done using the method of water shed and edge detection. Compared to TDS (total dermatoscopy score), PCA (main component analysis) provides 92% accuracy [3]. We found that the color diameter of the asymmetry border was the classification and segmentation of images. It is also possible to analyze patterns and texture. The origin of the image associated with a particular property is found. It includes filtering and sampling [4]. To indicate that melanoma disease can be detected using MATLAB, for early detection of melanoma, image segmentation method was used [5]. For subsequent observation or confirmation of the lesion, first, it captures and analyzes the label image of pigmented skin lesions. In 2002, Thomas Bayes rule was revealed with rather inconclusive results in concert with the skin lesions classification technique. Dermoscopic images, however, result in an increased view of skin lesions; the interpretation and accuracy of the diagnosis depend primarily on the viewer's experience. Multiple diagnostic models with similar reliability have been widely accepted by physicians, such as 7-point checking; Menzie rule is the most popular scoring system known as ABCD rule which determines and is further extracted which indicates the lesion property and characterized it usually which involve error-prone operation through all automated diagnostics. Therefore, an endless work of interest is introduced to these automated diagnostic systems which

are noninvasive tools to support diagnostics. A significant number of research and publications are focused on the identification and classification of melanoma images in the field of image analysis and pattern classification. Computer-aided melanoma diagnostic systems have developed a diagnostic accuracy of approximately 73–98% over the past 20 years [6]. Common classification strategies have been applied in the research, such as applied mathematics and rule-based strategies. In the analysis, the closest neighbor was used as another classification technique [7]. In 1985, in recognition of the need to educate doctors and the public to recognize melanoma in its 7 early clinical presentations, the New York University group coined the acronym for ABCD (asymmetry, irregularity of the border, color variety, diameter >6 mm). For the detection of melanoma skin lesions, ABCD features are most commonly used for the extraction of features based on a morphological analysis of the skin lesion dermoscopic image [8]. Earlier detection of lesion features compared to the ABCD rule was performed. Support vector machine detects the disease in this research, in which the feature is an input classifier. Image processing has a promising scope for early detection and classification of skin cancer using the image of the skin. Most of the research has been done in this area, including early detection of disease, malignant transformation, classification of cancer, and benign/malign discrimination. However, to make the system more useful and applicable, further research in this area should be carried out. To improve this technology, the authors propose several improvements [9].

3 Methodology

See Fig. 1.

3.1 Image Acquisition

The main stage of any vision system is image acquisition. Image acquisition is the process in which the creation of the digitally encoded representation is done. In this methodology, the skin lesion images were acquired by a digital camera. The image was in the JPEG form.

3.2 Image Preprocessing

After image acquisition, the second step is preprocessing. The image is given as input, the lesion image is checked, if so, it is preprocessed. Image Pre-Processing is the process in which the image is filtered which can affect the accuracy. So, the digital image is filtered for noise removal.

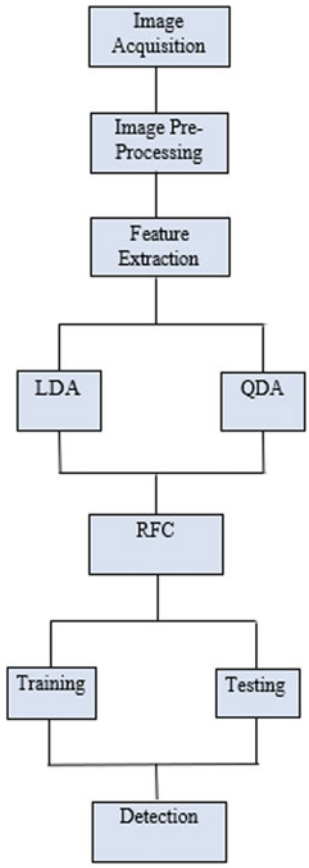


Fig. 1 Proposed architecture

3.3 Feature Extraction

Extraction of features plays an important role in the processing of digital images. Extraction of these features is a process to receive the features of the image to be classified. Extraction of features also plays a major role in image detection. The feature which was extracted on the image, in this research paper, was affected and unaffected by skin lesion. The original image was processed and the feature was extracted by making two classes.

3.4 LDA

Linear Discriminant Analysis (LDA) is a method used to find linear combination in statistics, pattern recognition, and machine learning which characterize or separate

multiple classes of objects or events. It is a generalization of Fisher's linear discriminant. LDA predicts by estimating the probability that each class has a new set of inputs. The most likely class is the output class and a forecast is made. The most commonly used technique by LDA is the size reduction method which is used for pattern classification and machine learning applications in image preprocessing. It maximizes class separation component axes.

3.5 QDA

This is about Quadratic Discriminant Analysis (QDA), which is a standard probabilistic classification method in statistics and machine learning. QDA is closely linked to Linear Discriminant Analysis (LDA) where measurements are normally assumed to be distributed. Unlike LDA, there is no assumption in QDA that the covariance of each class is the same. More computation and data are needed to estimate the parameters required in quadratic discrimination than in linear discrimination. If the group covariance matrices do not make a big difference, the latter will perform as well as quadratic discrimination. The general form of Bayesian discrimination is quadratic discrimination.

3.6 RFC

The random forest classification can be used for both classification and types of regression problems. In the same way, the higher number of trees in the forest results in the random forest classification being highly accurate.

3.7 Training

During feature extraction, the training file was made with two classes which are class 0 and class 1. To classify the image as normal skin and melanoma cancer lesion, the RFC classifier is selected. The RFC classifier was used to train the data for matching the lesion on the affected area using the Linear Discriminant Analysis.

3.8 Testing

Once the training of the data is done, the testing part is important. After applying testing on the data, the confusion matrix was prepared, and accuracy as well as the kappa coefficient was measured.

3.9 Detection

Detection was done on the skin lesion image which was that of a melanoma skin cancer disease.

4 Results

To measure the accuracy of the skin lesion disease for the image, LDA and QDA algorithms were applied. Random Forest Classifier was applied. LDA and QDA algorithms were used for effectiveness and accuracy of the melanoma skin cancer disease (Fig. 2).

Figure 3 shows the affected area which was classified with the RFC classifier. The affected area shows the lesion on the skin which is the melanoma skin disease. The accuracy for the LDA algorithm for the affected area image was 98.83% and the kappa coefficient was 0.96. Similarly, for the QDA algorithm, the accuracy for the affected area image was 98.49% and the kappa coefficient was 0.95.

Fig. 2 Original image

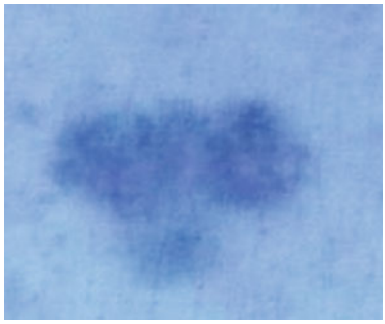
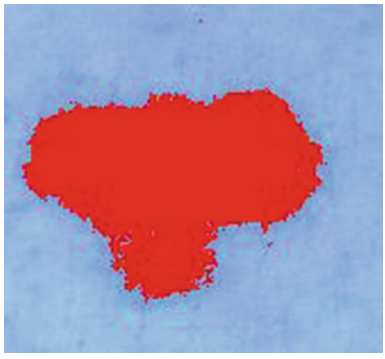


Fig. 3 Affected area



The confusion matrix was applied to the affected area image for the accuracy of the melanoma disease. A confusion matrix is a table that is often used to describe the performance of a classification model on a test dataset known for the true values. The confusion matrix is a two-by-two table with four binary classifier results. The matrix of confusion is derived from different measurements such as error rate, accuracy, specificity, sensitivity, and precision. In addition, several advanced measures are based on them, such as ROC and precision reminder. It is the tool to measure the accuracy of the given image. It consists of information in supervised learning to classify and misclassify data. Hence, the confusion matrix provides the accuracy that is effective for the detected image. Kappa coefficient is another method for the measurement of accuracy of the classifiers. This is basically to measure how well the classifier performed compared to how it will perform in a simple manner. In other words, a model will have a high kappa score if there is a big difference between precisions (Figs. 4, 5 and Table 1).

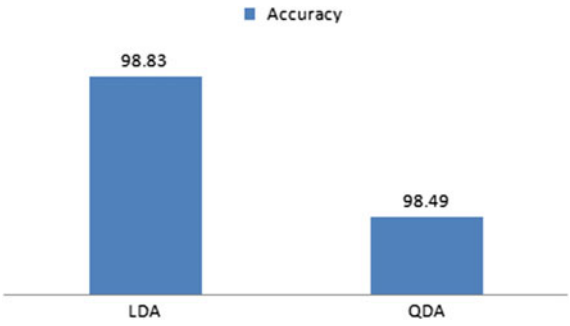


Fig. 4 Graphical representation of accuracy for different discriminant analysis

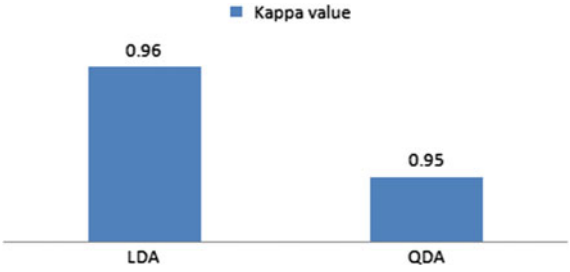


Fig. 5 Graphical representation of kappa value for different discriminant analysis

Table 1 Different discriminant analysis

Algorithms	Accuracy	Kappa value
LDA	98.83	0.96
QDA	98.49	0.95

5 Conclusion

This research paper proposes a method for early detection of melanoma. After applying the LDA and QDA algorithms, the disease was detected. The RFC classifier was developed in the research for the classification and the regression of the data. Basically, RFC gives higher accuracy results. These techniques work on the image so there is no physical contact with any part of the body, so this is noninvasive. The confusion matrix which was applied was efficient to measure the accuracy as well as the kappa coefficient for the skin lesion image.

6 Future Enhancement

The system's final output helps the dermatologist detect the injury and its type. He will therefore examine the patient with his knowledge to determine whether the lesion can be cured or not, e.g., by using medicines or ointments, or in any other way. The skin cancer detection system will assist dermatologists in the early diagnosis of melanoma. The best way to reduce the risk of melanoma is by exposure to high sunlight and other sources of ultraviolet light. Take care of all the following necessary measures: use clothes to protect the skin, wear hats, wear sunscreen, and stay in the shade.

Refrences

1. R.S. Soumya, S. Neethu, T.S. Niju, A. Renjini, R.P. Aneesh, Advanced earlier melanoma detection algorithm using colour correlogram, in *2016 International Conference on Communication Systems and Networks (ComNet)* (IEEE, 21 July 2016), pp. 190–194
2. R. Maurya, S.K. Singh, A.K. Maurya, *GLCM and Multiclass Support Vector Machine Based Automated Skin Cancer Classification* (IEEE, 2014)
3. N. Paliwal, Skin cancer segmentation, detection and classification using hybrid image processing technique. Int. J. Eng. Appl. Sci. **3**(4)
4. N. Oommachen, V. Vismi, S. Soumya, C.D. Jeena, Melanoma skin cancer detection based on skin lesions characterization. IOSR J. Eng. (IOSRJEN). e-ISSN.:2250-3021
5. N.S. Ramteke, S.V. Jain, ABCD rule based automatic computer-aided skin cancer detection using MATLAB. Int. J. Comput. Technol. Appl. **4**(4), 691–697
6. M.A. Sheha, A. Sharwy, M.S. Mabrouk, Automated imaging system for pigmented skin lesion diagnosis. Int. J. Adv. Comput. Sci. Appl. **1**(7), 242–254 (2016)
7. M.E. Celebi, H. Iyatomi, G. Schaefer, W.V. Stoecker, Lesion border detection in dermoscopy images. Comput. Med. Imaging Graph. **33**(2), 148–153 (2009)
8. H.T. Lau, A. Al-Jumaily, Automatically early detection of skin cancer: study based on neural network classification, in *International Conference of SOCPAR* (2009), pp. 375–380
9. H. Ganster, P. Pinz, R. Rohrer, Ernst Wildling, Michael Binder, and Harald Kittler. Automated melanoma recognition. IEEE Trans. Med. Imaging **20**(3), 233–239 (2001)

Design and Analysis of a Multilevel DC–DC Boost Converter



Divesh Kumar, Dheeraj Kalra and Devendra Kumar

1 Introduction

Increasing demand in energy conservation leads world to use LED bulbs and in the coming years, BLDC motors are going to replace the existing induction motor of ceiling fans. BLDC motor is widely used in E-rickshaws. The conversion from DC to AC and then AC to DC makes the solar PV system less efficient. There are a lot of conversion losses which can be saved if DC supply is directly given to the appliance in place of AC supply. To cope up with the different operating voltage levels of different applications, there is the need for such a device that can provide multilevel DC voltage. DC–DC converters are more popular nowadays. These converters are generally used in applications like digital systems, renewable power applications, hybrid electric vehicle systems and regulators. The grid system requirement is high power converters at low-frequency operation. Boost converter finds its applications in X-ray, hybrid electric vehicle, and DC drive system, renewable applications such as PV system, fuel cell, micro-grid applications, and HVDC system.

In high-gain converters, transformers are not feasible as they produce non-linearities in the output. Transformers used in the resonant (isolated) converters are push-pull, flyback, half-bridge DC–DC converter, full-bridge DC–DC converter, etc. These converters preferably operate at high frequencies so it reduces the size of passive component, inductor current and capacitor voltage ripple.

DC–DC converters are of two types, isolated and non-isolated. Non-isolated converters are operated at a high frequency. These converters do not have transformers. To get a high output voltage, larger value of the passive component is needed [1]. In conventional DC–DC converter topologies, to reduce the size of the passive component, it is to be operated at high frequencies [2]. The basic DC–DC converter is not suitable for high-power applications. A new topology is derived from the basic

D. Kumar (✉) · D. Kalra · D. Kumar
GLA University, Mathura, India
e-mail: divesh.kumar@gla.ac.in

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_108

DC–DC converter. It increases the output voltage magnitude by adding capacitors and diodes without changing the conventional boost converter [3].

Multilevel DC–DC boost converter consists of one switch, one inductor with *N* number of capacitors and diodes. The inductor is connected in series with the DC source. The proposed topology can give any output voltage by adding capacitors and diodes in the basic circuit. The main demerit of the proposed topology is the requirement of large inductors [4]. In a DC–DC converter, inductors and capacitors are the main components for bucking or boosting the voltage. In capacitor switching, only capacitors and diodes serve the purpose. It is done only by inductors and diodes in inductor switching. Merits include reduced ripple current, where reduced ripple voltage reduces and a single switch operates at high frequency [5]. Nowadays, requirement of power is increasing day by day. With conventional energy sources, we cannot meet the demand. The renewable energy sources such as photovoltaic (PV) system, fuel cell system, and wind system are an alternative choice. For all these renewable applications, converters are needed. For high-power applications, single input with multi-output like MBC is found to be more suitable. Merits of the presented work are high output voltage, no requirement of transformers and operate at high switching frequency. Demerits in the MBC topology are more number of capacitors and diodes required to produce high output voltage [6, 7].

1.1 Design Considerations

In DC–DC converters, the appropriate size of the inductor and capacitor selection is necessary. The paper describes the process to design a multilevel inverter optimally [8].

1.2 Modelling and Control of DC–DC Converters

Due to a sudden change in load, output parameters also change. In order to maintain constant output, controller design becomes necessary. The paper presents state-space modelling of the converter for a full-circuit model. Analysis of a full-circuit model for MBC is difficult. A reduced-order model analysis for MBC is also explained [9]. DC–DC converters are highly non-linear. Small-signal modelling of the DC–DC converter along with its controller design is presented [10–13]. DC–DC converters play an important role in the PV system. The shadow effect in the solar system reduces its power generation. It is eliminated by designing a suitable controller. MPPT controller design along with small signal dynamic modelling is examined [9, 13].

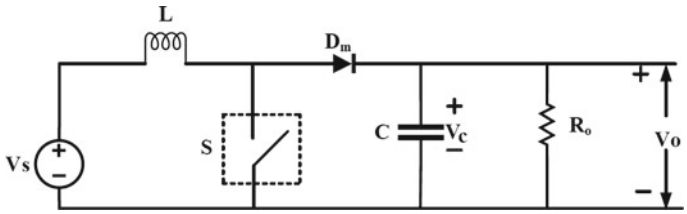


Fig. 1 Boost converter

1.3 Efficiency of DC–DC Converter

A DC–DC converter consists of components such as inductor, capacitors and semiconductor devices like diodes, MOSFETs and IGBT. The components do not operate ideally. A proper selection of components for modelling is described [14]. Conventional boost converter operates at a high duty period for higher gain. At high duty cycles, inductor may get saturated. Multilevel DC–DC boost converter operates at low duty cycle. Multilevel converter for fuel-cell application is presented in [15]. The need for the recent days is to reduce the uses of non-renewable sources. The transformation of petroleum-based vehicles to hybrid electrical vehicles and then to electric vehicles has been seen. Recent developments in DC–DC converters mark them suitable for electric vehicles. A number of DC supplies are needed to charge electric vehicles. A multilevel DC–DC converter is derived to serve the purpose [16]. Single fuel-cell voltage output is roughly 1.5 V. To increase the voltage, fuel cells are arranged in a stacked manner. The voltage is then fed to a DC–DC boost converter. Further the boosted voltage is fed to an inverter to achieve the desired ac voltage and frequency for utility grid [17].

1.4 Boost Converter

The conventional boost converter is shown in Fig. 1. The capacitor and inductor work as energy storing elements. Switch S may be any electronic switch, e.g., BJT and MOSFET. The capacitor maintains the constant voltage across the load. The inductor stores energy during the switch-on (DT_s) period and transfers the stored energy to the load resistance through the diode (D_m) during the switch-off period (1-D)T_s.

$$\text{Output voltage } V_o = \frac{V_s}{(1-D)} \tag{1}$$

$$\text{Source current } I_s = \frac{I_o}{(1-D)} \tag{2}$$

2 Design of Multilevel DC–DC Boost Converter

A multilevel DC–DC boost converter is shown in Fig. 2. It consists of N number of conventional boost converters. Multiple levels of output voltages are synthesized using one inductor, one electronic switch and multiple capacitors and diodes. To obtain N number of output levels, $2N - 1$ number of diodes and $2N - 1$ number of capacitors are connected in the conventional circuit. This flexibility makes the proposed topology more robust and useful.

Based on the state of the switch, there are two modes of operation of MBC. First is the switch-on condition ($T1 = DT_s$) and second is the switch-off condition ($T2 = (1–D)T_s$). The mode of operation of a three-level boost converter is explained for both the switching conditions.

2.1 Switch-ON Condition ($T_1 = DT_s$)

The inductor is connected in series with the source to provide constant input current and it stores the charge during on period as shown in Fig. 3a. When switch S is on, and the capacitor C_5 is charged having voltage greater than the voltage across C_4, then C_4 will be charged from C_5 through D_5 and switch S as shown in Fig. 3b. If

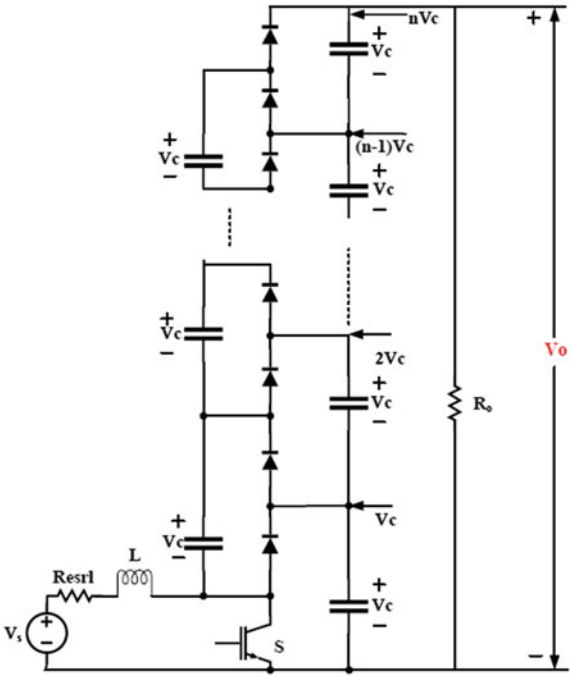


Fig. 2 DC–DC multilevel boost converter

the total voltage across C_5 and C_3 is more than the total voltage across C_4 and C_2, then both of the capacitors, C_4 and C_2, get charged from C_5 and C_3 through D_2 and switch S as shown in Fig. 3c.

2.2 *Switch-OFF Condition* [$T_2 = (1 - D)T_s$]

The inductor is charged and adds up with the source voltage and is used to charge the capacitor C_5 through D_5 during off period as shown in Fig. 4a. If the voltage across C_4 is greater than the voltage across capacitor C_4, D_3 will become forward bias and the capacitors C_3 and C_5 will be charged as shown in Fig. 4b. Similarly, upper stages work and the respective capacitor will be charged as shown in Fig. 4c.

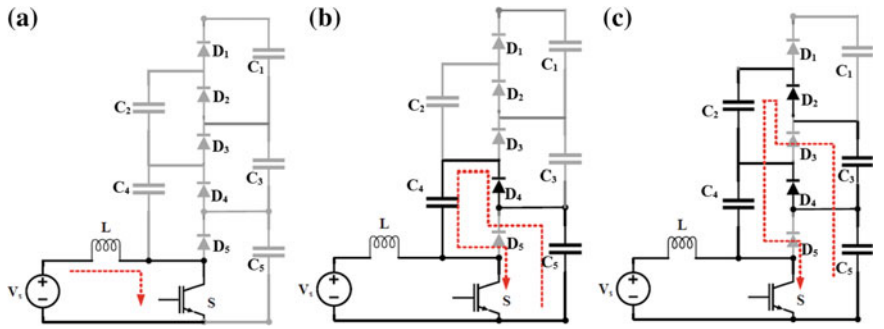


Fig. 3 **a** Inductor charging from source **b** charging of C_4 from C_5 **c** charging of C_2, C_4 from C_3, C_5

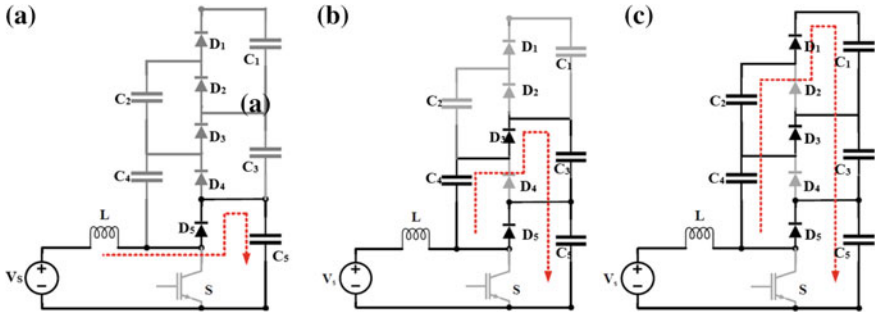


Fig. 4 Charging of capacitor C_5 **b** charging of capacitor C_3 and C_5 **c** charging of capacitors C_1, C_3 and C_5

2.3 *Effect of Parasitic Resistance* ($\mathbf{R_L}$)

Ideal conventional boost converter provides maximum boost ratio, but a practical converter has lower boost ratio. Parasitic resistance (internal resistance of inductor) is responsible for the limitation in the boost factor. As multilevel DC–DC boost converters operate at high switching frequencies, the size of inductor reduces and due to reduction in inductor size, internal resistance also reduces. At different parasitic resistance, gain of the converter is different at various duty periods, which is presented below.

2.4 *Analytical Expressions for Output Voltage and Source Current Without Parasitic Resistance*

From Fig. 3, the output voltage of the converter is equal to the sum of the capacitor voltages.

For N number of output side capacitors,

$$\text{Output voltage,} \quad V_o = N \times V_C, \quad I_L = I_S \tag{3}$$

During the on condition, inductor voltage is

$$V_L = V_S \tag{4}$$

During the off condition, inductor voltage is

$$V_L = (V_S - V_C) \tag{5}$$

Averaging the voltage across the inductor:

$$V_L|_{\text{on}} + V_L|_{\text{off}} = 0 \tag{6}$$

Using Eqs. (4, 5 and 6)

$$DV_S|_{\text{on}} + (1-D)(V_S - V_C)|_{\text{off}} = 0 \tag{7}$$

$$V_L = V_C(1-D) \tag{8}$$

From Eqs. (3) and (8)

$$V_o = \frac{N \times V_S}{(1-D)} \tag{9}$$

Equation (9) represents the voltage equation of *N*-level boost converter. Considering the lossless system,

$$V_S \times I_S = V_O \times I_O \tag{10}$$

From Eqs. (9) and (10)

$$V_S \times I_S = \frac{(N \times V_S)}{(1-D)} \times I_O \tag{11}$$

$$\mathrm{I_S} = \frac{N \times I_O}{(1-D)} \tag{12}$$

Equation (10) represents the source current equation for *N*-level boost converter. Equation (9) represents the multilevel DC–DC boost converter output voltage. For $N = 1$, output voltage represents the conventional boost converter output voltage. If input voltage(V_s) = 50 V, duty cycle (D) = 0.5, then conventional boost converter will provide the output voltage of 100 V, the same input parameters for three-level boost converter will provide the output voltage of 300 V in according to Eq. (9).

3 Results

The open-loop three-level MBC is shown in Fig. 5. Simulation parameters are chosen as follows.

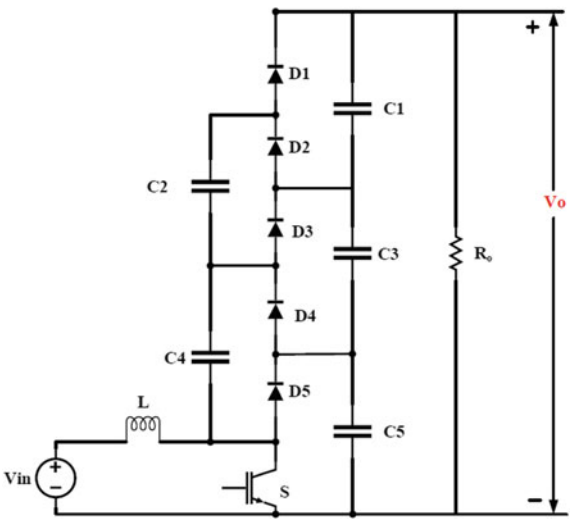


Fig. 5 Three-level boost converter

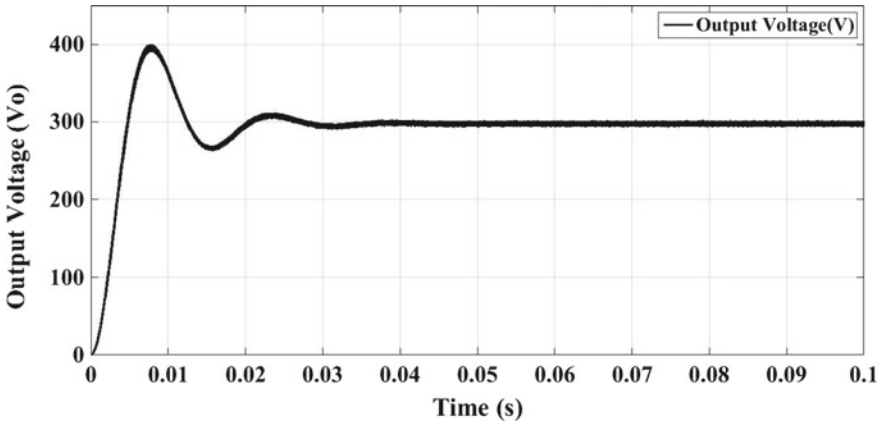


Fig. 6 Three-level boost converter output voltage (V)

Table 1 Three-level boost converter model specifications

Parameter	Specification
Input voltage (V_s)	50 V
Switching frequency	30 kHz
Capacitance (C)	200 μF
Load resistance (R_o)	30 Ω
Inductance (L)	1.33 mH
Duty cycle (D)	0.5

Figure 6 shows the output voltage of a three-level boost converter. The theoretical value of the output voltage can be calculated from Eq. 9 by putting the values from Table 1.

$$\text{Output voltage } V_o = \frac{N \times V_S}{(1-D)}, \text{ where } N = 3,\ V_s = 50\,\text{V and } D = 0.5.$$

After calculating, the theoretical value of output voltage is 300 V and also it is verified by the obtained result from the simulation.

Figure 7 shows the output current of the three-level boost converter. The output current is 10 A which is sufficient to operate the home appliance and devices used in home automation.

Figure 8 shows the response for the input current. From the simulation, the steady-state input current of 60 A is obtained, which validates the theoretical value that can be calculated from Eq. 12 as follows.

Input current is given as $I_S = \frac{N \times I_O}{(1-D)}$, where $N = 3$, $I_o = 10$ A and $D = 0.5$.

The theoretical value calculated from the above equation is 60 A which is also verified by the simulation result.

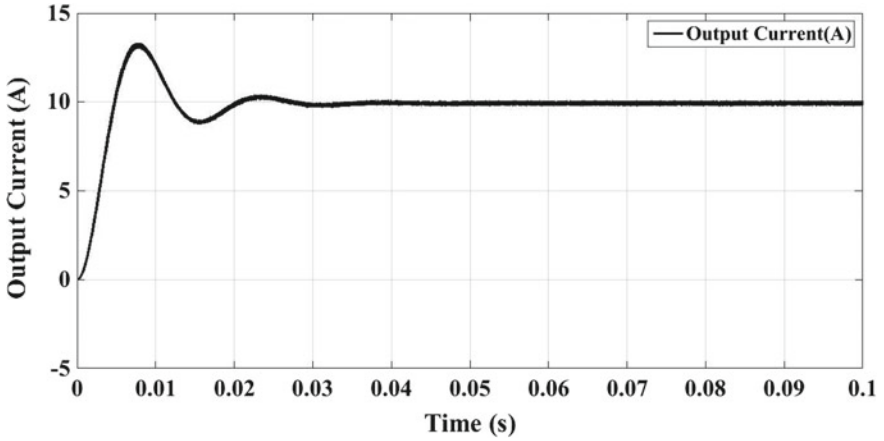


Fig. 7 Three-level boost converter output current (A)

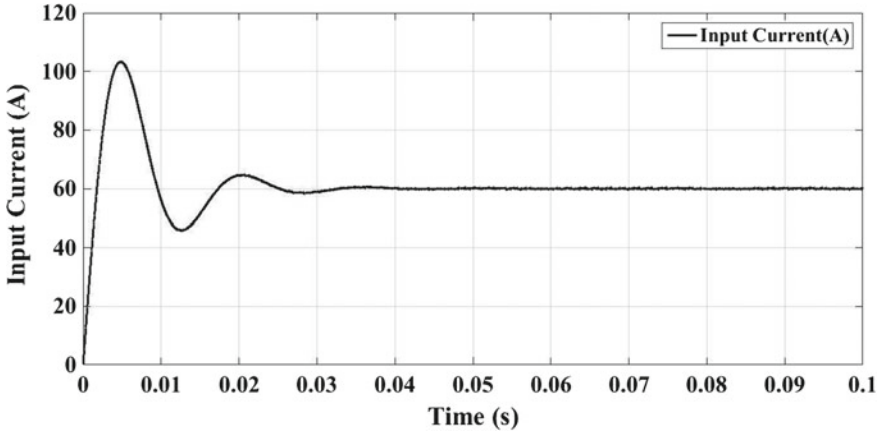


Fig. 8 Three-level boost converter output current (A)

4 Conclusion and Future Scope

Multilevel DC–DC boost converter is designed just by adding $(2N - 1)$ capacitors, $(2N - 1)$ diodes to the conventional boost converter to obtain N-level output voltage. The output voltage is boosted to a higher value without the use of transformer so that conversion losses may be reduced and converter size also reduces. The continuous conduction mode of the inductor current is verified by simulation results. The simulation results validate the mathematical design, and high voltage and low current are obtained by using a three-level boost converter. It is also observed that the output voltage is just an N multiple of the conventional converter output. Further,

the closed-loop analysis can be done to achieve different levels of output voltage and PWM can be used to compensate the internal losses. As this converter boosts the DC voltage, it can be used in charging the circuit of electric vehicles, and can also be used for PV systems and fuel cells, etc.

References

1. R.D. Middlebrook, Transformerless DC-to-DC converters with large conversion ratios. IEEE Trans. Power Electron. **3**(4), 484–488 (1988)
2. D. Zhou, A. Pietkiewicz, S. Cuk, A three-switch high-voltage converter in power electronics. IEEE Trans. **14**, 177–183 (1999)
3. B. Axelrod, Y. Berkovich, A. Ioinovici, Switched-capacitor/switched-inductor structures for getting transformerless hybrid DC–DC PWM converters. IEEE Trans. Circuits Syst. I Regul. Pap., **55**(2), 687–696 (2008)
4. J.C. Rosas-Caro, J.M. Ramirez, F.Z. Peng, A. Valderrabano, A DC–DC multilevel boost converter. IET Power Electron. **3**(1), 129 (2010)
5. M. Mousa, M.E. Ahmed, M. Orabi, New converter circuitry for high voltage applications using switched inductor multilevel converter, in *2011 IEEE 33rd International Telecommunications Energy Conference (INTELEC)* (2011)
6. M. Kasper, D. Bortis, J.W. Kolar, Classification and comparative evaluation of PV panel-integrated DC–DC converter concept. IEEE Trans. Power Electron. **29**(5), 2511–2526 (2014)
7. M. Bella, F. Prieto, S.P. Litrán, J. Manuel, E. Gómez, Combining single-switch nonisolated DC–DC converters for single-input, multiple-output applications. IEEE Ind. Electron. Mag. **10**(2), 6–20 (2016)
8. M. Mousa, M. Hilmy, M. Ahmed, M. Orabi, A.A. El-Koussi, Optimum design for multilevel boost converter, in *Fourteenth International Middle East Power Systems Conference (MEPCON)* (2010), pp. 234–239
9. J.C. Mayo-Maldonado, R. Salas-Cabrera, J.C. Rosas-Caro, J.De Leon-Morales, E.N. Salas-Cabrera, Modelling and control of a DC–DC multilevel boost converter. IET Power Electron. **4**(6), 693–700 (2011)
10. T.-S. Hwang, M.J. Tarca, S.-Y. Park, Dynamic response analysis of DC–DC converter with supercapacitor for direct borohydride fuel cell power conditioning system. IEEE Trans. Power Electron. **27**(8), 3605–3615 (2012)
11. J. Huusari, T. Suntio, Origin of cross-coupling effects in distributed DC–DC converters in photovoltaic applications. IEEE Trans. Power Electron. **28**(10), 4625–4635 (2013)
12. A. Diab-Marzouk, O. Trescases, SiC-based bidirectional Ćuk converter with differential power processing and MPPT for a solar powered aircraft. IEEE Trans. Transp. Electr. **1**(4), 369–381 (2015)
13. Y. Chen, S. Zhao, Z. Li, X. Wei, Y. Kang, Modeling and control of the isolated DC–DC modular multilevel converter for electric ship medium voltage direct current (MVDC) power system. IEEE J. Emerg. Sel. Top. Power Electron. **99**, 124–139 (2016)
14. A. Villarruel-Parra, A.J. Forsyth, Enhanced Average-value modeling of interleaved DC–DC converters using sampler decomposition. IEEE Trans. Power Electron. **32**(3), 2290–2299 (2017)
15. J.R. Rahul, A. Kirubakaran, D. Vijayakumar, A new multilevel DC–DC boost converter for fuel cell based power system, in *IEEE Students' Conference on Electrical, Electronics and Computer Science* (Mar 2012), pp. 1–5

16. S. Rezaee, E. Farjah, A DC–DC multiport module for integrating plug-in electric vehicles in a parking lot: topology and operation. IEEE Trans. Power Electron. **29**(11), 5688–5695 (2014)
17. B. Mangu, S. Akshatha, D. Suryanarayana, B.G. Fernandes, Grid-connected PV-wind-battery-based multi-input transformer-coupled bidirectional DC–DC converter for household applications. IEEE J. Emerg. Sel. Top. Power Electron. **4**(3), 1086–1095 (2016)

Improved Image Quality of Hybrid Underwater Image Restoration (IR) Using the Dark Channel Prior (DCP), Color Attribution Prior, and Contrast Stretching



Anuradha Vashishtha and Jamvant Singh Kumare

1 Introduction

UW vision is a standout among the most major parts in marine scientific research and ocean engineering, for example, UW imaging technology that causes a subsea investigation to think about marine biology and inspect geological conditions. In addition, the Autonomous Underwater Vehicle (AUV) depends on vision technology so that they can be controlled in a complicated situation. However, light attenuation poses a threat to high-quality underwater mages/video, which promotes slowing down like ambient environment for the UW imaging framework, and prevents most computer vision applications in the ocean conditions. The light attenuation is brought by the dissolution and ingestion Due to the presence of residue like particles gliding around in the water; UW pictures dependably experience the ill effects of the dissipating impact. As shown in Fig. 1, the light reflected from the surface of the object goes to the camera, and when light transmits light with suspended particles in the imaging medium, the dispersion is shaped. It is divided into backscattering and forward decoding. Backscattering occurs when the brightness involves the source; the light spreads in the line of sight (LOS) and receives the picture plane in the long run. Air light enters from air to the water, part of the light attenuates along the line of sight [1] which is shown in Fig. 1. The point x represents the nearest scene point to the camera in Fig. 1.

This causes clouds like UW for water and brings incredible visual contrast. Forward scattering is visible when a part of the light reflected with little edges in relation

A. Vashishtha (✉) · J. S. Kumare
Department of Computer Science & Engineering and Information Technology, Madhav Institute of Technology and Science, Gwalior, M.P, India
e-mail: Er.anuradhavashishtha30@gmail.com

J. S. Kumare
e-mail: jamvantsingh@mitsgwalior.in

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_109

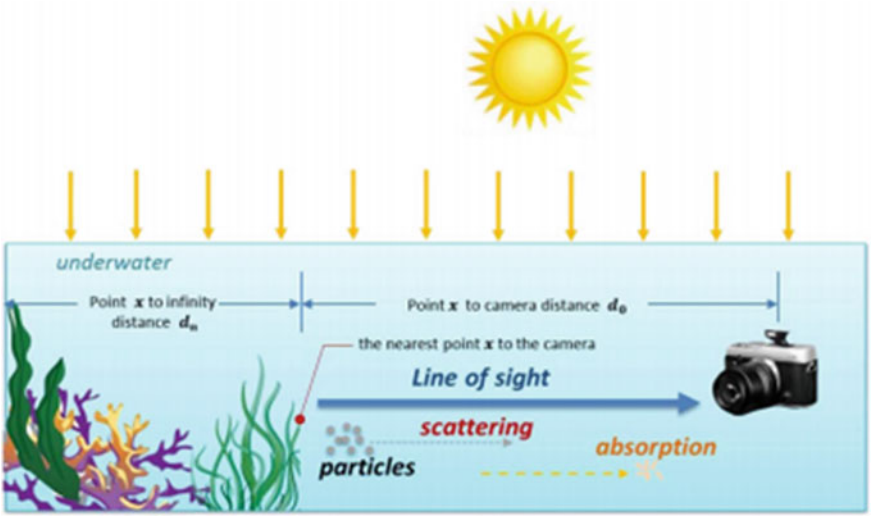


Fig. 1 Underwater scene of light propagation [1]

to the LOS spreads, which indicates the blurring of the picture. As the light gets penetrated in the ocean [1] which is shown in Fig. 2.

This causes clouds like UW for water and brings incredible visual contrast. Forward scattering is visible when a part of the light reflected with small angles in relation to the LOS spreads, which indicates the blurring of the picture.

Apart from this, as can be found in Fig. 2, the rate of light absorption in water varies with their wavelength so that the extraordinary shade of light will disappear slightly with the abundance of water. The water in most part absorbs red light due to the longest wavelength. On the other hand, the little wavelength is in blue shading; this can enter the longest separation of water in the water, so the UW pictures exhibit

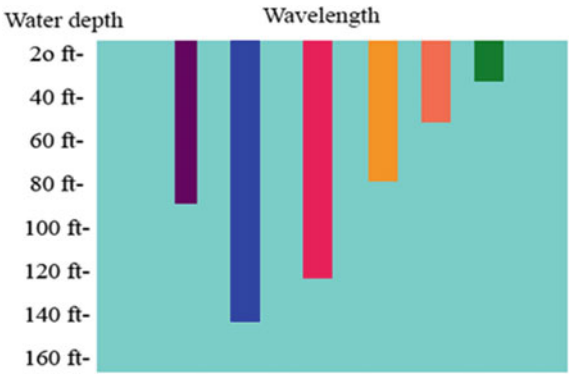


Fig. 2 Penetration of light with different wavelengths in the ocean [1]

some degree of blue or green tone. In this way, less complexity and shading mutations are two important issues that we need to handle [1].

A camera wavelength experiences shading divergence due to subordinate light absorption. In all, the red light starts separating from orange light, yellow light, purple light, yellow-green light, green light, and objects trapped by blue light. The principle behind this is the inspiration for why most of the UW pictures are ruled by blue or green tones. To address this issue, the classification of strategies has recently been proposed. Existing techniques can be sorted out into one of four general classifications: single UW picture enhancement strategy, single UW image restoration strategy, deep learning-based technique, and extra data strategy [2].

Further, this paper is sorted out as pursues. Section 2 presents various techniques and details of the proposed scheme, Sect. 3 presents a literature survey of the previous scheme, Sect. 4 presents proposed work, and Sect. 5 presents experiment result analysis and conclusions of the study are presented in Sect. 6.

2 Different Methods

In this paper, we have used three techniques which are described below.

2.1 Dark Channel Prior (DCP)

DCP depends on the insights of haze-free pictures of open air. In the vast majority of the non-sky patch, some pixels like some color channels (RGB) are an exceptionally low intensity on some pixels (called dark pixels). These dark pixels give estimates of haze transmission. This methodology is physically sufficient and works admirably in a dense haze. At the point when the visual objects are like air light, it is invalid. DCP depends on the assumption on outdoor haze-free images: in a large part of the non-sky patch, some pixels have low intensity, like at least one color channels. DCP is usually a wonderful first for single picture dehazing, yet it is invalid when there are objects that are comparable to the light of atmospheric light. The exact transmission factor for this situation is difficult to detect, and with these lines, the results show some shading rotations. In the dark channel, the low intensity is for the most part due to three components: shadows, colorful objects, and dark objects. After all, it uses a measurable perception of haze-free pictures, which is named "DCP" for estimating the depth of the scene [3].

Principle commitments of DCP as pursues:

(1) Separating a picture into diffuse and specular parts is a poorly presented issue because of lack of perceptions.

(2) The watched color of a picture is shaped from the spectral energy distributions of the light reflected by the surface reflectance, and the intensity of the color is dictated by the imaging geometry.
(3) The dark channel is taken from the lowest intensity an incentive among RGB channels at every pixel [4].

Correction of dark channel estimations
The DCP is straightforward yet amazing for single picture haze removal. While it is a sort of picture insights, it may not work for some specific pictures. What is more, the technique is handling each there diverts similarly. As the RGB color space is not intended to rough human vision, the progressions of its parts cannot reflect the impacts of human vision straightforwardly. At the point when the shadow of the visual objects is near the white and there is no dark channel area like a shadow or spectacular shading, then the intensity of the dark channel will generally be higher. Let's prior treat minorly shaded articles as cloud layers, which think a lot about the transmission map. The recovered picture will be oversized in these territories. As indicated by human visual qualities, individuals are increasingly touchy to changes in close white color. At the point when these articles have some slightly bright color, the oversaturation of the outcome is noticeable. This deformity of the prior may not be disregarded [5].

2.2 Color Attenuation Prior

Color attenuation prior is based on the difference between the brightness and the saturation of the pixels within the hazy image. Repairs transmission map and restores visibility. Color attenuation prior is based on the difference between the brightness and the saturation of the pixels within the hazy image. By creating a linear model for modeling the scene depth of the hazy image under this novel prior and learning the parameters of the model by using a supervised learning method, the depth information can be well recovered The need of feature selection is to select most sensitive features which make changes in image quality. The main goal is haze removal by depth map estimation.

The advantages of this prior are as follows:

- This simple and powerful prior can help to create a linear model for the scene depth of the hazy image.
- The bridge between the hazy image and its corresponding depth map is built selectively.
- With the recovered depth information it is possible to remove the haze from the single hazy image easily [6].

2.3 Contrast Stretching (CS)

CS technique is utilized to build up a picture by developing out the power esteems it contains to utilize full conceivable qualities in the range. CS is widely used for making pictures and there is a preprocessing phase in applications such as discourse receipts, construction surfaces, and some other video or picture handling applications. On this occasion, due to low light conditions, a picture of less complexity is acquired, if the dynamic cluster of the camera sensor is not correct, then the results of the CS process will be in a great picture [7]. CS system is one of the pictures upgrade procedures that are generally investigated to conquer the poor differentiation issue on intense leukemia slide pictures. Fractional Contrast Stretching is a picture improvement procedure that utilizes direct mapping capacity to upgrade the differentiation level and the splendor dimension of the picture [8].

Contrast stretching is a technique used to stretch the histogram of an image so that the full dynamic range of the image is filled. The minimum pixel value is 81, and the maximum pixel value is 127. Increasing the dynamic range entails "stretching" the histogram by applying a linear scaling function that maps pixel values of 81–0, maps pixel values of 127–255, and scales pixel intensities that lie within the range [82–126] accordingly. The contrast stretch algorithm essentially pulls the boundaries of the original histogram function to the extremes. Contrast Stretch is the simplest contrast stretch algorithm that stretches the pixel values of a low-contrast image or high-contrast image by extending the dynamic range across the whole image spectrum.

Algorithm for contrast stretching

- First, we browse an input image.
- Calculate the minimum value of the pixel.
- Then calculate the maximum value of the pixel.
- The slope of line joining point (0, 255) to (r min, r max).
- Repeat the process for Red, Green, and blue components.
- Concentrate all three components to find the final contrast image.

3 Literature Survey

Sun et al. (2018) Turbid UW condition presents extraordinary challenges for the utilization of vision advancements. One of the greatest difficulties is the entangled clamor conveyance of the UW pictures because of the genuine dispersing and ingestion. To minimize this problem, this task proposes a deep pixel-to-pixel system that displays the encoding–decoding system by configuring UW for image correction. It uses the convolution layers as encoding to filter the noise while using the devolutionary layers so that the missing subtleties can be re-examined and pixel helps to refine the picture [9].

Zhou et al. (2018) UW pictures ordinarily experience the ill effects of low visibility and serious colorcast because of dissipating and assimilation. In this letter, a novel strategy is proposed to deal with the dissipating and retention issues of light with various wavelengths dependent on the shading line show. We sift through picture fixes that display the qualities of the shading line earlier and recoup the color line of the patches. At that point, the neighborhood transmission for each fix is assessed depending on the counterbalances of the shading lines along the foundation light vector from the beginning. Trial results are exhibited to demonstrate that the proposed technique can create high caliber UW pictures with generally certified color, normal appearance, and improved contrast and visibility [10].

Peng et al. (2017) UW pictures regularly experience shading and the effects of low contrast, because when light passes through the water, the light is destroyed and consumed. We propose an abundant assessment technique for photo-blurred and light retention-based UW visuals, which can be used to re-install and improve the images immersed in the image creation model (IFM). The previous IFM-based IM technology gains are dependent on DCP or the biggest intensity of the past. These are often incorrect due to light conditions in UW paintings, which reflect the result of bad restoration [11].

Neha et al. (2017) Haze phenomenon degrades the permeability of a scene in the picture. DCP is one of the strategies used widely for expelling cloudiness. The techniques conveyed on this model evaluate the transmission map utilizing a constant transmission parameter. In this paper, we watch the yield acquired on shifting consistent transmission parameter for various channels. Results demonstrate that the expansion inconsistent transmission parameter prompts increment in the complexity of the scene. We additionally infer that the picture quality increments and commotion diminishes as the fix estimate increments in the DCP [12].

Li et al. (2016) Images caught UW are normally degraded because of the impacts of retention and dispersing. Degraded UW pictures demonstrate a few confinements when they are utilized for showcase and examination. To defeat those restrictions, an efficient submerged picture improvement technique which incorporates a UW picture dehazing algorithms and a contrast enhancement algorithm are proposed. The proposed technique can yield two variants of upgraded yield. One form with generally real shading and normal appearance is appropriate for presentation. The other rendition with high contrast and brightness can be utilized for removing increasingly important data and revealing more subtleties [13].

Zhang et al. (2016) In this paper, we propose an improved heterogeneous climate light estimation strategy and a novel depth estimation algorithm with Color Attenuation Prior (CAP) to a dehaze single picture. Right off the bat, it evaluates the environment light with mean pooling on the illuminance segment from HSV shading space. Furthermore, the scene depth is assessed by a nonlinear CAP show which can defeat the deformities of the event of negative scene profundities from the direct CAP display. Exploratory outcomes exhibit that the proposed algorithms outflank the best in class strategies in dehazing pictures [14].

4 Problem Statement and Propose a Methodology

4.1 Problem Statement

1. Air light estimation is poor.
2. It produces some Halo effects on the resultant images.
3. This technique is invalid when the scene object is similar to Air light like vehicle headlights, blanketed ground, and so forth.

4.2 Proposed Methodology

In this proposal, we have used CS, which is a simple image enhancement technique that attempts to improve the difference in a picture, in which there is a limit to the intensity values, in which the desired values are limited, for example, Full range of pixels for It assumes that the picture type allows related. This more sophisticated histogram is different from the equivalent in which this image can apply a linear scaling function in pixel esteem. After this "upgrade" is less inefficiency. Before stretching can be done, it is important to set the upper and lower pixel esteem restraint on which the image should be standardized. Regularly, these cutoff points will be the bus base and the most extreme pixel esteem, which are the type of permits related to the picture type.

First, we browse an original image from the dataset and then by reducing haze of original underwater image by applying transmission estimation in the haze-free image from the dataset. Then the image is measured by Salience weighed map on the transmission image using color attenuation prior to this salience image and also by applying Dark channel prior on the CAP image. NPEA is applied to improve DCP image and then we calculate UCIQE and PCQI value of the original image.

4.3 Propose Algorithm

Step. 1. First, we browse the underwater image from the dataset.
Step. 2. Then we reduce the haze of the original underwater image.
Step. 3. Use transmission estimation on this haze-free image.
Step. 4. Salience weighted map on this transmission image.
Step. 5 Apply color attenuation prior (CAP) on this salience image.
Step. 6. Dark channel prior on this CAP image.
Step. 7. Apply contrast stretching on DCP image.
Step. 8. Performance measurement of UCIQE and PCQI value.

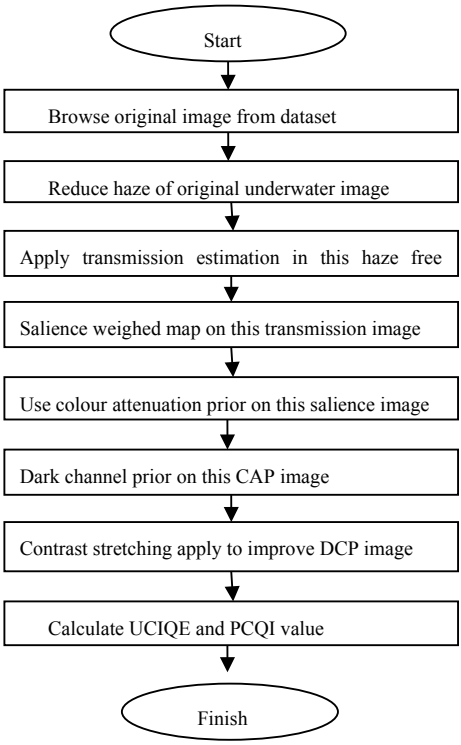


Fig. 3 Flow diagram of the proposed methodology

Flow Diagram

Flow diagram of the methodology adopted is shown in Fig. 3.

5 Experiment Result Analysis

In our experiment work UWIR technique based on DCP, CAP, and contrast stretching. First, take the original picture and decreasing its haze quality as it can improve the permeability and clarity by computerized picture handling. The average value of the brightest 1% in a haze picture is determined for climatic light estimation in transmission estimation. Salience weighted map-based visual attention to distinguishing obvious items in jumbled visual situations then we put our methods. The test examination is utilized UW picture rebuilding pictures for execution assessment. It takes shading pictures for assessment. To quantitatively think about the aftereffects of the previously mentioned techniques, we apply UW color picture quality evaluation (UCIQE) and the patch-based contrast quality index (PCQI) is connected to assess the complexity varieties. The calculation is structured on MATLAB2018(a) utilizing the image processing tool compartment. In this execution, this calculation is contrasted and distinctive calculations. As we saw in the test result. The output of

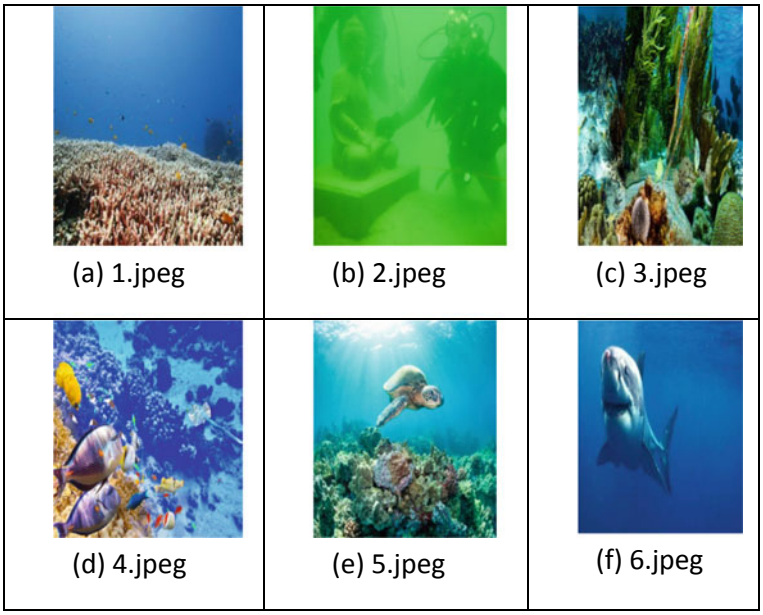


Fig. 4 Dataset images

all the above-mentioned techniques is compared on the basis of their corresponding UCIQE and PCQI values and the following table shows the output.

Images of the dataset are shown in Fig. 4, which are further used for applying various restoration methodologies.

The Graphical User Interface of the adopted methodology is shown in Fig. 5.

Various steps of the methodology adopted are applied to Image dataset which is shown in Fig. 6.

Figure 7. image (a) 2.jpeg, image (b) 5.jpeg, image (c) 3.jpeg, image (d) 4.jpeg, and image (e) 6.jpeg, has been used in research work which shows the different techniques used in the proposed methodology (Table 1, Figs. 8, 9).

6 Conclusion

In this paper, we have proposed a UWIR technique dependent on DCP, CAP, and CS utilizing another UW image formulation model. We additionally infer that the picture quality increments and decreases noise as the patch size increases in the DCP. Applying the prior to the haze imaging model, haze can be successfully removed. This extends the range of gray levels (contrast) that are close and packs the range of gray levels that are close to the histogram minima. For most pictures, the contrast is normally extended for most pixels, improving many picture features.



Fig. 5 Execution steps of methodology used

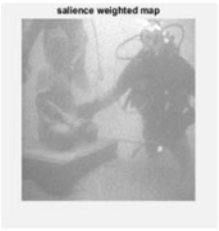
For future work, we expect to address the previously mentioned issues. Furthermore, we will build up a UW picture database for the advancement of UW picture and video investigation. Additionally, other enhancement strategies and learning procedures can be utilized to improve the preparing impact of the proposed techniques.



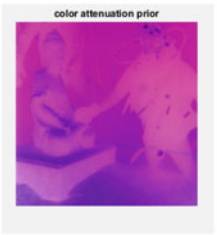
(a) Original image



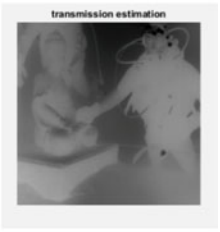
(b) reduce haze



(c) SWM



(d) CAP



(e) Transmission estimation



(f) DCP



(g) Contrast stretching

Fig. 6 Qualitative comparison of the image shown in Fig. 1. **a** Original underwater image **b** result of reducing haze **c** result of salience weighted map **d** result of CAP **e** result of transmission estimation **f** result of DCP **g** proposed result of contrast stretching

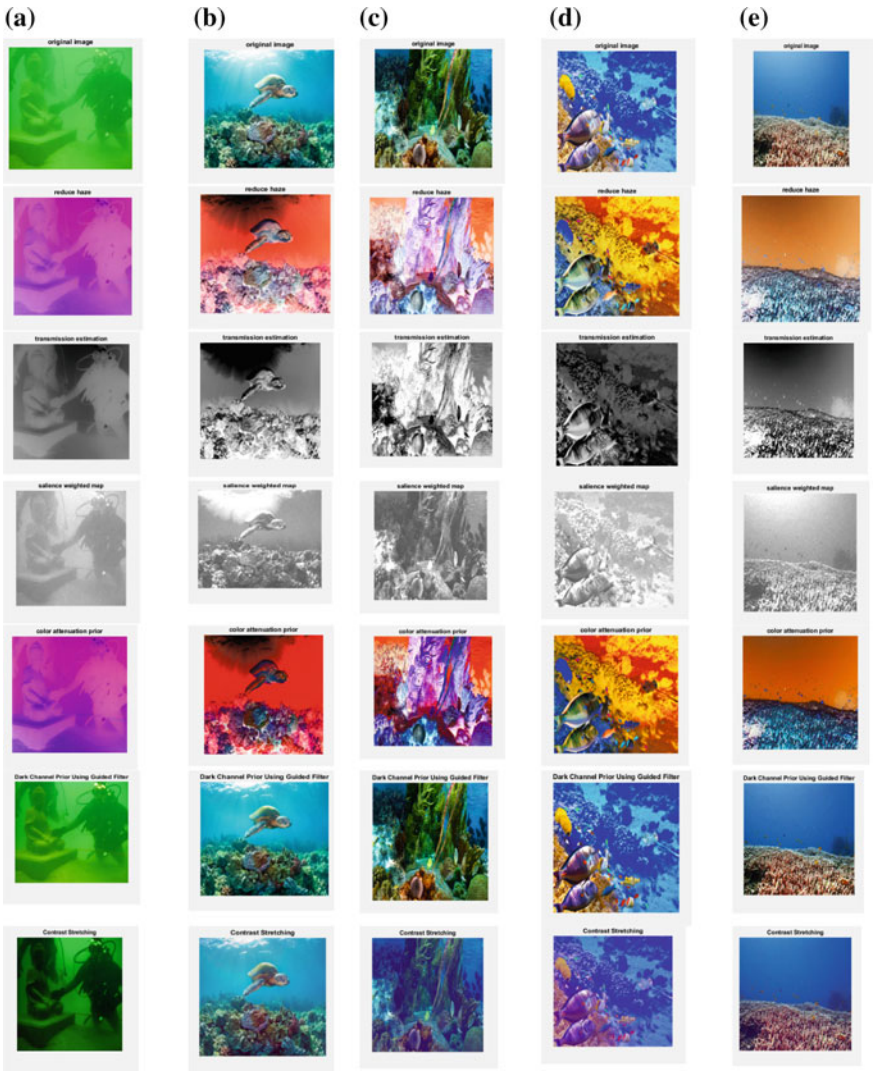


Fig. 7 The visual output images after applying proposed techniques on dataset

Table 1 Comparison of base [2] and propose of UCIQE and PCQI

Images	1.jpg		2.jpg		3.jpg		4.jpg		5.jpg	
Matrix	Base [2]	Proposed	Base [2]	Proposed	Base [2]	Proposed	Base [2]	Proposed	Base [2]	Proposed
UCIQE	1.4840e+03	1.4603e+06	1.1686e+03	5.5389e+04	1.4443e+03	3.3973e+05	1.4879e+03	2.4732e+06	1.1444e+03	4.6603e+05
PCQI	0.7239	0.9734	0.9068	0.9197	0.8263	0.9277	0.8311	0.9031	0.8359	0.8809

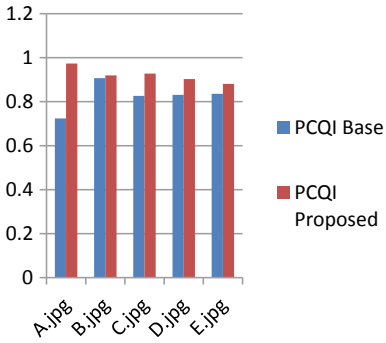


Fig. 8 Graph comparison of base and propose PCQI

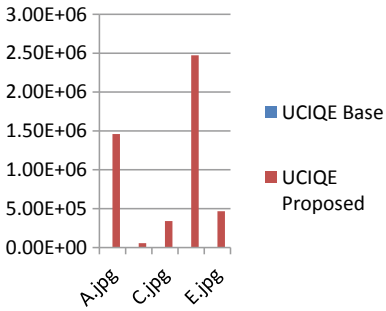


Fig. 9 Graph comparison of base and propose UCIQE

References

1. Y. Wang, H. Liu, L.P. Chau, Single underwater image restoration using adaptive attenuation-curve prior. IEEE Trans. Circuits Syst. i: Regul. Pap. 1549–8328 © IEEE 2017
2. M. Zhang, J. Peng, Underwater image restoration based on a new underwater image formation model, https://doi.org/10.1109/access.2018.2875344, IEEE Access https://doi.org/10.1109/access.2875344, IEEE Access (2018)
3. T.H. Kil, S.H. Lee, N.I. Cho, Single image dehazing based on reliability map of dark channel prior 978-1-4799-2341-0/13/$31.00 ©2013 IEEE 882 ICIP (2013)
4. R. Sathya, M. Bharathi, G. Dhivyasri, Underwater image enhancement by dark channel prior, in *IEEE Sponsored a 2nd International Conference on Electronics and Communication System (ICECS 2015)* 978-1-4788-7225-8/15/$31.00 © IEEE 1119.2015
5. F. Liu, C. Yang, *A Fast Method for Single Image Dehazing Using Dark Channel Prior*. 978-1-4799-5274-8/14/$31.00 © IEEE 2014
6. P.V. Angitha, K.A. Santhini, Improved colour attenuation prior based dehazing by edge attenuation method, in *International Conference on Electronics, Communication and Aerospace Technology ICECA* (2017)
7. B.H. Ramyashree, R. Vidhya, D.K. Manu, FPGA implementation of contrast stretching for image enhancement using system generator (Dept. of ECE KSSEM, Bangalore, India, 2015)
8. L.B. Toh, M.Y. Mashor, P. Ehkan, H. Rosline, A.K. Junoh, N.H. Harun, Implementation of high dynamic range rendering on acute leukemia slide images using contrast stretching. 978-1-5090-2160-4/16/$31.00 © IEEE 2016
9. X. Sun, L. Liu, Q. Li, J. Dong, E. Lima, R. Yin, Deep pixel-to-pixel network for underwater image enhancement and restoration (IEEE, 2018), https://doi.org/10.1049/iet-ipr.2018.5237

10. Y. Zhou, Q. Wu, K. Yan, L. Feng, W. Xiang, Underwater image restoration using color-line model, 1051–8215 (c) IEEE 2018
11. Y.T. Peng, P.C. Cosman, Underwater image restoration based on image blurriness and light absorption, IEEE Trans. Image Process. **26**(4), 1579 (2017)
12. Neha, R.K. Aggarwal, Effect of various model parameters on fog removal using dark channel prior, in *2nd IEEE International Conference on Recent Trends in Electronics Information & Communication Technology (RTEICT)* (19–20 May 2017, India) 978-1-5090-3704-9/17/$31.00 © 2017 IEEE 1764 2017
13. C. Li, J. Guo, R. Cong, Y. Pang, B. Wang, Underwater image enhancement by dehazing with minimum information loss and histogram distribution prior. IEEE Trans. Image Process. IEEE Trans. Image Process. (2016). https://doi.org/10.1109/TIP.2016.2612882
14. S. Zhang, C. Qing, X. Xu, J. Jin, H. Qin, Dehazing with improved heterogeneous atmosphere light estimation and a nonlinear color attenuation prior model. 978-1-5090-2526-8/16/$31.00 © IEEE 2016
15. S. Kaur, M. Bansal, A.K. Bathla, Dehazing image using analytical model and color attenuation prior, in *2016 2nd International Conference on Next Generation Computing Technologies (NGCT-2016)* (Dehradun, India, 14–16 Oct 2016)
16. Q. Zhu, J. Mai, L. Shao, A fast single image haze removal algorithm using color attenuation prior, in *IEEE Transactions on Image Processing IEEE Transactions on Image Processing* (2015), https://doi.org/10.1109/tip..2446191
17. S.S. Negi, Y.S. Bhandari, A hybrid approach to image enhancement using contrast stretching on image sharpening and the analysis of various cases arising using histogram, in *IEEE International Conference on Recent Advances and Innovations in Engineering (ICRAIE-2014)* (Jaipur, India, 09–11 May 2014)
18. P. Jagatheeswari, S. Suresh Kumar, M. Rajaram, Contrast stretching recursively separated histogram equalization for brightness preservation and contrast enhancement, 978-0-7695-3915-7/09 $26.00 © IEEE, https://doi.org/10.1109/act.2009.37 (2009)

Design of Low Voltage High-Speed Universal Logic Gates Using Different Models of CMOS Schmitt Trigger



Bhavika Khanna, Raghav Gupta, Cherry Bhargav and Harpreet Singh Bedi

1 Introduction

Schmitt trigger is a well-known digital circuit which is used for the shaping of input pulses, i.e., it is used for the conversion of any continuously time-varying signal to a stable state signal which in other words means that it is a wave-shaping circuit used for the conversion of analog signal into digital signal [1]. The name of Schmitt trigger itself suggests that it will cause the triggering of the output when the input signal shows a sufficient change. The reason for popularity of Schmitt trigger is the hysteresis property that can be analyzed from the DC transfer characteristics of Schmitt trigger, which shows that it has got two different working threshold values for the rising and the falling input signal [2]. Due to the presence of two different working threshold values, the Schmitt trigger will show the switching in the output when input crosses the upper and lower threshold triggering points, respectively. The Schmitt trigger is practically a comparator circuit, which has got positive feedback with it for the enhancement of the static noise margin of system. The feedback of the Schmitt trigger is for the purpose of facilitating two different working threshold values, which can be used for the improvement of sensitivity of the devices from noise. Schmitt trigger works as signal restoring circuit for the extraction of the original information from the input signal by the elimination of the content of noise present in the original signal [3]. The Schmitt trigger is basically used for providing a worthy solution for the improvement of sensitivity of device from various kinds of interference present in the system. The main application of Schmitt trigger is that it is extensively used in input buffers for increasing noise immunity. The best approach used for the improvement

B. Khanna (✉) · R. Gupta (✉) · C. Bhargav · H. S. Bedi
School of Electronics and Communication Engineering, Lovely Professional University, Phagwara, Punjab, India
e-mail: bhavikak54@gmail.com

R. Gupta
e-mail: raghav.22026@lpu.co.in

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_110

of Static Noise margin of a logic gate is by implementing the logic gates using the Schmitt trigger concept. As the gates implemented using the Schmitt trigger concept are having positive feedback in them which is the main factor used for improving the immunity making them insensitive to any kind of unwanted variations. As we know that Universal gates are the most important component of a digital system, as these are the gates which help to implement a digital design using only one kind of gate, i.e., NAND/NOR gate thereby making the system symmetric. We have used different techniques for the implementation of Schmitt trigger design, and using those versions of Schmitt trigger three designs for NAND gate and NOR gate has been implemented. The paper is organized in a manner where Sect. 2 represents the implemented designs and their description, Sect. 3 represents the results and the analysis obtained after simulation and finally, the conclusion and future objective have been presented in the last section of the paper.

2 Implementation and Circuit Description

The Schematic of a NAND gate using six transistors (6T) model of Schmitt trigger consists of two sections and two output transistors [4], which are connected as shown in Fig. 1. The conventional Schmitt trigger-based NAND circuit is basically composed of two subsections, Sects. 1 and 2, where Sect. 1 represents the circuit of PMOS, which is composed of five PMOS transistors and Sect. 2 represents the

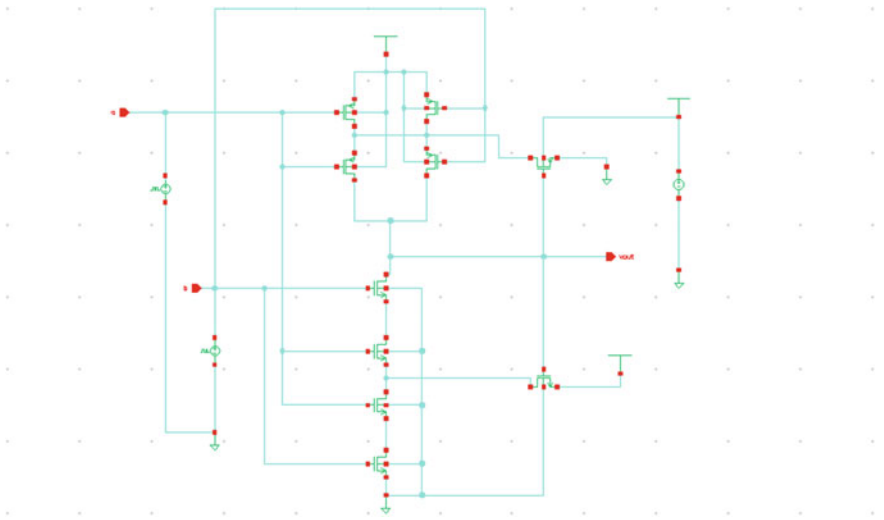


Fig. 1 NAND gate design using 6T Schmitt trigger design

circuitry of NMOS which is composed of five NMOS transistors that are responsible for the generation of the lower threshold switching voltage and upper threshold switching voltage VH.

The devices of PMOS and the NMOS at the output are connected for providing the feedback for elimination of noise signal from the output when the changes in the amplitude of input signal are not more than the switching threshold voltage difference. The conventional CMOS NAND gate uses ten transistors for providing the hysteresis property which in turn is used for the enhancement of the noise immunity of the system but when analyzed for various design parameters like delay and power the results obtained were not very much appropriate. So, a new design for NAND gate has been implemented for improving the design parameters of the system, which is shown in the figure below. The modified NAND gate using four transistors (4T) based Schmitt trigger design is composed of seven transistors PM0, PM1, NM0, NM1, NM2, NM3, and NM4 out of which the transistor NM4 is used for providing the positive feedback. As this modified circuit has got less number of transistors in P-section so the overall delay and power dissipation of circuit has been reduced when compared to the traditional CMOS Schmitt trigger design (Fig. 2).

The designs of NAND gate shown above are not much suited for applications requiring more immunity in terms of noise and high-speed operations so a new design using the Schmitt trigger implemented using body-biasing technique has been presented as shown in Fig. 3. The circuit implemented below is composed of two subsections, i.e., Sects. 1 and 2 where Sect. 1 is composed of four transistors PM0, PM1, NM0, and NM1 and Sect. 2 is composed of three inverter circuits (PM2, NM2,

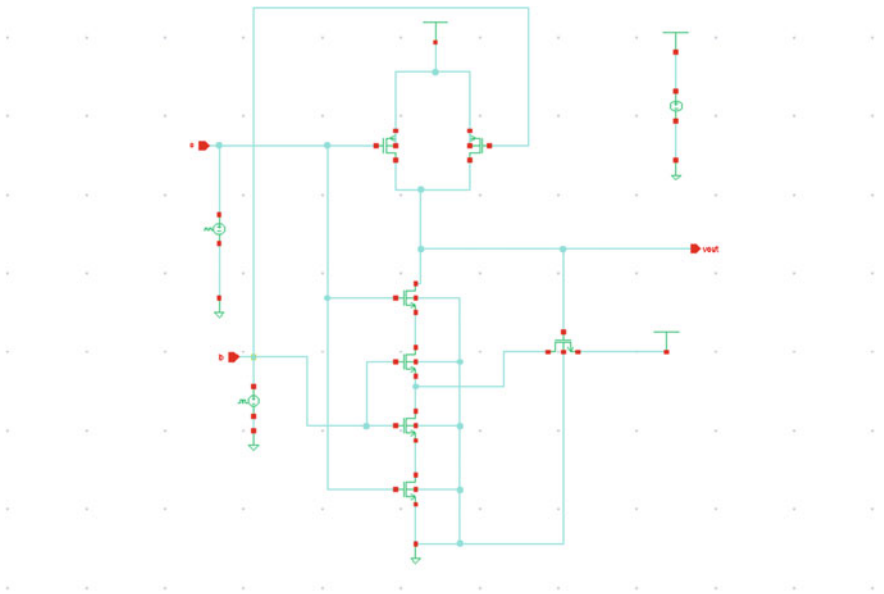


Fig. 2 Modified NAND gate design using 4T Schmitt trigger version

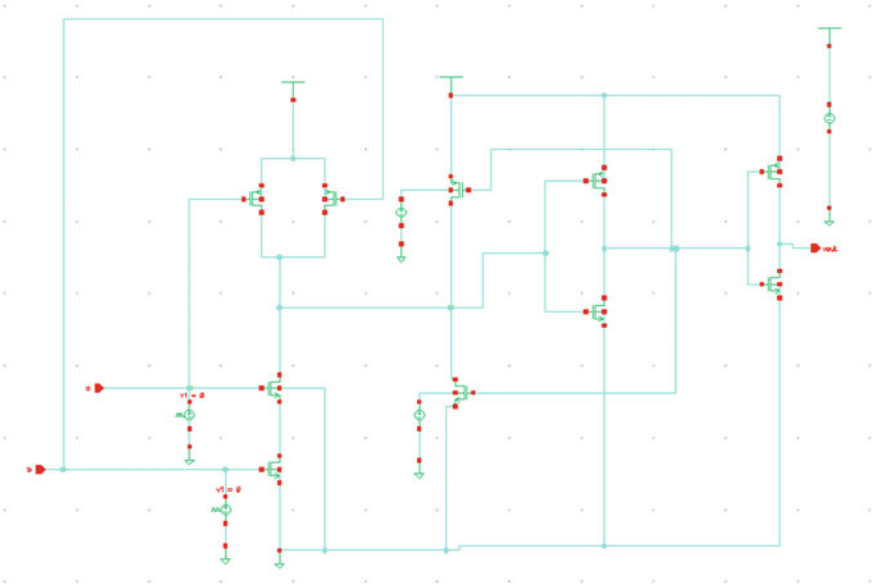


Fig. 3 NAND gate design using body-biasing-based Schmitt trigger version

PM3, NM3, PM4, and NM4) in which one inverter circuit's body is controlled using body-biasing technique for adjusting the threshold of the circuit and controlling the switching of the output in order to minimize the propagation delay of the circuit. Also by controlling the voltages at the body terminal of the transistors PM2 and NM2 the width of the hysteresis loop is adjusted thus making it much suited for the applications requiring the variable noise immunity. It is seen that the body-biasing voltages are used to control the switching action of the output voltage thus change the delay of the circuit. As in this case, the body biasing voltages are chosen in such a way that we get the minimum possible value of the delay. Also by controlling the voltages at the body terminal of the transistors PM2 and NM2 the width of the hysteresis loop is adjusted thus making it much suited for the applications requiring the variable noise immunity.

In a similar way, the designing of NOR gate using different models of Schmitt trigger has been shown in this section. The schematic of NOR gate using conventional model of Schmitt trigger consists of two sections and two output transistors which are connected as shown in Fig. 4.

The conventional six transistors (6T) Schmitt trigger-based NOR gate circuit has got two subsections, Sects. 1 and 2, where Sect. 1 represents the PMOS circuitry consisting of five PMOS transistors, which are responsible for the generation of the lower threshold switching voltage VL and the Sect. 2 represents the NMOS circuitry consisting of five NMOS transistors, which are responsible for the generation of the upper threshold switching voltage VH. The conventional CMOS NOR gate using six Schmitt trigger uses ten transistors for providing the hysteresis property which in

turn is used for the enhancement of noise immunity of the system but when analyzed for various design parameters like delay and power the results obtained were not very much appropriate. So, a new design has been implemented for improving the design parameters of the system, which is shown in Fig. 5. The modified NOR gate using 4T-based Schmitt trigger design is composed of seven transistors PM0, PM1, NM0, NM1, NM2, NM3, and NM4 out of which the transistor NM4 is used for providing the positive feedback. As this modified circuit has got less number of transistors in PMOS-section so the overall delay and power dissipation of circuit has been reduced when compared to the traditional CMOS Schmitt trigger-based NOR gate design.

The design of NOR gate implemented above is not much suited for applications requiring more immunity in terms of noise and high-speed operations so a new design of NOR gate using the Schmitt trigger implemented using body-biasing technique has been presented as shown in Fig. 6.

The implemented circuit is composed of two subsections, i.e., Sects. 1 and 2 where Sect. 1 is composed of four transistors PM0, PM1, NM0, and NM1 and Sect. 2 is composed of three inverters circuits (PM2, NM2, PM3, NM3, PM4, and NM4) in which one inverters circuit's body is controlled using body-biasing technique for adjusting the threshold of the circuit and controlling the switching of the output in order to minimize the propagation delay of the circuit. Also by controlling the voltages at the body terminal of the transistors PM2 and NM2 the width of the hysteresis loop is adjusted thus making it much suited for the applications requiring the variable noise immunity. It is seen that the body-biasing voltages are used to control the switching action of the output voltage thus change the delay of the circuit. As in this case, the body biasing voltages are chosen in such a way that we get the minimum possible value of the delay. Also by controlling the voltages at the body terminal of the transistors PM2 and NM2 the width of the hysteresis loop is adjusted thus making it much suited for the applications requiring the variable noise immunity.

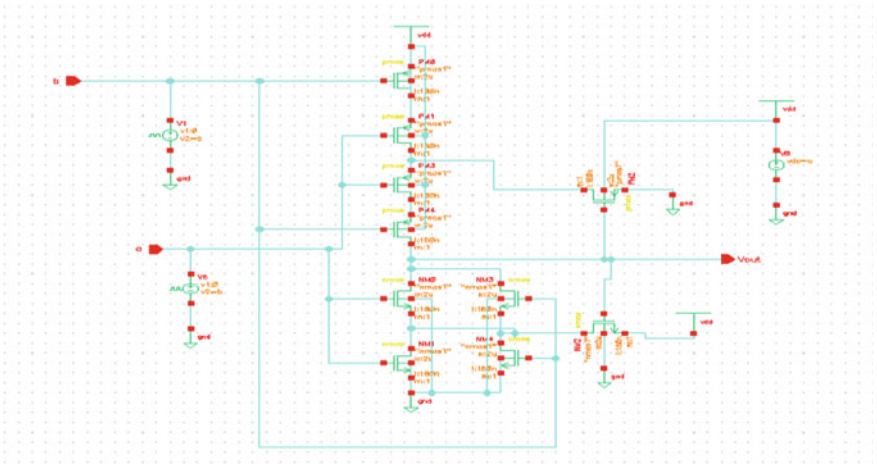


Fig. 4 NOR gate design using 6T Schmitt trigger design

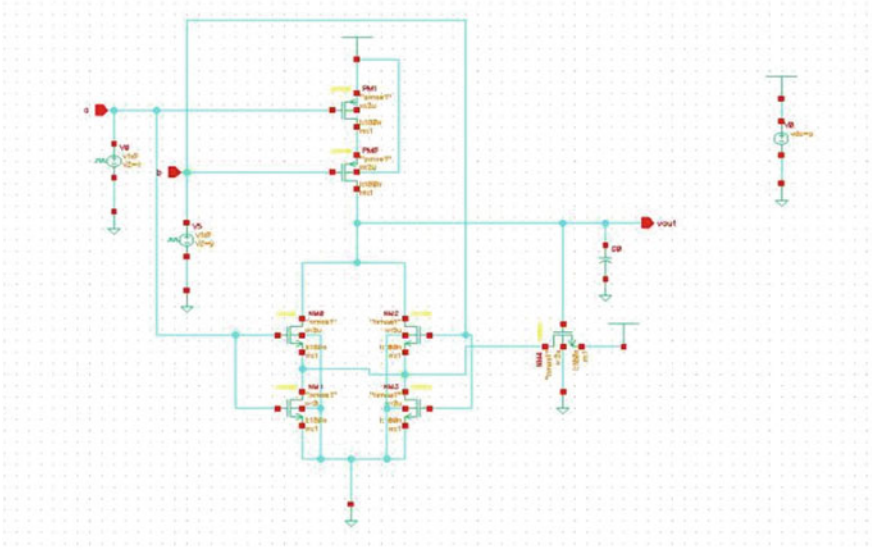


Fig. 5 NOR gate design using 4T-based Schmitt trigger design

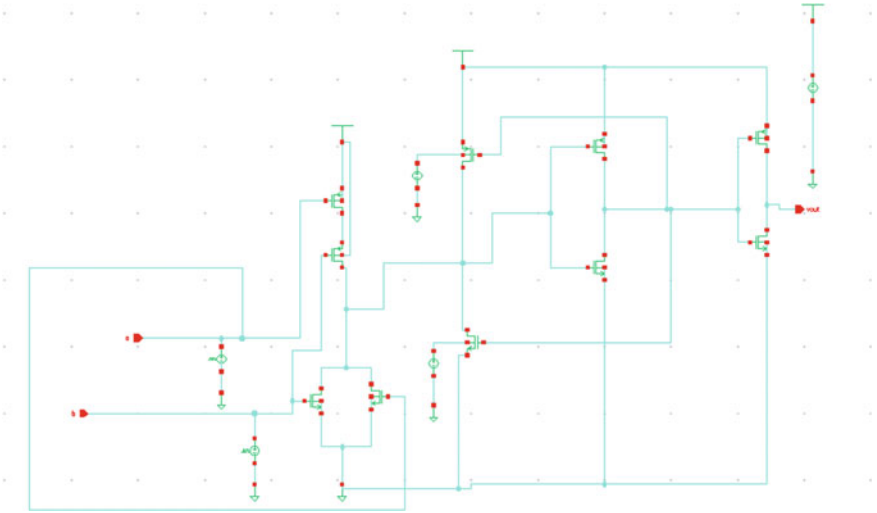


Fig. 6 NOR gate design using body-biasing-based Schmitt trigger design

3 Results and Discussion

The Schmitt trigger is a bistable circuit because of its capability to work on two different threshold values. This property of the Schmitt trigger also facilities the conversion of the analog signal into digital one thus resulting in the reshaping of the pulses. The figures below show the values of two different threshold voltages obtained for different designs of Schmitt trigger. The section below represents the simulation results of implemented design. The figure below shows transient response of NAND gate using 6T model of Schmitt trigger and it can be seen from the response that the output goes low when both the inputs are high otherwise the output remains high (Fig. 7).

The below figure shows the transient response of NAND gate using 4T model of Schmitt trigger and it can be seen from the response that the output goes low when both the inputs are high otherwise it remains high (Fig. 8).

The below figure shows transient response of NAND gate using body biasing model of Schmitt trigger and it can be seen from the above response that the output goes low when both the inputs are high otherwise the output remains high (Fig. 9).

The below figure shows the transient response of NOR gate using 6T model of Schmitt trigger and it can be seen from the response that the output goes high when both the inputs are low otherwise the output remains low (Fig. 10).

The below figure shows the transient response of NOR gate using 4T model of Schmitt trigger and it can be seen from the response that the output goes high when both the inputs are low otherwise it remains low (Fig. 11).

The below figure shows transient response of NOR gate using body-biasing model of Schmitt trigger and it can be seen from the response that the output goes high when both the inputs are low otherwise the output remains low (Fig. 12).

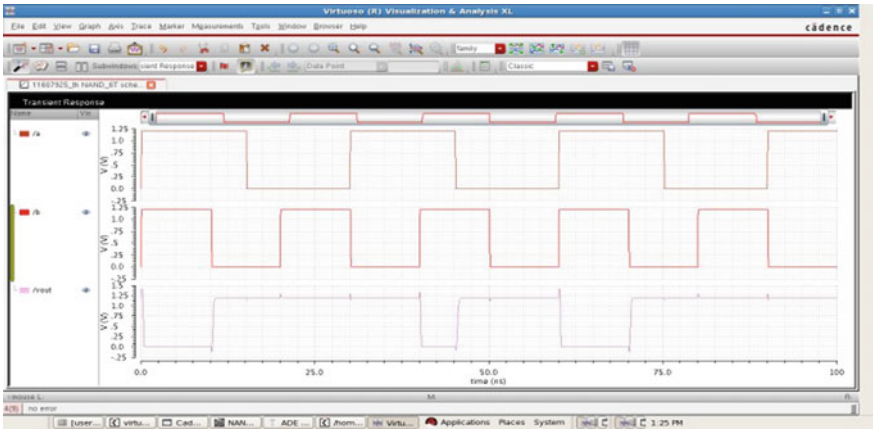


Fig. 7 Transient response for 6T Schmitt trigger-based NAND gate design



Fig. 8 Transient response for 4T Schmitt trigger-based NAND gate



Fig. 9 Transient response for body-biasing Schmitt trigger-based NAND gate

3.1 Analysis of Different Designs of Schmitt Trigger

In this section, the different designs of Universal logic gates have been compared on the basis of different design parameters such as delay, switching power, and so on. For delay analysis, I have added a load capacitance at the output of every kind of design. Different values of load capacitance were taken for the analysis purpose.



Fig. 10 Transient response for 6T Schmitt trigger-based NOR Gate



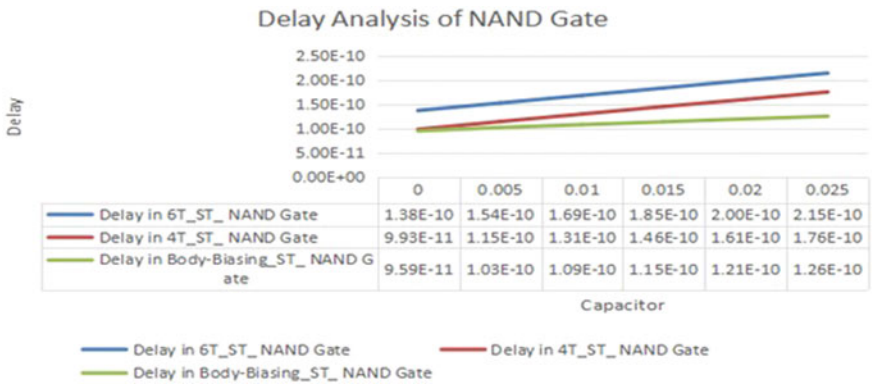
Fig. 11 Transient response for 4T Schmitt trigger-based NOR Gate

3.1.1 Analysis of 6T, 4T and Body-Biased-Based Schmitt Trigger NAND Gate

The figure below represents the delay analysis of 6T, 4T, and body-biased-based Schmitt trigger NAND gate. It has been found that out of all the three considered Schmitt triggers-based design of gates, body-biased circuit has minimum delay and 6T model of Schmitt trigger-based NAND gate is having maximum delay value when analyzed for different capacitive loads where the unit of the capacitance is pf (pico Farad) (Fig. 13).



Fig. 12 Transient response for body-biasing Schmitt trigger-based NOR Gate



	0	0.005	0.01	0.015	0.02	0.025
Delay in 6T_ST_ NAND Gate	1.38E-10	1.54E-10	1.69E-10	1.85E-10	2.00E-10	2.15E-10
Delay in 4T_ST_ NAND Gate	9.93E-11	1.15E-10	1.31E-10	1.46E-10	1.61E-10	1.76E-10
Delay in Body-Biasing_ST_ NAND Gate	9.59E-11	1.03E-10	1.09E-10	1.15E-10	1.21E-10	1.26E-10

Fig. 13 Delay analysis of 6T, 4T, and body-biased-based Schmitt trigger NAND gate

3.1.2 Analysis of 6T, 4T and Body-Bias-Based Schmitt Trigger NOR Gate

Below figure represents delay analysis of 6T, 4T, and body-biased Schmitt trigger-based NOR gate. It has been found that out of all the three considered Schmitt triggers-based NOR gates, body-biased-based circuit has minimum value delay and 6T model of Schmitt trigger-based NOR gate is having maximum delay value when analyzed for different capacitive loads where the unit of capacitance is in pf (Fig. 14).

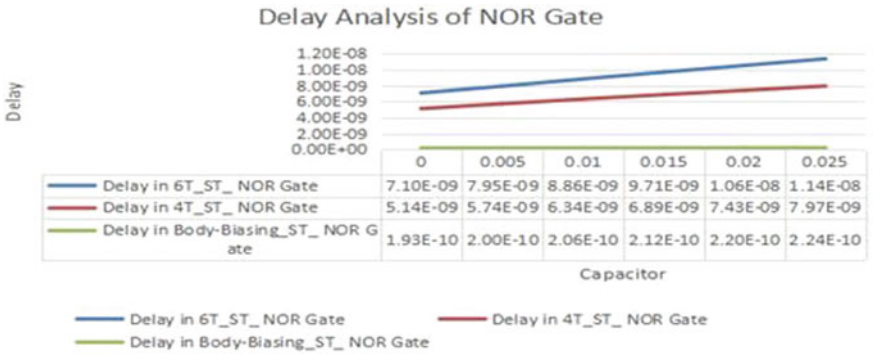


Fig. 14 Delay analysis of 6T, 4T, and body-biased-based models of Schmitt trigger NOR gate

3.1.3 Power Analysis of 6T, 4T, and Body-Bias Schmitt Trigger-Based Universal Logic Gates

Below tables shows the analysis on dynamic power of 6T, 4T, and body-biasing models of Schmitt trigger universal logic gates with different capacitive load at output. It is observed that power loss in body-biasing model is found to max for every case and as load capacitance value is increasing power also increases. 4T-based models have got the minimum amount of power dissipation among the three models. Also in 4T model as capacitive load increases, dynamic power also increases. It has been found that with increase in capacitive load dynamic power is not affected as compared to other models for same value of capacitive load for body-biasing model. Table 1 shows dynamic power analysis for different models of NAND gate at different capacitive loads (in pf).

Where P_D1 represents the dynamic power dissipation of 6T-based model of Schmitt trigger NAND Gate, P_D2 represents the dynamic power dissipation of 4T-based model of Schmitt trigger NAND Gate and P_D3 represents the dynamic power dissipation of body-bias-based model of Schmitt trigger NAND Gate in watts.

In a similar way, Table 2 shows dynamic power analysis for different models of NOR gate at different capacitive loads where the unit of capacitance is in pF.

Table 1 Dynamic power of 6T, 4T and body-bias model-based Schmitt trigger NAND gate

Capacitor	P_D1	P_D2	P_D3
0	8.23E–06	5.99E–06	3.40
0.005	9.05E–06	6.70E–06	3.40
0.01	9.81E–06	7.38E–06	3.40
0.015	1.06E–05	8.07E–06	3.40
0.02	1.13E–05	8.73E–06	3.40
0.025	1.21E–05	9.40E–06	3.40

Table 2 Dynamic power of 6T, 4T and body-bias model-based Schmitt trigger NOR gate

Capacitor	P_D1	P_D2	P_D3
0	9.91E–06	6.27E–06	3.40
0.005	1.13E–05	7.29E–06	3.40
0.01	1.26E–05	8.26E–06	3.40
0.015	1.40E–05	9.23E–06	3.40
0.02	1.53E–05	1.02E–05	3.40
0.025	1.67E–05	1.11E–05	3.40

Table 3 Hysteresis width analysis for different configurations of Schmitt trigger

Models	Hysteresis width (mV)
6T_ST-based NAND gate	419.541–778.2
4T_ST-based NAND gate	580.862–778.2
Body-biasing NAND gate	218.644–978.9
6T_ST-based NOR gate	420.218–776.45
4T_ST-based NOR gate	584.85–780.22
Body-biasing NOR gate	220.10–976.91

Where P_D1 represents the dynamic power dissipation of 6T-based model of Schmitt trigger NOR Gate, P_D2 represents the dynamic power dissipation of 4T-based model of Schmitt trigger NOR Gate and P_D3 represents the dynamic power dissipation of body-bias-based model of Schmitt trigger NOR Gate.

3.1.4 Hysteresis Analysis of Schmitt Trigger

The table presented below shows analysis of hysteresis width of different transistor models of Schmitt trigger. It is clear that body-biased Schmitt trigger-based NAND and NOR gate is giving best hysteresis width. Width of Schmitt trigger defines noise immunity of the circuit. More the width more is the noise immunity. Hence, it is required that a Schmitt trigger must provide maximum hysteresis width. Out of the all implemented models, body-biasing-based CMOS Schmitt trigger at Vn approaching to Vdd and Vp approaching to ground level is giving maximum hysteresis width (Table 3).

4 Conclusion and Future Scope

This paper represents the work on different models of Universal logic gates using different designs of Schmitt trigger like 4T model, 6T Model (Conventional model), and body-biased Schmitt Trigger design. This paper compares (in terms of graph) the propagation delay at the capacitance load of different models of Universal logic

gates thereby showing that body-bias models of Universal logic gates are best suited for low voltage and high-speed applications. Also, the body-biased voltages which are used for implementation purpose of designs are $V_n = 1.2$ V and $V_p = 0.5$ V. The main reason behind selecting this V_n and Vp is Schmitt trigger is giving maximum hysteresis width at this particular value. In the later section of the paper, analysis of different designs on the basis of power has also been performed. In future, working on the optimization of the design in terms of area utilized and power consumed will be performed. Also, digital circuits using hybrid design of Schmitt trigger will be implemented.

References

1. K.S. Vasundara Patel, H.N. Bhushan, K.G. Gadag, B.N. Nischal Prasad, M. Haroon, Schmitt trigger based SRAM using finfet technology-shorted gate mode. World Acad. Sci. Eng. Technol. Int. J. Comput. Electr. Autom. Control Inf. Eng. **8**(2) (2014)
2. J.S. Joseph, R. Shukla, V. Niranjan, Performance evaluation of low voltage Schmitt triggers using variable threshold techniques, in *2015 4th International Conference on Reliability, Infocom Technologies and Optimization (ICRITO) (Trends and Future Directions)* (IEEE, Sept 2015), pp. 1–5
3. A. Saxena, S. Akashe, Comparative analysis of Schmitt trigger with AVL (AVLG and AVLS) technique using nanoscale CMOS technology, in *2013 Third International Conference on Advanced Computing and Communication Technologies (ACCT)* (IEEE, Apr 2013), pp. 301–306
4. K. Kim, S. Kim, Design of Schmitt trigger logic gates using DTMOS for enhanced electromagnetic immunity of subthreshold circuits. IEEE Trans. Electromagn. Compat. **57**(5), 963–972 (2015)
5. J.P. Kulkarni, K. Roy, Ultralow-voltage process-variation-tolerant schmitttrigger-based SRAM design. IEEE Trans. Very Large Scale Integr. (VLSI) Syst. **20**(2), 319–332 (2012)
6. N. Lotze, Y. Manoli, A 62 mV 0.13 μm CMOS standard-cell-based design technique using Schmitt-trigger logic. IEEE J. Solid-State Circuits **47**(1), 4760
7. I.M. Filanovsky, H. Baltes, CMOS Schmitt trigger design. IEEE Trans. Circuits Syst. I: Fundam. Theory Appl. **41**(1), 46–49 (1994)
8. A.W. Kadu, M. Kalbande, Design of low power Schmitt trigger logic gates using VTCMOS, in *2016 Online International Conference on Green Engineering and Technologies (IC-GET)* (IEEE, Nov 2016), pp. 1–5
9. M. Kumar, P. Kaur, S. Thapar, S. Design of CMOS Schmitt trigger. Int. J. Eng. Innov. Technol. (IJEIT) **2** (2012)
10. Z. Chen, S. Chen, A high-speed low voltage CMOS Schmitt trigger with adjustable hysteresis, in *2017 IEEE/ACIS 16th International Conference on Computer and Information Science (ICIS)* (IEEE, May 2017), pp. 293–297
11. R. Sapawi, R.L.S. Chee, S.K. Sahari, N. Julai, Performance of CMOS Schmitt trigger, in *ICCCE 2008 International Conference on Computer and Communication Engineering 2008* (IEEE, May 2008), pp. 1317–1320
12. A. Suresh, A low power Schmitt trigger design using SBT technique in 180 nm CMOS technology, in *2014 International Conference on Advanced Communication Control and Computing Technologies (ICACCCT)* (IEEE, May 2014), pp. 533–536
13. N. Lotze, Y. Manoli, Ultra-sub-threshold operation of always-on digital circuits for IoT applications by use of Schmitt trigger gates. IEEE Trans. Circuits Syst. I Regul. Pap. **64**(11), 2920–2933 (2017)

14. N. Arjun, A. Marwah, S. Akashe, The alleviation of low power Schmitt trigger using FinFET technology. in *2015 International Conference on Communication Networks (ICCN)* (IEEE, Nov 2015), pp. 328–333
15. G. Hang, G. Zhu, A new Schmitt trigger with adjustable hysteresis using floating-gate MOS threshold inverter, in *2015 IEEE 11th International Conference on ASIC (ASICON)* (IEEE, Nov 2015), pp. 1–4
16. A. Nejati, Y. Bastan, P. Amiri, 0.4 V ultra-low voltage differential CMOS Schmitt trigger, in *2017 Iranian Conference on Electrical Engineering (ICEE)* (IEEE, May 2017), pp. 532–536
17. M. Moghaddam, M.H. Moaiyeri, M. Eshghi, Design and evaluation of an efficient Schmitt trigger-based hardened latch in CNTFET technology. IEEE Trans. Device Mater. Reliab. **17**(1), 267–277 (2017)
18. G.A. Raza, R. Mehra, Area and power efficient layout design of Schmitt trigger, in *2015 International Conference on Computer, Communication and Control (IC4)* (IEEE, Sept 2015), pp. 1–4
19. D. Sreenivasan, D. Purushothaman, K.S. Pande, N.S. Murty, Dual-threshold single-ended Schmitt-trigger based SRAM cell, in *2016 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC)* (IEEE, Dec 2016), pp. 1–4
20. S. Ahmad, M.K. Gupta, N. Alam, M. Hasan, Single-ended Schmitt-trigger based robust low-power SRAM cell. IEEE Trans. Very Large Scale Integr. (VLSI) Syst. **24**(8), 2634–2642 (2016)

Design and Analysis of 3-Bit Shift Register Using MTCMOS Technique



Bhupendra Sharma, Ashwani Kumar Yadav, Vaishali and Amit Chaurasia

1 Introduction

In the small channel high performance digital circuits, total power consumption depends on leakage currents [1]. Where the leakage power affects the available power, performance, and static power consumption during standby operation and thus reduces the battery life. To evaluate its effect many types of research have been done by various researches in the past few years. The ultra-low power design usually faces the challenges of increased delay and area. MTCMOS is really an effective technology to reduce power dissipation along with high reliability and lower leakage current. Different type of techniques and approaches used by the researcher is discussed in Sect. 1. Section 2 provides the theoretical aspects and architectural design of the targeted work. The detail of Methodology and the overall system design flow is discussed in Sect. 3. The working principle of the shift register is also explained in it. The general information regarding the simulation software HSPICE and waveform analyzer tool Cosmos Scope is explained in this section. Digital circuits in transistor level mode are also proposed. Results have been discussed in Sect. 4. Optimization in power, delay, and energy are also explained in this Section. The experiment performed on shift register with and without MTCMOS and comparison with the

B. Sharma (✉) · A. K. Yadav · Vaishali · A. Chaurasia
ASET, Amity University Rajasthan, Jaipur, India
e-mail: bhupesh19@gmail.com

A. K. Yadav
e-mail: ashwaniy2@gmail.com

Vaishali
e-mail: vaishaliyadav26@gmail.com

A. Chaurasia
e-mail: achaurasia2583@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_111

previous work has also been discussed. Finally the conclusion of the work is discussed in last segment. Thus, main emphasis is laid on low power design of digital circuits with less delay and optimized area.

2 Power Dissipation

Power dissipation in the circuits is mainly categorized into static and dynamic power dissipation [1–4].

2.1 Static Power Dissipation

Static power is the power that is dissipated when the circuit is idle, or not in switching states. Static power consumption is caused by the leakage currents, while the gates are idle. It is caused due to the high performance systems. Leakage currents affect the circuit both during active and idle mode of operations [5]. With continuous increase in leakage power in nanometer devices, it is evident that it will eventually exceed the active power if no leakage reduction scheme is used. Leakage power reduction during active mode of operation is investigated in this work.

2.2 Dynamic Power Dissipation

Dynamic power dissipation also effects the performance of the circuit, this type of dissipation depends on Switching (charging/discharging) of the output load capacitance to V_{DD} or ground voltage. Mathematically it can be expressed as:

$$P_{\text{dynamic}} = \alpha C_{\text{L}} V_{\text{DD}}^2 f + \sum_{i=1} n\alpha_i C_i V_{\text{DD}} (V_{\text{DD}} - V_T) \tag{1}$$

As per the above equation, this type of power dissipation mainly depends on the load capacitor.

Where V_{DD} = Supply Voltage, C_L = Load Capacitance, f = Operating Clock Frequency and α is the switching activity of gate [3].

2.3 Components Which Contribute to the Average Power Consumption (P_{avg})

- Static Power
- Short circuit power
- Leakage power
- Switching power

Leakage power is caused by substrate injection at pn junctions and subthreshold effects [6]. It contributes less than 1% of the average power consumption and can, therefore, be ignored [7].

2.4 Techniques for Reducing the Leakage Power

There are various techniques for leakage power reductions like Variable Threshold CMOS (VTCMOS) Circuits [8, 9], Multi-Threshold CMOS (MTCMOS) Circuit [8, 10], Power Gating [11], and Sleep Transistor [12].

2.5 Shift Registers

Shift registers are digital circuits that are used to shift data in several ways, with different storage capacities. Basic Building blocks used to design shift registers are flip-flops. They are designed in various sizes according to number of flip-flops used. In these types of digital circuits a common clock is used so that they all can be set or reset at the same time. The shift register prepares by all the type of flip-flop, for example, D flip-flop, SR flip-flop, and JK flip-flop. The shift register operates two types of function first one is the storage and the second one is the movement.

Shift register shifts the value when any active edge triggering occurs. A simple Shift Registers are designed with the help of D-type flip-flops, one flip-flop is used to store one bit the circuit diagram of 3-bit Shift Register is shown in Fig. 1. The output from one flip-flop is connected to the D input. Shift registers are used to

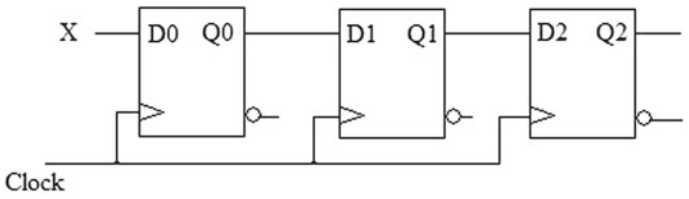


Fig. 1 3 bit shift register using D flip-flop

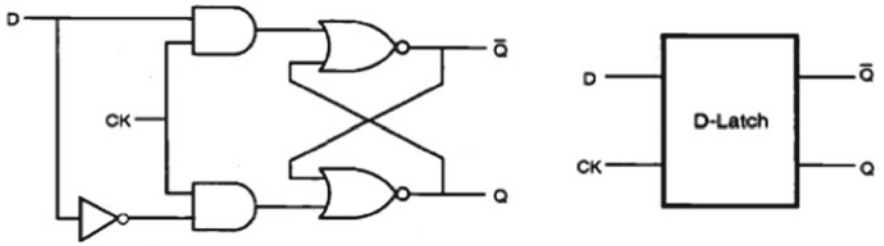


Fig. 2 Gate-level schematic and the block diagram view of the D-latch [8]

hold the data in their memory which can be moved or "shifted" to desired positions on each high clock signal. Each high clock signal shifts the contents of the register one-bit position to either the left or the right.

2.6 D Flip-Flop for Shift Register

The gate-level Schematic of the D-latch is designed with the help of a clocked NOR-based SR latch circuit with minute changes. The clocked D flip-flop having single input D, it is simply called Data flip-flop. It is directly linked to the set terminal of the input of the latch because there is a required only set and reset terminals so instead of another flip-flop it is much convenient in operation purpose. The schematic and block diagram illustrated as seen in Fig. 2. It can be seen from circuit that the output Q attains the input D when the clock is high, i.e., for CK = 1 and when the clock signal CK = 0 the output will preserve its state. It can offer a delay in digital circuitry as well as flip-flop change their state as soon as input D varies. D flip-flop used many types of application like for temporally storing the data and for delaying the input.

2.7 Schematic Diagram of CMOS D Flip-Flop Circuit

The transistor-level diagram consists of two tristate inverters; these inverters are driven by the clock pulse, where the one inverter is acting as a switch. This inverter is used to accept the input signal at the time where clock pulse is high. At the same time second inverter is in high impedance state. Both the inverters work alternatively as per the input clock and store the state until the next pulse reaches [8].

Table 1 Different parameters used for shift register

Technology (nm)	V_{DD} (V)	W (nm)	L (nm)	V_{th} for NMOS (V)	V_{th} for PMOS (V)	T_{clk} (n)
32	0.7	64	32	0.63	−0.58	0.02
45	0.9	90	45	0.46	−0.49	0.02

where
V_{DD} = supply voltage, W, L = width and length ratio in CMOS circuit, T_{clk} = Clock pulse, V_{th} = Threshold voltage

3 Different Parameters Used for Experimentation

The experiment is conducted for the 32 and 45 nm technology. As per the requirement of the low power design, the nominal V_{DD} is 1 v. The clock pulse for shift register is given 0.04 ns. The parametric variation is shown below in Table 1.

4 Methodology

The shift register is used with the proposed technique. The low V_{th} circuit is connected to the virtual ground and virtual V_{DD} in the power or ground gating mode. After examining the behavior of the simple shift register the same structure is connected with the MTCMOS technique shown in Fig. 3.

5 Description of Performance Measures

All parameters are measured in two different types of technology 32 and 45 nm.

- *Average and leakage Power with and without MTCMOS technique*

  Subthreshold leakage currents are the dominant source in the VLSI circuits.

$$I_{leakage} = e^{-V_{tn} \backslash (KT \backslash q)} \tag{2}$$

Here, V_{tn} represents the threshold voltage of NMOS, K is the process transconductance parameter and q is the total charge in the channel.

$$P_{leakage} = E_{Total} \backslash T \tag{3}$$

$P_{leakage}$ is known as the leakage power, E_{total} is the total energy of the circuit and T is the time period of the applied input.

- *Delay with and without MTCMOS technique*

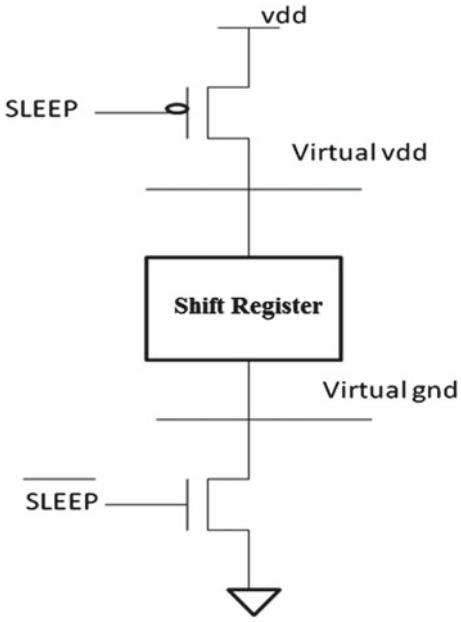


Fig. 3 MTCMOS technique in shift registers

- *Short Circuit and total energy with and without MTCMOS technique*

  The total energy measured by the following formula:

$$E_{\text{total}} = \left(E_{\text{Total_HL}} + E_{\text{Total_LH}} \right) /2 \tag{4}$$

where

$E_{\text{Total_HL}}$ Total energy at the time of high to low transitions
$E_{\text{Total_LH}}$ Total energy at the time of low to high transitions.

- *PDP with and without MTCMOS technique*

  Power and delay product of the circuit is defined as PDP.

5.1 Details of Experiment

In this work, the experiment is done for the shift register in the 32 and 45 nm technique.

Performance evaluation of shift register
For simulating the shift register first it is important to find the delay of the D-FF. Three D-FF connected in the serial in and serial out shift register and output from the

last flip-flop (FF3). The D-FF simulated in two PTM file one 32 nm and the second one is 45 nm. The input and output waveforms are shown in Fig. 4.

- **Shift Register in 32 nm**.

The Basic parameters for simulating the Shift register in 32 nm are: $V_{DD} = 0.7$ V, the threshold voltage for nmos is 0.63 V, threshold voltage for pmos is −0.58 V and *W/L* ratio is 2 for 32 nm. Where all the parameters are taken for 32 nm have been shown in Table 2.

The graph shown in Fig. 5 explains the comparison of the power dissipation of the shift registers in the 32 nm technology with respect to the time. It clearly shows the power dissipation of the circuit applied with and without MTCMOS technique. The MTCMOS technique provides the power dissipation is in uW range.

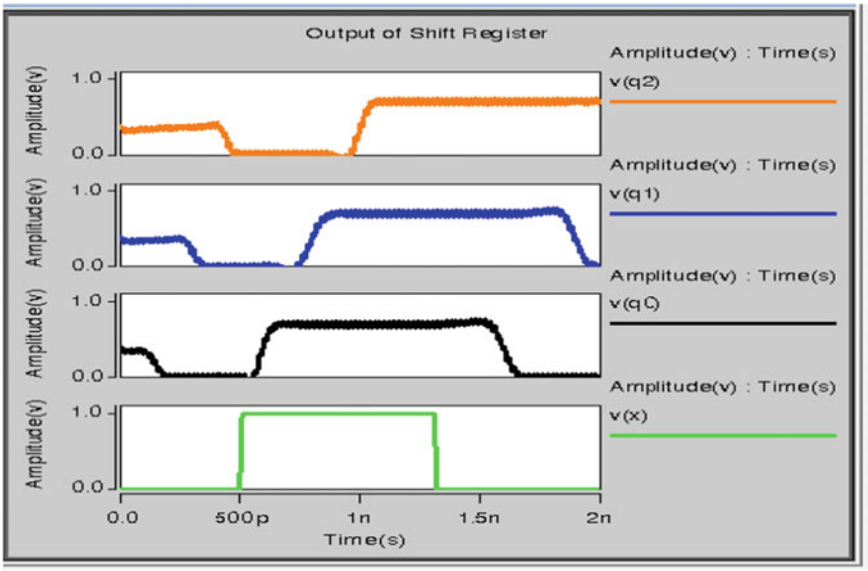


Fig. 4 Shift register input and output signal

Table 2 Performance parameters of shift register in 32 nm

Parameter	Without MTCMOS	With MTCMOS	Percentage reduction (%)
Average power	2.9×10^{-6} W	1.6×10^{-6} W	44
Delay	0.8 ns	0.76 ns	5
Short circuit energy	2.7×10^{-19} J	1.5×10^{-19} J	44
Total energy	71×10^{-18} J	58×10^{-18} J	18
Leakage power	35×10^{-7} W	29×10^{-7} W	17
Leakage current	50×10^{-7} A	41×10^{-7} A	18
PDP	28×10^{-16} J	22×10^{-16} J	21

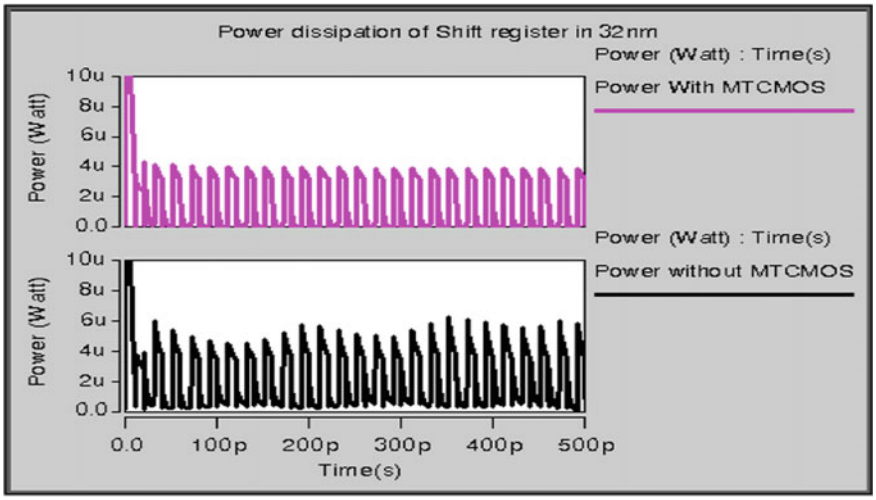


Fig. 5 Power dissipation in Shift registers in 32 nm technology

- **Shift Register in 45 nm**.

The next experiment shows the performance evaluation of shift register has been simulated with 45 nm technology and its transient analysis has been evaluated. The period of waveform is in 0.02 ns. Where $V_{DD} = 0.7$ V. various parameters observed have been shown in Table 3.

The above-presented graph shown in Fig. 6 explains the comparison of the power dissipation of the shift registers in the 45 nm technology with respect to the time. It clearly shows the power dissipation of the circuit applied with and without MTCMOS technique. The MTCMOS technique provides the power dissipation is in uW range.

Table 3 Performance parameters of shift register in 45 nm

Parameter	Without MTCMOS	With MTCMOS	Percentage reduction (%)
Average power	2.8×10^{-5} W	1.2×10^{-5} W	57
Delay	0.51 ns	0.4 ns	21
Short circuit energy	6.9×10^{-21} J	1.0×10^{-22} J	98
Total energy	5.9×10^{-17} J	1.1×10^{-20} J	99
Leakage power	2.9×10^{-6} W	5×10^{-9} W	99
Leakage current	3×10^{-6} A	5.5×10^{-9} A	99
PDP	1.4×10^{-15} J	2×10^{-18} J	99.7

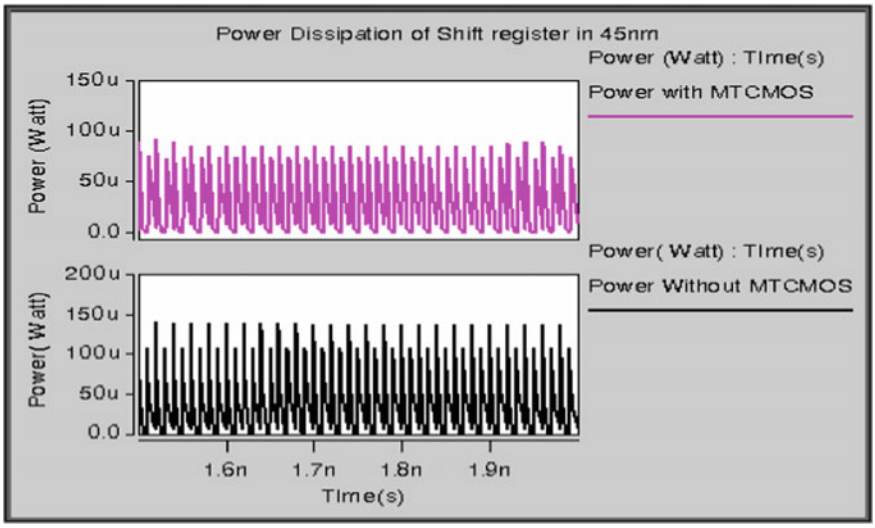


Fig. 6 Power dissipation in shift register in 45 nm technology

6 Result and Discussion

Thus the power consumption is less with the help of MTCMOS or in other words the parameter is optimized. All the experiments are conducted with the power and ground gating techniques for MTCMOS. The virtual voltage of the power line and ground line is found by the power supplied to the circuit. The reduction in the leakage current and power is 18 and 17% for 32 nm and 99 and 99% for 45 nm technique for shift register. In Shift register the 45 nm technique is proved better. Where the Percentage Optimization of Shift Registers for both technology has been shown in Fig. 7 for the comparison purpose.

7 Conclusion

To decrease power dissipation, a reduction in the leakage current is very important. There are many leakage current reduction techniques which reduce the leakage current. In all the described techniques the MTCMOS provides the high advantage and data retention capability. The behavior of the digital building blocks like a shift register is examined with or without MTCMOS technique. Simulation of shift register has been executed with 32 and 45 nm technology and experimental results are compared. Delay has also been reduced by 5% and 21% in 32 nm and 45 nm, respectively, and reduction in the leakage current is achieved by 18% and 99% for 32 nm and 45 nm in shift register.

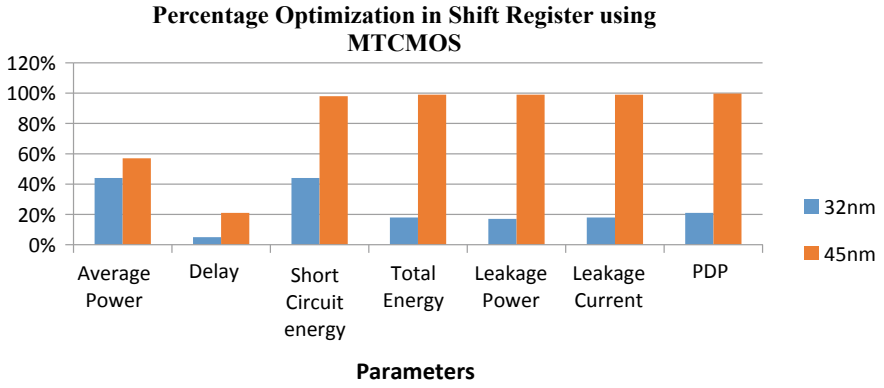


Fig. 7 Percentage optimization of shift register

References

1. S. Borkar, T. Karnik, Design and reliability challenges in nanometer technologies, in *IEEE 41st Proceedings of Design Automation Conference* (7–11 July 2004)
2. E.N. Shauley, CMOS leakage and power reduction in transistors and circuits. J. Low Power Electron. Appl. (JLPEA) **2** (2012)
3. S.M. Kang, Y. Leblebici, *CMOS Digital Integrated Circuits: Analysis and Design Book*, 3rd edn. (Tata McGraw Hill, 2003)
4. E. Macii, D. Liu, Power consumption of static and dynamic CMOS circuits: A comparative study, in *2nd International Conference on ASIC* (21–24 Oct 1996), pp. 425, 427
5. A.K. Yadav, K. Upadhyay, P. Gandhi, Vaishali, Various issues and considerations for the static power consumption in NANO-CMOS: design perspective, in *Materials Today: Proceedings* (published by Elsevier, 2019)
6. P. Saini, R. Mehra, Leakage power reduction in CMOS VLSI circuits. Int. J. Comput. Appl. **55**(8), 42–48 (2012)
7. A.K. Yadav, Y. Vaishali, Implementation of CMOS current mirror for low voltage and low power. IJECCE **3**(3), 620–624 (2012)
8. S.M. Kang, Y. lebelebici, *CMOS Digital Integrated Circuits Analysis and Design* (McGmlfl-Hill Higher Education). ISBN G-07-246053-9
9. H. Im, VTCMOS Characteristics and its optimum conditions predicted by a compact analytical model, in *Low Power Electronics and Design* (2001), pp. 123–128
10. J. Hailong, Noise mitigation in low leakage MTCMOS circuits. Ph.D. dissertation, Dept. Electron. Eng., Hong Kong University of Science and Technology, Hong Kong, 2012
11. M.T. Kumar, S.N. Pradhan, Power-gated FSM synthesis integrating partitioning and state assignment (Nov 2008) pp. 1–6 TENCON
12. H. Homayoun, S. Golshan, Post-synthesis sleep transistor insertion for leakage power optimization in clock tree networks, in *11th International Synopsis on Quality Electronics Design (ISQED)* (2010), pp. 499–507
13. B. Singh, A. Khosla, Power optimization of linear feedback shift register (LFSR) for low power BIST, in *2009 IEEE International Advance Computing Conference (IACC 2009)* (Patiala, India, 6–7 Mar 2009), pp. 311–314

Dynamic Power Reduction Techniques for CMOS Logics Using 45 nm Technology



Rajesh Yadav, Reetu and Rekha Yadav

1 Introduction

Power dissipation can be thought of as the rate at which energy is transferred from source to the drain. The thickness of the gate oxide and gate length are decreased due to technology scaling and hence the transistor density increases and the circuit delay is abridged [1, 2]. When the length of the gate is decreased, leakage power dissipation is increased due to the lower supply voltage and threshold voltage. It is explained that at 45 nm static power equals dynamic power. The GALEOR technique is proved effective for the system that spends substantial time in the sleep mode of the Power gating technique [3].

Here, four power reducing techniques are used in which power is reduced significantly, namely GALEOR, Power gating, Drain gating, and DFPH. Power is the least when DFPH is used. The power consumed using GALEOR is 1.896 μW, using Power gating is 1.700 μW, using Drain gating is 1.645 μW, and using DFPH is 1.573 μW.

R. Yadav (✉) · R. Yadav
ECE Department, HMRITM, Delhi, India
e-mail: raj_opyadav@yahoo.co.in

R. Yadav
e-mail: rekha_7r@yahoo.co.in

Reetu
ECE Department, Deenbandhu Chhotu Ram University of Science and Technology, Murthal, Sonipat 131039, India
e-mail: reetudalal3@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_112

In this paper, the various power reducing techniques are discussed. Section 2 consists of the proposed research methodology which describes various types of power dissipated by the CMOS circuits and the different power reducing techniques, Sect. 3 presents the simulation results and analysis. Section 4 concludes the results.

2 Proposed Methodology for Power Reduction

The power dissipation is basically divided into two major categories:

a. Dynamic power
b. Static power

The power dissipated by the circuit in an inactive state is considered to be static power. The power dissipated by the circuit in active state is referred to as dynamic power [4–7].

2.1 Dynamic Power

It can be further divided into

a. Switched Power
b. Short Circuit Power
c. Glitch Power

2.1.1 Switched Power

This type of power dissipation basically depends upon switching activity factor, timing, supply voltage (V_{dd}) given, and o/p capacitance of the circuit. The o/p capacitance of the circuit is charged and discharged continuously and repeatedly to transmit the data in the CMOS circuits [8]. The power consumed by the circuit is

$$P = f \cdot C \cdot V_{dd}^2 + f \cdot I_{Short} \cdot V_{dd} + I_{Leak} \cdot V_{dd} \quad (1)$$

where f is frequency, C is capacitance, V_{dd} is supply voltage given to the circuit, I_{Short} is short circuit current, and I_{leak} is leakage current.

2.1.2 Short Circuit Power Dissipation

When PMOS and NMOS both conduct simultaneously, the power dissipated at that time is known as short circuit power [9].

2.1.3 Glitch Power

The undesirable signal that is introduced in the circuit that does not have any useful information is known as a glitch and the power dissipated is known as glitch power [10–12]. Glitches are of two types: generated and propagated.

2.2 Static Power

Power consumed by the circuit in an inactive state is called static power. In this paper, various techniques have been described to lower the power in CMOS circuits. Here, they are explained by taking the example of a NAND gate and the results are compared with the previous results.

1. GALEOR,
2. Power gating,
3. Drain gating, and
4. DFPH.

These techniques are employed on a NAND Gate. NAND gate consists of two PMOS which are connected in parallel that forms a PULL-UP (PU) network and two NMOS which are joined in series that forms a PULL-DOWN (PD) network. The PU network and PD network are connected in series [13]. The schematic of the NAND gate is shown in Fig. 1.

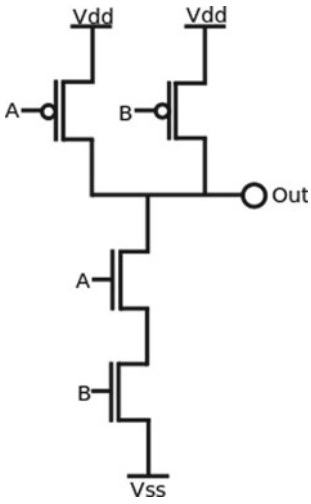


Fig. 1 Schematic of basic CMOS-based NAND gate

2.3 GALEOR Technique

This technique employs an NMOS and a PMOS connected in series between the PULL-UP and PULL-DOWN networks. GLT (Gated Leakage Transistor). The output voltage swing is decreased by this technique which is caused due to threshold voltage loss by an additional MOS transistor [14, 15]. The propagation delay of the circuit is enhanced by reduced voltage swing. The disadvantage of the GALEOR technique is that the low signal is much higher than 0 volts while the high signal is much lower than the supply voltage. The technique is shown in Fig. 2a.

2.4 Power Gating

In the Power gating technique a PMOS is inserted between the supply voltage and the PULL-UP network and an NMOS are injected between the ground and the PULL-DOWN network. The input of both the inserted MOSFETs is a complement to each other. The technique is shown in Fig. 2b.

2.5 Drain Gating

In this technique, a PMOS is inserted between the PULL-UP network and the output and an NMOS is inserted between the output and the PULL-DOWN network. The input of both the inserted MOSFET is a complement to each other. The leakage current is reduced by adding extra sleep transistors amid PULL-UP and PULL-DOWN networks. When in active mode, the resistance of the conducting paths is decreased because both the sleep transistors are ON, thus dropping the performance degradation [16]. The sleep transistors are turned OFF due to which a stacking effect is produced that decreases the leakage current by improving the path resistance from the voltage supply to the ground in the standby mode. The technique is represented in Fig. 2c.

2.6 Drain Footer and Power Header (DFPH) Technique

In this, a PMOS is inserted between the voltage supply (V_{dd}) and the PULL-UP network and an NMOS are injected between the output and PULL-DOWN network. The input of both the inserted MOSFETs is a complement to each other. The supply voltage given is 1 V. The technique is shown in Fig. 2d.

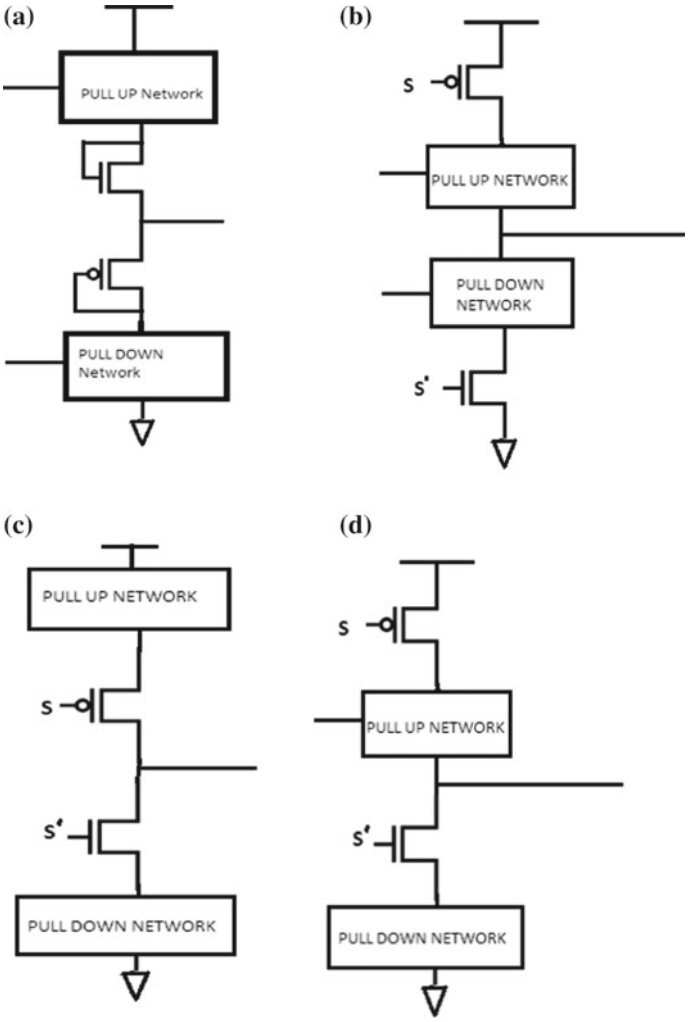


Fig. 2 Schematics of any digital function using different power reduction techniques **a** GALEOR, **b** Power gating technique, **c** Drain gating technique, and **d** DFPH technique

3 Result and Discussion

There are numerous techniques that can be employed to decrease the power consumption for the CMOS logic family. Here, the schematic of the NAND gate can be seen in Fig. 1. Figure 3a shows schematically using GALEOR Technique. Schematic of NAND gate using Power gating technique is shown in Fig. 3b. Figure 3c represents

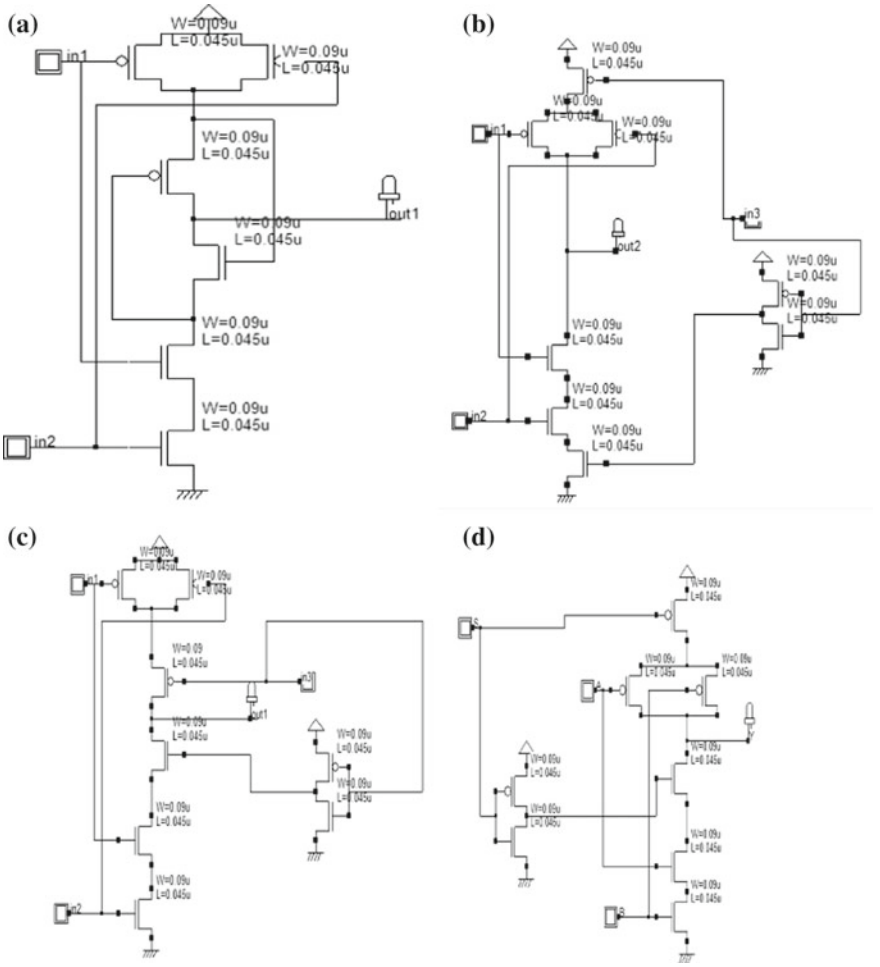


Fig. 3 Schematic NAND gate using **a** GALEOR technique, **b** Power gating technique, **c** Drain gating technique, and **d** DFPH technique

the schematic of the NAND gate using Drain gating technique. Figure 3d symbolizes the schematic of the NAND gate using the DFPH Technique.

Input Parameter Specification:

The input parameters like Length of MOS, Width of MOS, and Voltage supply are specified in Table 1. To design various low-power techniques based NAND gate, the 45 nm CMOS technology is used; for this, MOS transistors' channel length are 45 nm and MOS transistors' width are 90 nm.

Table 1 Input parameter specifications for different techniques

Sr. No.	Techniques used	Length of MOS (nm)	Width of MOS (nm)	Voltage (V)	Rise and fall time (ns)
1	Simple NAND gate	45	90	1	1
2	GALEOR technique	4	90	1	1
3	Power gating technique	45	90	1	1
4	Drain gating technique	45	90	1	1
5	DFPH technique	45	90	1	1

3.1 Simulation

Simulation of the NAND gate is done using different techniques. Figure 5a shows a simulation of the NAND gate at 1 V. Figure 5b shows simulation using the GALEOR Technique. Simulation using the Power gating technique can be seen in Fig. 5c. Figure 5d symbolizes simulation of the NAND gate using the Drain gating technique. Figure 5e shows a simulation using the DFPH Technique.

Layouts of the NAND gate can be seen in Fig. 4. Figure 4a shows the layout of the NAND gate, Fig. 4b is the layout of the NAND gate using the GALEOR Technique. Figure 4c shows the layout of the NAND gate using the Power gating technique. Figure 4d symbolizes the layout of the NAND gate using the Drain gating technique. Figure 4e represents the layout of the NAND gate using the DFPH Technique.

Transient analyses of various power reduction techniques for the NAND gate are shown in Fig. 5. A 1 V pulse voltage is given at the input of all gates which have 5 ns pulse width, 1 ns rise time, and 1 ns fall time. Figure 5 shows the maximum output level and propagation delay with respect to the input signal (Table 2).

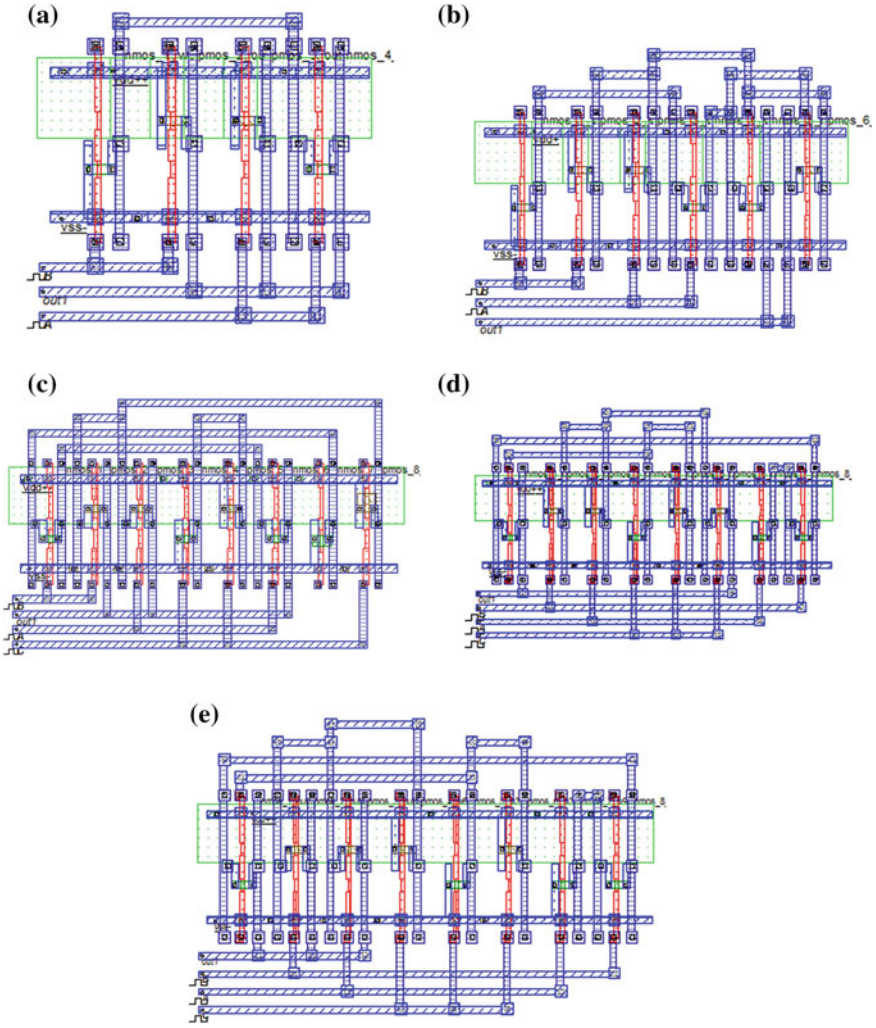


Fig. 4 Layout of basic NAND gate using **a** Static CMOS logic technique, **b** GALEOR technique, **c** Power gating technique, **d** Drain gating technique, and **e** DFPH technique

Table 2 A comparative analysis of different power reduction techniques

Sr. No.	Techniques	Power consumed (μW)	Average delay (ps)	Layout area (μm^2)
1.	Simple NAND gate	1.907	231	0.68
2.	GALEOR	1.896	210	1.09
3.	Power gating	1.700	3214.5	1.55
4.	Drain gating	1.645	6223	1.395
5.	DFPH	1.573	3210	1.55

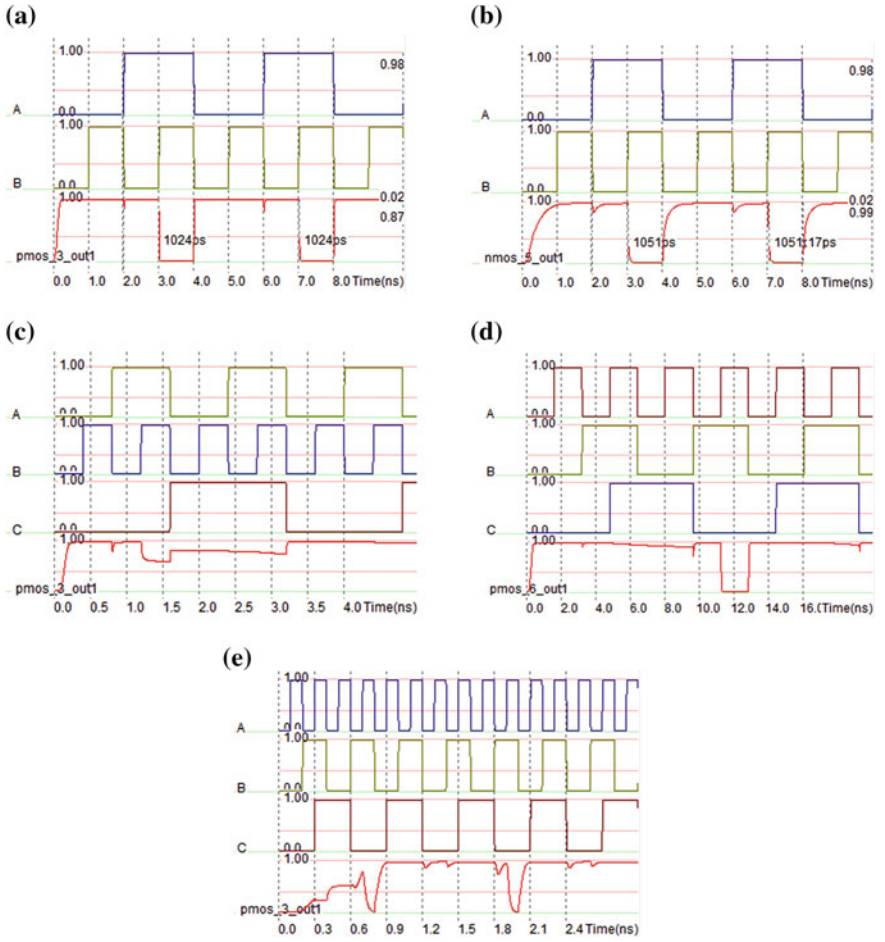


Fig. 5 Simulation of NAND gate using **a** CMOS logic technique, **b** GALEOR technique, **c** Power gating technique, **d** Drain gating technique, **e** DFPH technique

4 Conclusion

NAND gate is used in various circuits for implementing various functions. NAND gate using various techniques such as GALEOR, Power gating, Drain gating, and DFPH is designed and analysed in this paper using DSCH and MICROWIND software at the 45 nm technology. DSCH is used for drawing the schematic and MICROWIND is used for the layout and simulation. Power analysis of NAND gate is done in this paper and it is concluded that the DFPH technique is the most efficient technique for minimizing the power. The power dissipated using DFPH is 1.573 μW.

References

1. J.W. Chun, C.Y. Roger Chen, A novel leakage power reduction technique for CMOS circuit design, in *2010 International SoC Design Conference* (2010), pp. 119–122
2. M. Geetha Priya, K. Baskaran, D. Krishnaveni, Leakage power reduction technique in deep submicron technologies for VLSI applications, in *International Conference on Communication Technology and System Design*, vol. 30 (2012), pp. 1163–1170
3. N.B. Romli, K.N. Minhad, M.B.I. Reaz, M.D.S. Amin, An overview of power dissipation and control techniques in CMOS technology. J. Eng. Sci. Technol. **10**(3), 364–382 (2015)
4. S. Panwar, M. Piske, A. Vivek Madgula, Performance analysis of modified drain gating techniques for low power and high speed arithmetic circuits. Hindawi (VLSI Design) 380362 (2014)
5. Reetu, R. Yadav, Dynamic power reduction of VLSI circuits: a review. Int. J. Adv. Res. Electron. Commun. Eng. **7**(3), 245–259 (2018)
6. K. Vinothkumar, P. Karthikeyan, Study of outpouring power diminution technique in CMOS circuits. Int. J. Comput. Sci. Mob. Comput. **3**(11), 137–143 (2014)
7. R.K. Patlani, R. Yadav, Design of low power ring VCO and LC-VCO using 45 nm technology. IJISET Int. J. Innov. Sci. Eng. Technol. **1**(4) (2014)
8. M. Manoranjani, T. Ravi, R. Ramya, Effect of leakage power reduction techniques on combinational. ARPN J. Eng. Appl. Sci. **10**(7)
9. J. Deshmukh, K. Khare, Implementation and analysis of SC-LECTOR CMOS circuit using cadence tool. Int. J. Eng. Sci. Technol. **2**(5), 1250–1252 (2010)
10. A. Nagda, R. Prasad, T.N. Sasamal, N.K. Vyas, Leakage power reduction techniques: a new approach. Int. J. Eng. Res. Appl. **2**(2), 308–312 (2012)
11. J.C. Park, V.J. Mooney, Sleepy stack leakage reduction. IEEE Trans. Very Large Scale Integr. (VLSI) Syst. **14**(11), 1250–1263 (2006)
12. N. Hanchate, N. Ranganathan, LECTOR: a technique for leakage reduction in CMOS circuits. IEEE Trans. Very Large Scale Integr. (VLSI) Syst. **12**(2), 196–205 (2004)
13. S. Katrue, D. Kudithipudi, GALEOR: leakage reduction for CMOS circuits, in *International Conference on Electronics, Circuits and Systems* (2008), pp. 574–577
14. P.F. Butzen, L.S. da Rosa Jr, E.J.D.C. Filho, A.I. Reis, R.P. Ribas, Standby power consumption estimation by interacting leakage current mechanisms in nanoscaled CMOS digital circuits. Microelectron. J. **41**(4), 247–255 (2010)
15. W.M. Elgharbawy, M.A. Bayoumi, Leakage sources and possible solutions in nanometer CMOS technologies. IEEE Circuit Syst. Mag. **5**(4), 1–17 (2005)
16. N. Ekekwe, R. Etienne-Cummings, Power dissipation sources and possible control techniques in ultra-deep submicron CMOS technologies. Microelectron. J. **37**(9), 851–860 (2006)

Design of an Optimal Integer Frequency Synthesizer for 5 GHz Frequency at 45 nm CMOS Technology



Rekha Yadav and Sakashi Kaushik

1 Introduction

Recently, CMOS frequency synthesizers using PLL are widely used in radio communication applications. These frequency synthesizers are found as an essential block in the transceiver of cellular phones and various wireless applications such as domestic radio and television. It operates as a professional radio frequency equipment like a signal generator, spectrum analyzers, and much more [1, 2]. In this paper, an integer frequency synthesizer is designed using the 45 nm technology. To optimize the integer frequency synthesizer, a third-order Chebyshev filter is used and current starving VCO is used. In this proposed frequency synthesizer, the value of the integer is 2 and the divider is designed using the TSPC technology. The schematic of a different block of the integer frequency synthesizer is shown in Fig. 1.

The frequency synthesizer is an electronic circuit that is used to generate the frequency range from a single frequency. The frequency range can also be generated with the help of an oscillator. But there are many disadvantages of oscillators like there is a need for control heater for stabilization. And if the oscillator is operated in a non-temperature control environment, then the frequency is changed. That is why nowadays, frequency synthesizer is used for generating the frequency range.
The main contributions of the paper are as follows:

- A novel CMOS frequency synthesizer for radio communication applications operating at frequency 5 GHz is proposed.
- The parameters such as phase noise, output power, output noise, and power dissipation of the above-proposed circuit are analyzed with all its scaling issues.
- Performance parameters are optimized for radio communication applications.

The organization of this paper is discussed as follows: Sect. 2 describes the proposed methodology in which different blocks of the integer frequency synthesizer are defined. And Sect. 2.3 introduces the mathematical modeling in which the output

R. Yadav (✉) · S. Kaushik
ECE Department, Deenbandhu Chhotu Ram University of Science and Technology, Murthal, Sonepat 131039, Haryana, India
e-mail: rekha_7r@yahoo.co.in

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_113

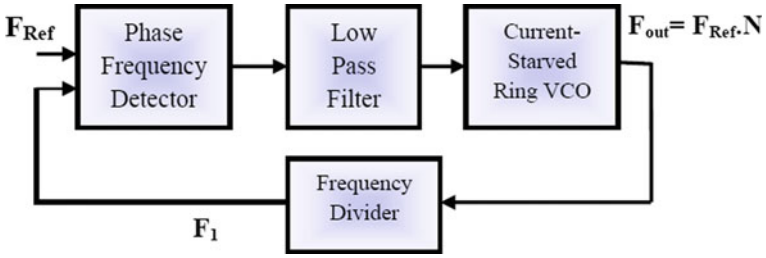


Fig. 1 Basic block diagram of the integer frequency synthesizer

frequency of VCO is calculated. Simulation of the synthesizer is described in Sect. 3 and in Sect. 4, the conclusion is discussed.

2 Design Methodology

In this paper, a range of frequencies is generated with the help of PLL. To achieve synchronization, PLL is a basic building block in all communication systems. The main aim to use PLL in a frequency synthesizer is to produce a signal which has the same phase as that of the reference signal. The output from the phase frequency detector is passed to a low-pass filter or loop filter and then the output of the low-pass filter is fed into VCO (voltage control oscillator), and at last output of VCO goes to a frequency divider (divide by 2). The VCO with frequency divider generates a signal which is then compared by the reference signal using the Phase frequency detector. Then PFD generates a signal which is corresponding to the phase difference of two signals. This error which is generated by PFD goes to LPF (low-pass filter) which eliminates the high-frequency component and works as a control signal for VCO. Here, Fig. 1 shows the block diagram of a frequency synthesizer using PLL [3, 4]. Frequency synthesizer consists of four blocks:

– Phase frequency detector (PFD),
– Low-pass filter (LPF),
– Current-starved ring VCO, and
– Frequency divider.

2.1 Phase Frequency Detector

The phase frequency detector is the main component of the phase-locked loop. Phase frequency detector generates a signal which is corresponding to the phase difference between two signals. One advantage of using phase frequency detector over the phase detector (using XOR gate) is that whenever the phase difference is in the middle of

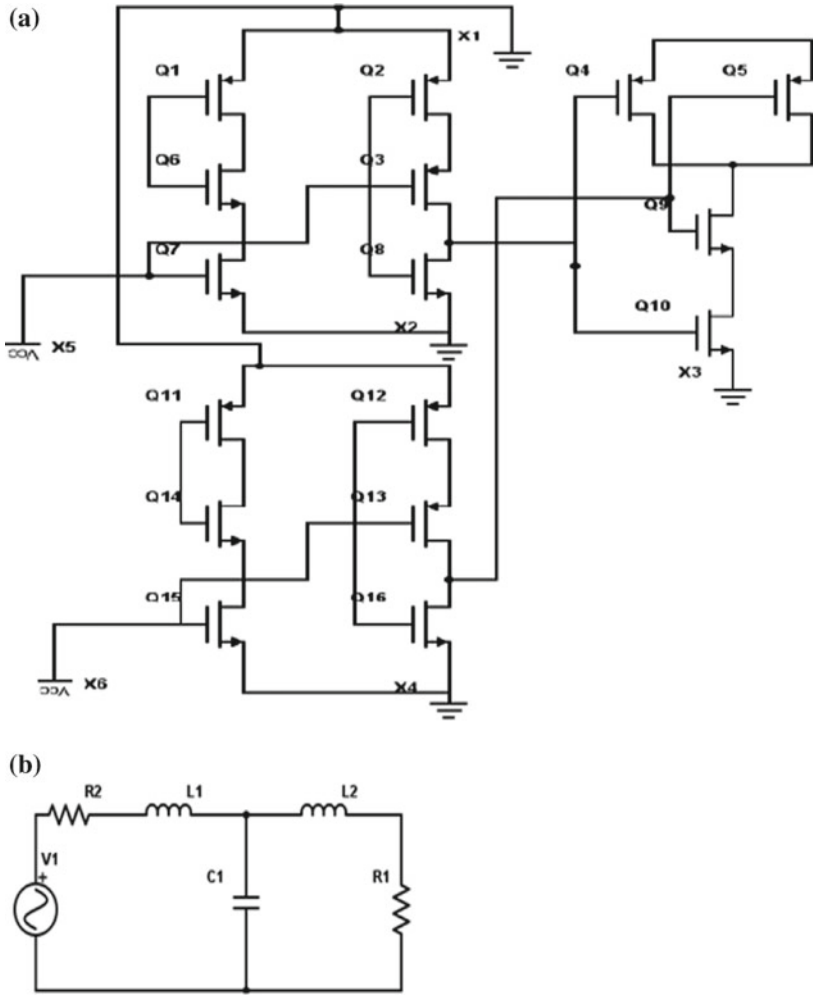


Fig. 2 **a** Schematic of phase frequency detector, and **b** Low-pass filter using T network of LC circuit

± 180, then the voltage is generated corresponding to the phase difference. Whenever a loop is out of lock, then no AC component is to be produced. In this paper, PFD is designed using a D flip-flop. The schematic of PFD is as shown in Fig. 2a.

2.2 Low-Pass Filter

The error signal which is produced by a phase frequency detector is filtered out by a low-pass filter. The low-pass filter is used to remove the unwanted AC component because this ac component cannot be used as input to VCO. In this paper, we proposed three-order Chebyshev's filters to modify the existing LPF. The design of the low-pass filter is shown in Fig. 2b.

2.3 Voltage Controlled Oscillator

VCO is a key building block for the phase-locked loop. VCO is an oscillator whose oscillation frequency is controlled by a control voltage. The output of LPF is to act as a control voltage for VCO. In this paper, ring VCO is used because of its high stability as compared to others. Ring VCO has many advantages as compared to other oscillators. Ring VCO has low cost, small chip area, and good performance. The schematic of the ring VCO is shown in Fig. 3a [5].

The output of a ring VCO is fed into the phase frequency detector through a frequency divider. The frequency divider takes an input signal of a frequency Fin and generates an output signal of frequency F_{out}, where F_{out} is given as follows:

$$F_{out} = \frac{F_{in}}{N} \tag{1}$$

There are two types of frequency dividers:

– Fixed frequency divider and
– Programmable frequency divider.

In this paper, we are using a fixed frequency divider where the output of VCO is divided by 2. The schematic of the frequency divider is shown in Fig. 3b.

The output frequency generated by this current-starved ring VCO depends upon propagation time (T) and control voltage (V_{ctrl}). And propagation time (T) depends upon the delay of the number of stages (n) of current-starved CMOS inverter and it is defined as follows:

$$PropagationDelay(T) = 2n \cdot t_d \tag{2}$$

Here, t_d is the delay of single stage and N is the number of inverter stages used in VCO. At minimum control voltage, the frequency of oscillation can be calculated as follows [6–9]:

$$F_{osc} = 1/T \tag{3}$$

From Eq. (2),

$$T = \frac{1}{2n \cdot t_d} \tag{4}$$

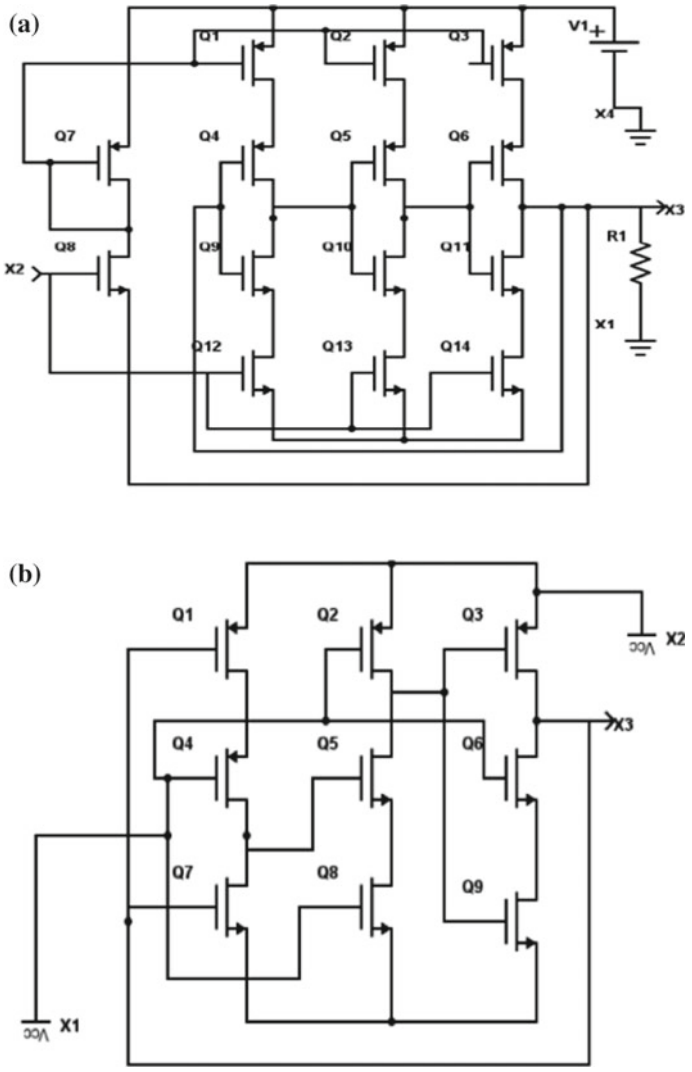


Fig. 3 **a** Schematic of CMOS current-starved ring oscillator, and **b** Schematic of the frequency divider (divided by 2)

where t_d is expressed as [8, 9]

$$T_d = \frac{4C_{Load}V_{tn}}{\beta_n(V_{gs}-V_{tn})^2} + \frac{C_{Load}}{\beta_n(Vdd-V_{tn})} ln(\frac{2(Vdd-V_{tn}}{Vdd} - 1) \tag{5}$$

Thus, the output frequency (F_{out}) of the CMOS current-starved ring VCO is calculated as [10, 11]

$$F_{out} = F_{osc} + K_{vco} \cdot V_{ctrl} \tag{6}$$

Here, K_{VCO} is known as gain/sensitivity of the ring VCO, C_{load} is known as the total capacitance at the output of each stage, i.e., total equilibrium load capacitance between input and output terminal of inverter stages. Power dissipation across this current-starved ring VCO is written as

$$P_{VCO} = 2 \cdot \eta \cdot q_{max} \cdot V_{dd} \cdot F_{osc} \tag{7}$$

where q_{max} is the maximum charge stored at each of the stages and η is the efficiency of VCO. The frequency divider divides that frequency with a fixed/constant parameter N and generates a sinusoidal frequency F1. The phase frequency detector compares this F1 frequency with the reference frequency (Fref) and gives the differential signal to the low-pass filter which generates a control voltage (V_{ctrl}). If $F_{ref} > F_1$, then V_{ctrl} goes down its level and the output frequency of current-starved ring VCO is decreased. And, if $F_{ref} <_1$, then V_{ctrl} goes up its level and the output frequency of the current-starved ring VCO is increased. However, the low-pass filter is used to pass low-frequency signal, stop RF signal, and provide a transfer function F(s). Open-loop gain of this optimized CMOS frequency synthesizer is expressed as Open-loop gain:

$$G(s) = \frac{K_\phi \cdot F(s) \cdot K_{VCO}}{N \cdot s} \tag{8}$$

where K_ϕ, $\frac{K_{VCO}}{s}$ is the gain of phase frequency detector and the current-starved ring VCO.

3 Results and Discussion

There are many aspects that can be combined in the design to achieve a better result. Here, a frequency synthesizer using PLL is designed using the 45 nm CMOS technology simulated on Cadence. Here, the proposed VCO is designed for generating 5.36 GHz frequency. The overall power consumption of the frequency synthesizer is 490 mw with supply voltage 1.5 V. Here, Table 1 shows the dimension of different blocks of the integer frequency synthesizer (PFD, LPF, VCO, and frequency divider). The output of a different block of the frequency synthesizer is shown in the below figures. Figure 4a shows the output of PFD where v_{in1} and v_{in2} are input and v_{out} is the output of PFD. And Fig. 4b shows the output of the low-pass filter and current-starved VCO. Here, the cut-off frequency of the low-pass filter is 40 MHz and the frequency generated at the output of VCO is 5.36 GHz.

Figure 5a shows the output and input frequency response of the frequency divider. CMOS current-starved VCO generates 5.31 GHz oscillation frequency as shown in Fig. 5b which shows the range of oscillation frequency and its settling time, i.e., nearly 10 ns.

Phase noise and output noise response at the output of the proposed frequency synthesizer are shown in Fig. 6a, b and PSS (periodic steady state) of the frequency

Table 1 Input parameter specification for the proposed CMOS frequency synthesizer

Device name	Component name	Component value
Phase frequency detector	MOS size	45 nm × 120 nm × 1
	Vdd	1 V
Low-pass filter	Resistor (R1)	50 Ω
	Inductor (L1)	198 nH
	Capacitor (C1)	160 pF
	Inductor (L2)	198 nH
	Resistor (R2)	160 pF
Voltage controlled oscillator	Load resistance	50 Ω
	Vdd	1.5 V
	NMOS size	45 nm × 45 nm × 7
	PMOS size	45 nm × 150 nm × 7
Frequency divider	N	2
	MOS size	45 nm × 45 nm × 1

Table 2 Comarative analysis of the proposed design with the previous work

Serial No.	Parameters name	Proposed design	[8]	[12]	[13]
1.	Technology used	45 nm CMOS	250 nm CMOS	350 nm CMOS	180 nm CMOS
2.	Reference frequency	40 MHz	5 MHz	20 MHz	5 MHz
3.	Supply voltage	1.5 V	1.2 V	3.0 V	0–2 V
4.	Output frequency	5–11 GHz	100–200 MHz	350 MHz–16 GHz	5.15–5.25 GHz
5.	Power consumption	490 mW	367 mW	7.6 W	–
6.	Phase noise of VCO	−100 dBc/Hz	–	–	–

synthesizer is shown in Fig. 7. The phase noise of integer frequency synthesizer is −100 dBc/Hz at 1 MHz offset and output noise at that frequency is 1.0298 e^{-19} V/sqrt (Hz).

Table 2 shows a comparison between the previous work and the proposed one.

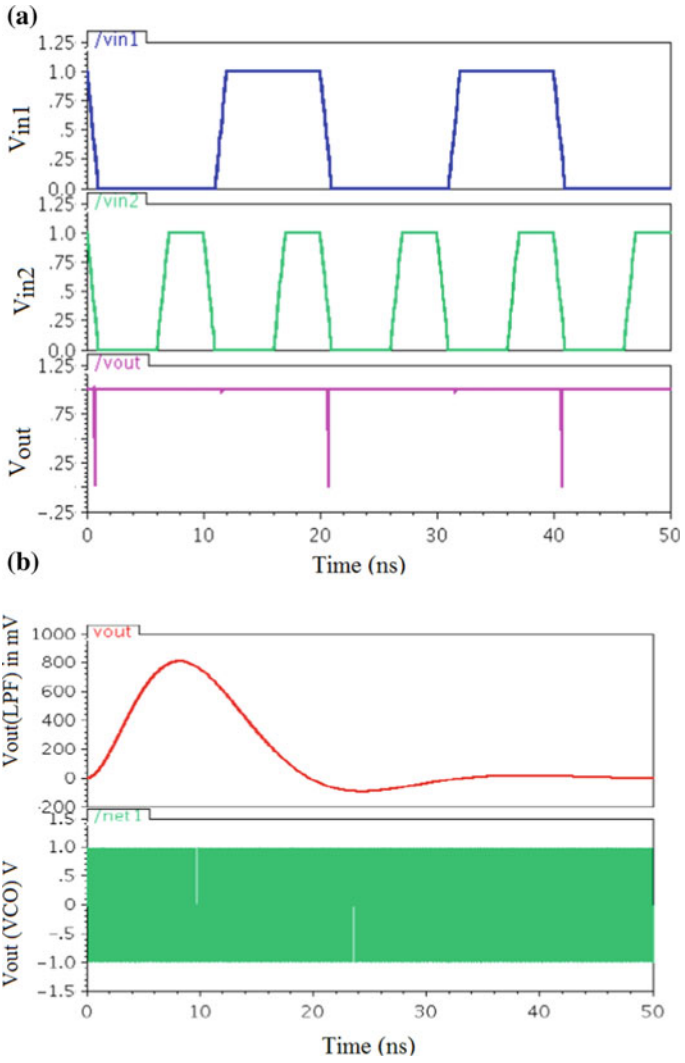


Fig. 4 **a** Output and input of phase frequency detector, and **b** Output of the low-pass filter and current-starved VCO

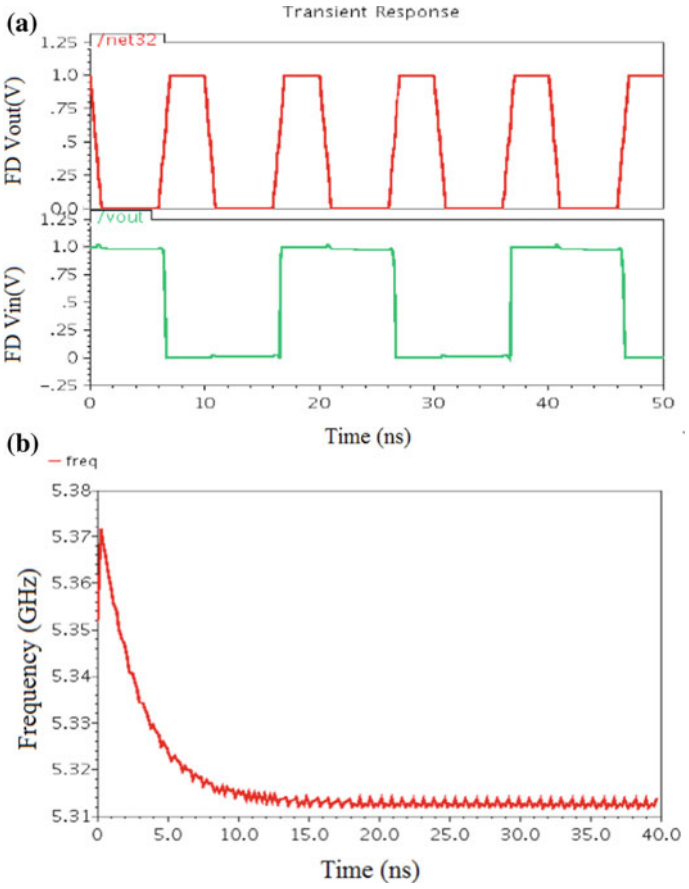


Fig. 5 **a** Frequency response at the output and input terminal of the frequency divider, and **b** The range of oscillation frequency and its settling time of the current-starved VCO

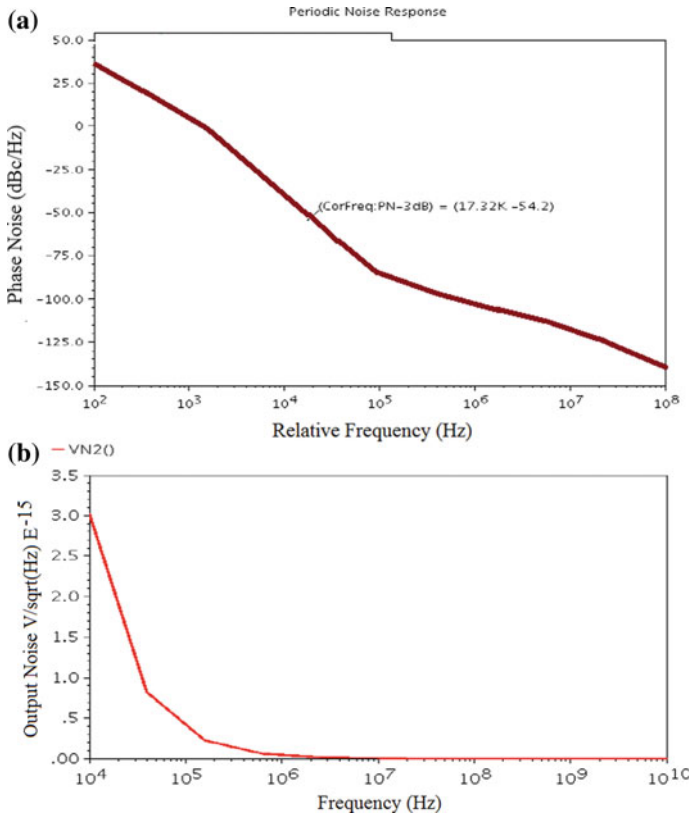


Fig. 6 **a** Phase noise and **b** Output noise response at the output of the proposed frequency synthesizer

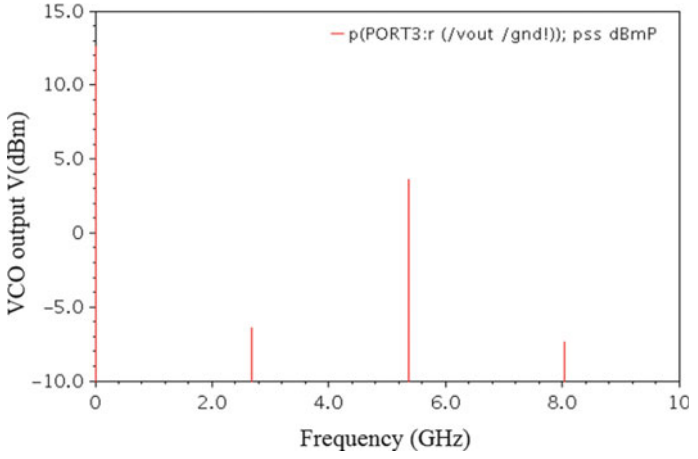


Fig. 7 PSS analysis of frequency synthesizer which shows the output power of the proposed frequency synthesizer

4 Conclusion

Frequency synthesizer using PLL covers many areas of electronics engineering such as signal analysis, digital circuit design, and communication theory. Integer frequency synthesizer has been designed using a small-size chip, low output noise using the 45 nm CMOS technology for high performance and is simulated by the Cadence tool. Various characteristics of the circuit are measured and verified in this paper. The frequency generated by the proposed methodology is 5.37 GHz and PSS (periodic steady state) at the frequency 5.36 GHz is 4.5 dbm. Output noise at offset frequency 1 MHz is $1.0298\,e^{-19}$ V/sqrt (Hz) and its phase noise is -100 dBc.

References

1. M. Thamsirianunt, T.A. Kwasniewski, CMOS VCO's for PLL frequency synthesis in GHz digital mobile radio communications. IEEE J. Solid State Circuits **32**(10), 1511–1524 (1997)
2. H. Lee, J. Cho, K. Lee, I. Hwang, T. Ahn, K. Nah, B. Park, A Σ-Δ fractional-N frequency synthesizer using a wideband integrated VCO and a fast AFC technique for GSM/GPRS/WCDMA applications. IEEE J. Solid State Circuits **39**(07), 1164–1169 (2007)
3. S.T. Moon, A.Y. Valero-López, E. Sánchez-Sinencio, Fully integrated frequency synthesizes: a tutorial. Int. J. High Speed Electron. Syst. **15**(02), 353–375 (2005)
4. R.K. Patlani, R. Yadav, Design of low power ring VCO and LC-VCO using 45 nm technology. IJISET Int. J. Innov. Sci. Eng. Technol. **1**(4) (2014)
5. S. Biabanifard, S.M. Largani, S. Asadi, Delay time analysis of combined CMOS ring oscillator. Electr. Electron. Eng. Int. J. (ELELIJ) **4**(2), 53-64 (2015)
6. K.W. Tang, S. Leung, N. Tieu, P. Schvan, S.P. Voinigescu, Frequency scaling and topology comparison of millimeter-wave CMOS VCOs, in *2006 IEEE Compound Semiconductor Integrated Circuit Symposium*, Nov 2006
7. P.M. Farahabadi, H. Miar-Naimi, A. Ebrahimzadeh, A new solution to analysis of CMOS ring oscillators. Iran. J. Electr. Electron. Eng. **5**(1), 32–41 (2009)
8. P. Patil, V.K. Verma, Integer N frequency synthesizer using phase lock loop. IJCTT **32**(01) (2016)
9. B. Razavi, *Design of Analog CMOS Integrated Circuits* (McGraw- Hill, 1973)
10. C.-L. Wei, S.-I. Liu, A digital PLL using oversampling delta-sigma TDC. IEEE Trans. Circuits Syst. II Exp. Briefs **63**(07), 633–637 (2016)
11. R. Yadav, P.K. Dahiya, R. Mishra, A scalable millimetre-wave differential CMOS cross-coupled VCO for automotive radar application. J. Circuits Syst. Comput. **27**(10), 1850158 (1997)
12. V.L. Kumar, Simulation of integer N frequency synthesizer. Int. J. Eng. Trends Technol. **04**(06), 2662–2665 (2013)
13. K. Rajasekhar, S. Adilakshmi, Design of high performance phase locked loop for multiple outputs with ultra low power sub threshold logic. Int. J. Eng. Res. Appl. (IJERA) **020**(2), 46–52 (2012)

Performance Evaluation of Hybrid RF/FSO Communication System in High-Speed Train



Ankita Chauhan and Pankaj Verma

1 Introduction

High-speed trains (HSTs) have become a global choice of transportation (HSTs) for a large number of long-distance travelers. Currently, HST has attained the highest speed of 500 km/h [1]. The increase in the use of Internet-connected digital devices poses a large demand for reliable and fast Internet access in high-speed trains. To support these demands of wireless communication applications, there is a perpetual need for higher bit rates to drive the developing technologies in the field of Wireless Optical Communication (WOC) and Millimeter Wavelength radio frequency (MMW-RF). WOC can be further categorized as free-space optical (FSO) communications, Visible Light Communications (VLC), and Ultra-Violet (UV) Communications. The FSO technology is widely used for symbol transference due to higher bit rate, higher transmission reliability, large unlicensed spectrum, and cheaper and speedy deployment [2, 3]. Due to high velocity of the train, the system capacity or throughput of free-space optical systems for high-speed trains may be much lower than stationary point-to-point (P2P) free-space optical communication systems [4]. However, FSO technology cannot replace current RF communication systems. The requirement of line-of-sight limits the implementation of FSO links [5, 6]. An attractive solution is to build a wireless system including both FSO and reliable lower rate RF links, forming Hybrid FSO/RF communication systems. This paper discusses the Hybrid FSO/RF system in HSTs using FSO and MMW-RF communication technologies. Both the technologies, FSO and MMW-RF, offer huge capacity (up to a few Gbps) by operating in the unlicensed free-space optical band and 70 GHz frequencies, at

A. Chauhan (✉) · P. Verma
Department of Electronics and Communication Engineering, National Institute of Technology Kurukshetra, Kurukshetra, India
e-mail: ankitachauhanece237@gmail.com

P. Verma
e-mail: pankaj@nitkkr.ac.in

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_114

a very small cost [7–9]. Although both links have many common features, they are not affected in the same way in case of weather and meteorological conditions [10]. For FSO link, fog and scintillation losses are the main attenuation factors; similarly, RF links (70 GHz) can withstand heavy rain conditions, and fog does not cause any particular effect [8, 10]. The adjunct nature of MMW-RF and FSO links led to various approaches in Hybrid RF/FSO information transmission systems. Such type of system maintains high data rate transmissions as well as shows excellent data link reliability and performance in all weather conditions. To analyze the performance of this system, two types of switching have been used, first, individual switching and second, integrated switching. Individual switching schemes define when to switch between the two links for different visibility conditions to obtain maximum capacity. In this scheme, the total available power is utilized by one link (either FSO or RF) [11–15]. Authors in [11] obtain the error probability and throughput expression for the Hybrid RF/FSO scheme with two levels of threshold. In a time-varying fading channel for Hybrid RF/FSO link, the authors in [12] present a closed-form throughput expression. In [13], feasibility of RF/FSO is investigated and a hybrid link with the different decoders and encoders is considered. In [14], data rate and link reliability of Hybrid RF/FSO scheme are considered. Hybrid RF/FSO communication system can also be established by multiplexing the data over both the links (FSO and RF) for high-speed trains (HSTs). But, in the above case, the power is equally divided between both FSO and RF links. This is also called integrated switching [16–18].

In our previous work [19], we have proposed an integrated switching scheme in Hybrid RF/FSO for high-speed trains and compared its performance with the individual switching scheme. Now, in this paper, we have extended our work and analyzed the effect of scintillation loss and fog on the performance of Hybrid RF/FSO communication system in HST. The system contains rotating transceiver to minimize the impact of handover and to combat the effect of atmospheric conditions. Our contribution in this work is as follows:

(1) In this paper, rotational transceivers are installed on the ground base stations and on top of the train to increase the effective coverage area of every ground base station.
(2) Different switching methods for the high-speed train Hybrid RF/FSO communication system have been used. Three transceivers are installed on the train (one is rotating and two are fixed). During handover, a fixed transceiver (FSO system) is used to switch from source to target base station and rotating transceiver (RF/FSO link) is used to make continuous links between the ground and train. Thus, the users sitting inside the train do not face a handover problem. On the other side, when visibility is less than the RF link on the train is used to communicate with the ground base station.
(3) Performance of integrated RF/FSO scheme with RF and FSO link is compared.
(4) A MATLAB simulator is used to justify the result.

Further, the manuscript is categorized as follows. In Sect. 2, the network architecture of HSTs is introduced. The system model for the Hybrid RF/FSO scheme

is introduced in Sect. 3. Section 4 presents the simulation results, followed by the conclusion in Sect. 5.

2 Network Architecture of High-Speed Train

Figure 1 shows stratified two-hop network architecture which consists of radio frequency/free-space optical transceivers and is used to provide high-speed data connection for high-speed train. Inside the train, handheld devices are connected to access points and the access points are further connected to a mobile router by fiber-optic link. Mobile router and three transceivers (rotating RF/FSO transceiver and fixed transceiver) are connected to the central controller (CC) of the rail via fiber link. CC is liable for controlling the system to maximize performance of the network. The three transceivers (two fixed and one rotating) are fixed at the same place and position on the train. So ME, MR, APs, CC, and three transceivers create a mobile network. Every base station located near the railway track is connected to a rotating transceiver (one FSO and one RF). When HST moves along the railway track and when the weather is clear, FSO is used to connect to the base station otherwise RF is used.

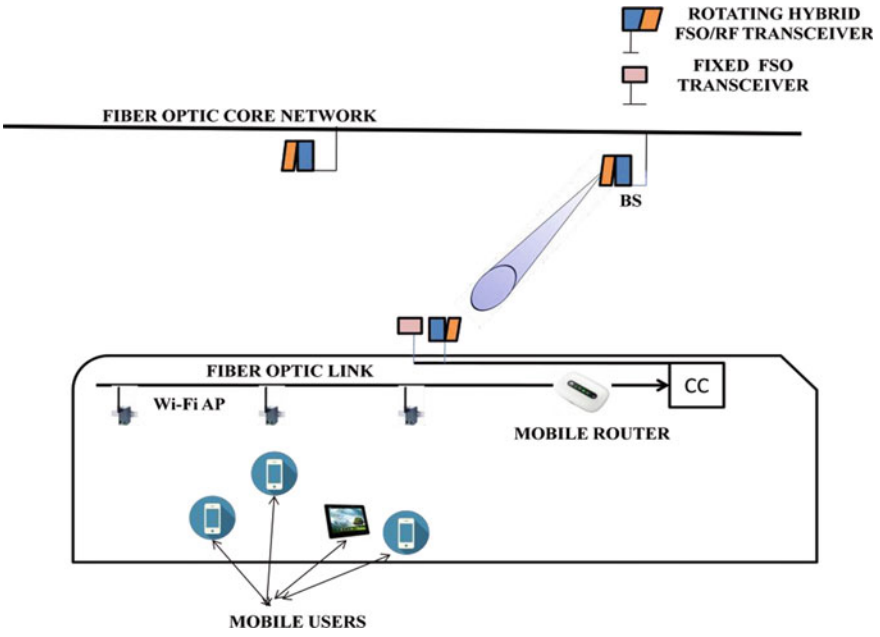


Fig. 1 Network designing of HST

3 System Model for Hybrid RF/FSO System

3.1 Received Signal for the Individual Scheme

In this case, the power is fully drained either on the radio frequency (RF) link or on free-space optical (FSO) link (Fig. 2).

Received Signal For FSO System. We first develop the system model in which radio frequency link is used as a backup link when the quality of the free-space optical link degrades. In the individual scheme, at any given time instant, only one link (either FSO or RF) is in on mode and the other one is in idle mode. Received signal for the FSO link can be written as [11]

$$\Psi_0 = \mathrm{R} X_o \mathrm{J} + N_o \tag{1}$$

Here, J denotes the information symbol having average power V_t^o, N_o is Additive white Gaussian Noise (AWGN) with mean zero and standard deviation Δ_{n0}, and responsivity of the receiver is represented by R which is assumed to be 1. X_o is the intensity gain and it is defined as

$$X_o = X_{fo} X_{lo} \tag{2}$$

where X_{fo} and X_{lo} are the atmospheric turbulence and path loss induced fading coefficients. The path loss induced fading coefficient can be written as

$$X_{lo} = \left(\mathrm{erf}\left(\frac{\sqrt{F}}{\psi Z} \right) \right)^2 \quad e^{-\mu_0 Z} \tag{3}$$

where $\mathrm{F} = \frac{\pi D^2}{4}$ is the aperture area of the receiver having the diameter denoted by D, divergence angle of the beam is denoted by ψ (in radian), link distance is denoted by Z in km and weather-dependent coefficient is denoted by μ_0 in dB/km.

For optical wireless link, fog is the main attenuation factor that cannot be ignored. Line-of-sight limits the communication through FSO link. But as the visibility distance increases, the signal attenuation due to fog will decrease and the capacity of

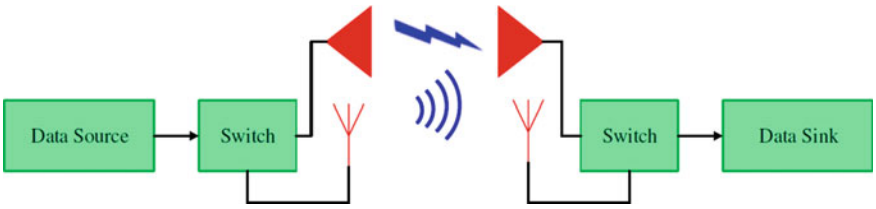


Fig. 2 Individual FSO/RF communication system

the system will also increase. Attenuation coefficients due to fog can be defined by [20, 21]

$$\mu_0 = \frac{10\log(L\%)}{L(\text{km})}\left(\frac{\gamma_1}{\gamma_0}\right)^{-p} \tag{4}$$

Here, L denotes the detectable or visibility distance (in km); in the clear sky, the transference of air drops is denoted by L%, optical wavelength is represented by γ_1 (in nm), γ_0 is reference wavelength whose value is 550 nm and p is a function which depends upon visibility range and can be obtained by the Kim and Kruse model. For the Kruse model,

$$\begin{array}{rll} \text{p} = 1.6 & \text{if} & \text{L} > 50\,\text{km} \\ 1.3 & \text{if} & 6\,\text{km} < \text{L} < 50\,\text{km} \\ 0.585\,L^{1/3} & \text{if} & \text{L} < 6\,\text{km} \end{array}$$

For the Kim model,

$$\begin{array}{rll} \text{p} = 1.6 & \text{if} & \text{L} > 50\,\text{km} \\ 1.3 & \text{if} & 6\,\text{km} < \text{L} < 50\,\text{km} \\ 0.16\text{L} + 0.34 & \text{if} & 1\,\text{km} < \text{L} < 6\,\text{km} \\ \text{L} - 0.5 & \text{if} & 0.5\,\text{km} < \text{L} < 1\,\text{km} \\ 0 & \text{if} & \text{L} < 0.5\,\text{km} \end{array} \tag{5}$$

Though, in the atmosphere, a ray of light is defined by its amplitude, phase, and propagation direction that fluctuate along its channel due to scintillation loss which occurs because of environmental turbulence, the change in turbulent channel along with pressure, humidity, and local temperature causes uncertainty in the air refractive index. This can be calculated using the refractive index structure constant (RISC) I_n^2 that also the amplitude and phase of FSO link. The increase in wave frequency leads to increase in speed and intensity of the fluctuations, i.e., scintillations frequency. The scintillation variance for a plane wave with low turbulence is given by Eq. (6), where channel length is denoted by h in km, wavelength of transmitter is γ_0, RISC is denoted by I_n^2 in $m^{-2/3}$. For RF wave and optical wave, the RISC does not have the same value. RF waves are easily affected by uncertainty in the humidity while in FSO, refractive index is the main parameter which is dependent on temperature [22].

$$\mu_0 = 2\sqrt{23.167\left(\frac{2\pi}{\gamma_0}10^9\right)^{7/6} I_n^2 h^{11/6}} \quad [\text{dB/km}] \tag{6}$$

Parameter X_{fo} is a random variable having gamma distribution [11], i.e.,

$$y_{X_{fo}}(X_{fo}) = \frac{2ed^{\frac{e+d}{2}}}{\Gamma(e)\Gamma(d)} X_{fo}^{\frac{e+d}{2}-1} k_{e-d}(2\sqrt{edX_{fo}}) \tag{7}$$

$k_q(\cdot)$ is the modified second order Bessel function of order "q", gamma function is denoted by $\Gamma(.)$, parameters e and d are atmosphere-dependent random variables [23]

$$\mathrm{e} = \left[\exp\left(\frac{0.49\Delta_d^2}{1 + 1.11\Delta_d^{12/5}} \right) - 1 \right]^{-1} \tag{8}$$

$$\mathrm{d} = \left[\exp\left(\frac{0.51\Delta_d^2}{\left(1 + 0.69\Delta_d^{12/5}\right)^{5/6}} \right) - 1 \right]^{-1} \tag{9}$$

where Δ_d^2 is Rytov variance.

Received signal For RF System. Received signal for the RF scheme is written as

$$\Psi_0 = X_R \mathrm{J} + N_R \tag{10}$$

Here, J denotes the information symbol having average power V_t^R and N_R is AWGN with mean zero and variance Δ_{nR}^2. X_R is the channel coefficient which can be written as

$$X_R = \sqrt{X_{lR} X_{fR}} \tag{11}$$

in which, X_{lR} and X_{fR} are the path loss coefficient and channel fading coefficient having Rayleigh distribution, respectively, and X_{lR} can be written as

$$X_{lR} = S_t + S_R - 20\log_{10}\left(\frac{4\pi Z}{\gamma_R} v \right) \tag{12}$$

S_R and S_t are the receiver and transmitter antennas gain, respectively, and Z is the separation between the receiver and transmitter (km) and wavelength of RF channel is represented by γ_R, velocity of the train is denoted by v.

3.2 Capacity for the Integrated RF/FSO Scheme

Overall bit rate of the integrated RF/FSO scheme is defined by

$$\mathrm{C} = \delta_O C_O + \delta_R C_R \tag{13}$$

in which, C_R and C_O are the bit rates of RF and FSO channels defined as [24] (Fig. 3).

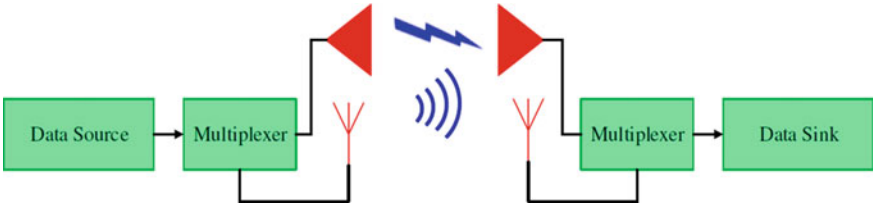


Fig. 3 Integrated RF/FSO communication system

$$C_O = B_0 \log_2\left(1 + \frac{V_t^O X_o^2}{\Delta_{no}^2}\right) \tag{14}$$

$$C_R = B_R \log_2\left(1 + \frac{V_t^R X_R^2}{\Delta_{nR}^2}\right) \tag{15}$$

δ_O and δ_R can be 0 or 1 depending upon which link is active at that time instant.

4 Simulation Result

This section aims to present analytical and numerical results of the system described in Sect. 3. We developed a MATLAB code to study the performance evaluation of

Table 1 Parameters of RF and FSO subsystems

FSO subsystem		
Parameter	Symbol	Value
Wavelength	γ_1	1550 nm
Beam divergence angle	Ψ	2 m rad
Noise variance	Δ_{no}^2	$10^{-14}\mathrm{A}^2$
Receiver aperture diameter	D	20 cm
Refractive index structure parameter	I_n^2	$10^{-15}\mathrm{m}^{\frac{3}{2}}$
Responsivity	R	1 A/W
RF subsystem		
Noise PSD	Δ_{nR}^2	−114 dB/MHz
Receiver antenna gain	S_R	44 dBi
Bandwidth	B_R	100 MHz
Transmitter antenna gain	S_t	44 dBi
Carrier frequency	f_c	60 GHz

the defined network. In Table 1, we discuss the necessary relevant parameters for RF and FSO systems.

Results are obtained for all links, FSO, RF, and integrated RF/FSO individually. The length of the channel is assumed to be 1.5 km. We consider a point-to-point (P2P) integrated RF/FSO system and the separation between the receiver and the transmitter is denoted by Z and bandwidth of the FSO system is assumed to be 1 Gbps and for RF channel, it is 100 Mbps. The velocity of the train is assumed to be 350 km/h.

4.1 Maximum Capacity Versus Link Distance Between Transmitter and Receiver

It is shown in Fig. 4 that the maximum capacity for the integrated RF/FSO system is greater than that of the individual case. As the distance between the transmitter and the receiver increases, losses due to scintillation also increase. Hence the maximum capacity due to scintillation losses decreases. For a given condition, an integrated RF/FSO system always outperforms the individual system.

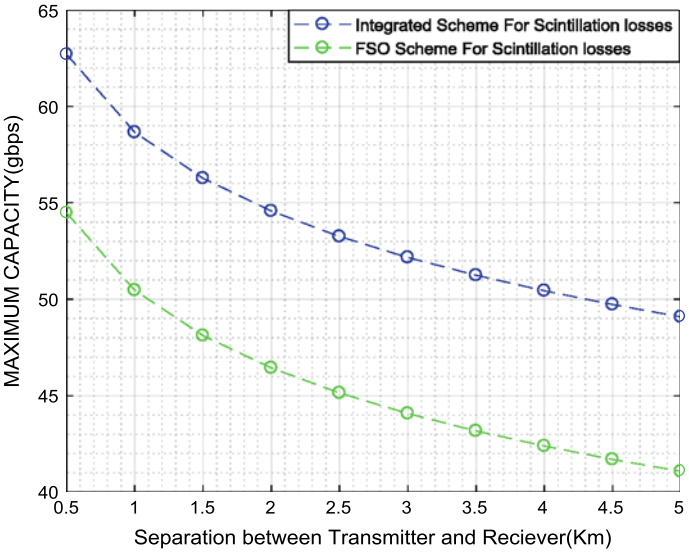


Fig. 4 Maximum capacity versus link distance between the transmitter and receiver (for Z = 3 km and Pt = 0.5 km)

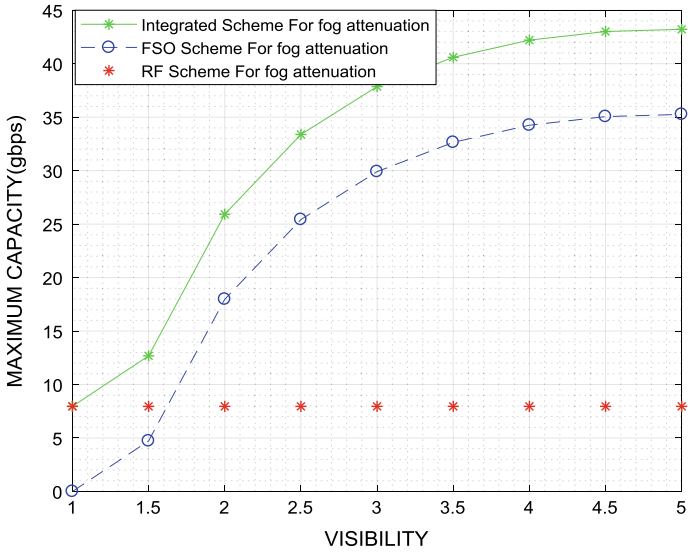


Fig. 5 Maximum capacity versus visibility under foggy condition (for Pt = 0.5 W and Z = 3 km)

4.2 Maximum Capacity Versus Visibility Under Foggy Condition

It is shown in Fig. 5 that with the increase in visibility the signal attenuation due to fog will decrease and the maximum capacity of both schemes increases as the line of sight between the transmitter and the receiver improves. Below 1 km, the capacity of the FSO system is almost zero because of attenuation due to fog.

5 Conclusion

The complementary nature of millimeter wavelength radio frequency link and free-space optical link led to various ways in developing the Hybrid RF/FSO system for transmission of data. The effect of fog and scintillation loss on the maximum capacity of the high-speed train Hybrid RF/FSO communication systems is discussed in this paper. Simulated results explain the supremacy of the integrated RF/FSO system over the individual system which is so remarkable that if the scintillation losses and attenuation due to fog increase, the integrated system always outperforms the individual system.

References

1. Top Ten Fastest Trains in the World, http://www.railwaytechnology.com/features/feature-toptenfastest-trains-in-the-world/. Accessed 01 May 2016
2. V.K. Jagadeesh, I.S. Ansari, V. Palliyembil, P.M. Nathan, K.A. Qaraqe, Channel capacity analysis of a mixed dual-hop RF-FSO transmission system with Málaga distribution. IET Commun. **10**(16), 2119–2124 (2016)
3. T. Rakia, H.-C. Yang, M.-S. Alouini, F. Gebali, Outage analysis of practical FSO/RF hybrid system with adaptive combining. IEEE Commun. Lett. **19**(8), 1366– 1369 (2015)
4. H. Urabe, et al., High data rate ground-to-train free-space optical communication system. Opt. Eng. **51**(3), 031204 (2012)
5. I.S. Ansari, F. Yilmaz, M.-S. Alouini, Performance analysis of free-space optical links over Málaga (M) turbulence channels with pointing errors. IEEE Trans. Wireless Commun. **15**(1), 91–102 (2016)
6. Y. Tang, Hybrid free space optical and RF wireless communication. Ph.D. Dissertation, Department of Electrical Engineering, University of Virginia, 2013
7. F. Giannetti, M. Luise, R. Reggiannini, Mobile and personal communications in the 60 GHz band: a survey. Wireless Pers. Commun. **10**, 207–243 (1999)
8. L. Andrews, R.L. Philips, C.Y. Hopen, *Laser Beam Scintillation with Applications* (SPIE Press, 2001)
9. A.A. Farid, S. Hranilovic, Outage capacity optimization for free-space optical links with pointing errors. J. Lightwave Technol. **25**(7), 1702–1710 (2007)
10. H. Wu, M. Kavehrad, Availability evaluation of ground-to-air hybrid FSO/RF links. Int. J. Wirel. Inf. Netw. **14**(1), 33–45 (2007)
11. J. Li, M. Uysal, Optical wireless communications: system model, capacity and coding, in *IEEE Vehicular Technology Conference* (2003), pp. 168–172
12. N.D. Chatzidiamantis, G. K. Karagiannidis, E.E. Kriezis, M. Matthaiou, Diversity combining in hybrid RF/FSO systems with PSK modulation, in *International Conference on Communication (ICC)* (2011), pp. 1–6
13. H. Kazemi, M. Uysal, F. Touati, Outage analysis of hybrid FSO/RF systems based on finite-state markov chain modeling, in *3rd International Workshop Optical Wireless Communications (IWOW)* (IEEE, 2014), pp. 11–15
14. M.M. Abadi, Z. Ghassemlooy, S. Zvanovec, M.R. Bhatnagar, Y. Wu, Hard switching in hybrid FSO/RF link: investigating data rate and link availability, in *IEEE ICC Workshops* (2017), pp. 463–468
15. S. Sharma, A.S. Madhukumar, R. Swaminathan, Switching-based Hybrid FSO/RF transmission for DF relaying system, in *IEEE Wireless Communications and Networking Conference (WCNC)* (2018), pp. 1–6
16. A. Eslami, S. Vangala, H. Pishro-Nik, Hybrid channel codes for efficient FSO/RF communication systems. IEEE Trans. Commun. **58**(10), 2926–2938 (2010)
17. Y. Wu, Q. Yang, D. Park, K.S. Kwak, Dynamic link selection and power allocation with reliability guarantees for hybrid FSO/RF systems. IEEE Access 13654–13664 (2017)
18. M.N. Khan, S.O. Gilani, M. Jamil, A. Rafay, Q. Awais, Maximizing throughput of hybrid FSO-RF communication system: an algorithm. IEEE Access 30039–30048 (2018)
19. A. Chauhan, P. Verma, Throughput maximization in high speed train using hybrid RF/FSO communication system, in *3rd International Conference, OWT* (Springer, 2019)
20. M. Uysal, J. Li, M. Yu, Error rate performance analysis of coded free-space optical links over gamma-gamma atmospheric turbulence channels. IEEE Trans. Wirel. Commun. **5**(6) (2006)
21. S. Sheikh Muhammad, P. Köhldorfer, E. Leitgeb, Channel modeling for terrestrial free space optical links, in *ICTON* (2005)
22. P.S. Ray, Broadband complex refractive indices of ice and water. Appl. Opt. **11** (1972)

23. M. Karimi, M. Uysal, Novel adaptive transmission algorithms for free-space optical links. IEEE Trans. Commun. **60**(12), 1808–3815 (2012)
24. H. Moradi, M. Falahpour, H.H. Refai, P.G. LoPresti, M. Atiquzzaman, On the capacity of hybrid FSO/RF links. in *Global Telecommunications Conference (GLOBECOM)* (2010), pp. 1–5

Real-Time Analysis of Low-Cost Software-Defined Radio Transceiver Using ZigBee Protocol



Nikhil Marriwala, O. P. Sahu and Anil Vohra

1 Introduction

The Substitute for usual Hardware and Software realization of various wireless communication systems is SDR, which is the latest area of research. Conventional wireless devices were designed and developed to deliver all in one communication service for a particular standard [1]. With the unagitated process of new wireless services and standards 3G, 4G, one resolution device with devoted constituent resources can no longer meet the needs of users. Due to the rapidly changing technology nowadays new standards and their upgrades are launched very fast. To keep our system updated to these changes is a costly affair. The solution lies with the SDR technology which offers to make flexible and user-friendly communication systems [2, 3].

The signal processing tasks in an SDR are implemented all the way through the software, at the transmitter side or at the receiver [4]. The basic idea of an SDR is to provide software control of radio functionality. Traditional radios consist of fixed analog or digital components hence they are custom built for each application. After comparing SDR technology with other existing technologies one can say that SDR offers inherent flexibility, a total reconfigurable environment and serves as the main incentive to engage in this methodology [5, 6].

N. Marriwala (✉)
Electronics and Communication Engineering Department, University Institute of Engineering and Technology, Kurukshetra University, Kurukshetra, India
e-mail: nikhilmarriwala@gmail.com

O. P. Sahu
Electronics and Communication Engineering Department, NIT Kurukshetra, Kurukshetra, India
e-mail: opsahu@nitkkr.ac.in

A. Vohra
Electronic Science Department, Kurukshetra University, Kurukshetra, India
e-mail: anilvohra@kuk.ac.in

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_115

The paper is organized in different sections as follows: Sect. 1 gives the Introduction, Sect. 2 provides an overview of Software-Defined Radio, Sect. 3 tells about the related work done in the field of SDR, Sect. 4 highlights the Generic Transceiver Module designed for ZigBee-based SDR, Sect. 5 provides an overview of the XBee-PRO Hardware used, Sect. 6 shows the analysis of the experimental results based on Eye Diagram, BER, SNR, and Constellation Diagram, Sect. 7 highlights the discussion of result and lastly in Sect. 8 conclusion for the work done has been drawn.

2 About Software-Defined Radio

SDR technology talks about providing a reconfigurable and flexible platform through software instead of conventional hardware [7]. The software is generally preferred over process signal as with the help of the software quick to compilation and loading can be done, hence giving iterative development a much higher cycle rate. An SDR technique enables the developers and researchers working on new technologies and their upgrades to quickly prototype solutions with different approaches and tests them for real-world solutions [8]. The generic software developed for SDR has the advantage that it could be easily shared and adopted by several users working in different areas of wireless communication thus allowing the new developments made in the field of wireless to be reusing the existing code [6, 8]. We have always known that specific hardware used is faster in performance but it is a costly affair, hence SDR provides a general-purpose platform, which has a solution for all the radio applications through minimal use of hardware this, in turn, reduces the cost involved. The inherent features of SDR like flexibility and reconfigurability have made SDR's prime focus area to be used in military and catastrophic conditions [9].

2.1 ZigBee and IEEE 802.15.4

ZigBee is an important technology being used in the area of home automation in the Home Area Network (HAN) [10]. ZigBee is a wireless transmission protocol technology which helps in achieving longer battery life, offers low data rate and low power. The state of devices can be easily changed by the use of ZigBee, which is its biggest advantage. ZigBee helps us to control locks of the door, monitor and adjust thermostat monitor and respond to body changes wirelessly [11, 12]. The IEEE 802.15.4 defines the ZigBee standard which uses a low data rate for transmission through wireless personal area network (LR-WPAN). The clear aim of ZigBee is to create a network with a low data rate protocol for different applications using low power [13].

2.2 Different Uses of ZigBee

The main use of the ZigBee network is to create interoperable products for home automation and industrial use [14]. Some of the applications of ZigBee devices are the communication of Home Automation, Smart Energies, Health Care, Building Automation, Telecommunication Services, Green Power, etc. With the use of ZigBee, one can easily establish interoperability between different standards provided by different manufacturers. For example, ZigBee is being used by Siemens in their system for Floor Level Network Controller named APOGEE, ZigBee is used to provide Interface to heaters, fans, AC units, and lighting through field-level controllers. Kwikset is using ZigBee for SmartCode to lock and unlock doors using a mobile device. ZigBee is being used in Philips 431643 Hue for Personal Wireless Lighting, Starter Pack. ZigBee is also being used in Nest Learning Thermostat—2nd Generation T200577 and ADT are using it for Monitoring Home Security Alarm Systems.

3 Related Work

For any dependable communication system, secured communication is the main part of it. By the use of FEC codes and adding redundant bits to the information to be sent from the transmitter helps the receiver to correct errors. Convolution codes and Block codes are the two categories of FEC codes. These Block codes are usually denoted by n and k where

n total number of coded bits
k number of input bits

The coding theory came into existence with the seminal work of Claude Shannon [15] on "The mathematical theory of communication". Shannon proved it mathematically that we could have an error-free transmission if we can make the bit rate smaller than the capacity of the channel. The error-free transmission can be achieved by the usage of "appropriate" codes. One way of classifying codes is to distinguish between *block codes*, where the redundancy is added to blocks of data and convolutional codes, where redundancy is added continuously [16]. Burst errors can be easily corrected by the use of Block codes something that frequently occurs in wireless communications. By the use of interleaving techniques, one can convert error bursts into random errors. The advantage of the Convolutional codes is that these codes can be decoded easily using a Viterbi decoder. The same algorithm can be used for equalization and joint decoding in convolutional codes. Turbo codes [17] and LDPC codes easily fit into this categorization.

The growth of turbo codes has revolutionized the complete coding theory. The performance of Turbo codes comes out to be very near to the Shannon limit when we use iterative decoding algorithms [18].

4 General Transceiver for ZigBee-Based SDR

This research paper presents a working design in the form of a generic platform that makes use of ZigBee Protocol. This generic platform can further be modified and used for many more protocols at later stages. This generic design was built using NI software, LabVIEW. The investigational setup is shown in Fig. 1. One of the ideal software for testing real-time applications is LabVIEW as it provides a graphical user interface (GUI) environment for the user. LabVIEW helps us to build and analyze the design in a very less amount of time as the software provides integration of all the tools required to build a wide range of applications.

LabVIEW is thus a development environment for solving tedious problems. The tool also provides enhanced productivity, incessant innovation and offers matchless integration with existing legacy software, Internet Protocols (IP), and different hardware. The generic design proposed in this paper ZigBee protocol for SDR provides a low Error Probability for Coherent Receivers and is future-ready for innovation in a faster and more effective way. The BER values achieved through this design are as low as 10^{-5} for real-time transfers which are quite good as compared to other designs. The Block diagram in Fig. 2 shows the VI hierarchy for the Transmitter and Receiver section.

4.1 Transmitter Section

The generic design presented in this paper shows that a fully reconfigurable environment has been designed for the ease of the user. The entire module is defined through the software hence can be termed as an SDR. A fully reconfigurable multi-mode/multi-modulation environment has been provided in the proposed SDR system. For the ease of usage the front panel is loaded with many flexible control which is defined and controlled through the software. To prepare the transmitter for transmitting the data the resource name of the device connected is selected via a com port. To

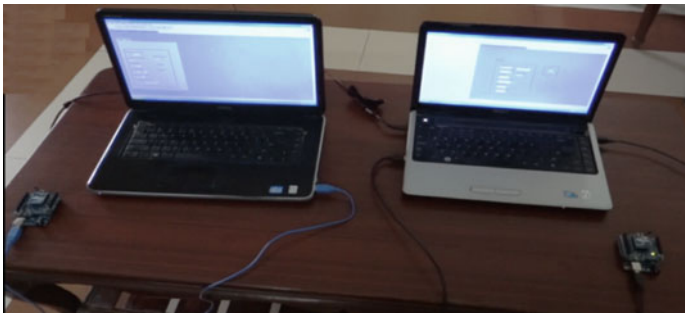


Fig. 1 Experimental setup of generic transceiver module for ZigBee

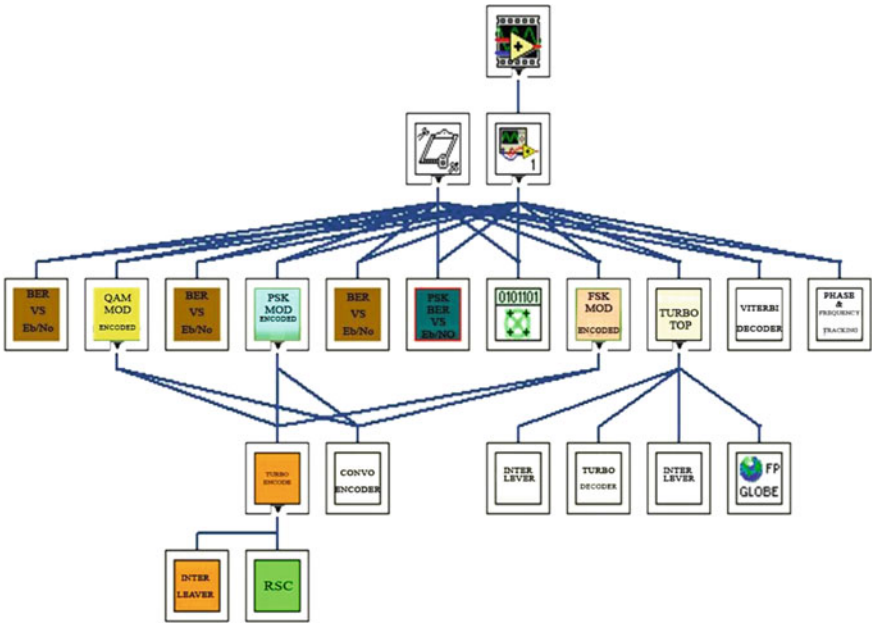


Fig. 2 VI hierarchy of generic transceiver module transmitter and receiver

effectively communicate between the transmitter and receiver the console speed is defined. The proposed SDR transceiver system uses a console speed of 115200/8/N/1, which means that the baud rate for the transmission is 115200 bits/sec, data bits field is 8, parity field is none, and the stop bits field is 10 which is equivalent to 1 in the proposed system. Timeout is set to 10000 for low data and for large data, such as an image it is set to 100000. During the transmission, the Return count shows the bits written by the serial port. To start the writing of the bits to be transmitted, the 'Write' button is turned on. The proposed SDR transceiver system provides the option to select the type of the file to be transmitted from the 'File type' window in the transmitter front panel. There are three different options, i.e., Random binary data, text and image indicating the type of data that needs to be transmitted. The instance used here maps bits to complex-valued symbols for digital modulation schemes such as PSK, FSK, and QAM. The flowchart for the transmitter section is shown in Fig. 3.

4.2 Receiver Section

The flowchart of the receiver section is shown in Fig. 4. On the receiver side a pulse-shaped filter is applied to the input bitstream comprising of PSK, FSK, and QAM modulated symbols. The precise matched-filter coefficients are applied to the input containing the complex I/Q baseband waveform. The output obtained from

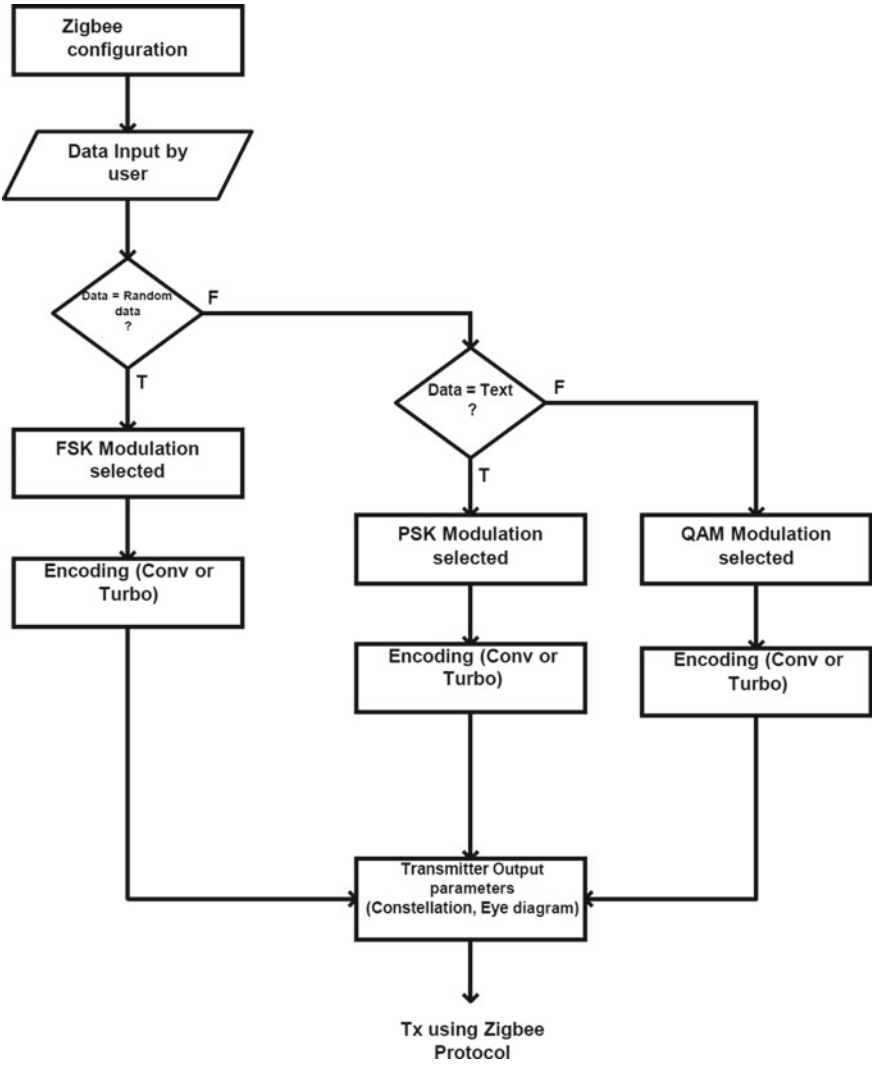


Fig. 3 Flow chart of transmitter section using ZigBee protocol

the matched filter is returned with the period equal to that of an integer number of symbols. The receiver section then looks for the first incidence of the best symbol timing moment in the input complex waveform obtained from the matched filtered. For this, a symbol timing VI has been used in the design for the acquirement of symbols. The mapping of complex-valued modulated symbols generated for PSK, FSK, and QAM is done to a stream of bits generated as output through a user-specified symbol map. The required error correction codes, i.e., the Viterbi decoding algorithm or the Turbo decoding algorithm is applied to retrieve back the original information.

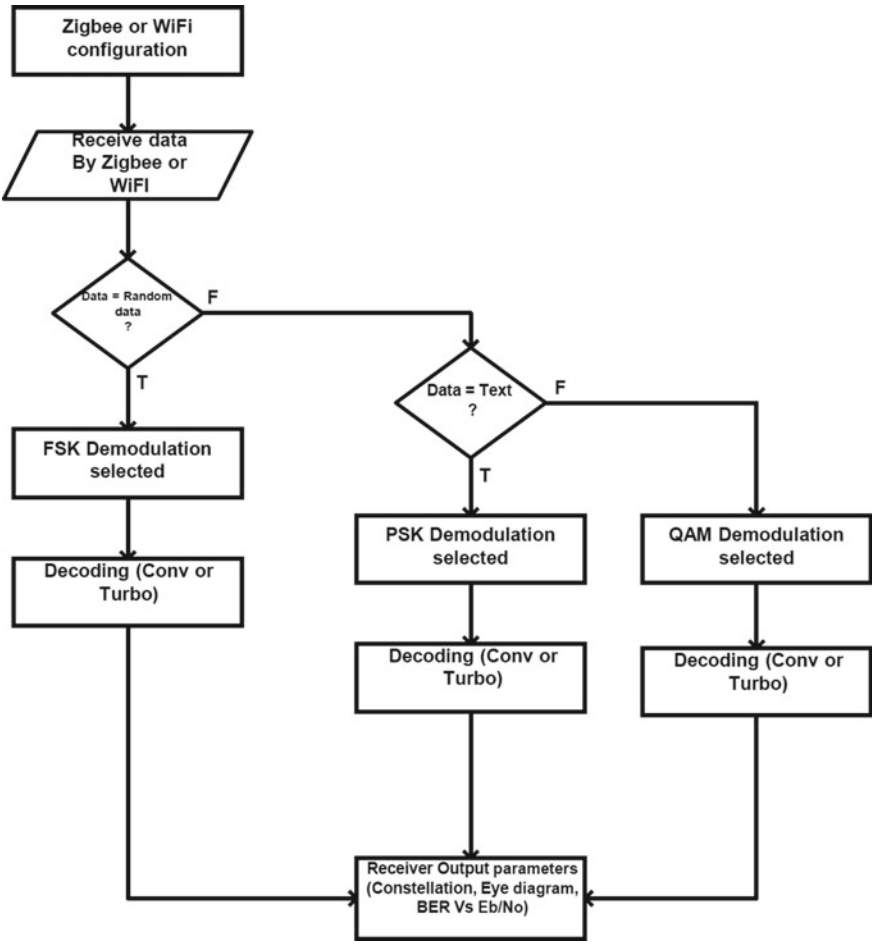


Fig. 4 Flow chart of transmitter section using ZigBee protocol

Figure 5 shows the Front Panel Transmitter Section (a) Input message Tab (b) Modulation Tab for the proposed design highlighting that it provides a fully reconfigurable environment to the user. Many reconfigurable controls and indicators have been provided in the design for the ease of the user. The Visa resource name window is used to select the communication port of the ZigBee transmitter device. The Baud rate for real-time transfer of data is set to 112500 and the Data bits are given the value 8. The user has been given an option to select the Parity accordingly, i.e., we can set it to 0 for odd and 1 for even parity. After giving the input from ZigBee test jig the user needs to select File type, i.e., Random Data, Text, and Image so that the correct file is decoded.

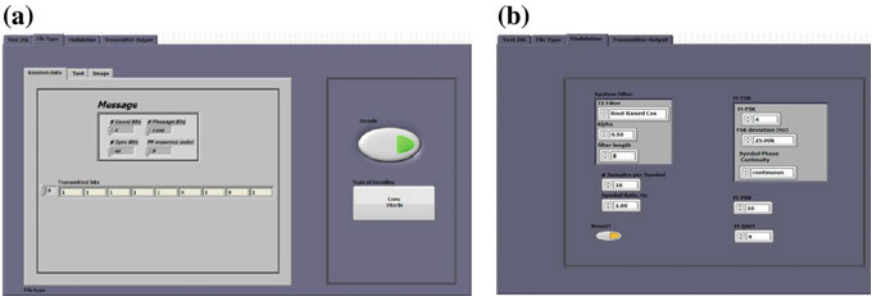


Fig. 5 Front panel transmitter section **a** input message tab, **b** modulation tab

5 About the Hardware Used XBEE-PRO

The XBee-PRO RF Modules support ZigBee protocol transfer and can be easily interfaced with the generic design through the USB port. These modules support the sole need of low-cost and low-power wireless sensor networks. The hardware module used for real-time testing of the SDR requires minimal operating power (3 V) and provides a reliable transfer of data between remote devices. The specifications of XBee-PRO are as given in Table 1.

Table 1 Specifications of XBee-PRO

Specifications	XBee-PRO
Indoor range	Up to 90 m
Outdoor RF line-of-sight range	Up to 3200 m
Output power (Transmit)	50 mW (+17 dBm)
Data rate (RF)	250,000 bps
Data throughput	up to 35000 bps
Data rate (Serial Interface)	1200 bps–1 Mbps
Receiver sensitivity	−102 dBm
Input voltage	3.0–3.4 V
Operating frequency band	ISM 2.4 GHz
Antenna options available	Embedded PCB antenna, integrated whip antenna
I/O interface	3.3 V CMOS UART (not 5 V tolerant), DIO, ADC
Network topologies supported	Point-to-point, point-to-multipoint, peer-to-peer, and mesh
Number of channels available	14 direct sequence channels
Available addressing options	PAN ID and addresses, cluster IDs

Any user application can be updated by making use of the boot loader through the "over-the-air" configuration. The XBee RF Module has the capability to interface many host devices through an asynchronous serial port [19]. The XBee-PRO RF Module used maintains small data buffers to accumulate the serial and RF data at the receiver. The successive characters received are stored by the sequential receive buffer before being processed. The data received through the Radio Frequency (RF) link is stored by the sequential transmit buffer which is further transmitted through UART. The data received serially is made to enter the RF module from the DIN Pin (pin 3). Clear to Send (CTS) flow control is used to evade overflow of the receive buffer if large data is sent from the XBee-PRO RF Module. The Application Program Interface gives an optional way of building modules and steering data at the host application layer. The data frames are sent by the ZigBee module to the application comprising of the status packets and payload information obtained from the data packets received. The module uses a mode called the Command Mode through which one can modify or read parameters of the RF Module. In this mode, the serial characters coming in are interpreted as commands.

6 Analysis of the Experiment Results Using Eye Diagram, BER and Constellation Diagram

The analysis of the experimental results is done through the constellation diagram, Eye diagram and plots between BER and E_b/N_o curves which are shown in Figs. 6, 7 and 8a, b, c, d respectively. The constellation plot shows a repetitive "snapshot", with values only at the decision points. The analysis of phase errors and amplitude errors, at these points, is done with the help of the illustration of the decision points. The constellation diagram is used to analyze different types of signal impairments. Figure 6a–d represents constellation diagrams for 4-QAM, 8-QAM, 32-QAM, 64-QAM modulation scheme using convolution coding for different values of SNR. The common parameters used during the simulation of the proposed SDR transceiver system are, time out = 100000 s, baud rate = 112500 bps, file type, data bits = 8, number of samples per symbol = 16, filter length = 8 bits, and symbol rate = 1 kHz. Based on these input parameters BER versus E_b/N_o curve for the proposed SDR transceiver using different values of (M = 2, 4, 8, 16, 32, 64, 128, 256) have been simulated using channel codes convolutional codes and Turbo Codes with raised cosine filter and root raised cosine filter. The output graphs plotted between BER versus E_b/N_o shown ahead for different M-ary modulation schemes shows the required value of E_b/N_o in decibels to achieve the desired BER of 10^{-5} for the selected M-ary modulation method.

The digitally modulated signal is analyzed by the eye diagram. Different eye diagrams for M-ary modulation schemes have been generated; as shown in Figs. 7 and 8a–d one for I-channel data and Q-channel data using PSK modulation for Turbo and convolution coding.

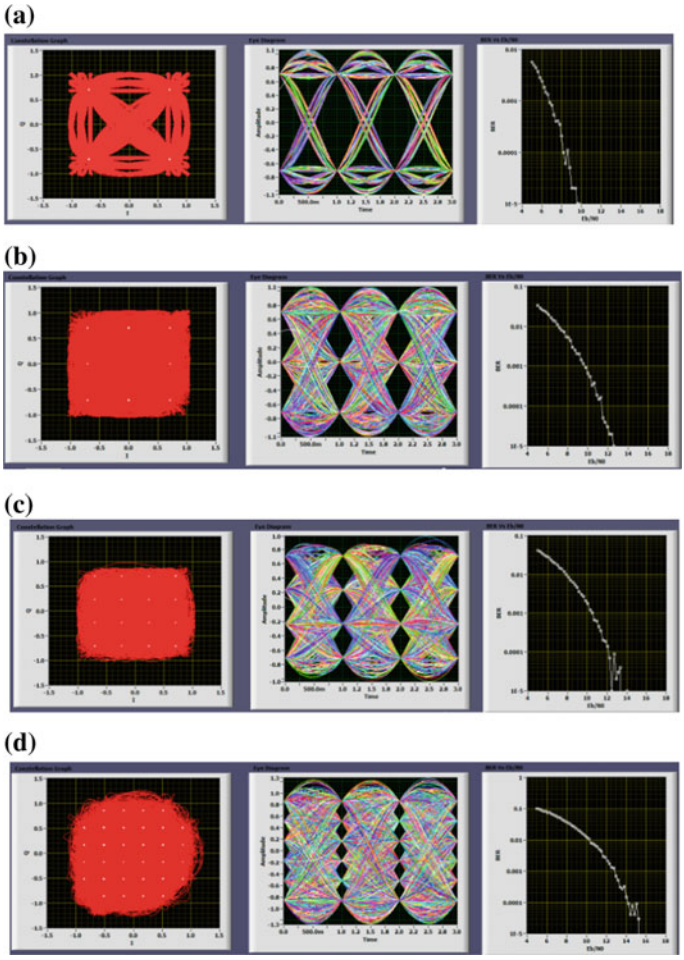


Fig. 6 Outputs for M-QAM using convolution coding shown with the help of constellation diagram, eye diagram, BER versus E_b/N_o: **a** 4QAM, **b** 8 QAM, **c** 16 QAM, **d** 32 QAM

An eye diagram effectively shows the I and Q magnitude with respect to time. An "eye" is shaped according to the symbol decision times whereas the I & Q magnitudes are shown individually. The eye diagram is used to locate optimal decision points. The plots represented through the eye diagram and the constellation diagrams, help to analyze the complete signal.

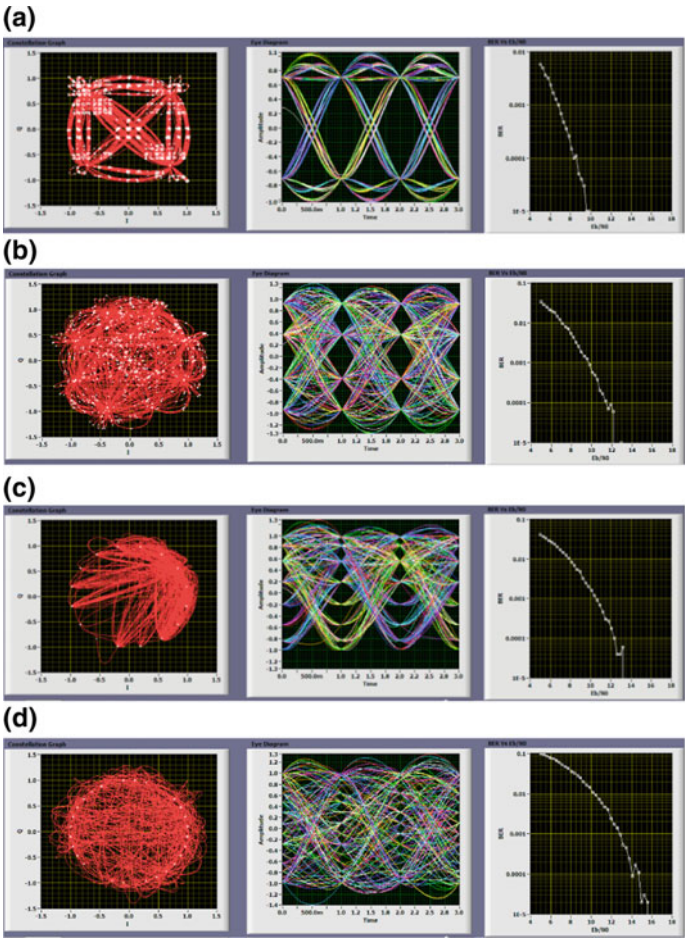


Fig. 7 Output curve for M-PSK using turbo coding represented through constellation diagram, eye diagram, BER versus E_b/N_o: **a** 4PSK, **b** 8 PSK, **c** 16 PSK, **d** 32 PSK

7 Result Discussion

In this section, the analyses of results for different M-ary transfers performed using different M-ary modulation schemes (M-FSK, M-PSK, M-QAM) for simulated and real-time transfer of information. Through the graphs shown in Fig. 9, we see that at 15 dB of E_b/N_o our designed system gives a BER of 10^{-5} for 4-bit QAM in simulation (without coding). The output for real-time transfer of data using 4-bit QAM with Convolution coding gives a BER of 10^{-5} at 9 dB whereas we obtained a BER of 10^{-5} at 10 dB for real-time transfer of data using 4-bit QAM with Turbo Coding. Similarly for real-time transfer of data using 4-bit PSK without coding gives a BER of 10^{-5} at 18 dB, whereas for real-time transfer of data using 4-bit

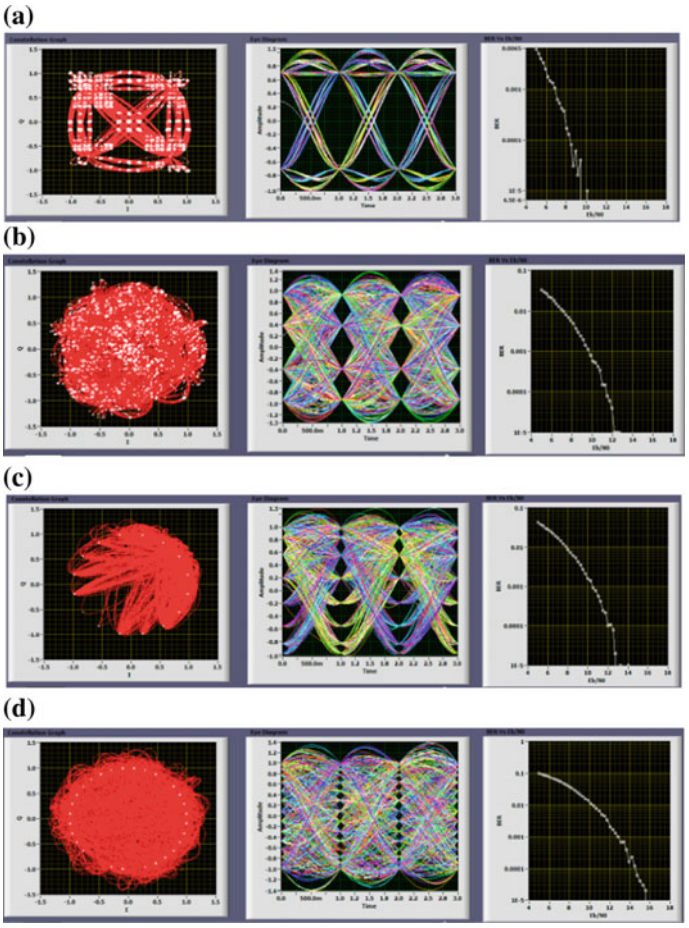


Fig. 8 Output curve for M-PSK using convolution coding represented through constellation diagram, eye diagram, BER versus E_b/N_o output curve for **a** 4 PSK, **b** 8 PSK, **c** 16 PSK, **d** 32 PSK

PSK with Convolution coding we have obtained a BER of 10^{-5} at 14 dB and a real-time transfer of data using 4-bit PSK with Turbo Coding gives a BER of 10^{-5} at 12 dB. Through this experimental setup, it can be seen successfully that BER achieved for secured real-time transfers of information is much lower than achieved previously which was of the order of 10^{-3} at >20 dB. We have been able to design a low-cost solution for the reconfigurable multi modulated SDR transceiver. The Eye diagram as shown in Figs. 6, 7 and 8a–d also supports that for higher order M-ary digital transfers such as 64, 128, 256 there is minimal signal distortion. With the increase in E_b/N_o ratio, BER tends to decrease which can be clearly seen through the plotted values of BER versus E_b/N_o and the eye diagram. The BER keeps on increasing for higher M-array techniques. By increasing the E_b/N_o ratio we can, in turn, increase the signal power w.r.t. noise energy. In high-order M-ary transfers,

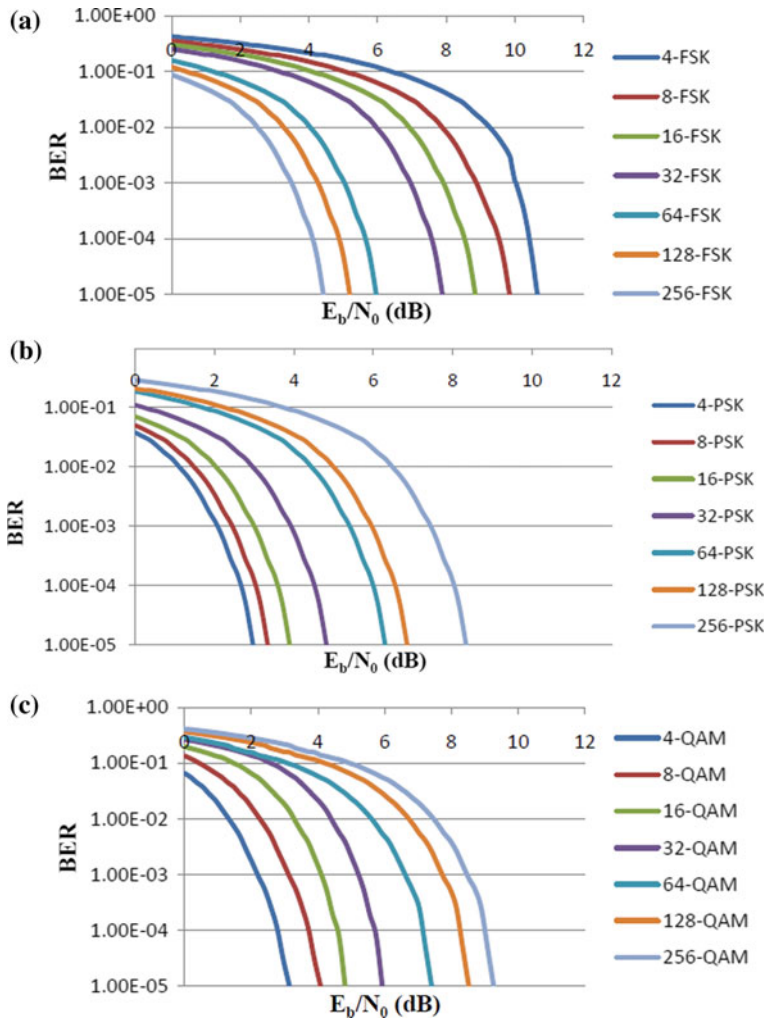


Fig. 9 Comparison graph, between BER versus E_b/N_o for simulated values: **a** M-FSK, **b** M-PSK, **c** M-QAM

we have to use more bits for making a symbol thus they are packed closely in the signal constellation which is also shown by the Figs. 6, 7 and 8a–d. With the help of the result analysis, we are able to prove that the use of FEC coding techniques offers higher data rates for long-distance communication in less amount of time. By analyzing Figs. 9 and 10 one can say that Turbo Coding provides a better BER to E_b/N_o ratio than the Convolution coding in real-time transfer. The detailed analysis of the designed generic system shows us that as we select higher order M-ary digital modulation techniques, i.e., 32, 64, 128, 256 the BER keeps on increasing which is clearly shown in Figs. 9 and 10. The distance between the receiver and Transmitter

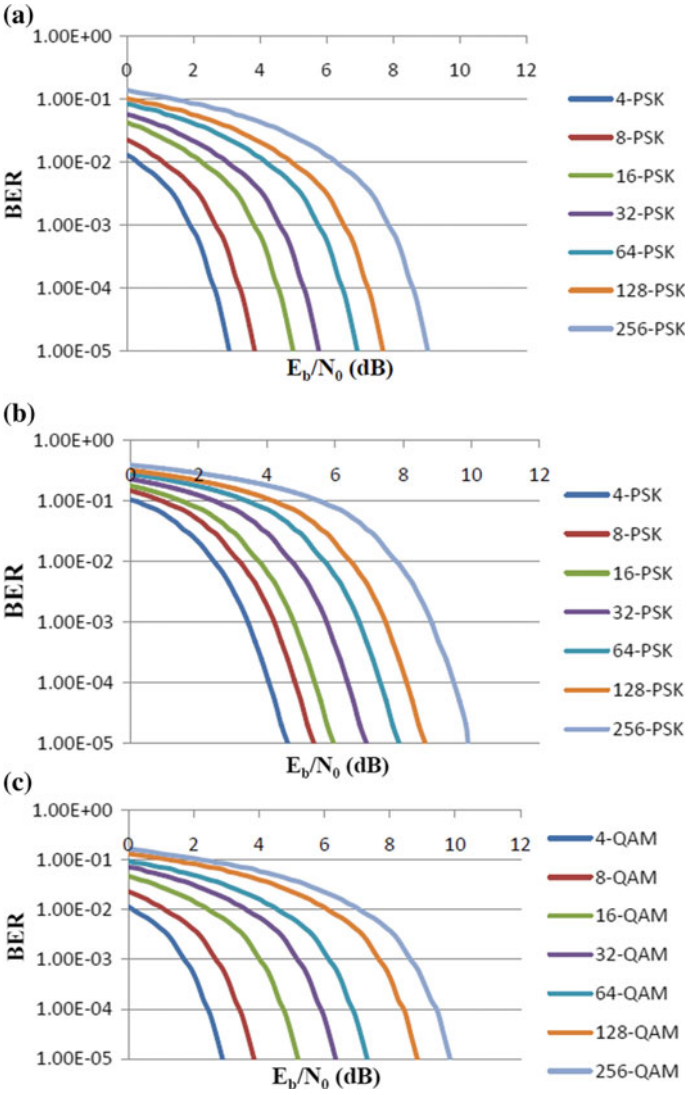


Fig. 10 Comparison graph, between bit error rate versus E_b/N_o for real time communication using XBee-Pro: **a** output graph for M-PSK and turbo coding, **b** output graph for M-PSK and convolution coding, **c** output graph for M-QAM and convolution coding

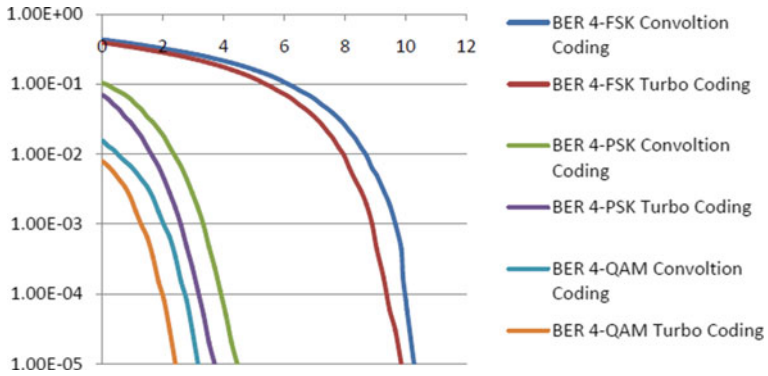


Fig. 11 Comparison between outputs for 4-QAM, 4-PSK, and 4-FSK with convolution and turbo coding used in the proposed SDR

has been kept at 10 m while taking these results. Figure 11 highlights the comparison between outputs for 4-QAM, 4-PSK, and 4-FSK with convolution and turbo coding for the proposed SDR. Table 2 gives the **c**omparison of the features of the proposed SDR transceiver system using the XBee-Pro RF module simulated using LabVIEW for 4-QAM, 4-PSK, and 4-FSK. Table 3 shows the comparison table depicting the features of different Existing SDR Systems with the proposed SDR system. Through this table, it is quite evident that the proposed SDR system shows a much improved performance in terms of BER and E_b/N_o than the other existing systems.

8 Conclusion

In this paper, the design and testing of a low-cost transceiver for SDR systems using ZigBee protocol has been presented. In this generic design, real-time data in the form of (Binary bits, Text and image) has been transmitted and received to verify the functionality of the designed system using three digital modulation techniques M-PSK, M-FSK and M-QAM, and the data transfer is secured using the FEC coding techniques namely the Convolution and the Turbo codes. The main of this design was to bridge the gap between the virtual and real world and also to ensure that measurement of real error rate, for SDR systems can be made easily and effectively. This system completely fulfills the standard of SDR as we have used universal hardware to test the different M-ary digital modulation schemes as the software designed which is the heart of the wireless transmission through the hardware has been made entirely reconfigurable which can be adjusted accordingly by the user on the basis of its requirements. The platform used to build, test and analyze the entire system is totally reconfigurable and user-friendly.

In the designed SDR transceiver system the signal parameters can be changed by the user according to the requirement and the data can be sent in a secured fashion.

Table 2 Comparison of the features of proposed SDR transceiver system using XBee-Pro RF module simulated using LabVIEW for 4-QAM, 4-PSK, and 4-FSK

Features	Refs.					
	Proposed SDR transceiver system using XBee-PRO test jig for real-time data transfer for 4-QAM		Proposed SDR transceiver system using XBee-PRO test jig for real-time data transfer for 4-PSK		Proposed SDR transceiver system using XBee-PRO test jig for real-time data transfer for 4-FSK	
	Convolution coding	Turbo coding	Convolution coding	Turbo coding	Convolution coding	Turbo coding
Bandwidth	10 kHz–2.4 GHz	10 kHz–2.4 GHz	10 kHz–2.4 GHz	10 kHz–2.4 GHz	10 kHz–2.4 GHz	10 kHz–2.4 GHz
Data rate (Mbps)	1	1	1	1	1	1
Application	Wide application	Wide application	Wide application	Wide application	Wide application	Wide application
BER	10-May	10-May	10-May	10-May	10-May	10-May
E_b/N_o (dB)	3.65 dB	2.88 dB	5.76 dB	3.46 dB	10.4 dB	10 dB
Reliability	High reliability	High reliablity	High reliablity	High reliablity	High reliablity	High reliablity
Hardware complexity	Easy modification	Easy modification	Easy modification	Easy modification	Easy modification	Easy modification
Hardware cost	Very low	Very low	Very low	Very low	Very low	Very low
Implementation on FPGA	Very easy implementation	Very easy implementation	Very easy implementation	Very easy implementation	Very easy implementation	Very easy implementation
Modulation	FSK, PSK, QAM	FSK, PSK, QAM	FSK, PSK, QAM	FSK, PSK, QAM	FSK, PSK, QAM	FSK, PSK, QAM

Table 3 Comparison table depicting the features of different existing SDR systems with proposed system

Features	Refs.								
	Existing SDR systems								Proposed SDR transceiver system
	U. Ramacher [20]	Bhalchandra B. Godbole [21]	Radek Martinek et al. [22]	Fumiyuki Adachi [23]	O. Jignesh, Y. Patel, P. Trivedi [24]	Jonathan Roth [25]	Qiyue Zou [26]	Mohamed Al Wohaishi [27]	N. Marriwala et al.
Bandwidth	7 GHz	2.5 MHz	20 MHz	–	2.5 GHz	–	2.5 GHz range	20 MHz	2.4 GHz range
SNR	25 dB		–	–	20 dB	–	27.5 dB	25-30	15 dB
Data rate (Mbps)	3	25 Mbps	132 Mb/s	64 Kb/s	25	–		132 Mb/s	1 Mbps
Application	–	–	Real measurement of error rate	–	Video image	–	–	–	Real measurement of error rate
E_b/N_o dB	18 dB	5-12 dB	28 dB (32-PSK); 21 dB (64-QAM)	40 dB	35 dB	2.7 dB	–	6 dB	3 dB
BER	10^{-5}	10^{-5}	10^{-4}	10^{-4}	10^{-4}	10^{-1}	10^{-4}	10^{-3}	10^{-5}
Reliability	High	–	–	–	High	Less	–	–	
Hardware complexity	Non-complex	Highly complex	Complex	Complex	Low bandwidth	Complex	Less complex	–	Less complex
Hardware cost	Low	High	High	High	Low	High	Low	–	Low
Implementation on FPGA	–	–	–	–	–	–	–	–	Yes
Modulation	OFDM	BPSK, QPSK, 16-QAM, 64-QAM	M-QAM, M-PSK	QDPSK	FSK, QPSK	QPSK	QPSK	M-QAM	M-QAM, M-PSK, M-FSK
Coding techniques	–	–	–	–	–	–	–	–	FEC coding (Convolution and Turbo Coding)

The proposed system demonstrates that without replacing the existing hardware components newer protocols can be accommodated by merely making a few changes in the software of the system. This paper shows that how effectively an SDR transceiver system can be built entirely through software. The developed SDR system is further put to test for flexibility by altering the different protocols and physical layer application in real-time. XBee-PRO test jig supporting ZigBee module has been used for testing the SDR system for real-time transmission and reception of binary data, text and image. The design proposed in this paper can be easily reconfigured accordingly in a very small time thus increasing the efficiency of the modified system. This generic platform is universal in nature and can be further programmed for different protocols. The main aim of this proposed system was to bridge the gap between the virtual and real world and also to ensure that measurement of real error rate, for SDR systems can be made easily and effectively. This system completely fulfills the standard of SDR as universal hardware has been used to test the different M-ary digital modulation schemes as the proposed software which is the heart of the wireless transmission through the hardware has been made entirely reconfigurable which can be adjusted accordingly by the user on the basis of its requirements. The proposed SDR system offers a multi-modulation, low-cost, and low-power solution. The data transmitted from the transmitter end has been recovered with very low BER at the destination. Analysis of experimental results show that, compared to a conventional transceiver with similar functions, the proposed SDR transceiver system achieves a very low BER of the order of 10^{-5} at around $E_b/N_o = 3$ dB for 4-PSK and $E_b/N_o = 2.6$ dB for 4-QAM using convolution coding, and for turbo coding, respectively, and the BER of 10^{-5} is achieved at around $E_b/N_o = 2.1$ dB for 4-PSK and $E_b/N_o = 1.9$ dB for 4-QAM using ZigBee protocol for real-time data transfer.

References

1. J. Rohde, T.S. Toftegaard, Adapting cognitive radio technology for low-power wireless personal area network devices. Wirel. Pers. Commun. **58**(1), 111–123 (2011)
2. J. Mitola, Technical challenges in the globalization of software radio. IEEE Commun. Mag. **37**(February), 84–89 (1999)
3. M. Grimm, M. Allen, J. Marttila, M. Valkama, R. Thoma, Joint mitigation of nonlinear RF and baseband distortions in wideband direct-conversion receivers. IEEE Trans. Microw. Theory Tech. **62**(1), 166–182 (2014)
4. N. Marriwala, O.P. Sahu, A. Vohra, Novel design of a low cost flexible transceiver based on multistate digitally modulated signals using Wi-Fi protocol for software defined radio. Wirel. Pers. Commun. **87**(4), 1265–1284 (2016)
5. M.N.O. Sadiku, C.M. Akujuobi, Software-defined radio: a brief overview. IEEE Potentials **23**(4), 14–15 (2004)
6. N. Marriwala, LabVIEW based design implementation of M-PSK transceiver using multiple forward error correction coding technique for software defined radio applications. J. Electr. Electron. Eng. **2**(4), 55 (2014)
7. N. Marriwala, O.P. Sahu, A. Vohra, Design of a hybrid reconfigurable software defined radio transceiver based on frequency shift keying using multiple encoding schemes. Egypt. Inform. J. **17**(1), 89–98 (2015)

8. N. Marriwala, O.P. Sahu, A. Vohra, 8-QAM software defined radio based approach for channel encoding and decoding using forward error correction. Wirel. Pers. Commun. **72**(4), 2957–2969 (2013)
9. H.-K. Song, S.-J. Yu, W.-J. Choi, Efficient decoding scheme for cooperative communication using hierarchical modulation in the mobile communication systems. Wirel. Pers. Commun. (2015)
10. T. Mangir, L. Sarakbi, H. Younan, Analyzing the impact of Wi-Fi interference on zigbee networks based on real time experiments. Int. J. Distrib. Parallel Syst. **2**(4), 10 (2011)
11. L. Mraz, V. Cervenka, D. Komosny, M. Simek, Comprehensive performance analysis of zigbee technology based on real measurements. Wirel. Pers. Commun. **71**(4), 2783–2803 (2013)
12. S. Ouni, Z.T. Ayoub, Cooperative association/re-association approaches to optimize energy consumption for real-time IEEE 802.15.4/zigbee wireless sensor networks. Wirel. Pers. Commun. **71**(4), 3157–3183 (2013)
13. V. Nithya, B. Ramachandran, V. Bhaskar, Energy efficient coded communication for IEEE 802.15.4 compliant wireless sensor networks. Wirel. Pers. Commun. **77**(1), 675–690 (2014)
14. P. De Valck, I. Moerman, D. Croce, F. Giuliano, Exploiting programmable architectures for WiFi/zigbee inter- technology cooperation. EURASIP J. Wirel. Commun. Netw. **212**, 1–13 (2014)
15. C.E. Shannon, A mathematical theory of communication. Bell Syst. Tech. J. **27**(3), 379–423 (1948)
16. C. Fleming, A tutorial on convolutional coding with Viterbi decoding. Spectr. Appl. 1–28 (2002)
17. P.S. Bhanubhai, M.G. Shajan, U.D. Dalal, Performance of turbo encoder and turbo decoder for LTE. Int. J. Eng. Innov. Technol. **2**(6), 125–128 (2012)
18. K.M. Borle, Y. Zhao, B. Chen, A software radio design for communications in uncoordinated networks, in *IEEE International Workshop on Signal Processing Advances in Wireless Communications* (2014), pp. 254–258
19. H. Shiba, T. Shono, K. Uehara, S. Kubota, Design and evaluation of software radio prototype with over-the-air download function, in *Vehicular Technology Conference* (2001), pp. 2466–2469
20. U. Ramacher, Software-defined radio prospects for multistandard mobile phones. IEEE Comput. Soc. **40**(10), 62–69 (2007)
21. B.B. Godbole, D.S. Aldar, Performance improvement by changing modulation methods for software defined radios. Int. J. Adv. Comput. Sci. Appl. **1**(6), 72–79 (2010)
22. R. Martinek, M. Al-Wohaishi, J. Zidek, Software based flexible measuring systems for analysis of digitally modulated systems, in *2010 9th IEEE RoEduNet International Conference (RoEduNet)* (2010), pp. 397–402
23. F. Adachi, K. Ohno, BER performance of QDPSK with postdetection diversity reception in mobile radio channels. IEEE Trans. Veh. Technol. **40**(1), 237–249 (1991)
24. J. Oza, et al., Optimized configurable architecture of modulation techniques for SDR applications, in *International Conference Computer and Communication Engineering ICCCE'10*, no. May (2010), pp. 11–13
25. J. Roth, N. Manjikian, S. Sudharsanan, Performance optimization and parallelization of turbo decoding for software-defined radio. Can. J. Electr. Comput. Eng. **34**(3) (2009)
26. Q. Zou, M. Mikhemar, A.H. Sayed, Digital compensation of cross-modulation distortion in software-defined radios. IEEE J. Sel. Top. Signal Process. **3**(3), 348–361 (2009)
27. M. Al Wohaishi, J. Zidek, R. Martinek, Analysis of M state digitally modulated signals in communication systems based on SDR concept, in *Proceedings of 6th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems*: *Technology and Applications, IDAACS'2011*, vol. 1, no. September (2011), pp. 171–175

Comparative Overview of Profit-Based Unit Commitment in Competitive Electricity Market



Ayani Nandi and Vikram Kumar Kamboj

1 Introduction

The development of the modern economy of various sector depends on electric power which plays an exceedingly important role in our daily life and increases several numbers of power plants with their capacities. Thus, power transmission lines are also increased consequently to connect load centers to generating stations. The electrical energy can be obtained by transfiguration from fossil fuels like coal, natural gas, oil, nuclear and hydro sources. Each element of the system which contains solutions of power plant technology, generator, transformer, power electronic device, power lines, data acquisition, supervisory control, etc. These types of several control elements are important so the system stability, system operation, and optimization system balance and settling is really needed.

To determine the scheduling operation at every interval of a hour of the generating unit with varying loads under different conditions including some different constraints and different environments, the optimization of Unit Commitment Problem (UCP) is required which deals with some optimum amount of effective time in order to meet the load demand per hour basis and the generating unit should be operated. A generating unit scheduling needs satisfaction to a number of operating constraints for achieving the reduction of minimum total production cost. The constraints are restricted for individual generating units, generation limit, capacity limits, minimum uptime, minimum downtime for the first and last hour and also limited some constraints such as spinning reserve constraint, power balance, group constraints, ramp rate, etc. The aim of UCP is to better identification for the time period of medium term (days and weeks), associated with in terms of the generating units of their status including committed or uncommitted and status related to power output also. To maintain minimize the total cost for production this agenda has

A. Nandi · V. K. Kamboj (✉)
School of Electronics and Electrical Engineering, Lovely Professional University, Punjab, India
e-mail: vikram.23687@lpu.co.in

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_116

to propitiate the demand under the system environmental, technical and operating constraints. In order to calculate the power for each unit which is associated within the previous phase to solve UCP in online period including hours and minutes and short-term period, the problem of economic load dispatch is solved easily to meet the system demand in real time operation. One of the major problems can be solved by unit commitment by proper production of energy scheduling. The main objective is to decide a blend of the electrical generators which are available in the unit and their respective power output must be scheduled to satisfy the demand of load at total minimum production cost in a distinguishing period of time which basically changes from 24 h to 1 week.

In order to deal with maximum profit, it is possible to supply power including a small amount of fuel consumption and the least amount of losses with the help of the optimization technique. The market competition provides better opportunities for financial resources and reliability of electricity at cheaper cost through many power companies that are growing through their proper roles, objectives, and utilities. One of the main problem for the power suppliers to improve their own profit by optimal generation scheduling is known as Profit-Based Unit Commitment (PBUC) problem. In restricted environment, the scheduling problem is constructed with two important aspects such as solving the problem of the condition of commitment for available units and reserve capacity and generation allocation of the committed unit is determined. So the main objective of PBUC problems is nonlinear in nature and complicated when compared with the counterpart of traditional unit commitment. Through all of these the main function is the differences in PBUC problems with traditional problems is to introduce for change in constraint modeling of the system. Such constraints relevant to load satisfaction and capacity of reserve of the system which acts as capacity of generation in markets of energy and markets of reserve. Some generation company uses such procedures for their security against for such kind of multiple uncertainties of energy markets including trading of financial sector and physical sector. The cost minimization problem in traditional aspects is transformed into profit-based maximization problem. The problem of scheduling has to be solved by consideration of production cost and by satisfied the total system of operating constraints which reduce the independence the selection of the units for starting up or shutting down. Not only this scheduling but also those constraints need to be satisfied the status of the units for the minimum uptime and minimum downtime to the production of power limit and limit of the capacity to increase the maximum ramp-up rate and to maximum ramp-down rate to such other operating characteristics and spinning reserve.

In the type of earlier research which was focused basically for the market of spot power and market in reserve capacity which is based on the energy market price. PBUC problem is also known as complex problem which increases nonlinearity and dimensionality of optimization associated with several types of constraints. In past, the PBUC included many techniques of optimization aspects and such kind of considerable interest which are proposed for the solution of PBUC problem in power sector. To analyze the maximum profit, the Profit-Based Unit Commitment appraises reserve and power which can be proposed for acceptance in market. To solve PBUC

problem, the combination of stochastic search algorithm and conventional algorithm including hybrid method and bio-inspired algorithm have proposed. For the market of power sector deregulation of power involved a composition among some companies for generation of power which is profitable. These offers must be decided in duration for short time period. The most algorithm support solution including more time for computation. The main profit of GENCOs can be done by bringing up to date the Lagrange multipliers and using hybrid algorithm which gives better solution also. To match the power demand with supply can be done by independent system operator (ISO), which create the competitions among power generation companies where supplies will itinerary their power generation units for maximizing their profit as per prediction of the price of electricity.

2 Literature Review

Unit commitment is an important optimization problem, which is used to obtain the optimal schedule for operating the generating units in the most economical manner to meet the required load demand and at the same time to accomplish the constraints' requirements. Unit commitment has attracted many researchers over past few years. Most of the researchers had explored the solution of single area unit commitment problem. The researchers are looking forward to finding new techniques and methodologies for solving unit commitment problem. The first paper in the field of unit commitment was introduced by Baldwin et al. [1] in 1959. Since 1959, appreciable research work has been done in the area of single area unit commitment problem. Lots of research papers are available related to single-objective and multi-objective or multi-criterion optimization problems and single area unit commitment problem. A brief review of the literature has been presented in the subsequent sections:

Reddy et al. [2] proposed a new technique in scheduling for generation in thermal power plants rigged with Amine-based post-combustion carbon capture technology, which develops the sensitivity of resource of coal and fuel used in fuel combustion or coal combustion through interpreted model for planning of operation the unit in thermal generators over unit commitment. A new methodology is constructed by Yang et al. [3] on analysis of gap theory to appraise the strategy of operation of profitability for united power and heat units in a liberalistic market of electricity. The level of risk can be evaluated using this methodology which is taken into discussion whether there is any kind of risk arises for generating company. The sale price is taken into consideration of uncertainty when there is implemented some information about decision of gap theory to model its vaporization around predicted value. Aghaei et al. [4], Bavafa et al. [5] designed a heuristic methodology on Wind-Thermal Unit Commitment of Bi-objective Probabilistic Risk-Based. To minimize the operational risk and cost this model is implemented using risk-based unit commitment model. A new formulated method, i.e., cycle-based unit commitment algorithm presented a new method of process of power re-dispatch to satisfy the ramp rate constraints including up-time and down-time. An analysis of hypothetical unit commitment is created by

Wang and Hobbs [6] in the market of real time for Flexi-ramp or ramp capability. This technique proposed about need for flexible generator due to an increase in the penetration of renewable energy where Flexi-ramp is defined as the capacity of reserve committed unit to accommodate abrupt ramp. To get improved optimization technique while integrating large and complicated number of constraints including meet the load demand, power generating limit, minimum up-time and minimum down-time and spinning reserves, Saravanan et al. [7] implemented a technique which includes Some of the units in generating station have their convinced ranges where restricted operation is included limitations in component of machines and insecurity due to vibration occurred in shaft bearings or steam valve.

Wang et al. [8] invented a new methodology for tackle some issues created by non-dispatch able generation of wind power including the reliable operation and considered a variation of wind including the robust unit commitment (RUC). The most common drawback for these kinds of dispatch which increased load demand on adaptability of resources which is also associated with cost. Thus, the requirement of ramp should be decreased and increased the capability of ramp-up for wind power generation and it also increased the profit which is gained by wind power generation by as long as capabilities of ramp-up with other services of auxiliary unit. A program is designed by Abdollahi and Moghaddam [9] which include load management (LM) and its expansion and demand response programs (DRPs) which is divided into mainly two sections, such as time-based rate (TBR) and incentive-based programs (IBPs). The model is derived which is based on price of elasticity demand and benefits of customer. A review of Stochastic Optimization technique for unit commitment is constructed by Zheng et al. [11]. In industry of power sector the optimization technique has been broadly used to support the process of decision making for dispatch and scheduling the generation of electric power resources which is known as unit commitment which has two major types of innovations on research based on unit commitment and real time operations.

Anand et al. [10] implemented a new integrated optimization process using civilized swarm optimization (CSO) and binary successive approach (BSA) technique for solving the profit-based unit commitment (PBUC) problem to get maximum profit which is obtained by generating companies (GENCOs). PBUC problem deals with continuous and binary variables. These optimization techniques are tested on some system units including various types of operated constraints like minimum up and downtime, balance of load and ramp rate of the unit, etc. Constrained emission for PBUC which are defined as bi-objective function of optimization for transition of global warming and variation of climate in environment by emission of greenhouse gases from the unit of thermal power plant is invented by Senthilvadivu et al. [12]. Using algorithm of Exchange Market (EM) is used to solve the two major problems including reduction of emission of environment and maximize the profit. This access has more capability and stability to solve PBUC problem. After testing this methodology on testing system of 10 units within 24 h of IEEE 39 bus and its results are substantiated about the effectiveness of the methodology for the solution of constraints of emission PBUC in an ambitious market of electricity. Reddy K.

et al. [13, 14] presented binary whale optimization algorithm (BWOA) for solving no convex, complicated, complex, constrained and binary nature of problem of PBUC. The variants including hyperbolic tangential, sigmoid and inverse tangent transfer functions are introduced by BWOA which is tested by systems with various mechanisms of electricity market, i.e. a reserve and energy market and only energy market with different types of methods of reserve payment. The simulated outputs including quality of solution, consistency, and characteristics are compared and discussed with other approaches. The solution quality [18] is taken into consideration to deciding the status of commitment and make profit obtain by GENCOs. Nowadays in the market of deregulation there is freedom of GENCOs for generator scheduling based on profit in energy market. This methodology can be tested for the various reserve market participated scenarios of the unit system for thermal power plants. Lagrange Relaxation (LR)-Differential Evolution (DE) algorithm is used [19] for solving PBUC problem by using LR is used to get solution to the problem of unit commitment and the algorithm DE is implemented to apply up-gradation of LR multiplication. Lagrange Relaxation (LR) with Evolutionary Programming (EP) [29] which used to solve PBUC problem. From the comparison of other research work to increase the profit in generation of electricity market, this study is done by including losses and it should be added the revenue of power. For real time operation in the market of power deregulation, forecasting of electricity market and PBUC problems are most important. To solve PBUC problem, the approach on binary grey wolf optimizer (BGWO) is also presented in another article that is used in unit of thermal power plant to change the ON or OFF state through the procedure using self-scheduling. A heuristic algorithm [15] for optimization on environmental economic unit commitment for combined cooling, heat, and power (CCHP) thermal power system which is integrated with production unit of power to meet the market demand, traditional separate cooling and heating application for valve- point effect for turbine. Binary Variant of Sine Cosine Algorithm [16] to solve PBUC problem in competitive market of electricity. In a day, market of energy as well as reserve market by thermal unit scheduling and commitments of thermal units, the PBUC problem can be solved by GENCOs with aim of maximize profit for forecasts of load. Another technique named Evolutionary Particle Swarm Optimization (EPSO) used to solve PBUC problem [17] for maximum profit in the environment of deregulation and it's done by applying forecast demand of electrical energy and other additional services containing supply also. The design of these methodology shows for 10th units of thermal power plant with GENCOs that this implemented methodology should be performed for quality solution and characteristics of convergence which are superior to the algorithm based on classic PSO methodology. Modified Dynamic Programming using particle swarm optimization (PSO) to solve PBUC problem which is divided into some kind of subproblems like interior and exterior including continuous and discrete nature, respectively. This methodology helps to find out the maximum profit in electricity market about how much power must be taken in for sale and reserve [21]. Swarm Intelligence [27] with limitations of emission for solving PBUC problem. This modified technique is used for controlling and modification purposes or sometimes replacements of generating units.

For emission one of the vital contributions of greenhouse gases to the environment through increasing the usages of fossil fuel in power plants. Singhal et al. [20] presented a new method on dynamic economic dispatch (DED) and binary fish swarm algorithm (BFSA) for increasing the profit of GENCOs to get solution on PBUC problem consisting power reserve and power generation through the whole day in competitive electricity market. The effectiveness of validation of these techniques on 100 units and 10 units of thermal power plant in a whole day in power generation market for increase time of computation and the rate of profit of GENCOs. Venkatesan et al. [22] implemented a new method to get solution on multi-area unit commitment (PBMAUC) problem using evolutionary algorithm on particle swarm optimization (EPPSO) technique which has an operation depend on cost-saving. This technique was implemented an IBM PC through which in a reasonable time period large types of systems can proceed. Morales-España et al. [23] designed basic action in power based unit commitment of capabilities of quickly start and slow unit using some basic constraints for power generation limit minimum down-time and up-time, start-up and shut down of electric power. The schedule of power represents the step for energy schedule which leads unattainable delivery of electrical energy. The imperialist competitive algorithm (ICA) is used to solve PBUC problem in reorganizing energy market [24]. This technique is basically upgraded version of the evolutionary algorithm which reduces complexity in computation. Prakash and Yuvaraj [25] are designed a table based on Improved Pre-prepared Power Demand (IPPD) to solve PBUC problem. The generating companies scheduled generators to get maximum profit rather than get satisfaction in power demand in the environment of deregulation of power market. PBUC problem under electricity deregulated market [26] in sector of power system to increase efficiency of production of electricity and distribution of electric power at low price with high quality and more reliable and provide security for product. This methodology is tested on system of IEEE 30 bus with 6 numbers of power generating units. Columbus and Sishaj [28] Simon created parallel nodal ant colony optimization (PNACO) technique which follows the method of intelligence of ant which is used to get decision for non-committed generating units in power plant. The problem including economic load dispatch which is followed by using parallel artificial bee colony (PABC) technique represents the behavior like bees and described a technique used to perform operation of parallel distribution for solving the problem of PBUC. Columbus and Simon [30] invented an algorithm based on a parallel artificial bee colony (PABC) using cluster of workstation to solve PBUC problem and maximize the profit. The effectiveness to use this method in power sector to compute the resources to reducing the complexity of time period for system use for large-scale power generation. Power systems ranging from 10 to 1000 generation units are tested for this technique to get the solution of quality and complexity of time with respect to number of the formation of cluster can be analyzed thoroughly. Delarue et al. [31] invented a model associated with PBUC which are utilized and developed by implementing Mixed Integer Linear Programming (MILP) to get expected profit in price based unit commitment problem when inappropriate price projection is applied. This technique is useful to find out the relation between Mean Absolute Percentage Error (MAPE) for price projection and

profit or loss. PBUC in practical field with some limitations [32] of emission due to fossil fuel in power plant. This methodology address the profitable market of power including emission in practically with the help of multi-objective optimization (MO) problem. So, between emission and profit, the trade curve is defined for profiles of different types of energy files. Chandram and Subrahmanyam [33] invented a new methodology to solve PBUC problem using Muller method which have two steps, at firstly information is taken from committed unit of a power plant using table of the units which are committed and secondly the problem of economic load dispatch including nonlinearity are solved by using Muller method which had already tested with three to ten numbers of generating units in power system. Catalão et al. [34] presented a multi-objective method to determine PBUC problem with struggle in profit in a competitive energy market and objectives of emission. This methodology can be tested on standard IEEE 30 bus system. The brief review of aforesaid literature is presented in Table 1.

3 Profit-Based Unit Commitment Problem Formulation

The main objective of PBUC, from power scheduling of the committed generating units in the market of deregulation to get maximum profit of GENCO's is formulated as

Maximum Profit (PF) = Revenue (RV) – Total Operating Cost (TOC)

$$\text{Maximize PF(i, t)} = \text{RV(i, t)} - \text{TOC(i, t)} \tag{1}$$

From the thermal generating units by selling the power where the maximum Profit (PF) can be obtained and the total operating cost is represented as TOC and the RV represent the generated revenue at an appropriate price of energy by selling the generated power which is given as

$$\text{RV} = \sum_{t=1}^{T} \sum_{i=1}^{N} \partial_t \times P_{t,i} \times U_{t,i} \tag{2}$$

when

∂t energy price at tth sub-interval in \$/MW h
Pt, i the generation of power at ith unit at tth sub-interval in MW
Ut, i status of the unit of ith unit at tth sub-interval
T the total interval of time
N the total number of generating units

The summation on startup cost and fuel cost of the thermal unit can be obtained as total operating cost (TOC) which is given below:

Table 1 Brief review of unit commitment and profit based unit commitment problem

Sl. No.	Authors name	References	Year	Brief review	Algorithm	Constraints
1	C. J. Baldwin, K. M. Dale, R. F. Dittrich	[1]	1959	Appreciable research work has been done in the area of single area unit commitment problem	Priority list	Start-up and shut-down constraints
2	Srikanth Reddy K., Lokesh Panwar, B. K. Panigrahi, Rajesh Kumar	[2]	2018	Development of the sensitivity of resource of coal and fuel used in fuel combustion or coal combustion through the interpreted model for planning of operation the unit in thermal generators over unit commitment	Post carbon capture and storage (CCS) Technology	Start-up and shut-down ramp constraints, Spinning reserve constraint, emission constraints
3	Zhile Yang, Kang Li, Yuanjun Guo, Shengzhong Feng, Qun Niu, Yusheng Xue, Aoife Foley	[3]	2018	Constructed a new methodology on analysis of gap theory to appraise the strategy of operation of profitability for united power and heat units in a liberalistic market of electricity	Binary particle swarm optimization (BPSO) combining the GA and PSO with lambda iteration	Generation limit, charging/discharging power limit, power demand limit, Power reserve limit and Minimum up and downtime constraint
4	Jamshid Aghaei, Vassilios G. Agelidis, Mansour Charwand, Fatima Raeisi, Abdollah Ahmadi, Ali Esmaeel Nezhad and Alireza Heidari	[4]	2016	To minimize the operational risk and cost this model is implemented using risk-based unit commitment model	Information gap decision theory	Minimum up and downtime constraint, Start-up and shut-down ramp constraints, Power capacity limits, Heat capacity limits

(continued)

Table 1 (continued)

Sl. No.	Authors name	References	Year	Brief review	Algorithm	Constraints
5	Farhad Bavafa, Taher Niknam, Rasoul Azizipanah-Abarghooee, Vladimir Terzija	[5]	2016	A new formulated method i.e. cycle based unit commitment algorithm presented a new method of process of power re-dispatch to satisfy the ramp rate constraints including up-time and down-time	Non-dominated sorting backtracking search optimization (NSBSO)	Start-up and shut-down constraints, Minimum up and downtime constraint, random integer generation
6	Beibei Wang and Benjamin F. Hobbs	[6]	2015	This technique proposed about need of flexible generator due to increase the penetration of renewable energy where Flexi-ramp is defined as the capacity of reserve committed unit to accommodate abrupt ramp	Stochastic programming	Minimum up and downtime constraint, Start-up and shut-down constraints and ramp constraints
7	B. Saravanan, C. Kumar, D. P. Kothari	[7]	2015	Some of the units in generating station have their convinced ranges where restricted operation is included limitations in component of machines and insecurity due to vibration occurred in shaft bearings or steam valve	Fireworks algorithm (FWA)	Power balance constraint, power Generation limits, Minimum up and downtime limit, Spinning reserve

(continued)

Table 1 (continued)

Sl. No.	Authors name	References	Year	Brief review	Algorithm	Constraints
8	C. Wang, F. Liu, W. Wei, S. Mei, F. Qiu, J. Wang	[8]	2016	A new methodology for tackle some issues created by non-dispatch able generation of wind power including the reliable operation and considered variation of wind including the robust unit commitment (RUC)	Column and constraint generation (C&CG) algorithm	Minimum up and downtime constraint, Start-up and shut-down constraints and ramp constraints
9	Amir Abdollahi, Mohsen Parsa Moghaddam, Masoud Rashidinejad and Mohammad Kazem Sheikh-El-Eslami	[9]	2012	Designed a program including load management (LM) and its expansion is done including demand response programs (DRPs) which is divided into mainly two sections, such as time-based rate (TBR) and incentive-based programs (IBPs)	Incentive-based programs (IBPs) and time-based rate (TBR) programs	Minimum up and downtime constraint, start-up and shut-down constraints and Emission constraints
10	Himanshu Anand, Nitin Narang, J. S. Dhillon	[10]	2018	Using hybrid optimization technique to solve profit based unit commitment problem based on civilized swarm optimization (CSO) and binary successive approach (BSA) technique	Integrated technique of society civilized algorithm (SCA) and technique of Particle Swarm Optimization (PSO)	Minimum up and downtime constraint, Spinning reserve constraint, Start-up and shut-down ramp constraints, Power demand constraint and Power inequality constraint

(continued)

Table 1 (continued)

Sl. No.	Authors name	References	Year	Brief review	Algorithm	Constraints
11	Qipeng P. Zheng, Jianhui Wang and Andrew L. Liu	[11]	2014	A review of Stochastic Optimization technique for unit commitment. In the industry of power sector, the optimization technique have been broadly used to support the process of decision making for dispatch and scheduling the generation of electric power	Stochastic Programming and Mixed integer programming (MIPs)	Minimum up and downtime constraint, Spinning reserve constraint, Start-up and shut-down and risk constraints
12	A. Senthilvadivu, K. Gayathri, K. Asokan	[12]	2018	Considering the transition of global warming and emission of environment in the competition of electricity generation market used to solve the problems including reduction of emission of environment and maximize the profit	Exchange market algorithm (EMA)	Emission constraints, limit constraints of generator, Spinning reserve constraint and Minimum up/downtime constraints
13	Srikanth Reddy K., Lokesh Panwar, B. K. Panigrahi & Rajesh Kumar	[13]	2018	A new methodology is invented to solve profit based unit commitment (PBUC) problem using optimization technique based on the binary whale to get the results as binary in nature of PBUC problem	Binary whale optimization algorithm (BWOA)	Load constraint, spinning reserve constraint and thermal unit constraints

(continued)

Table 1 (continued)

Sl. No.	Authors name	References	Year	Brief review	Algorithm	Constraints
14	Srikanth Reddy Ka, Lokesh Kumar Panwarb, B. K. Panigrahia, Rajesh Kumarc, Ameena Alsumaitid	[14]	2018	In this paper the outputs including solution quality and consistency are compared with others approaches using Nature-inspired optimization and Constrained optimization	Binary Gray Wolf Optimizer (BGWO)	Load constraint, spinning reserve constraint, generation limit constraints, minimum up/downtime constraints and ramp up-down constraints
15	Javad Olamaei1, Mohammad Esmaeil Nazari, Sepideh Bahravar	[15]	2018	An integrated methodology is applied with production of power in thermal unit to meet the market demand and also applied for traditional separate cooling and heating application for valve-point effect for turbine	Heuristic optimization algorithm including the valve-point effect of the system based on combined cooling, heat, and power (CCHP) thermal power system	Generation capacity, minimum up-downtime limit, and spinning reserve
16	K Srikanth Reddy1, Lokesh Kumar Panwar, BK Panigrahi, Rajesh Kumar	[16]	2017	A new application to solve the problem of PBUC in competitive generation market of electricity by GENCOs with aim of maximize profit for forecasts of load	A binary variant of sine cosine algorithm (BSCA) is used to solve such kind of PBUC problem	Generation limit, power demand limit and power reserve limit, minimum up/downtime limit, hot and cold start-up and ramp limit

(continued)

Table 1 (continued)

Sl. No.	Authors name	References	Year	Brief review	Algorithm	Constraints
17	Adline Bikeri, Peter Kihato, Christopher Maina	[17]	2017	This methodology shows for 10th thermal units of power plant with implemented the methodology which should performed for quality solution	Evolutionary particle swarm optimization (EPSO) algorithm is used to solve PBUC problem	Generator limit, Ramp up–down time, Minimum up/down time and Power balance for bilateral contracts
18	K. Srikanth Reddy, Lokesh Kumar Panwar, Rajesh Kumar, B. K. Panigrahi	[18]	2016	The solution quality from this paper is taken into Consideration for deciding the status of commitment and make profit obtain by GENCOs	A binary coded fireworks algorithm	Load constraint, spinning reserve constraint, Thermal unit constraints, Minimum up/downtimes and Ramp up/down rates
19	A. V. V. Sudhakar, Chandram Karri, A. Jaya Laxmi	[19]	2016	For real time operation in the market of power deregulation, forecasting of the electricity market and PBUC problems are most important	Lagrange Relaxation (LR)-Differential Evolution (DE) algorithm is implemented here	Generator and reserve, Minimum up/downtimes, Reserve constraint and Power demand constraint
20	Prateek Kumar Singhal, Ram Naresh, Veena Sharma	[20]	2015	The effectiveness of these techniques on 100 unit and 10 unit of power plant for thermal unit in a whole day in power generating market for increase the computational time and the rate of profit of GENCOs	A binary fish swarm algorithm (BFSA) and dynamic economic dispatch (DED) method	Load demand constraint of a system, spinning reserve, Unit generation limit, Reserve Generation limit constraint and minimum up and downtime constraints

(continued)

Table 1 (continued)

Sl. No.	Authors name	References	Year	Brief review	Algorithm	Constraints
21	Anup Shukla, Vivek Nandan Lal, and S. N. Singh	[21]	2015	This methodology helps to find out the maximum profit in the electricity market about how much power must be taken in for sale and reserve	Modified dynamic programming using particle swarm optimization (PSO)	Demand constraint of load, Reserve constraints, Power and Reserve limit and Minimum Up and Downtime constraint
22	K. Venkatesan1, Selvakumar, C. C. A. Rajan	[22]	2015	The operation of this technique depends on cost-saving. This technique was implemented an IBM PC through which in a reasonable time period large types of system can be proceed	Evolutionary programming based particle swarm optimization (EPPSO) algorithm is used	Power balance constraint, Spinning reserve constraint in each area, Generator limits of each unit and Thermal units including Minimum Up and Downtime constraint
23	Germán Morales-España, Claudio Gentile, Andres Ramos	[23]	2015	This paper basically discuss quickly start and slow unit using some basic constraints which help in power scheduling and represent the step for schedule in energy to leads unattainable delivery of electrical system	Mixed-integer programming (MIP) is used	Generating limits, minimum up and downtimes and startup and shut down the power
24	M. Jabbari Ghadi, A. Baghramian, M. Hosseini Imani	[24]	2015	This technique is basically an upgraded version of the evolutionary algorithm which reduces complexity in computation to get solution in PBUC problem	Imperialist competitive algorithm (ICA) is implemented	Load demand constraint, generation limits, minimum up and downtime constraint and ramp unit constraints

(continued)

Table 1 (continued)

Sl. No.	Authors name	References	Year	Brief review	Algorithm	Constraints
25	A. Prakash, M. Yuvaraj	[25]	2014	The generating companies scheduled generators to get maximum profit rather than get satisfaction in power demand in the environment of deregulation of the power market	Improved pre-prepared power demand (IPPD) and a genetic algorithm is used here	Power balance, reserve constraint, limits of output powers of units and minimum up and downtime constraint
26	I. Jacob Raglend, Rohit Kumar, S. Prabhakar, Karthikeyan, K. Palanisamy, D. P. Kothari	[26]	2014	This paper is all about PBUC problem under electricity deregulated market in the sector of power system to increase efficiency of production of electricity and distribution of electric power in low price with high quality and more reliable and provide security for product	Dynamic Programming (DP) Technique is applied	Thermal units including Minimum Up and Downtime constraint, startup and shutdown power and ramp constraints
27	D. Sam Harison, T. Sreerengaraja	[27]	2013	This modified techniques used to control and modify purpose or sometimes replacements of generating units and it is also used for emission in one of the vital contributions of greenhouse gases to the environment through increasing the usages of fossil fuel in power plant	The traditional algorithms based on Swarm Intelligence	Power balance constraint, power generation limits, spinning reserve constraints and minimum up and downtime constraint

(continued)

Table 1 (continued)

Sl. No.	Authors name	References	Year	Brief review	Algorithm	Constraints
28	C. Christopher Columbus, Sishaj P. Simon	[28]	2013	This technique represents the behavior like bees and described a technique used to perform the operation of parallel distribution for solving the problem of PBUC	Parallel Nodal Ant Colony Optimization (PNACO) which is followed by using a parallel artificial bee colony (PABC) technique	Thermal units including minimum up and downtime constraint, power balance constraint, spinning reserve constraints and startup and shut down power
29	S. Chitra Selvi, M. Bala Singh Moses, C. Christober Asir Rajan	[29]	2013	In this paper from the comparison of other research work to increase the profit in the generation of electricity market, this study is done by including losses and it should be added the revenue of power to solve PBUC problem	Algorithm based on Lagrange relaxation (LR) with evolutionary programming (EP)	Real power constraints, reserve constraints, real and reserve power operating limits and minimum up and downtime constraint
30	C. Christopher Columbus, Sishaj P. Simon	[30]	2012	In power system ranging from 10 to 1000 generation units are tested for this technique to get the solution of quality and complexity of time with respect to the number of the formation of cluster can be analyzed thoroughly	Parallel artificial bee colony (PABC) algorithm is used	Demand constraints, unit power limit and minimum up and downtime constraints

(continued)

Table 1 (continued)

Sl. No.	Authors name	References	Year	Brief review	Algorithm	Constraints
31	Erik Delarue, Pieterjan Van Den Bosch, William D'haeseleer	[31]	2010	This technique is useful to find out the relation between Mean Absolute Percentage Error (MAPE) for price projection and profit or loss	Algorithm based on Mixed Integer Linear Programming (MILP) is introduced	Thermal units including minimum up and downtime constraint, power demand constraints and reserve constraints
32	J. P. S. Catalão a, S. J. P. S. Mariano, V. M. F. Mendes, L. A. F. M. Ferreira	[32]	2009	This methodology addresses the profitable market of power including between emission and profit, the trade curve is defined for profiles of different types of energy files	Multiobjective optimization (MO) technique	Hourly generation constraints, Cumulative constraints and thermal unit constraints
33	K. Chandram, N. Subrahmanyam, M. Sydulu	[33]	2008	A new approach in PBUC in the practical field with some limitations of emission due to fossil fuel in power plant. This methodology address the profitable market of power including emission in practically with the help of multiobjective optimization (MO) problem	Subproblem of economic dispatch using nonlinear programming is solved by Muller method	Power demand constraint, reserve constraint, generator constraints, minimum up and downtime constraint
34	J. P. S. Catalão, S. J. P. S. Mariano, V. M. F. Mendes, and L. A. F. M. Ferreira	[34]	2007	This method is used to determine the PBUC problem with the struggle in profit in a competitive energy market and objectives of emission	Multiobjective optimization (MO) technique	Thermal Constraints, Global Constraints, and Local Constraints

$$\text{TOC} = \sum_{t=1}^{T} \sum_{i=1}^{N} F\left(P_{t,i}\right) \times U_{t,i} + STC_{t,i} \tag{3}$$

where

$F(Pt, i)$	the cost of fuel of the ith generating units at tth sub-interval in \$/h
Pt, i	the generation of power of the ith unit at tth sub-interval in MW
Ut, i	the status of the unit of ith unit at tth sub-interval
$STCt, i$	the starting cost of ith generating units at tth sub-interval in \$/h

When the valve point loading effect which introduces a sinusoidal function in the fuel cost of the generator is neglected then the cost of fuel of thermal unit $F\left(P_{t,i}\right)$ is treated as a quadratic function which is indicated as

$$\begin{aligned} F(Pt, i) = {} & A_i + B_i \times P_{t,i} + C_i P_{t,i}^2 \\ & + \left| D_i \sin(E_i \times \left(P_i^{MIN} - P_{t,i}\right)) \right| i \in [1, N], t \in [1, T] \end{aligned} \tag{4}$$

where

A_i, B_i, C_i, D_i & E_i	are the coefficients of cost of the ith generating units
Pt, i	the generation of power at tth subinterval of the ith unit in MW
P_i^{MIN}	the minimum generation of power at ith unit in MW

The starting cost $\left(STC_{t,i}\right)$ of thermal units depends on the sage of the operating generating unit which includes hot startup and cold startup are indicated as

$$STC_{t,i} = \begin{cases} HSU_i, & T_{t,i}^{OFF} \leq T_i^{DW} + T_i^{COLD} \\ CSU_i, & T_{t,i}^{OFF} > T_i^{DW} + T_i^{COLD} \end{cases} i \in [1, N], t \in [1, T] \tag{5}$$

where

HSU_i	the cost of Hot Start-Up for ith generating unit in \$/h
CSU_i	the cost of Cold Start-Up for ith generating unit in \$/h
$T_{t,i}^{OFF}$	the time of OFF condition for ith unit at tth sub-interval in an hour
T_i^{DW}	the time of DOWN condition for ith unit at tth subinterval in hour
T_i^{COLD}	the time taken for COOLING condition for ith unit at tth subinterval in hour

PBUC problem Structure based on Unit Constraints:

The prospective of PBUC problem exposed to some standard types of operating constraints like spinning reserve, power balance, and power limit of generators, emission constraints and minimum ON time and minimum OFF time. To turns on the generating units is required a specific temperature and pressure. The details discussion about those constraints are given below:

Constraints depend on the generation of power:

There are three types of constraints which are depending on power generation. Such as constraint of power inequality, constraint of power demand and generation of ramp limit.

Constraint of power inequality:

Each of the committed units on thermal power can generate power in a particular limit at tth subinterval and it can be implemented as

$$P_i^{MIN} \leq P_{t,i} \leq P_i^{MAX} \; i \in [1, N] \tag{6}$$

where

P_i^{MIN} minimum power of ith unit
P_i^{MAX} maximum power of ith unit

Constraint of Power demand:

For the market of deregulated electric power, the power generation from the generated unit which is committed at tth sub-interval must be equal to or smaller than the demand of load for forecast (PDt) can be written as follows:

$$\sum_{i=1}^{N} P_{t,i} \times U_{t,i} \leq PD_t \; t \in [1, T] \tag{7}$$

where

$P_{t,i}$ the generation of power of the ith unit at tth sub-interval in MW
$U_{t,i}$ the status of the unit of ith unit at tth sub-interval
PD_t demand of power at tth of sub-interval

Generation of ramp limit:

The power which is generated cannot transfer after exceed its rate of ramp limit which is indicated as

$$P_{t,i} - P_{(t-1),i} \leq UPR_i \; i \in [1, N], \; t \in [1, T] \tag{8}$$

$$P_{(t-1),i} - P_{t,i} \leq DWR_i \; i \in [1, N], \; t \in [1, T] \tag{9}$$

where

UPR_i upper limit of ramp rate (RAMP-UP) of ith unit
DWR_i down limit of ramp rate (RAMP-DOWN) of ith unit

Constraints depend on Unit Commitment:

Including some important constants are depending on various kinds of rules and regulations extracted at time of scheduling by the council of reliability, pool of power and individual power system, etc.

Constraints for Minimum Up-time and Down-time:

When the unit is in running condition then immediately it will not be turned off and there is required minimum time for recommitted a unit once the unit is decommitted (Fig. 1).

$$U_{t,i} = \begin{cases} 1, \ t^{ON}_{(t-1),i} \leq t^{UP}_i \\ 0, \ t^{OFF}_{(t-1),i} \leq t^{DW}_i \quad i \in [1, N], \ t \in [1, T] \\ 1 \, or \, 0, \ otherwise \end{cases} \tag{10}$$

where

$U_{t,i}$ the status of the unit of ith unit at tth sub-interval
t^{UP}_i at hth hour minimum time for ON condition for ith unit
t^{DW}_i at hth hour minimum time for OFF condition for ith unit
$t^{ON}_{(t-1),i}$ from its Starting-Up time the ON time of ith unit at hth hour at tth sub-interval
$t^{OFF}_{(t-1),i}$ from its Shutting-Downtime the OFF time of ith unit at hth hour at tth sub-interval

The spinning power reserve is required to maintain the reliability of the system and at tth sub-interval it is given as:

Constraints for Spinning Reserve:

The maximum available power should not be more than power demand along with spinning power reserve at tth sub-interval and is given as

$$\sum_{i=1}^{N} P^{MAX}_i \times U_{t,i} \leq PD_t + SRP_t \quad t \in [1, T] \tag{11}$$

where

P^{MAX}_i maximum power of ith unit
$U_{t,i}$ the status of the unit of ith unit at tth sub-interval
PD_t demand of power at tth of sub-interval
SRP_t at tth sub-interval requirement of Spinning Reserve

The mechanism to handle the spinning reserve constraints requirement is shown in Fig. 2.

Constraints for Thermal Unit:

Requirements of the thermal unit depend on a group of operators when turn-on and turn-off are required. To bring the unit online, it takes some time in hour due to change in temperature. Take consideration of some constraints for the operation of the thermal power plants due to these kinds of restrictions.

Start-Up Constraints:

The generation of power which is known as power-constrained during the process of start-up conditions should have to take ST^{UH}_i hour to reach its power in maximum limit and is implemented as

$$PC_{h,i} = MIN\left(P^{MAX}_i, \ UPR_i \times (h - n)\right) \quad h \in \left[t, \ t + ST^{UH}_i\right],$$

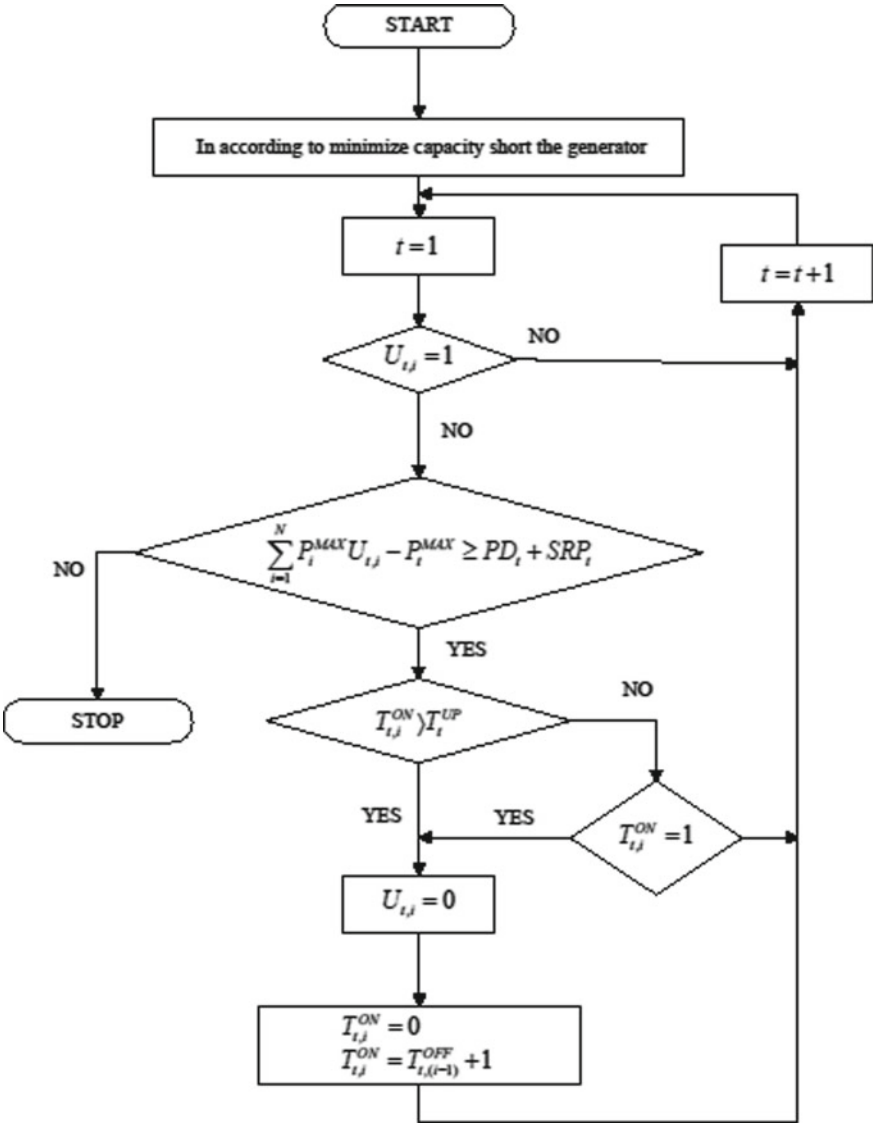


Fig. 1 Decommitment of the excessive generating unit while considering minimum up and minimum downtime constraints

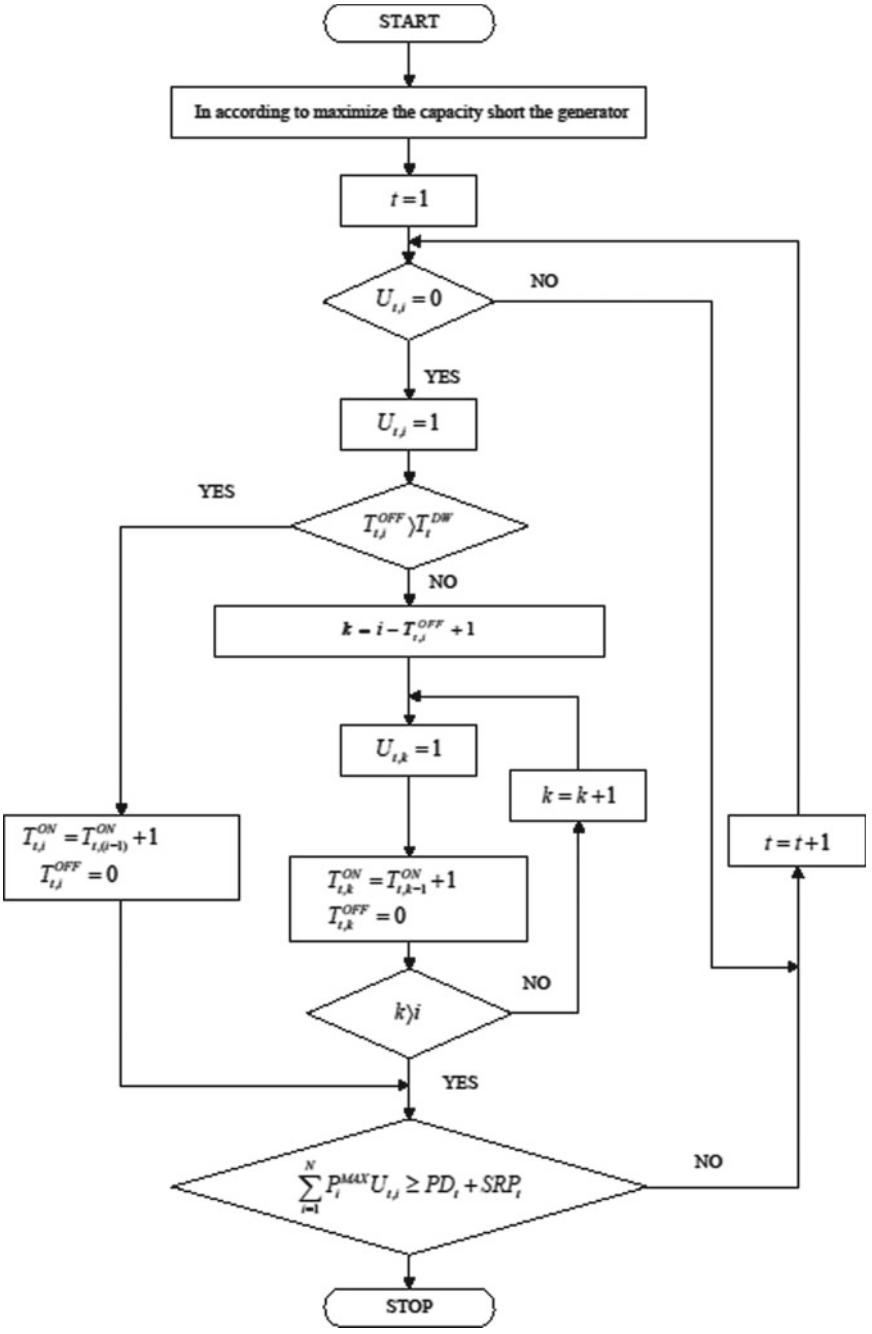


Fig. 2 Constraints handling strategy for spinning reserve requirement

$$n = t,\ i \in [1, N],\ t \in [1, T] \tag{12}$$

where

$PC_{h,i}$ at hth hour Power Constrained of ith unit
P_i^{MAX} maximum power of ith unit
UPR_i upper limit of ramp rate (RAMP-UP) of ith unit
ST_i^{UH} the status of starting-up time of ith unit at hth hour

Shut-Down Constraints:

In case of the process of Shutting-Down, the power constraints of the unit should lower than the maximum power of the unit and ST_i^{DH} hour is required to achieve the zero generated power which can be defined as given below:

$$PC_{h,i} = P_i^{MAX} - DWR_i \times (h - n)\ h \in \left[t^1 - ST_i^{DH}, t\right],$$
$$n = \left[t^1 - ST_i^{DH}\right],\ i \in [1, N],\ t^1 \in [1, T] \tag{13}$$

where

$PC_{h,i}$ at hth hour power constrained of ith unit
DWR_i down limit of ramp rate (RAMP-DOWN) of ith unit
P_i^{MAX} maximum power of ith unit
ST_i^{DH} the status of shutting-downtime of ith unit at hth hour

4 Conclusion

In **t**he proposed research, the authors have efficaciously represented the preliminary concept of profit based unit commitment problem. Also, the mathematical formulation of profit based unit commitment has been explored to understand the fundamental correlation within different parameters of profit based unit commitment problem. Further, the flowchart of constraints handling mechanism has been explored to describe the solution methodologies of existing PBUCP. A brief review of profit-based unit commitment problems including classical unit commitment problem has also been reconnoitered staring from 1959 to recent years. The research study will be beneficial for new researchers, who are working in the expanse of profit-based unit commitment problem.

References

1. C. J. Baldwin, K.M. Dale, R.F. Dittrich, A study of economic shutdown of generating units in daily dispatch. AIEE Trans. Power Apparatus Syst. PAS-78. 1272–1284

2. S. Reddy K., L. Panwar, B.K. Panigrahi, R. Kumar, Low carbon unit commitment (LCUC) with post carbon capture and storage (CCS) technology considering resource sensitivity. J. Clean. Prod. (2018)
3. Z. Yang, K. Li, Y. Guo, S. Feng, Q. Niu, Y. Xue, A binary symmetric based hybrid meta-heuristic method for solving mixed integer unit commitment problem integrating with significant plug-in electric vehicles. Energy **170**, 889–905 (2019)
4. J. Aghaei, V.G. Agelidis, M. Charwand, F. Raeisi, A. Ahmadi, A.E. Nezhad, Optimal robust unit commitment of CHP plants in electricity markets using information gap decision theory. **8**(5), 2296–2304 (2017)
5. F. Bavafa, T. Niknam, R. Azizipanah-abarghooee, V. Terzija, A new biobjective probabilistic risk-based wind-thermal unit commitment using. IEEE Trans. Ind. Inform. **13**(1), 115–124 (2017)
6. B. Wang, B.F. Hobbs, Real-time markets for flexiramp : a stochastic unit commitment-based analysis. 1–15 (2015)
7. B. Saravanan, C. Kumar, D.P. Kothari, Electrical power and energy systems a solution to unit commitment problem using fire works algorithm. Int. J. Electr. Power Energy Syst. **77**, 221–227 (2016)
8. C. Wang, F. Liu, W. Wei, S. Mei, F. Qiu, J. Wang, Robust unit commitment considering strategic wind generation curtailment, in *2016 IEEE Power & Energy Society General Meeting*, no. 51321005, pp. 1–5 (2016)
9. A. Abdollahi, M.P. Moghaddam, Investigation of economic and environmental-driven demand response measures incorporating UC. **3**(1), 12–25 (2012)
10. H. Anand, N. Narang, J.S. Dhillon, Profit based unit commitment using hybrid optimization technique. Energy **148**, 701–715 (2018)
11. Q.P. Zheng, J. Wang, A.L. Liu, Stochastic optimization for unit commitment—a review. 1–12 (2014)
12. A. Senthilvadivu, K. Gayathri, K. Asokan, Exchange market algorithm based profit based unit commitment for GENCOs considering environmental emissions. **13**(21), 14997–15010 (2018)
13. S. Reddy K., L. Panwar, B.K. Panigrahi, R. Kumar, Binary whale optimization algorithm: a new metaheuristic approach for profit-based unit commitment problems in competitive electricity markets. Eng. Optim. 1–21 (2018)
14. S. Reddy K, L.K. Panwar, B.K. Panigrahi, R. Kumar, A. Alsumaiti, Binary grey wolf optimizer models for profit based unit commitment of price-taking GENCO in electricity market. Swarm Evol. Comput. (2018)
15. J. Olamaei, M.E. Nazari, S. Bahravar, Economic environmental unit commitment for integrated CCHP-thermal-heat only system with considerations for valve-point effect based on a heuristic optimization algorithm. Energy (2018)
16. K.S. Reddy, L. Kumar, P. Bk, P. Rajesh, A new binary variant of Sine–Cosine algorithm : development and application to solve profit-based unit commitment problem. Arab. J. Sci. Eng. (2017)
17. A. Bikeri, P. Kihato, C. Maina, Profit based unit commitment using evolutionary particle swarm optimization. 1137–1142 (2017)
18. K.S. Reddy, L. Kumar, R. Kumar, B.K. Panigrahi, Electrical power and energy systems binary fireworks algorithm for profit based unit commitment (PBUC) problem. Int. J. Electr. Power Energy Syst. **83**, 270–282 (2016)
19. A.V.V. Sudhakar, C. Karri, A.J. Laxmi, Engineering science and technology, an international journal profit based unit commitment for GENCOs using Lagrange relaxation – differential evolution. Eng. Sci. Technol. Int. J. **20**(2), 738–747 (2017)
20. P.K. Singhal, R. Naresh, V. Sharma, Binary fish swarm algorithm for profit-based unit commitment problem in competitive electricity market with ramp rate constraints. **9**, 1697–1707 (2015)
21. A. Shukla, V.N. Lal, S.N. Singh, Profit-based unit commitment problem using PSO with modified dynamic programming. 1–6 (2015)

22. K. Venkatesan, G. Selvakumar, C.C.A. Rajan, EP based PSO method for solving profit based multi area unit commitment problem. **10**(4), 442–460 (2015)
23. G. Morales-españa, A. Ramos, C. Gentile, Tight MIP formulations of the power-based unit commitment problem. OR Spectr. **37**(4), 929–950 (2015)
24. M.J. Ghadi, A. Baghramian, M.H. Imani, Appl. Soft Comput. J. (2015)
25. A. Prakash, M. Yuvaraj, Profit based unit commitment using IPPDT and genetic algorithm. **2**(1), 1053–1061 (2014)
26. R. Billinton, R. Mo, Deregulated environment. Power **20**(1), 485–492 (2005)
27. D.S.H.T. Sreerengaraja, Swarm intelligence to the solution of profit-based unit commitment problem with emission limitations. 1415–1425 (2013)
28. C.C. Columbus, S.P. Simon, Profit based unit commitment for GENCOs using parallel NACO in a distributed cluster. Swarm Evol. Comput. **10**, 41–58 (2013)
29. S.C. Selvi, M.B.S. Moses, C.C.A. Rajan, LR-EP approach for solving profit based unit commitment problem with losses in deregulated markets (11), 210–213 (2013)
30. C.C. Columbus, S.P. Simon, Profit based unit commitment: a parallel ABC approach using a workstation cluster q. Comput. Electr. Eng. **38**(3), 724–745 (2012)
31. E. Delarue, P. Van Den Bosch, D. William, Effect of the accuracy of price forecasting on profit in a price based unit commitment. Electr. Power Syst. Res. **80**(10), 1306–1313 (2010)
32. J.P.S. Catalão, S.J.P.S. Mariano, V.M.F. Mendes, L.A.F.M. Ferreira, Electrical power and energy systems a practical approach for profit-based unit commitment with emission limitations. Int. J. Electr. Power Energy Syst. **32**(3), 218–224 (2010)
33. K. Chandram, N. Subrahmanyam, New approach with Muller method for profit based unit commitment (2008)
34. J. P. S. Catalão, S. J. P. S. Mariano, V. M. F. Mendes, L. A. F. M. Ferreira, Profit-based unit commitment with emission limitations : a multiobjective approach. 1417–1422 (2007)

THD Analysis of New Multilevel Inverter Topology with Different Modulation Techniques



Nikhil Agrawal, Praveen Bansal and Niraj Umale

1 Introduction

Solid-state devices which convert the DC power into AC power at desired frequency and desired voltage is called inverter. The voltage source inverter is two-level inverter as supply Dc voltage is V_{dc} and it produces an output voltage either zero or $\pm V_{dc,}$ i.e., square wave. The two-level inverter has approx. 48% total harmonic distortion which is much higher than the IEEE 519 standard and not appropriate in power quality point of view. To obtain a less distorted output they require high-frequency pulse-width modulation techniques. These two-level inverters have a limitation in high power applications. To overcome these problems associated with two-level inverter, a multilevel inverter is introduced. The multilevel is introduced during 1975 [1]. MLI is modified and extended version of two-level inverter [2].

Multilevel inverter has less distortion factor as level increases and produces output voltage nearer to sinusoidal output voltage and having low rate of change of voltage and less voltage stress on switch in comparisons of two-level inverter [3, 4]. Some disadvantages associated with multilevel inverter, i.e., power semiconductor devices and respective gate driver circuit requirements are more and cost increase in comparison to two-level inverter. It is important here the compromise between economic and power quality term, the multilevel inverter topology is challenging today which balances both technical and economic issues. The conventional multilevel inverter divided into two groups one is monolithic multilevel inverter and second is modular

N. Agrawal (✉)
SR Group of Institutions, Jhansi, UP, India
e-mail: nikhilag.agrawal@gmail.com

P. Bansal
MITS, Gwalior, India

N. Umale
Surabhi College of Engineering & Technology, Bhopal, India

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_117

multilevel inverter. The monolithic converters are as flying capacitor and diode-clamped multilevel inverter and modular converter is cascaded H-bridge multilevel inverter. Diode-clamped inverter (DC-MLI) was introduced by Nabae et al. in 1981. DC-MLI has a common DC bus for all phase. The common DC bus requires less number of capacitors by which back-to-back topology is possible in diode-clamped inverter. But it requires huge clamping diodes as number of levels is high [5]. Flying capacitor inverter minimizes the clamping diode problem but a huge number of storage capacitors required as number of level is increased [6]. The monolithic converter requires large number of components in terms of clamping diode and capacitor, which produces lot of problems in terms of nonuniform voltage stress across the clamping diodes and this makes design of converter complicate and also raises the cost. The problem of monolithic converter is overcome by modular converter as it consists of series combination of single-phase full wave bridge subsystem. Cascade H-bridge multilevel inverter requires less number of components [7] with compared to monolithic converter and best part of this inverter is switching losses and stress on devices can be reduced by applying soft-switching techniques. This paper uses new MLI topology [2] and applies the pulse-width modulation techniques on new MLI topology and compares with other MLI topology based on total harmonic distortion and number of switches against number of output level.

2 Proposed Multilevel Inverter Topology

The basic unit for the proposed multilevel inverter is shown in Fig. 1. It consists of two parts: one is level generation part and second is polarity part [8]. Polarity part is fixed for all voltage level and level generation part increases as number of DC side voltage level increases and provides staircase wave with less harmonic distortion. Figure 1 shows the seven-level proposed multilevel inverter. The middle part of the

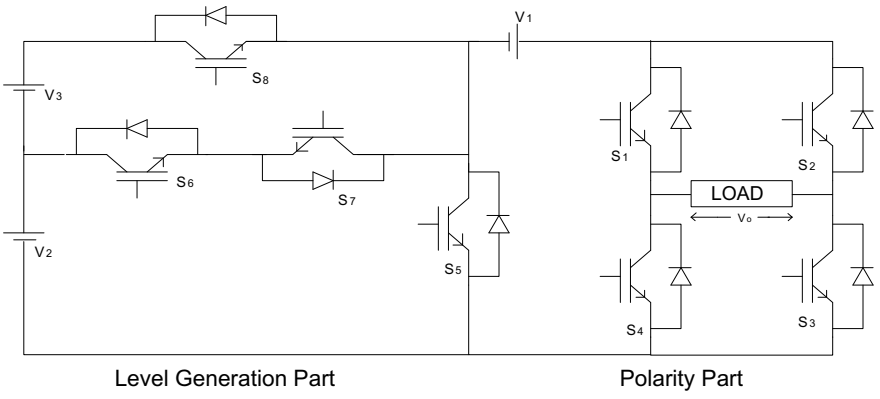


Fig. 1 Proposed topology 7-level MLI

level generation unit modified as number of voltage levels increases so this is best part of proposed MLI topology. It is constructed for any number of levels. Table 1 shows switching scheme for 7 levels.

Figures 2, 3, and 4 show the circuit diagram of proposed MLI topology of 9, 11, and 15 levels, respectively. These figures show if the level iof inverter is increased, the level generation part is improved by the connecting voltage source and bidirectional switch. The switching scheme for 7-level MLI is shown in Table 1 and count of components for different topology is shown in Table 2 (Fig. 5).

Table 1 Switching scheme of the proposed 7-level MLI

Output voltage	Switching scheme							
	SW 1	SW 2	SW3	SW4	SW 5	SW 6	SW 7	SW 8
+3 V	ON	OFF	ON	OFF	OFF	OFF	OFF	ON
+2 V	ON	OFF	ON	OFF	OFF	ON	ON	ON
+1 V	ON	OFF	ON	OFF	ON	OFF	OFF	OFF
0	ON	ON	OFF	OFF	OFF	OFF	OFF	OFF
−1 V	OFF	ON	OFF	ON	ON	OFF	OFF	OFF
−2 V	OFF	ON	OFF	ON	OFF	ON	ON	OFF
−3 V	OFF	ON	OFF	ON	OFF	OFF	OFF	ON

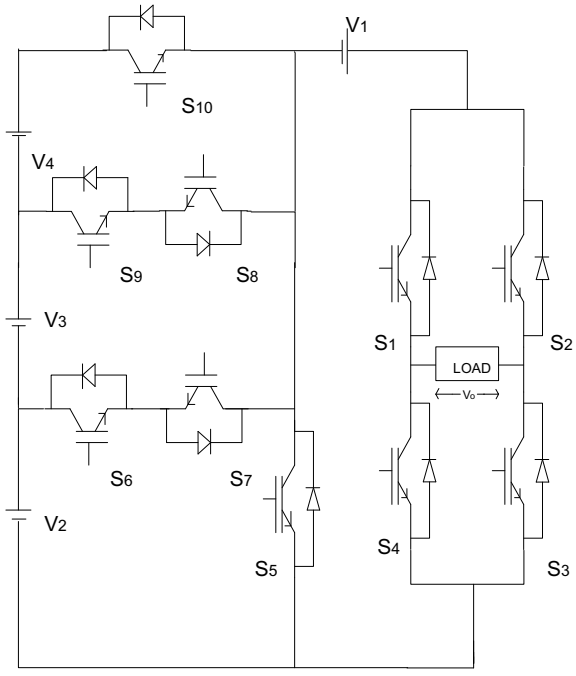


Fig. 2 Circuit of the 9-level proposed topology

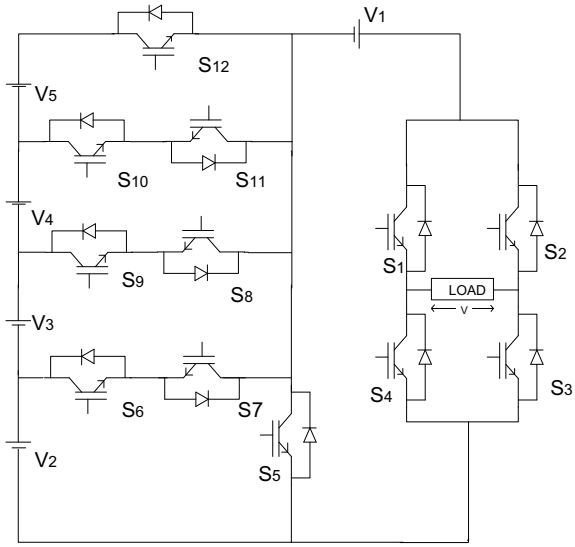


Fig. 3 Circuit of the 11-level proposed topology

Table 2 shows the total number of components used in conventional multilevel inverter topology, the topology presented in [2] and proposed topology. For 7-level MLI, CHBMLI requires total component are 15, FC-MLI requires 34, DC-MLI requires 49 components, for topology presented in [2] require 13 and for proposed topology require 11 components.

3 Modulation Strategies

Modulation technique is used in the inverter to obtain the output at the desired voltage and frequency. There are different modulation techniques used in inverter circuit. The basic concept of modulation is compared of sinusoidal wave signal to high-frequency carrier signal with some logic circuit. This paper uses multicarrier pulse-width modulation techniques to obtain the output at low THD value. Multicarrier pulse width modulation techniques are classified as carrier Disposition pulse width modulation (CD PWM) and phase-shifted modulation (PS PWM) and carrier-based modulation techniques are classified as in Fig. 6. In the carrier disposition techniques for L level MLI (L − 1) carriers required [9, 10] (Figs. 7 and 8).

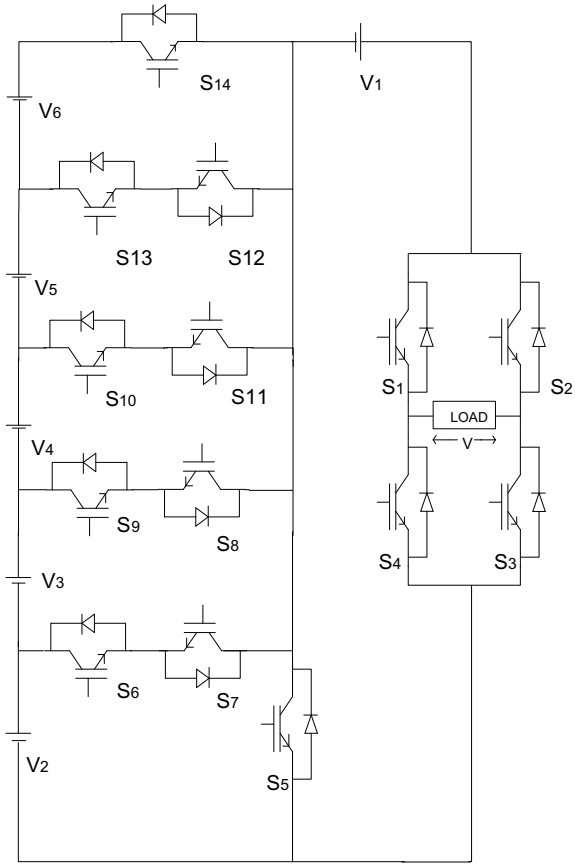


Fig. 4 Circuit of the 13-level proposed topology

Table 2 Comparison of power component in different MLI topologies

MLI structure component	CHBMLI	FC-MLI	DC-MLI	MLI topology [2]	Proposed topology
Main device	2 (L − 1)	2 (L − 1)	2 (L − 1)	L + 3	L + 1
Clamping diode	–	–	(L − 1) * (L − 2)	–	–
Clamping capacitors	–	(L − 1) * (L − 2)/2	–	–	–
DC split capacitor	–	(L − 1)	(L − 1)	–	–
DC source	(L − 1)/2	1	1	(L − 1)/2	(L − 1)/2
Total	5 (L − 1)/2	$(L^2 + 3L - 2)/2$	L^2	(3L + 5)/2	(3L + 1)/2

L Number of levels

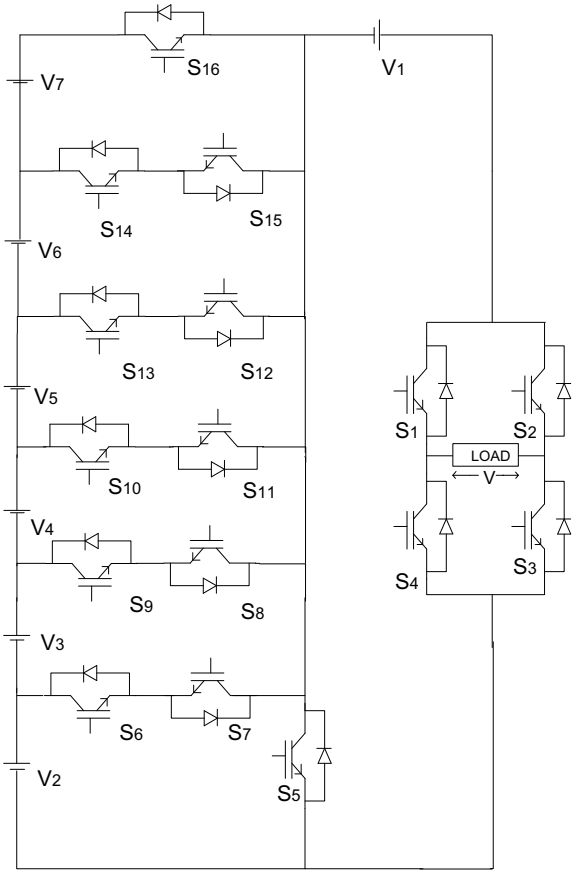


Fig. 5 Circuit of the 15-level proposed topology

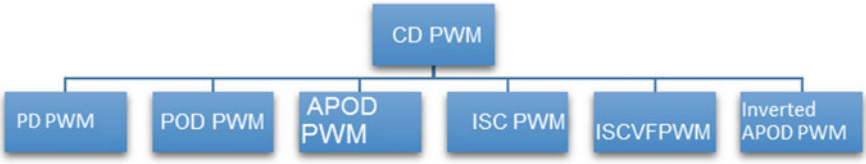


Fig. 6 Classification of carrier disposition PWM techniques

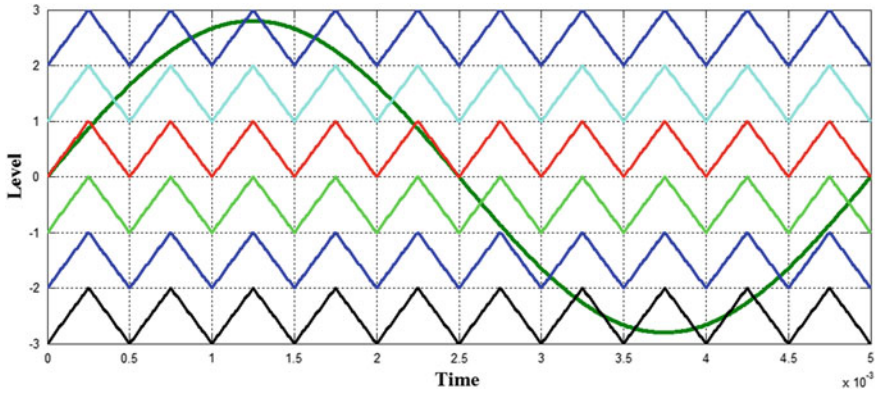


Fig. 7 Arrangement of carrier in PDPWM

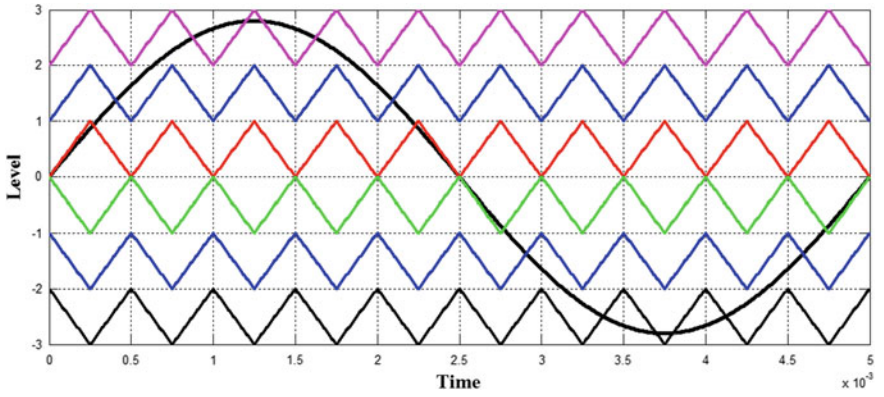


Fig. 8 Arrangement of carrier in PODPWM

3.1 Phase Disposition Modulation (PD PWM)

In this technique, all carriers are in same phase with same frequency and amplitude in both positive and negative groups. Positive and negative groups are the upper and lower portion at zero reference line.

3.2 Phase Opposition Disposition Modulation (PODPWM)

In the PODPWM [11] technique, positive group and negative group carriers have same frequency and amplitude. It has phase shift of 180°.

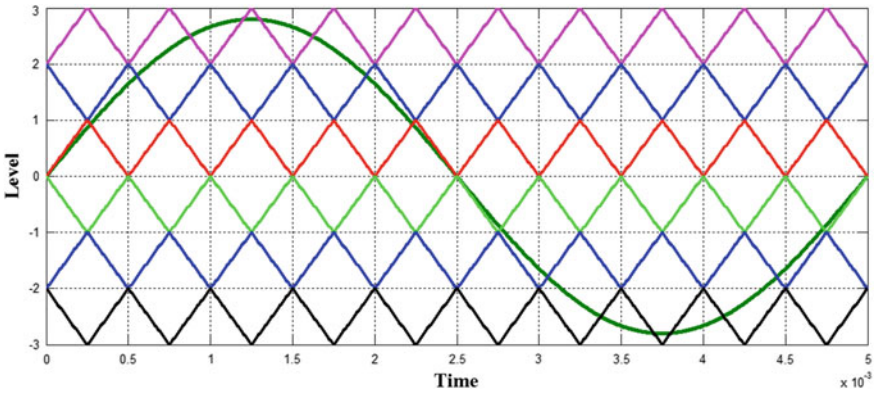


Fig. 9 Arrangement of carrier in APODPWM

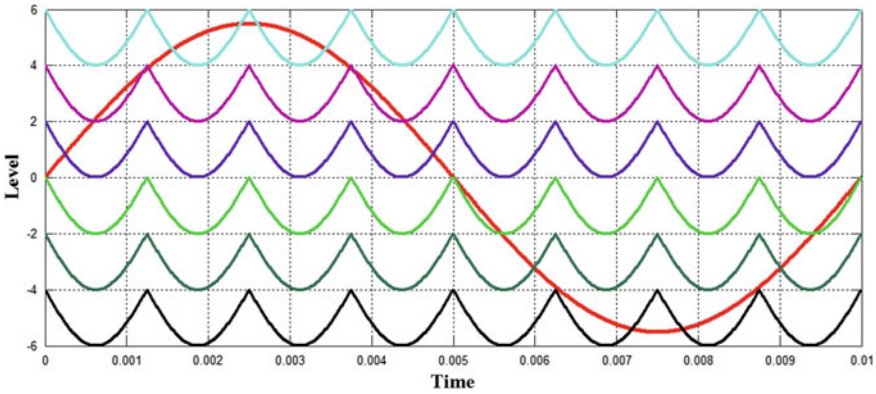


Fig. 10 Arrangement of carrier in ISCPWM

3.3 Alternate Phase Opposition Disposition Modulation (APODPWM)

In the APODPWM techniques, the positive group carrier is in 180° phase-shifted to each other and same for negative carrier group [12]. The carrier signal arrangement is as shown in Fig. 9.

3.4 Inverted Sine Carrier PWM (ISCPWM)

In this technique, the carrier is inverted sine wave shape with same frequency and magnitude as shown in Fig. 10.

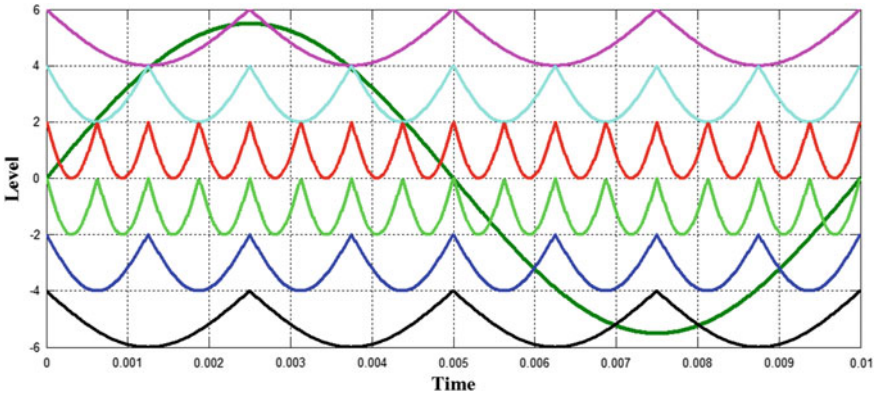


Fig. 11 Arrangement of carrier in ISCVF

3.5 Inverted Sine Carrier with Variable Frequency PWM (ISCVF)

ISCVF PWM techniques are similar to the ISCPWM technique, here the carrier is inverted sine but with a different frequency as shown in Fig. 11.

3.6 Inverted Alternate Phase Opposition Disposition PWM Techniques (Inverted APOD)

Inverted APOD PWM techniques are similar to the APOD modulation techniques as shown in Fig. 12.

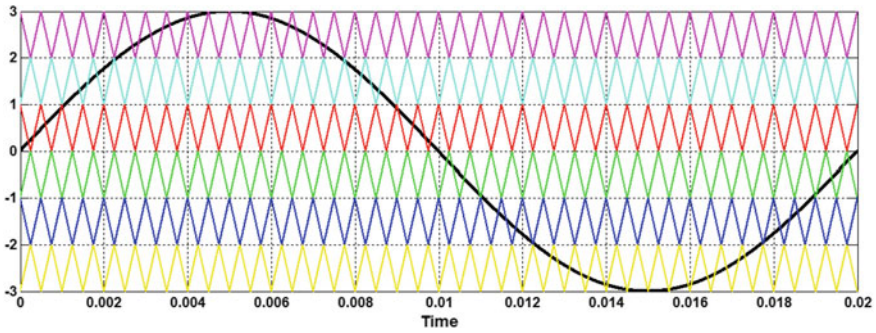


Fig. 12 Arrangement of carrier in inverted APODPWM

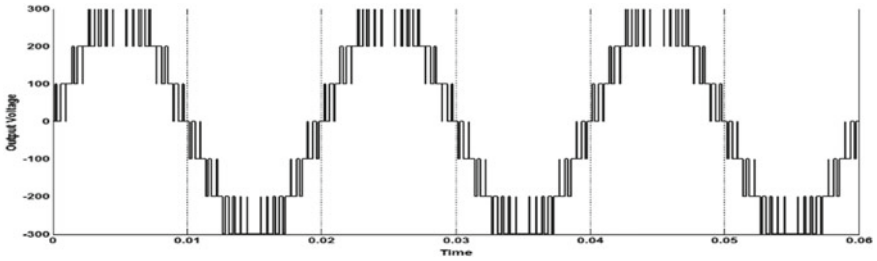


Fig. 13 Output waveform of the 7-level MLI

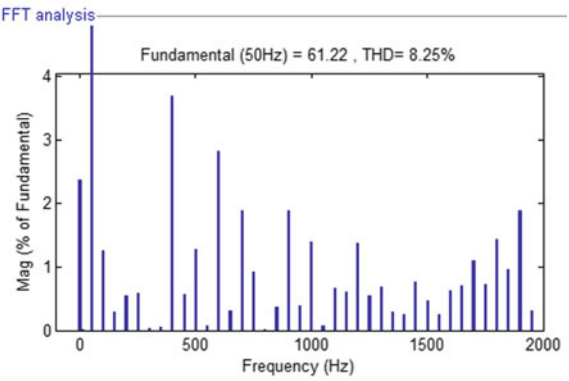


Fig. 14 FFT analysis for the 9-level MLI with POD

4 Simulation Result

Simulation has been done by MATLAB/SIMULINK R2010b. The result is taken by Simulation on 7 Level, 9 Level, 11 Level, 13 Level, and 15 Level of the mentioned above-proposed topology. Figure 13 shows the output waveform of 7-level proposed multilevel inverter topology (Figs. 14, 15, 16 and 17)

5 Conclusion

Simulation results of proposed MLI topology for single-phase 7-Level to 15-Level with carrier disposition modulation techniques and with under and over modulation index is calculated in MATLAB/Simulink. The total harmonic distortion decreases as number of level increases are shown in Table 3. A comparison made among proposed topology and conventional multilevel inverter topology and one another topology in [2] are shown in Table 2. In [13], Tables 2 and 3, above these result it is verified that the topology presented in this paper has superior performance in terms of number

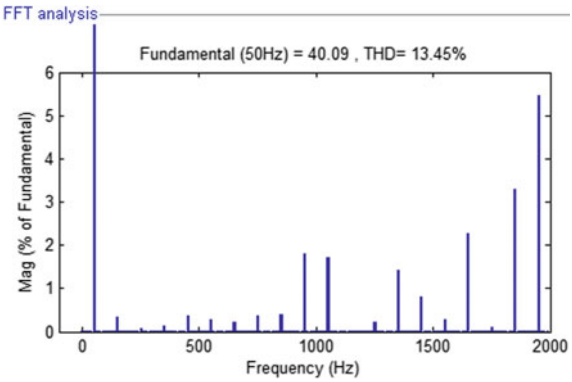


Fig. 15 FFT analysis for the 13-level

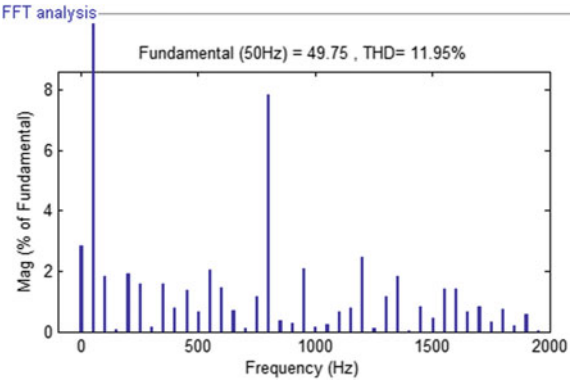


Fig. 16 FFT analysis for the 11-level with ISC

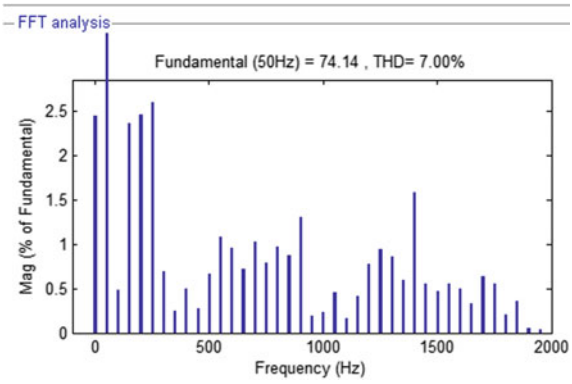


Fig. 17 FFT analysis for the 15-level with ISCVF

Table 3 Comparision of THD result with different PWM techniques and different levels

Level	Modulation index	Modulation techniques					
		PD	POD	APOD	ISC	ISCVF	Inverted APOD
7 Level	0.8	24.36	24.23	24.18	23.57	24.40	24.54
	0.9	22.42	22.15	22.08	20.35	20.12	22.74
	1.0	18.19	17.95	18.25	17.81	19.13	18.12
	1.1	15.67	15.41	15.65	15.94	17.68	15.69
9 Level	0.8	17.15	16.91	16.98	17.38	18.02	17.33
	0.9	16.80	16.76	16.92	18.18	16.86	16.67
	1.0	13.73	13.45	14.18	14.26	12.38	13.26
	1.1	11.91	11.68	12.82	11.74	12.89	10.93
11 Level	0.8	13.73	13.46	13.26	14.26	14.59	14.18
	0.9	13.05	12.93	12.07	13.18	13.64	13.96
	1.0	11.06	10.76	11.06	11.95	10.30	11.05
	1.1	10.06	9.50	9.97	10.17	10.13	10.15
13 Level	0.8	12.46	12.10	12.79	13.41	12.49	12.12
	0.9	10.80	10.32	10.72	10.40	11.46	10.88
	1.0	9.27	8.57	9.31	10.32	8.25	9.25
	1.1	8.06	7.78	8.31	8.25	8.72	7.95
15 Level	0.8	10.60	10.10	10.42	10.31	10.78	10.78
	0.9	8.80	8.29	8.99	9.05	8.35	8.71
	1.0	7.87	8.62	8.12	8.17	8.36	7.69
	1.1	7.35	8.23	7.14	7.56	7.00	7.63

of components and total harmonic distortion. For 15 level proposed MLI with 1.1 modulation index and with ISCVF modulation scheme, the THD is 7% which is under the maximum permissible limit according to EN 50160 standard.

References

1. R.H. Baker, High-voltage converter circuits. U.S. patent Number 4, 203, 151, May 1980
2. N. Agrawal, S. Singh, P. Bansal, A multilevel inverter topology using reverse-connected voltage sources, in *International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS 2017)*, Chennai, India, pp. 1290–1295
3. J. Rodriguez, L. Jih-Sheng, F. Zheng, Multilevel inverters: a survey of topologies, controls, and applications. IEEE Trans. Ind. Electron. **49**, 724–738 (2002)
4. I. Colak, E Kabalci, R. Bayindir, Review of multilevel voltage source inverter topologies and control schemes. Energy Convers. Manag. **52**(2), 1114–1128 (2011)
5. A. Nabe, I. Takahashi, H. Akagi, A New neutral point clamped PWM inverter. IEEE Trans. Ind. Appl. **1A-17**, 518–523 (1981)

6. T.A. Meynard, H. Foch, Multilevel choppers for high voltage applications, in *Proceeding of European Conference on Power Electronics and applications* (1992), pp. 45–50
7. G. Sinha, T.A. Lipo, A four level rectifier-inverter system for drive applications, in *31St IEEE IAS Annual Meeting Conference Record*, vol. 2 (1996), pp. 980–987
8. H. Samsami, A. Taheri, R. Samanbakhsh, New bidirectional multilevel inverter topology with staircase cascading for symmetric and asymmetric structures. IET Power Electron. (2017)
9. M. Manjrekar, G. Venkataramanan, Advanced topologies and modulation strategies for multilevel inverters, in *Proceedings of IEEEPESC'96*, Baveno, Italy, June 1996, pp. 1013–1018
10. D.G. Holmes, B.P. McGrath, Opportunities for harmonic cancellation with carrier-based PWM for two-level and multilevel cascaded inverters. IEEE Trans. Ind. Appl. **37**, 574–582 (2001)
11. E. Najafi, A.H.M. Yatim, A.S. Samosir, A new topology-reversing voltage (RV) for multilevel inverters, in *2nd International Conference on Power and Energy (PECon 08)*, Dec 2008, Malaysia, pp. 604–608
12. E. Najafi, A. Halim, M. Yatim, Design and implementation of a new multilevel inverter topology. IEEE Trans. Ind. Electron. **59**(11) (2012)
13. S.K. Sahoo, A. Ramulu, J. Prakash, Deeksha, Performance analysis and simulation of five level and seven level single phase multilevel inverters, in *Third International Conference on Sustainable Energy and Intelligent System (SEISCON 2012)*, VCTW, Tamil Nadu, India, 27–29 Dec 2012

Sensitivity-Based Adaptive Activity Mapping for Optimal Camera Calibration



Shashank and S. Indu

1 Introduction

Computer vision is the most emerging field of research that combines image processing, image understanding, and system designing to gain a high-level understanding of an event and taking appropriate control measures for optimizing the functionality of the system. There are multiple challenges in the domain like camera reconfiguration, camera calibration, occlusion handling, optimized tracking, detection, and pose estimation. This paper aims to focus on camera calibration based on parameter reconfiguration for optimized activity detection. This paper introduces adaptive pixel sensitivity for allotment of priority of each pixel in the field of view (FOV) of a camera and hence on this basis, the camera center can be redefined so that the activity can be monitored in maximum resolution and maximum information about the event. Adaptive background subtraction and foreground detection have been utilized for activity detection in the camera field of view followed by adaptive multilayered thresholding to reduce false activity detection. Normalized Gaussian distribution associates the importance of an event in the FOV with time lapse. Cumulative importance of all the Gaussian weighted frames when normalized represents the sensitivity of each pixel on the basis of past activities in FOV.

Shashank (✉) · S. Indu
Delhi Technological University, Delhi, India
e-mail: shashank.ece.dtu@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_118

1.1 Adaptive Background Subtraction

Background subtraction is one of the most commonly utilized methods for object detection. Traditionally, background subtraction defines a background by averaging initial frames in the time domain and subtracting the upcoming frames from the background for obtaining the activities in the camera FOV. The major drawback with the background subtraction is that it is only relative to one background defined at initialization and hence any activity if stopped cannot be determined, as the frame is still different from the background. This problem is addressed by introducing an adaptive background subtraction for activity detection in this paper. Each preceding frame behaves as a background for the upcoming frame and the difference is utilized for activity detection. Suppose, V(n) is the video frame set containing n video frames with respect to time and D(n) is the difference set obtained from adaptive background subtraction, then D(n) can be obtained as

$$\text{For n} = 2 \text{ to upcoming frames, } \mathrm{D(n)} = \mathrm{V(n)} - \mathrm{V(n-1)} \tag{1}$$

1.2 Normalized Half-Gaussian Distribution

The Gaussian distribution curve is a probability distribution curve that fits almost all the natural phenomenon. The curve can be represented by the equation given below:

$$\mathrm{F}\left(\mathrm{x}|\mu, \alpha^2\right) = \left(1/\left(\mathrm{sqrt}\left(2\pi\alpha^2\right)\right) * \exp\left((\mathrm{x}-\mu)/2\alpha^2\right)\right) \tag{2}$$

where μ represents the mean value and α represents the standard deviation of the function. Coding notations have been used for writing the equations with * representing multiplication, **/** represents division, **sqrt** representing square root, **exp** representing exponential, and **ln** representing natural log, respectively.

We utilize only half of the Gaussian distribution with mean as reference or present time frame and unity value at the mean to allocate maximum importance to the present frame. The past frames are allocated importance in accordance with the Gaussian distribution. A half wave half maxima (HWHM) is used for defining the limits of samples for assigning the importance of difference frames. HWHM can be obtained as follows:

$$\mathrm{HWHM} = (\mathrm{sqrt}(2 * \ln(2)) * \alpha)/2 = 1.1799 * \alpha \tag{3}$$

The HWHM normalized Gaussian distribution for 300 frames as an importance factor is as shown in Fig. 1.

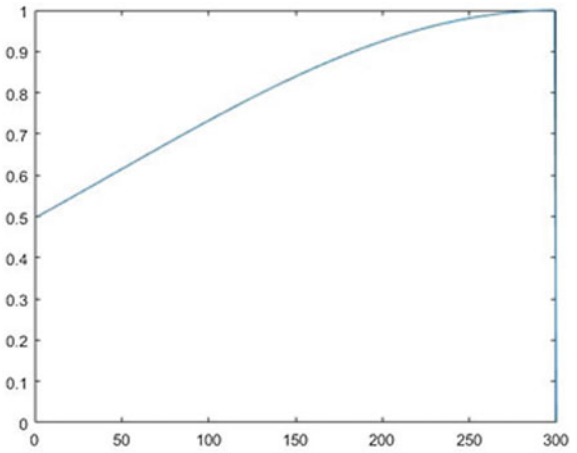


Fig. 1 Half wave half maxima Gaussian distribution

1.3 Multilayered Thresholding

For minimization of false activity detection, multilayered thresholding has been introduced. Cumulative global thresholding has been used first after adaptive thresholding and later for sensitivity allotment of the pixel. In the first stage, all the nonzero pixel values from all the difference frames D(n) are averaged to obtain the global threshold for adaptive background subtraction (Th^1) and binarization of the difference frames is done on the basis of the threshold value.

$$Th^1 = \sum \text{all non-zero values from all frames/total number of pixels in all frames} \tag{4}$$

The second layer of adaptive thresholding (Th^2) is applied to the weighted activity map for the removal of noise and normalizing the activity map as

$$\text{Normalized pixel value (post mapping)} = \text{pixel value at map/maximum map value} \tag{5}$$

The maximum map value is hence assigned unity (datum) and all the rest of the activity values are declared with reference to it. The normalized pixel value after activity mapping is used as the sensitivity value of the pixel.

2 Related Work

Most of the computer vision camera placement problems assume prior prioritization of surveillance area [1–5] and camera placement is done for maximizing the coverage of high priority areas. In [1], Indu et al. proposed an optimal camera placement mechanism for surveillance of large spaces predefining the priority areas. Zhang et al. [2] also proposed a sensor placement algorithm for orientation optimization of the surveillance system. In 2017, da Silva et al. [3] proposed a multi-UAV agent based pedestrian surveillance based on activity recognition but lacked initial camera optimal placement for optimized resolution of observed information. In [4], Jaenen et al. proposed distributed three-dimensional camera alignment with predefined priority areas. Shiang et al. [6] and McHugh et al. [7] introduced adaptive background subtraction for object detection. Jamshed et al. [5] introduced prioritization of tasks on the basis of decrease or increase in the frequency of events using dynamic voltage frequency scaling. As of now, there is not a single appropriate method that relates the reconfiguration of camera parameters in accordance with the past activities or events so that the data can be extracted in the maximum resolution bearing optimum information for improved processing and control system development.

3 Proposed Methodology

For activity mapping of the pixels and sensitivity allotment to each active pixel, a novel framework for this purpose has been introduced. The methodology for desired functionality is as shown in Fig. 2. The proposed system is tested on a real surveillance dataset to ensure its functionality.

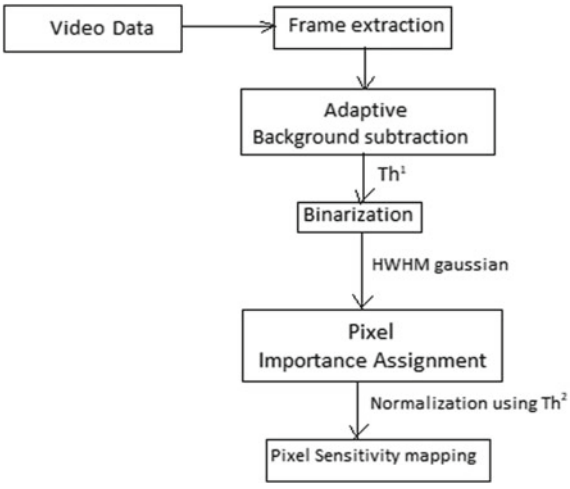


Fig. 2 Proposed methodology for sensitivity allotment of active pixels

Initially, frames are extracted from the video data which undergo adaptive background subtraction. A 10 s traffic surveillance video dataset at 30 fps with 360×640 pixel resolution has been utilized for the same. After adaptive background subtraction, the false activity removal can be used by binarization using Th^2 threshold as mentioned in Sect. 1. The binarized data is then multiplied by the HWHM Gaussian distribution to obtain pixel the importance factor. Normalization of the pixel importance with the maximum value of importance assigns the sensitivity on the basis of which activity mapping is done.

4 Results Obtained

The video dataset is a 10 s 30 fps traffic surveillance video. Some of the results after adaptive background subtraction are given in Fig. 3.

The total number of nonzero pixels obtained per frame (average) in a 360×640 resolution video frames is 20,733

$$\text{Total number of pixels in a frames: } 340 * 640 = 2, 30, 400.$$

The threshold obtained (Th^1) = (20,733/2,30,400) * 255 = 22.94 or 23 (rounded off). Some of the binarized frame results are shown in Fig. 4. The binarization is done using adaptive thresholding Th^1. Figure 5 shows the initial activity map utilizing

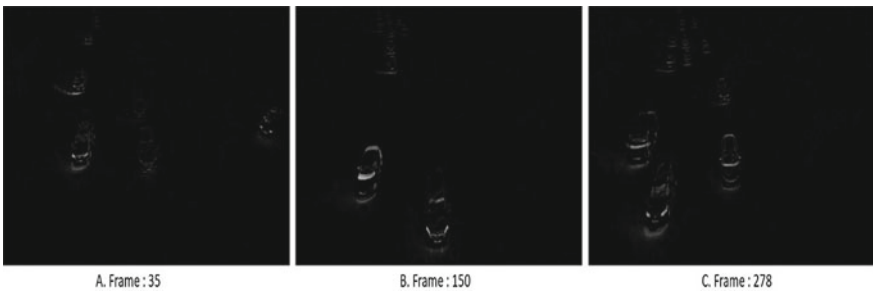


Fig. 3 Adaptive background subtraction at different frames of video dataset

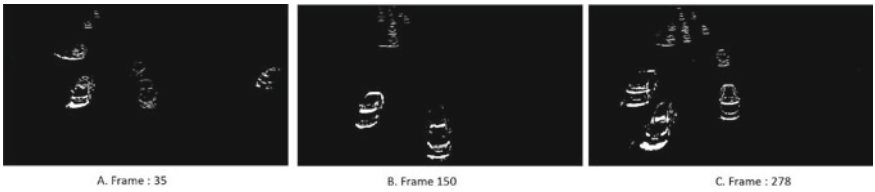


Fig. 4 Binarized activity of some frames from the dataset



Fig. 5 Initial activity map

binarizing. It can be seen that the activity map has many false positive detection values that occur due to the disturbance caused due to wind or camera distortion; hence this activity map is revised using Th^2. Figure 6 shows a normalized map after Th^2 displays more appropriate and correct activity mapping. Pixel sensitivity is then obtained by normalized activity mapping after Th^2. Figure 7 shows the pixel sensitivity allocation.



Fig. 6 Activity map post thresholding Th^2



Fig. 7 Sensitivity allocation

5 Conclusion

The paper introduced a novel approach for sensitivity allotment to every pixel in the camera FOV for prioritization of FOV on the basis of past and present activities. This work can be utilized for re-centering of the camera and calibrating intrinsic and extrinsic parameters of the camera such as focal length, pan, tilt, and zoom levels so that the highest priority area can be assigned at the center of FOV and the activities can be monitored with maximum information.

References

1. S. Indu, S. Chaudhury, N.R. Mittal, A. Bhattacharyya, Optimal sensor placement for surveillance of large spaces, in *2009 Third ACM/IEEE International Conference on Distributed Smart Cameras (ICDSC)* (IEEE, 2009), pp. 1–8
2. G. Zhang, B. Dong, J. Zheng, Visual sensor placement and orientation optimization for surveillance systems, in *2015 10th International Conference on Broadband and Wireless Computing, Communication and Applications (BWCCA)* (IEEE, 2015), pp. 1–5
3. L.C.B. da Silva, R. Maroquio Bernardo, H.A. de Oliveira, P.F.F. Rosa, Multi-UAV agent-based coordination for persistent surveillance with dynamic priorities, in *2017 International Conference on Military Technologies (ICMT)* (IEEE, 2017), pp. 765–771
4. U. Jaenen, M. Huy, C. Grenz, J. Haehner, M. Hoffmann, Distributed three-dimensional camera alignment in highly-dynamical prioritized observation areas, in *2011 Fifth ACM/IEEE International Conference on Distributed Smart Cameras* (IEEE, 2011), pp. 1–6
5. M.A. Jamshed, M.F. Khan, K. Rafique, M.I. Khan, K. Faheem, S.M. Shah, A. Rahim, An energy efficient priority based wireless multimedia sensor node dynamic scheduler, in *2015 12th International Conference on High-capacity Optical Networks and Enabling/Emerging Technologies (HONET)* (IEEE, 2015), pp. 1–4

6. H.-P. Shiang, M. van der Schaar, Online learning in autonomic multi-hop wireless networks for transmitting mission-critical applications. IEEE J. Sel. Areas Commun. **28**(5), 728–741 (2010)
7. J.M. McHugh, J. Konrad, V. Saligrama, P.-M. Jodoin, Foreground-adaptive background subtraction. IEEE Signal Process. Lett. **16**(5), 390–393 (2009)

Spectral–Spatial Active Learning with Attribute Profile for Hyperspectral Image Classification



Kaushal Bhardwaj, Arundhati Das and Swarnajyoti Patra

1 Introduction

Hyperspectral remote-sensing images acquire measurements in a large number of contiguous spectral bands with a narrow spectral bandwidth. Because of its ability to detect narrow absorption features, hyperspectral data are helpful in the classification of specific vegetation physiochemical characteristics, ocean's biological constituents, soil's physical and chemical properties, mineral composition, urban establishment, and snow characteristics [5]. The classification of such huge data is challenging and the availability of training samples in small numbers makes it more challenging. To handle such a situation, two different machine learning strategies are available in the literature. First is semi-supervised learning and the other is active learning (AL). Semi-supervised learning improves its decision boundaries by employing both labeled and unlabeled samples in the training phase [3]. However, the AL techniques optimize the classification performance by iteratively identifying informative samples for training. The AL process starts with a few labeled samples and iteratively appends informative samples (chosen by evaluating a query function) to the training set. The classifier model is retrained and the process is continued until the classification results are stable [10, 11].

The part of the AL procedure that is of prime importance is the definition of a query function. The query function is responsible to select informative samples from the unlabeled pool for a manual class assignment. For doing so, the query function uses one or more criteria to judge the information content of each sample. In the literature, several query functions are presented based on uncertainty, diversity, cluster assump-

K. Bhardwaj · A. Das · S. Patra (✉)
Department of Computer Science and Engineering, Tezpur University,
Tezpur 784028,
India
e-mail: swpatra@tezu.ernet.in

K. Bhardwaj
e-mail: kauscsp@tezu.ernet.in

A. Das
e-mail: arundha@tezu.ernet.in

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_119

tion, query-by-bagging, etc. [11–13, 15, 16]. The family of uncertainty criteria aims at discovering the samples whose class assignment is most ambiguous [4, 9], while the family of diversity criteria aims at avoiding the selection of redundant training samples by analyzing the dissimilarity among the unlabeled samples and selecting a set of most dissimilar samples [2, 16]. In the literature, the combination of these two criteria is suggested for better results [9]. Other methods such as entropy-based query-by-bagging [16] and cluster assumption with histogram thresholding [2] are also state-of-the-art AL methods. All these AL methods are totally based on spectral values of the HSI. Nonetheless, the spatial content of the image may play a significant role in class discrimination of pixels. This is because the neighboring pixels are typically correlated. In the literature, some methods are presented to integrate spectral and spatial information [8]. Morphological attribute profiles are a popular choice for this purpose [1, 7, 8].

In this paper, we propose a spectral–spatial active learning model where the spectral and spatial information content of the HSI is integrated with the help of attribute profiles. The state-of-the-art AL methods based on uncertainty, diversity, cluster assumption, and entropy-based query-by-bagging are used to demonstrate the effectiveness of the proposed model. In order to construct the AP for the HSI, the dimension of the HSI is reduced using the principal component analysis (PCA). On each PC, the multiscale attribute filtering results are obtained and stacked together to form an extended AP (EAP). Such constructed EAP for the HSI contains rich spectral–spatial information which is useful in recognizing the class uncertainty. Experiments are carried out on two real HSI datasets where the proposed model using spectral–spatial information outperforms the classic model that uses spectral measurements alone.

2 Proposed Spectral–Spatial AL Model

The proposed spectral–spatial AL method works in two stages. In the first stage, a spectral–spatial profile (extended attribute profile) is constructed for the given HSI and each pixel of the HSI is replaced by EAP features to be stored in an unlabeled pool. In the second stage, the AL procedure is executed to identify the informative samples (pixels) for labeling. The two-stage procedure is explained in Fig. 1. After the construction of the unlabeled pool in the first stage, a few samples from each class are randomly chosen and assigned into a labeled pool after labeling and the rest of the samples are kept in the unlabeled pool. Next, the classifier model is trained using the initial set of samples in the labeled pool. Then iteratively, a batch of samples are identified from the unlabeled pool, assigned the class labels, and appended to the labeled pool. The iterative procedure stops when the classifier model starts producing stable results. The two stages of the proposed spectral–spatial AL model are described below.

Construction of extended attribute profiles: The dimension of the given HSI is reduced using PCA. For each PC, separate APs are constructed and are concatenated together to form an EAP. An AP is constructed by combining the original image with

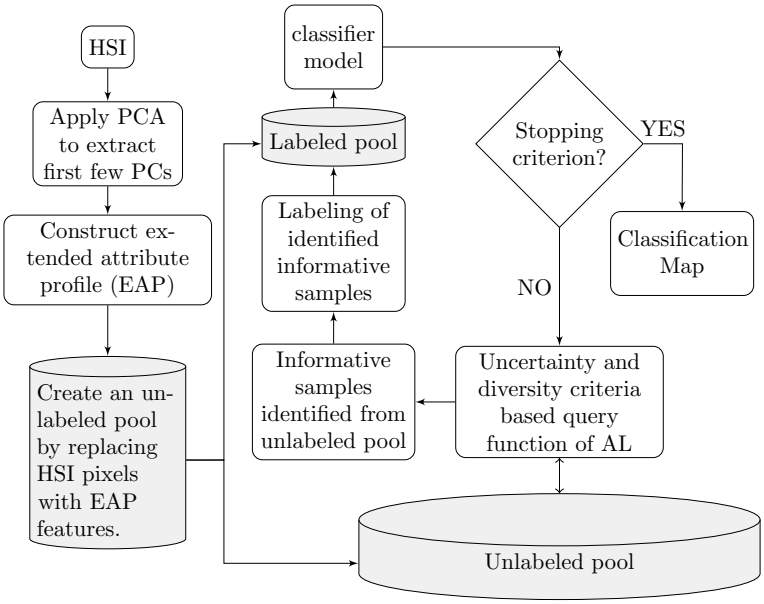


Fig. 1 Proposed spectral–spatial active learning framework. An EAP is constructed for the HSI to represent spectral–spatial information and stored in the unlabeled pool. A few unlabeled samples are stored in the labeled pool after labeling and the classifier is trained. In each iteration, a batch of informative samples is identified until stable results are achieved

its attribute filtering results [8]. Next, we demonstrate the construction of AP and EAPs based on attribute filtering.

Attribute filtering: The given image is represented by a tree representation called max-tree [14]. In this representation, the root has pixels with the lowest intensity and the children are the nested connected components. The leaf nodes have the components with maximum intensity. Similarly, a min-tree can also be created where the leaf nodes have the components with minimum intensity. Each node of such a tree represents a connected component in the image and its children represent the nested connected components. For each node, an attribute value (such as area or standard deviation) is calculated and stored at the node. Then, the attribute values at the nodes of the tree are compared against a given threshold. The node that does not satisfy the criteria is merged with their parents. Finally, restitution of the filtered tree is carried out to get a new gray-scale image corresponding to the filtered tree [14]. By considering multiple threshold values, different filtered images can be obtained. Such attribute filtering on an image considering a max-tree filters out the bright objects which do not satisfy the criterion and are called the attribute-thinning operation whereas min-trees are used to filter dark objects and are called attribute-thickening operations [8].

Attribute profile: Multiple thinning and thickening results obtained considering a sequence of threshold values are stacked together to form an attribute profile [8]. For a given image I, an AP can be defined as follows.

$$AP(I) = \{\phi^{T_L}(I), \phi^{T_{L-1}}(I), \ldots, \phi^{T_1}(I),\ I,\ \gamma^{T_1}(I), \gamma^{T_2}(I), \ldots, \gamma^{T_L}(I)\} \quad (1)$$

Here, $\phi^{T_i}(I)$ and $\gamma^{T_i}(I)$ represent thickening and thinning operations of image I with ith threshold. L is the number of thresholds considered for creating the profile. For L thresholds, the size of created AP will be $2 \times L + 1$.

Extended attribute profile: For an HSI, APs are constructed for each band in its reduced dimension and are stacked together to form an EAP [1, 7]. This is formulated for an HSI H considering ℓ PCs as follows.

$$EAP(H) = \{AP(PC_1), AP(PC_2), \ldots, AP(PC_\ell)\} \quad (2)$$

Selection of informative samples: For selecting the most informative samples to train the classifier, the AL technique starts with a small number of training samples (e.g., three samples per class). Let U be the set of all available unlabeled samples and L be the initial set of labeled samples. In each iteration of AL procedure, h unlabeled samples are selected for manual labeling and are appended to the existing labeled set L. This process is iterated until the classification results get stabilized. For selecting informative samples in each iteration, a query function is employed which exploits one or more criteria. These criteria can be based on *uncertainty*, *diversity*, *cluster assumption*, *query-by-bagging* or their combinations. In this paper, we discuss a few of these criteria and their combinations.

Uncertainty criteria: Several uncertainty criteria exist in the literature [4, 9]. Here, we discuss two state-of-the-art uncertainty criteria based on the support vector machine classifier. The first is margin sampling (MS) and the second is multi-class level uncertainty (MCLU). In case of MS, the AL procedure tries to identify the samples nearest to the separating hyperplane. The samples having the lowest classification certainty (CC) are selected for labeling. The MS can also be used with multiple classes where for c different classes, the distances from c hyperplanes are recorded and the minimum distance is considered for CC. This can be formulated for a sample s as follows.

$$CC(s) = \min_{i=1,2,\ldots c} \{|f_i(s)|\} \quad (3)$$

The MCLU criterion aims at identifying the sample that has maximum difference among the distances from the farthest separating hyperplanes. For this, the distance from each separating hyperplane is recorded; the two largest distances are noticed and the difference of these distances are recorded as CC. For a sample s, MCLU can be formulated as follows.

$$\begin{aligned} r_{mx1} &= \arg \max_{i=1,2,\ldots,c} f_i(s) \\ r_{mx2} &= \arg \max_{\substack{j=1,2,\ldots,c \\ j \neq r_{mx1}}} f_j(s) \\ CC(s) &= f_{r_{mx1}}(s) - f_{r_{mx2}}(s) \end{aligned} \tag{4}$$

Diversity: In the literature, several diversity criteria exist based on angle, closest support vector, and clustering [2, 16]. Here, we present two state-of-the-art diversity criteria: one is angle-based diversity (ABD) and the other is extended cluster based diversity (ECBD). In cluster-based diversity (CBD), h clusters are formed out of unlabeled samples for selecting h samples (one from each group). In case of ECBD, the clustering is done in a *kernel* space. In case of ABD, angles between unlabeled samples are computed and the h samples with maximum angle are selected for labeling. For two random unlabeled samples s_i and s_j, the angle Ang^{ABD} is obtained by employing a kernel function $K(s_i, s_j)$ shown as follows.

$$Ang^{ABD}(s_i, s_j) = \frac{K(s_i, s_j)}{\sqrt{K(s_i, s_i)K(s_j, s_j)}} \tag{5}$$

For the combination of both criteria, there can be two approaches. In the first approach, for each sample both, the uncertainty and diversity criteria are calculated and a weighted average is done. The weight is normally 0.5 for both. In the second approach (used in our work), v ($v > h$) samples are selected using uncertainty criteria and h samples are selected out of the v samples based on diversity. This approach is widely accepted. In this work, we present two state-of-the-art combinations namely MS-ABD and MCLU-ECBD [9].

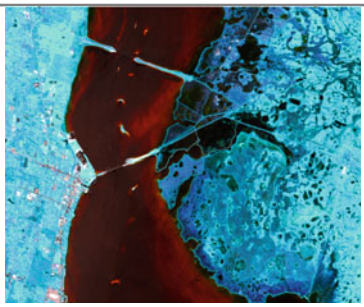
Entropy-based query-by-bagging and cluster assumption based criteria: In the literature, an entropy-based query-by-bagging (EQB) method is suggested that attempts to select h samples which have maximum disagreement between the committee of classifiers obtained by bagging [2]. A cluster assumption based histogram thresholding (CAHT) [16] is presented in the literature that attempts to select the samples from the low density region of the feature space. These methods perform better than many AL methods in the literature.

3 Experimental Results

Hyperspectral datasets and experimental setting: In the experiments, two benchmark real hyperspectral datasets[1] are being used. The false color image and information related to them are shown in Table 1. In the experimental analysis, the classic AL

[1] Available online: http://www.ehu.eus/ccwintco/index.php?title=Hyperspe-ctral_Remote_Sensing_Scenes.

Table 1 The two widely used hyperspectral datasets, related information, and false color image

Name	Description	False color image
KSC	Place: Kennedy Space Center (KSC) Merritt Island, Florida, USA Size: 512 × 614 Bands when acquired: 224 Bands available for processing: 176 Resolution: 18 m No. of classes:13 Available labeled samples: 5211	
University of Pavia	Place: University of Pavia, Italy Size: 610 × 340 Bands when acquired: 115 Bands available for processing: 103 Resolution: 1.3 m No. of classes: 9 Available labeled samples: 42,776	

model based on spectral values alone is compared to the proposed spectral–spatial model. For this, the results are obtained for the state-of-the-art methods, namely EQB, CAHT, MS-ABD, and MCLU-ECBD considering only spectral values and also by considering the EAP constructed as described in the proposed model. The methods considering EAP are named as EAP-EQB, EAP-CAHT, EAP-MS-ABD, and EAP-MCLU-ECBD, respectively. The initial labeled set comprises three samples randomly selected from each class. The batch size h is kept as 20 for the experiment. The value of v is kept as $3 \times h$ in the experiments of MS-ABD, MCLU-ECBD, EAP-MS-ABD, and EAP-MCLU-ECBD where first, the v samples are selected using uncertainty criterion and the final h samples are selected out of v samples using diversity criterion. For the construction of the EAP, the dimension of HSI is reduced using PCA and the first five PCs corresponding to maximum variance are considered. The EAP is constructed using the *area* attribute and the considered threshold values are $\{100, 500, 1000, 5000\}$. The threshold values are the same as those considered in [7]. The size of the EAP constructed for the HSI considering five PCs and four threshold values is 45 (nine for each PC).

All the implementations are done in MATLAB (R2015a). A one-Vs-all SVM classifier is used with the RBF kernel implemented with the help of the LIBSVM library [6]. The parameters of SVM are obtained following five-fold cross-validation

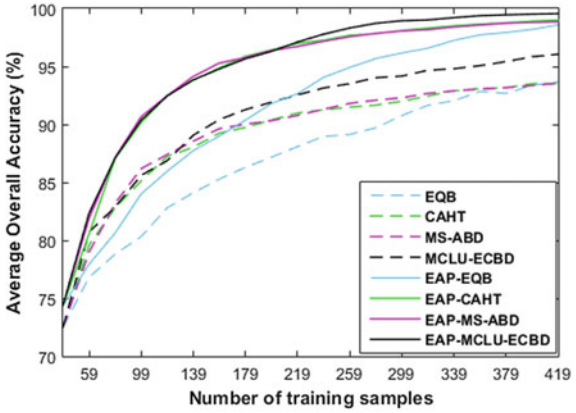


Fig. 2 Average overall accuracy obtained against increasing number of training samples obtained by different AL methods with and without EAP on the KSC dataset

in a grid search fashion. For simplicity, during the AL iterations, these parameter values were not changed.

Results on KSC dataset: The first experiment is conducted on the KSC dataset where the AL methods EQB, CAHT, MS-ABD, MCLU-ECBD, EAP-EQB, EAP-CAHT, EAP-MS-ABD, and EAP-MCLU-ECBD are applied ten times considering different randomly selected initial training samples. The results reported in Table 2 show that the AL methods working in the classic model depending on spectral values alone lead to poor results in comparison to the proposed model. One can observe from the table that EQB is the worst performing method which when executed in the proposed spectral–spatial model (EAP-EQB), outperforms the best performing MCLU-ECBD in the classic model. Moreover, the EAP-MCLU-ECBD has outperformed all other state-of-the-art methods. The performance of the proposed model is also visible from the plotted graph for all the methods in Fig. 2. The AL methods have started with a sharp increase in accuracy and went for stabilizing near 99%. Therefore, the proposed model is robust for the classification of HSI with a limited labeled samples.

Results on University of Pavia dataset: The second experiment is conducted on the University of Pavia dataset. Analogues to the previous experimental results, again the AL methods in the proposed model outperform their own results in the classic model. Table 3 demonstrates the results obtained for the University of Pavia dataset. All the EAP-based methods have better results than the best AL methods using spectral values alone. This can also be observed from Fig. 3 where the average overall accuracies for the proposed model is always above the best method in the classic model. This confirms the spectral–spatial content of EAP as well as the robustness of the proposed method in identifying the informative pixels for classification of HSI having limited training samples.

Table 2 Class-wise average accuracies, average overall accuracy, its related standard deviation, and average kappa coefficient obtained after ten runs of the experiment on KSC dataset. The best values are in bold face

Class	Methods							
	EQB	CAHT	MS-ABD	MCLU-ECBD	EAP-EQB	EAP-CAHT	EAP-MS-ABD	EAP-MCLU-ECBD
Scrub	94.599	98.108	97.845	97.937	99.698	99.842	99.934	**100.00**
Willow swamp	87.942	89.136	93.498	89.835	94.486	97.984	**98.560**	**98.560**
Cabbage palm hammock	58.516	95.234	95.039	95.195	97.461	98.359	**99.297**	99.063
Cabbage palm/-Oak hammock	91.706	55.595	56.865	85.595	93.175	96.944	98.214	**98.492**
Slash pine	78.634	73.851	72.050	86.894	95.714	93.851	94.224	**94.845**
Oak/-Broadleaf hammock	84.367	71.354	75.852	81.354	99.345	99.214	99.127	**99.782**
Hardwood swamp	90.476	90.857	81.619	91.524	98.095	95.048	98.286	**99.429**
Graminoid marsh	94.548	94.849	95.545	97.494	99.466	99.304	99.791	**99.907**
Spartina marsh	98.962	99.173	98.962	99.404	98.865	99.231	99.615	**99.885**
Cattail marsh	98.292	97.995	98.614	98.317	98.243	98.366	96.262	**99.356**
Salt marsh	99.403	99.212	99.189	98.807	99.451	99.594	99.714	**99.857**
Mud flats	99.304	96.859	95.408	97.972	99.404	99.662	97.237	**100.00**
Water	98.857	99.914	99.310	99.126	99.633	99.946	99.676	**100.00**
$\overline{OA}$	93.665	93.539	93.546	96.053	98.584	98.994	98.858	**99.553**
kappa	0.9295	0.9280	0.9281	0.9560	0.9842	0.9888	0.9873	**0.9950**
std.	0.8602	0.4038	0.5462	0.2784	0.2611	0.1134	0.4282	**0.0405**

Table 3 Class-wise average accuracies, average overall accuracy, its related standard deviation and average kappa coefficient obtained after ten run of experiment on University of Pavia dataset. The best values are in bold face

Class	Methods							
	EQB	CAHT	MS-ABD	MCLU-ECBD	EAP-EQB	EAP-CAHT	EAP-MS-ABD	EAP-MCLU-ECBD
Asphalt	91.436	89.036	89.469	91.081	98.726	99.180	99.570	**99.646**
Meadows	71.397	96.936	97.240	97.162	90.016	98.871	**99.616**	99.198
Gravel	39.967	70.276	67.132	69.238	85.607	94.859	98.166	**98.718**
Trees	86.172	92.249	92.288	92.428	95.193	97.836	99.070	**99.523**
Metal sheets	98.684	97.755	98.401	98.379	99.829	99.822	99.814	**99.963**
Bare soil	82.913	86.874	86.107	86.244	86.711	97.763	98.694	**99.254**
Bitumen	45.256	71.774	66.376	73.677	93.293	98.090	99.030	**99.880**
Self-blocking bricks	94.864	85.970	86.896	86.279	93.346	98.398	98.862	**99.223**
Shadows	97.276	99.155	98.828	99.250	99.736	**99.989**	99.873	99.863
$\overline{OA}$	78.012	91.233	91.116	91.644	92.044	98.507	99.319	**99.336**
kappa	0.7201	0.8834	0.8817	0.8888	0.8961	0.9802	0.9910	**0.9912**
std	3.3212	**0.1760**	0.4065	0.4138	3.2337	0.2649	0.1900	0.3337

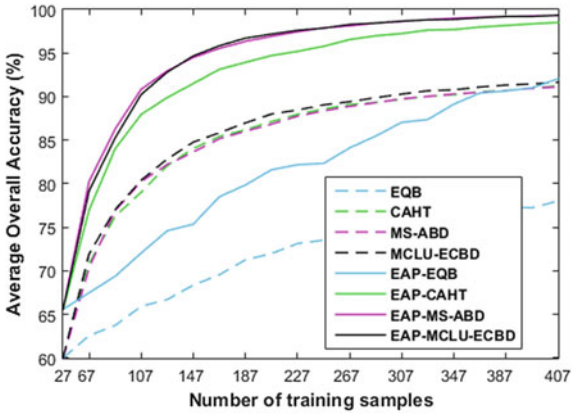


Fig. 3 Average overall accuracy obtained against increasing number of training samples obtained by different AL methods with and without EAP on University of Pavia dataset

4 Conclusion

This paper presents the state-of-the-art literature of the active learning techniques available for the classification of HSI with limited number of training samples. The paper also proposes a spectral–spatial model for the AL techniques which comprises constructing an expended attribute profile for HSI which is a state-of-the-art method for integrating spectral and spatial information. The state-of-the-art AL methods EQB, CAHT, MS-ABD, and MCLU-ECBD are executed considering spectral values alone and in the proposed spectral–spatial model. The AL methods executing in the proposed spectral–spatial model outperformed the best method in the classic model considering the spectral values alone. This confirms that EAP is rich in spectral–spatial information as well as the proposed AL model is effective enough for classification of HSI with limited training samples.

References

1. K. Bhardwaj, S. Patra, An unsupervised technique for optimal feature selection in attribute profiles for spectral-spatial classification of hyperspectral images. ISPRS J. Photogr. Remote Sens. **138**, 139–150 (2018)
2. K. Brinker, Incorporating diversity in active learning with support vector machines, in *Proceedings of the 20th International Conference on Machine Learning (ICML-03)* (2003), pp. 59–66
3. L. Bruzzone, M. Chi, M. Marconcini, A novel transductive SVM for semisupervised classification of remote-sensing images. IEEE Trans. Geosci. Remote Sens. **44**(11), 3363–3373 (2006)
4. C. Campbell, N. Cristianini, A. Smola, et al., Query learning with large margin classifiers, in *ICML* (2000), pp. 111–118
5. G. Camps-Valls, D. Tuia, L. Bruzzone, J.A. Benediktsson, Advances in hyperspectral image classification: earth monitoring with statistical learning methods. IEEE Sig. Process. Mag. **31**(1), 45–54 (2014)
6. C.C. Chang, C.J. Lin, LIBSVM: a library for support vector machines. ACM Trans. Intell. Syst. Technology (TIST) **2**(3), 27 (2011)
7. M. Dalla Mura, J.A. Benediktsson, B. Waske, L. Bruzzone, Extended profiles with morphological attribute filters for the analysis of hyperspectral data. Int. J. Remote Sens. **31**(22), 5975–5991 (2010)
8. M. Dalla Mura, J.A. Benediktsson, B. Waske, L. Bruzzone, Morphological attribute profiles for the analysis of very high resolution images. IEEE Trans. Geosci. Remote Sens. **48**(10), 3747–3762 (2010)
9. B. Demir, C. Persello, L. Bruzzone, Batch-mode active-learning methods for the interactive classification of remote sensing images. IEEE Trans. Geosci. Remote Sens. **49**(3), 1014–1031 (2011)
10. S. Patra, K. Bhardwaj, L. Bruzzone, A spectral-spatial multicriteria active learning technique for hyperspectral image classification. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. **10**(12), 5213–5227 (2017)
11. S. Patra, L. Bruzzone, A fast cluster-assumption based active-learning technique for classification of remote sensing images. IEEE Trans. Geosci. Remote Sens. **49**(5), 1617–1626 (2011)
12. S. Patra, L. Bruzzone, A batch-mode active learning technique based on multiple uncertainty for SVM classifier. IEEE Geosci. Remote Sens. Lett. **9**(3), 497–501 (2012)

13. S. Patra, L. Bruzzone, A novel SOM-SVM-based active learning technique for remote sensing image classification. IEEE Trans. Geosci. Remote Sens. **52**(11), 6899–6910 (2014)
14. P. Salembier, A. Oliveras, L. Garrido, Antiextensive connected operators for image and sequence processing. IEEE Trans. Image Process. **7**(4), 555–570 (1998)
15. A. Singla, S. Patra, A fast partition-based batch-mode active learning technique using SVM classifier. Soft Comput. **22**(14), 4627–4637 (2018)
16. D. Tuia, F. Ratle, F. Pacifici, M.F. Kanevski, W.J. Emery, Active learning methods for remote sensing image classification. IEEE Trans. Geosci. Remote Sens. **47**(7), 2218–2232 (2009)

Secrecy Performance Analysis of Hybrid-Amplify-and-Decode-Forward (HADF) Relaying Scheme Under Multi-hop Scenario



Shweta Pal and Poonam Jindal

1 Introduction

Nowadays, many technological advancements are taking place in the field of wireless communication networks that help in bringing out the most thrust areas of research. Communication nowadays becomes more prone to various security breaches that make the network insecure. Physical Layer Security (PLS) has been emanated as a new archetype which has gained a lot of importance among research areas for providing security to wireless communication networks. A secure communication without any key exchange is illustrated in a more precise manner by a Gaussian wiretap channel [1], provided the capacity of the main channel is better than that of the eavesdropper's channel. Cooperative communication also plays a vital role in improving the secrecy performance of any wireless network by providing strong signal strength between two relaying nodes. Commonly used protocols for relaying are Amplify and Forward (AF) and Decode and Forward (DF) [2]. For DF protocol, relaying takes place in two transmission slots. During the first transmission phase, the relay node decodes the transmitted information signal and in the next slot, re-encoding and retransmission of the received signal take place. Similar to DF, AF also took two time slots performing amplification during the first phase and retransmission and forwarding of information signal during the second. While the AF protocol results in amplification of noise along with relay nodes, the DF protocol on other hand results in a severe error propagation during incorrect decoding of the signal. This is the reason to consider a technique that can easily overcome these drawbacks of AF and DF. Hybrid-Amplify-and-Decode-Forward scheme was

S. Pal (✉) · P. Jindal
National Institute of Technology Kurukshetra, Haryana, India
e-mail: shwetapal6729162@gmail.com

P. Jindal
e-mail: poonamjindal81@nitkkr.ac.in

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_120

proposed in [3, 4] that collectively utilizes the benefits of both AF and DF by toggling between the two accordingly. The hybrid relaying scheme is further extended by authors in [5] to multiple relaying case and results are obtained based on their performance analysis. For single relay case, hybrid relaying for fixed AF and DF schemes based on SNR value is proposed [6]. Multi-hop wireless communication system over α-κ-μ and α-κ-μ-extreme fading channels for DF scheme is evaluated in [7] based on their probability density function (pdf) and cumulative density function (cdf) expressions of the received SNR value. Here in this paper, a multi-hop environment is implemented for a hybrid scheme that can adaptively choose between AF and DF based on SNR conditions. Secrecy performance of this proposed HADF scheme for multi-hop scenario is analyzed based on the distance between nodes. The rest of the manuscript is categorized as follows. In Sect. 2, system model for the multiple-hop environment employing the HADF protocol is described. In Sect. 3, performance analysis of different relaying schemes in terms of secrecy capacity is discussed. Simulated results are presented in Sect. 4 followed by conclusion in Sect. 5.

2 System Model

We consider a 3-hop multi-relaying system model employing the HADF relaying scheme in the presence of a passive eavesdropper. The system model contains a single source (T), destination node point (D), one passive eavesdropper node (E), and two trusted relays R_1 and R_2 as depicted in Fig. 1 [8]. Noise present at each of the individual network nodes is assumed to be Additive White Gaussian Noise with mean and variance as $(0, \sigma^2)$. Channel gains related from T to E, T to R, R to E, R to D, B to R, and B to T are represented as H_{TE}, H_{TR}, H_{RE}, H_{RD}, H_{BR}, and H_{BT}, respectively. Fading experienced by channel is Rayleigh fading and has access to Channel State Information (CSI) of the communicating nodes. A confidential message w(N) is sent by the source node T toward the destination node D via relays R_1 and R_2 under half-duplex mode in the presence of an eavesdropper. Direct links between T–D and T–E are assumed to be available. After receiving the desired signal from the prior adjoining node, each of the relay nodes transmits a jamming signal.

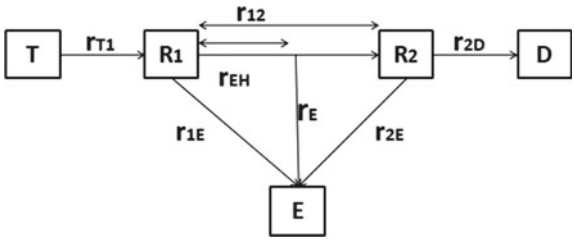


Fig. 1 Multi-hop relaying model

2.1 Cooperative Relaying Schemes

Different cooperative relaying schemes are used by relay nodes to process the signal after receiving the message signal from the source end. Fundamental schemes used to forward the message signal toward destination are DF and AF. Besides these schemes, a new adaptive relaying scheme known as Hybrid-Amplify-and-Decode-Forward (HADF) is introduced that cooperatively chooses between AF and DF based on decoding capabilities of the relay. The scheme is introduced by combining the benefits of both AF and DF. It switches between DF and AF based on SNR value associated with each participating node.

2.1.1 DF Protocol

It is a decode and forward relaying protocol involving a two-stage process. During the first transmission stage, the relay decodes the received information signal w(N) transmitted from T to R_1 and re-encodes and forwards it toward the destination D during the second transmission phase (Fig. 2). Perfect decoding of the signal is possible only when SNR of the received signal exceeds a certain threshold value. Signals arrived at both the destination and eavesdropper are given as [9]

$$y_{RD} = \sqrt{P_R} H_{RD} \mathrm{w(N)} + n_D \tag{1}$$

$$y_{RE} = \sqrt{P_R} H_{RE} \mathrm{w(N)} + n_E \tag{2}$$

Here, H_{RD} and H_{RE} are the channel fading coefficients from R-D and R-E, respectively, P_R is the transmitted power by the relay node, w(N) is the transmitted information signal using the DF cooperative scheme, and n_D and n_E are AWGN noise at D and E with mean and variance as (0, 1), respectively.

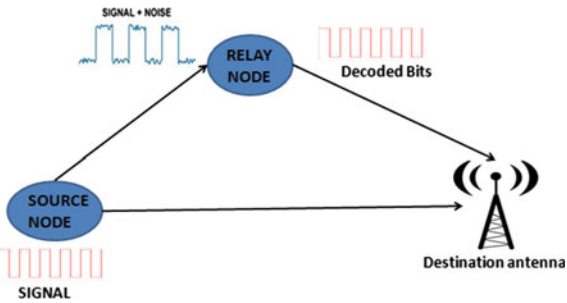


Fig. 2 DF protocol

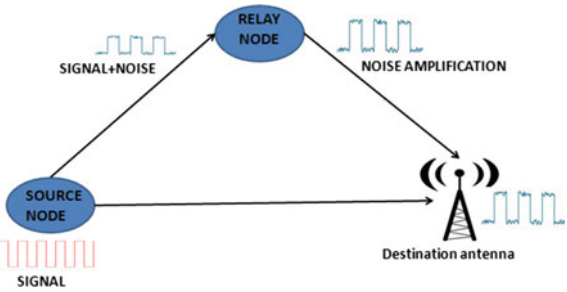


Fig. 3 AF protocol

2.1.2 AF Protocol

It is also a dual-stage process similar to the DF scheme. In the first stage, T sends a signal to R which is overheard by E simultaneously. The signal is first amplified by the relay node during the first transmission phase and then forwarded to the destination D during the second transmission phase (Fig. 3). But the disadvantage of AF relaying is that along with the message signal, it also amplifies the noise. Signals received at D and E are given as [9]

$$y_{RD} = \sqrt{P_R} H_{RD} v(\mathrm{N}) + n_d \tag{3}$$

$$y_{RE} = \sqrt{P_R} H_{RE} v(\mathrm{N}) + n_E \tag{4}$$

Here, $v(\mathrm{N}) = \frac{Y_{TR} H_{TR}}{\sqrt{P_R} |H_{TR}|^2}$ is the amplified signal, $v(\mathrm{N})$ is re-encoded signal received at R and n_d, n_E are AWGN noise at D and E, respectively, with mean, variance as (0, 1).

2.1.3 HADF Protocol

The DF protocol is used when perfect decoding of the relayed signal is possible and the relay and destination nodes are near to each other while when the relay node is far away from the source; AF gives better results as compared to DF. A new hybrid relaying protocol having features of both AF and DF is proposed to obtain the benefits of both the relaying schemes. The scheme is known as Hybrid-Amplify-and-Decode-Forward (HADF) relaying scheme. Whenever the signal propagates from one node to another and value of SNR between those nodes becomes greater than the predefined threshold value, the DF protocol outperforms AF, i.e., HADF is equivalent to DF for the case when signal can be decoded impeccably; else, AF is used (Fig. 4).

$$\begin{aligned} \mathrm{HADF} &= \mathrm{DF}; \text{If impeccable decoding of signal by relay.} \\ &= \mathrm{AF}; \text{else} \end{aligned}$$

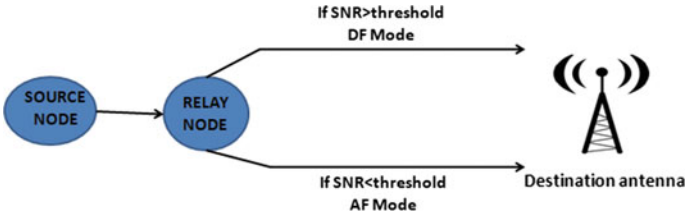


Fig. 4 HADF protocol

3 Secrecy Rate Analysis

Here, the secrecy rate has been evaluated as a performance metric for multi-hop environments. The secrecy rate is characterized as the rate of information transfer between source and receiver end through a secure and reliable link. According to Shannon's theorem, it is generally evaluated in terms of secrecy capacity which is given by the following expression [10]:

$$\mathrm{C} = 0.5 \log_2(1 + \mathrm{SNR}) \tag{5}$$

Overall secrecy capacity for all relaying schemes (both AF and DF) is given as [11]

$$C_S = [C_T - C_E]^+ \tag{6}$$

where C_E and C_T are eavesdropper channel capacity and data transmission capacity, respectively, and $[\mathrm{X}]^+ = \max(\mathrm{X}, 0)$.

3.1 DF Relaying Protocol

In the DF scheme, decoding of the received message signal is done initially by the relay node followed by re-encoding and retransmission.

SNR from source T to R_1, R_1 to R_2, and R_2 to D can be calculated as

$$\mathrm{SNR_{T1}} = \mathrm{P_T}\, \gamma_{T1} \tag{7}$$

$$\mathrm{SNR_{12}} = \mathrm{P_1}\, \gamma_{12} \tag{8}$$

$$\mathrm{SNR_{2D}} = \mathrm{P_2}\, \gamma_{2D} \tag{9}$$

where P_T, P_1, and P_2 are power transmitted to source, relay R_1, and relay R_2, respectively, and $\gamma_{T1} = \frac{|H_{T1}|^2}{\sigma^2}$, $\gamma_{12} = \frac{|H_{12}|^2}{\sigma^2}$, $\gamma_{2D} = \frac{|H_{2D}|^2}{\sigma^2}$ where H_{T1}, H_{12}, and H_{2D} are channel fading coefficients from T to R1, R_1 to R_2, and R_2 to D, respectively.

Thus, secrecy capacity at destination D using Eqs. (7), (8), and (9) is

$$C_T = 0.5 \log_2(1 + P_2 \gamma_{2D}) \tag{10}$$

Now SNR from source T to E, R_1 to E, and R_2 to E can be calculated as

$$\text{SNR}_{TE} = \frac{P_T \gamma_{TE}}{1 + P_{2J} \gamma_{2E}} \tag{11}$$

$$\text{SNR}_{1E} = \frac{P_1 \gamma_{1E}}{1 + P_{TJ} \gamma_{TE}} \tag{12}$$

$$\text{SNR}_{2E} = \frac{P_2 \gamma_{2E}}{1 + P_{1J} \gamma_{1E}} \tag{13}$$

where P_{TJ}, P_{1J}, and P_{2J} are jamming signal power, jamming power of relay R_1, and jamming power of relay R_2, respectively, and $\gamma_{TE} = \frac{|H_{TE}|^2}{\sigma^2}$, $\gamma_{1E} = \frac{|H_{1E}|^2}{\sigma^2}$, and $\gamma_{2E} = \frac{|H_{2E}|^2}{\sigma^2}$ where H_{TE}, H_{1E}, and H_{2E} are channel fading coefficients from T to E, R_1 to E, and R_2 to E, respectively.

Thus, secrecy capacity at the eavesdropper E using Eqs. (11), (12), and (13) is

$$\text{C}_E = 0.5 \log_2(1 + \text{SNR}_{TE} + \text{SNR}_{1E} + \text{SNR}_{2E}) \tag{14}$$

Overall secrecy capacity for DF is thus given by Eq. (6).

3.2 AF Relaying Protocol

In the AF scheme, amplification of the message signal is done by each of the relay nodes and then it is transmitted toward the destination node D. For the AF protocol, secrecy capacity at D is the same as that represented in Eq. (10) except the additional presence of an amplification factor G, where G is given as $G = \frac{1}{\sqrt{P_1 . |H_{1e}|^2 + N_0}}$

Thus secrecy capacity at D is given by

$$C_T = 0.5 \log_2\left(1 + G^2 (P_2 \gamma_{2D})\right) \tag{15}$$

Secrecy capacity at the eavesdropper E is the same as that given in Eq. (14) but with modification in SNR from R_1 to E and R_2 to E as

$$\text{SNR}_{1E} = \frac{G(P_1 \gamma_{1E})}{1 + P_{TJ} \gamma_{TE}} \tag{16}$$

$$\mathrm{SNR}_{2E} = \frac{G^2(P_2\gamma_{2E)}}{1 + P_{1J}\gamma_{1E}} \tag{17}$$

Overall secrecy capacity for AF also given by Eq. (6).

3.3 HADF Relaying Protocol

In HADF relaying, the DF mode is opted if decoding of the signal is perfect, i.e., ($SNR_{T,R} > threshold$), else if decoding is erroneous ($SNR_{T,R} < threshold$), AF mode is opted. Here threshold is given as target transmission rate.

For hybrid relaying, channel capacity is given as [4]

$$C_{\mathrm{HADF}} = \mathrm{Prob}(SNR_{T,Ri} > \mathrm{threshold})C_{\mathrm{DF}} + \mathrm{Prob}(SNR_{T,Ri} < \mathrm{threshold})C_{\mathrm{AF}} \tag{18}$$

where C_{DF} is the secrecy capacity for the DF mode and C_{AF} is the secrecy capacity for the AF mode for ith relay.

$$\mathrm{Prob}(\mathrm{SNR}_{\mathrm{T,R}i} > \mathrm{threshold}) \approx 1 - \pi_{i=1}^{N} \exp\left(\frac{-\mathrm{threshold}}{\mathrm{SNR}_{\mathrm{T,Ri}}}\right) \tag{19}$$

$$\mathrm{Prob}(SNR_{\mathrm{T,Ri}} < \mathrm{threshold}) \approx \pi_{i=1}^{N} \exp\left(\frac{-\mathrm{threshold}}{\mathrm{SNR}_{\mathrm{T,Ri}}}\right) \tag{20}$$

4 Simulated Results

Numerical results to investigate the secrecy performance of the HADF relaying protocol for multi-hop environments are presented in this section. The performance of this system employing a 3-hop multi-relaying network is evaluated in the form of secrecy capacity and is then compared with conventional AF and DF schemes. It is assumed that all significant nodes including T, D, R_1, and R_2 are located in LOS (Line of Sight) configuration as shown in Fig. 1, whereas the eavesdropper is located vertically away from the LOS configuration. Also, all nodes are employed with the Equal Power Allocation (EPA) scheme. For investigating the impact of various distances on secrecy capacity, channel between any of the two nodes will assume to follow the LOS model, i.e., its channel gain can be modeled as $r^{-C/2}\ \mathrm{e}^{\mathrm{j}\Phi,}$ where "r" denotes the distance between any two nodes while "Φ" is the random phase distributed evenly within the range of $[0, 2\pi)$ [12]. In Fig. 1, $r_{T1,}$ r_{12}, and r_{2D} represent the distances from T-R_1, R_1-R_2, and R_2-D, respectively. Furthermore, the distance from relay R_1 to E (r_{1E}) and the distance from relay R_2 to eavesdropper E (r_{2E})

are computed with the help of $r_{1E} = \sqrt{r_E^2 + r_{EH}^2}$ and $r_{2E} = \sqrt{r_E^2 + (r_{12} - r_{EH})^2}$, respectively. All the basic simulation parameters are shown in Table 1.

All the simulation results are obtained using MATLAB R2014a software. Threshold value for SNR comparison can be evaluated as

Threshold $= 2^{R\text{th}} - 1$, where Rth denotes the target transmission rate.

Figure 5 represents the plot of the secrecy capacity with respect to r_E for the HADF relaying scheme under a multi-hop environment. As the distance r_E increases, the eavesdropper is moving away from the LOS communication and hence secrecy improves. It has been observed that the secrecy capacity for the HADF relaying protocol comes out to be better than that of both AF and DF protocols by about 85.33% better as compared to DF and about 50% better as compared to AF (Table 2). Distance parameters for multi-hop environment are $r_{T1} = 25$ m, $r_{12} = 25$ m, $r_{2D} = 30$ m, and $r_{EH} = 0$ m.

Figure 6 illustrates the plot of secrecy capacity versus relay-to-destination distance, i.e., r_{2D} for HADF scheme under a multi-hop scenario. As the relay is moving away from the destination node, signal strength between them becomes weak and thus secrecy capacity for the HADF scheme degrades. It is about 68.8% better as compared to DF and about 24.59% better as compared to AF (Table 3). Distance parameters for multi-hop model are $r_{T1} = 25$ m, $r_{12} = 25$m, $r_E = 45$ m, $r_{EH} = 15$, and $c = 3.5$. The plot shows that when the distance r_{2D} increases, secrecy capacity tends to decrease.

Table 1 Parameters used for simulation

Path loss	3.5
Target transmission rate	2
No. of relay nodes	2
Relay network topology	LOS
Channel	Rayleigh fading
Total transmit power	30 dBm
Noise power	−30 dBm
r_{12}	25
r_{2D}	30
r_E	35
r_{EH}	15
r_{T1}	25

Table 2 Comparison table of different relaying schemes at $r_E = 24$m

Relaying scheme	Secrecy rate (bps/Hz)
DF	0.0198
AF	0.6879
HADF	1.3758

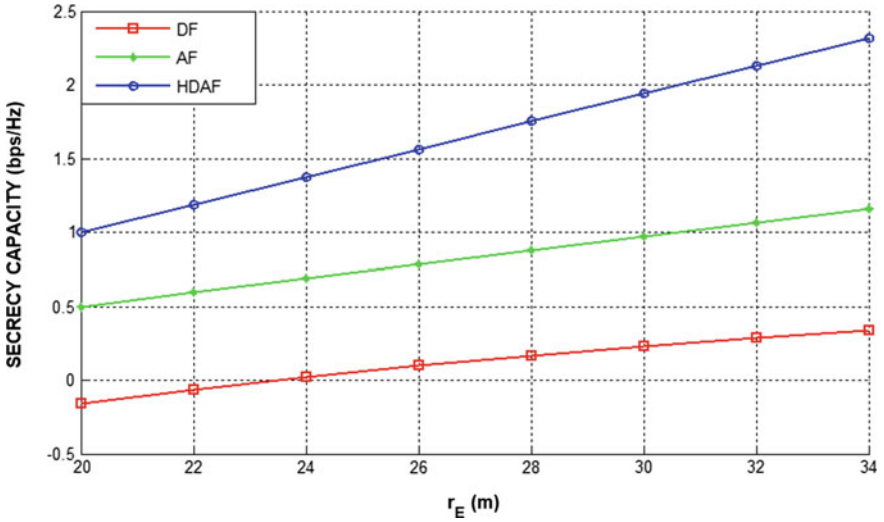


Fig. 5 Secrecy-capacity versus distance r_E

Table 3 Comparison table of different relaying schemes at $r_{2D} = 17$ m

Relaying scheme	Secrecy rate (bps/Hz)
DF	0.9459
AF	2.2871
HADF	3.0331

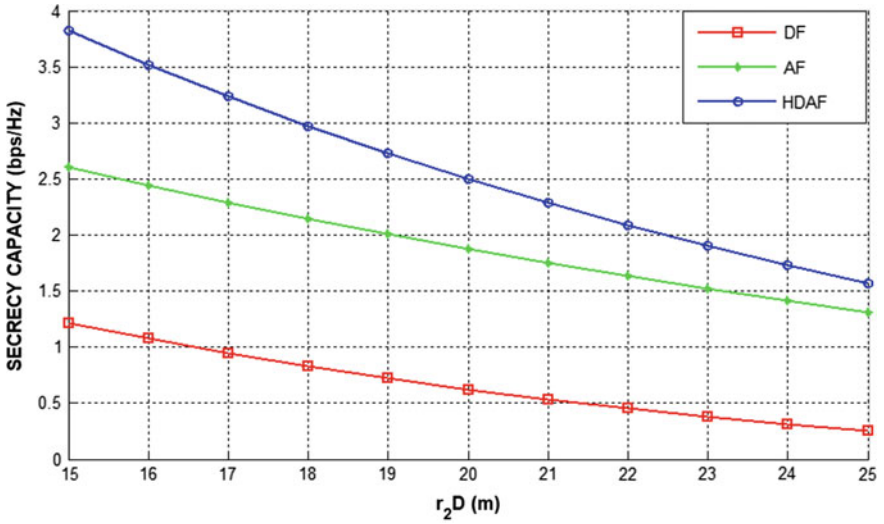


Fig. 6 Secrecy capacity versus distance r_{2D}

5 Conclusion

Performance analysis of the HADF relaying scheme based on SNR is evaluated under a multi-hop environment in the presence of two relay nodes. If ($SNR_{T,R} > threshold$), then perfect decoding of the signal takes place and DF is chosen else decoding is erroneous and the AF mode is opted. Multi-hop has many advantages in terms of improved coverage extension, lesser transmission power per participating node, and lower interference level when compared to single-hop networks. In this work, the HADF relaying protocol under a multi-hop environment is simulated. Results are obtained showing the comparison of secrecy capacity for the proposed HADF relaying scheme with the conventional multi-hop AF and DF techniques. Simulation results show that the secrecy capacity for HADF improves by about 85.33% and 50%, respectively, as compared to the DF and AF protocols as the distance r_E increases, and capacity degrades by about 68.8% and 24.59%, respectively, as compared to DF and AF scheme with an increase in r_{2D}. Thus, secrecy performance of the HADF scheme outperforms both DF and AF schemes under a multi-hop environment.

References

1. C. Xing, S. Ma, Z. Fei, Y. Wu, H.V. Poor, A general robust linear transceiver design for multi-hop amplify-and-forward MIMO relaying systems. IEEE Trans. Signal Process. **61**(5), 1196–1209 (2013)
2. J. Laneman, D. Tse, Cooperative diversity in wireless networks: Efficient protocols and outage behavior. IEEE Trans. Inf. Theory **50**, 3062–3080 (2004)
3. D. Thatha, K.K. Gurrala, S. Das, Performance analysis of hybrid decode-amplify-forward (HDAF) relaying for improving security in cooperative wireless network, in *2015 Global Conference on Communication Technologies (GCCT)* (2015), pp. 682–687
4. K.K. Gurrala, S. Das, Maximized channel capacity based power allocation technique for multi relay hybrid decode-amplify-forward cooperative network, in *Wireless Personal Communications*, vol. 87, Issue 3 (Springer Science, 2016), pp. 663–678
5. T.Q. Duong, H.J. Zepernick, Hybrid decode-amplify-forward cooperative communications with multiple relays, in *Proceedings of IEEE Wireless Communications and Networking Conference*, Budapest (2009), pp. 1–6
6. H. Chen, J. Liu, C. Zhai, L. Zheng, Performance analysis of SNR based hybrid decode amplify-forward cooperative diversity networks over Rayleigh fading channels, in *Proceedings of Wireless Communications and Networking Conference (WCNC)* (2010), pp. 1–6
7. T.R. Rasethuntsa, S. Kumar, M. Kaur, A Comprehensive Performance Evaluation of a DF-Based Multi-Hop System over α-κ-μ and α-κ-μ-extreme Fading Channels (2019)
8. L. Sanguinetti, A.A. DAmico, Y. Rong, A tutorial on the optimization of amplify-and-forward MIMO relay systems. IEEE J. Sel. Areas Commun. **30**(8), 1331–1346 (2012)
9. Y. Zou, X. Wang, W. Shen, Optimal relay selection for physical layer security in cooperative wireless networks. IEEE J. Sel. Area Commun. **31**(10) (2013)
10. J. Li, A.P. Petropulu, S. Weber, On cooperative relaying schemes for wireless physical layer security. IEEE Trans. Signal Process. **59**(10), 4985–4997 (2011)
11. J.H. Lee, Full-duplex relay for enhancing physical layer security in multi-hop relaying systems. IEEE Commun. Lett. **19**(4), 525–528 (2015)
12. J.-H. Lee, Full-duplex relay for enhancing physical layer security in multi-hop relaying systems. IEEE Access **19**(4) (2015)

A Review on Photonic Crystal Fibers



Arati Kumari Shah and Rajesh Kumar

1 Introduction

Optical fiber is a coaxial form of media of transmission that makes the optical channel more reliable and versatile. The most essential characteristics for any system of communication are signal-to-noise ratio and bandwidth as they determine the channel's capacity. For multimode, optical fiber connection currently has a loss of 0.2 dB/km at 1550 nm of wavelength. Single fiber optical link bandwidth is about 50 THz. Thus, the OFC system forms the backbone of the modern telecommunication system [1, 2].

But OFC is unable to provide flexibility in design. Thus, an attractive alternative fiber came into being. Photonic crystal fiber offers exceptional flexibility in design by simply varying its geometric dimensions. Photonic crystal fiber is a special type of optical fiber that consists of a micro-structured arrangement that surrounds low index material with higher refractive index material [3]. PCF's background material is usually pure silica. It is also recognized as holey fibers because it composed of a core region of defect surrounded by multiple air holes that run along the entire length of the fiber. The core/cladding index profile is the fundamental difference between PCF and traditional fiber. In particular, PCFs can be obtained with outstanding design flexibility by changing the geometry of the fiber cross-section air holes, i.e., their position or dimension. There is zero dispersion at 1310 nm in OFC. Using PCF can be specifically designed to obtain zero dispersion at separate operating wavelength dispersion compensating fiber. Photonic crystal fiber gives more flexibility in designing

A. K. Shah (✉) · R. Kumar
Department of ECE, NERIST, Itanagar, Arunachal Pradesh, India
e-mail: arati.shah20@gmail.com

R. Kumar
e-mail: itsrk2006@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_121

optical features such as birefringence, dispersion, confinement loss, effective area, and nonlinearity than conventional fibers [4].

2 Review on Literature

There has been a lot of research on Photonic Crystal Fiber, and the PCF technology has been modified by those researches. In 1978, the Bragg fiber idea was revolutionizing telecommunications with component sensors and filters, but there were the main disadvantages: no large modes, enormous size, and greater losses [4]. Later in 1992, the fiber design included the Total Internal Reflection method with good perforation in telecommunications except with a few problems such as limited choice of material, restricted core diameter for single-mode operation [3]. In 1996, the photonic coated fiber was manufactured with additional characteristics such as increased durability, high design strength, high-temperature resistance according to use in nuclear radiation, harsh chemical environments, and medical applications. In 1997, single-mode PCF with no higher order modes regardless of optical wavelength, low nonlinearity, and low confinement loss were used as filtering mode, sensors, interferometers, etc. [5]. In 1999, PCF with photonic bandgap air core was implemented as a different variety of wave-guide structures in the core of an array of air holes for various purposes [6].

In 2000, PCF of highly birefringent with different diameters of air hole along with the two orthogonal axes or high data rates and fiber loop production due to uneven core design. In the same year, supercontinuum generation has been generated by high nonlinear PCF and Zero Dispersion Wavelength applications in Pulse Compression, Laser Sour Spectroscopy, and WDM [7]. Later in 2001, manufacture of Bragg fiber eventually found uses in optical sensor, fiber laser, and PCF laser with double cladding (ytterbium-doped double-clad) provided high power was found by Fabry Perot specification. PCF with ultra-flattened dispersion was implemented in 2002 in which Zero Dispersion was acquired at a much broader wavelength range of 1–1.6 μm and was used primarily for supercontinuum generation. Bragg fiber with air core and silica was presented in 2003, reducing the loss of nonlinearity propagation and furthermore filled in as a model to consider the nonlinear optical stage materials [8]. Chalcogenide Photonic Crystal Fibers (CPCF) developed in 2004 offered a number of unique optical properties such as a transmission window that extends far into the infrared spectral region and demonstrates an extraordinarily high nonlinear refractive-index coefficient.

In 2005, Kagome Lattice PCF was implemented with a gas-filled hypo-cycloid fiber containing three very strong bandgaps overlapping to provide low loss at a very large wavelength range. The pressure and temperature of the gas can be observed and also the gas significant contribution to the refractive index, which was used to design bright temporally coherent optical sources [9]. Furthermore, in the year 2006, the creation of Hybrid Photonic Crystal Fiber, a type of PCF made up of air holes and germanium–silica rods prepared around an undoped silica core, which guides

light inside a core by Total Internal Reflection (TIR) and antiresonant reflection guidance, was used. Later in 2007, Silicon Double Inversion was used to produce photonic crystal polymer templates that were an intermediate approach where silica was produced at room temperature via Atomic Level Deposition (ALD). Hollow Core Photonic Band Fiber (HC-PCF) which was free of surface modes was developed in 2009. Due to the complete elimination of surface modes, there will be a substantial increase in fiber bandwidth and a reduction in dispersion may easily lead to more carrying capacity [10]. In 2013, the Double Cladding Seven Photonic Crystal Fiber was implemented in which each core was made to transmit only the basic mode known as the super mode and offered great support in making a multicore fiber with proper guiding properties for high-power supercontinuum generation [11].

Another very effective nanodisplacement sensor, which can work directly for horizontal as well as vertical displacement, was acquired in 2014 as a PCF based on a slightly different sensitivity nanodisplacement sensor. For mid-infrared supercontinuum generation, Photonic Crystal Fibers—with an equiangular 8 mm long PCF was intended in 2015. This would generate laser pulses with a high power of 500 W [12]. The PCFs were later integrated into a Fiber Laser. For high-power applications, a monolithic fiber with a 40 μm core with Ytterbium-doped PCFs amplifier configuration that generated up to 210 W average powers at 1064 nm was introduced. Helically twisted photonic crystal fibers (PCFs) were analyzed in 2016 based on the Helical Bloch theory. In this twisted periodic "space", cause spiral light across the fiber axis and will include dips in the transmission spectrum, and core-less PCF may have low loss guidance [13].

3 Types of Photonic Crystal Fibers

Photonic crystal fiber (PCF) can be illustrated as a structure comprising a core and clad, ensuring the propagation law of total internal reflection as in usual fiber. Periodic nanostructures influence photon motion as this ionic lattice affects electrons in solids. It occurs naturally in the shape of coloring the structure [4]. The core of this particular fiber is made of silica as a solitary material and can either be solid or hollow. The core is encompassed via air holes which experiences the fiber so it is called "holey" or micro-structured fiber and because of this structure, the light is restricted and transferred through the core which goes about as a cavity.

It tends to be composed of two kinds as indicated by various structures:

- Index guiding photonic crystal fiber
- Photonic bandgap fiber.

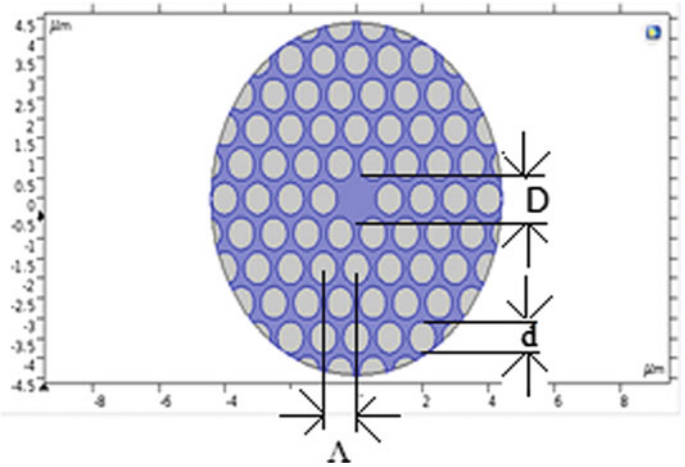


Fig. 1 Solid photonic crystal fiber [14]

3.1 Index Guiding Photonic Crystal Fiber

In index guiding PCF, light is centered by the total internal reflection between the solid core and various air-gap cladding. The solid core of the fiber controlling the PCF with a miniaturized scale basic exhibit of air gaps is encompassed by unadulterated silica cladding with refractive index of 1.462. Because of the huge refractive index differentiation between air (1.000) and silica (1.462), the light is centered by Total Internal Reflection which is totally an element of wavelength. Effective Refractive Index fundamentally measures the stage delay per unit length in PCF with respect to stage delay in vacuum.

In Fig. 1, the PCF comprises a missing air hole in the center of diameter meant by "D" and the pitch is named as "Λ" which estimates the separation between the focuses of the neighboring air holes. The hole measure is named as "d" [9].

3.2 Photonic Bandgap Fiber

Photonic bandgap fiber is obtained by the structure formed as if the core part of the air holes array is simply replaced by a much larger hole of much larger diameter in comparison to the surrounding holes. There is an adjustment in its optical properties because of the deformity of the broken structure of periodicity [15]. No electromagnetic modes are permitted to have a recurrence in the hole. Its impact is displayed in photonic crystal bandgap fiber where the wavelength controls light in a low index core region. The light controlling wonder in the fiber depends on the recurrence of the outside light if it matches the bandgap recurrence; the light gets limited in the

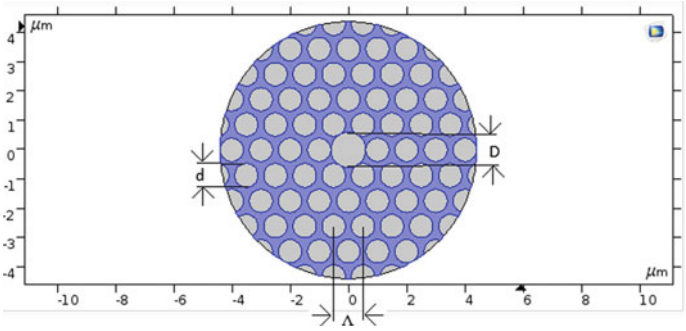


Fig. 2 Photonic bandgap fiber with a hollow cavity in the center [14]

holes and in like manner is guided all through the length of the fiber. So there is no prerequisite higher refractive index of the center as shown in Fig. 2 [15].

4 Analysis of Optical Properties

The properties of PCF such as birefringence, chromatic dispersion, confinement loss, effective mode area, nonlinearity, and zero-dispersion wavelengths are discussed as follows.

4.1 Birefringence

Birefringence is an important parameter in fiber optics and many detecting gadgets where light needs to hold a straight polarization area, regularly requiring high birefringence. Normally, materials with uniaxial anisotropy—the hub of symmetry is called the optical pivot of a specific material that has no comparable hub in the plane opposite to it—displays this optical phenomenon [16]. Linearly polarized light beam in parallel and opposite heading will express uneven effective refractive indices n_e and n_o for unexpected and normal developing beams separately. At the point when an un-captivated light emission goes through the material with a nonzero intense edge to the optical hub, the oppositely spellbound segment may confront refraction at an edge according to the ordinary law of refraction and its contrary part at a nonstandard point appears by the distinction between the two compelling refractive records called as the birefringence extent [17].

$$\Delta n = n_e - n_o \tag{1}$$

The distinction between the actual part values of the effective indices of the noticeable core own modes along with x-axis and y-axis—LP_{01x} and LP_{01y}.

$$B = \left| Re\left(n_{xeff}\right) - Re\left(n_{yeff}\right) \right| \tag{2}$$

4.2 Chromatic Dispersion

The total of waveguide and material dispersion adds to the chromatic dispersion or total dispersion. The material dispersion is a trademark for the use of material to create the fiber though the waveguide dispersion which can be fluctuated by changing the parameter of the waveguide in this way all out scattering is permitted to be modified. The material dispersion can be ignored when nm (λ) winds up steadily and the genuine piece of the effective index of refraction n_{eff} contains the scattering data [17].

$$D(\lambda) = -\frac{\lambda}{c}\frac{d^2 \text{Re}\left[n_{yeff}\right]}{d\lambda^2} \tag{3}$$

where c is the velocity of light in vacuum and λ is the operating wavelength.

4.3 Confinement Loss

The occurrence of limited air holes in the center region causes the optical mode to leak from the inner core region to the outer air holes and this is inevitable, resulting in confinement losses. Fundamental mode is used to calculate confinement loss from the imaginary part of the complex effective index n_{eff}, using

$$L_c = \frac{40\pi}{ln(10)\lambda} Im\left(n_{eff}\right) \tag{4}$$

Confinement loss is the leakage of light from the material of the core to the material of the external matrix. It can be changed according to the parameters like number of air holes, number of layers, air-hole diameter, and the pitch [17].

4.4 Effective Mode Area

Effective mode area (A_{eff}) of the PCF follows the below equations:

$$A_{\text{eff}} = \frac{\left(\iint |E|^2 \mathrm{d}x\mathrm{d}y\right)^2}{\iint |E|^4 \mathrm{d}x\mathrm{d}y} \tag{5}$$

Since E is the electric field amplitude. The integration is cross over the center zone, yet over the entire plane surface. A significant impact of a small effective mode area is that the optical intensities for a given power level are high, making nonlinearities important [18].

4.5 Nonlinearity

Nonlinear coefficient of PCF represents a very significant parameter during SCG analysis. Nonlinear coefficient (γ) is directly corresponding to nonlinear refractive index (n_2) and contrarily proportional to the effective area (A_{eff}) [18]. The nonlinear coefficient of PCF follows the below equations:

$$\gamma = \frac{2\pi n_2}{\lambda A_{eff}} \tag{6}$$

4.6 Zero Dispersion Wavelengths (ZDW)

For optical fibers, ZDW is the wavelength where the group delay dispersion (second-order dispersion) is zero. For PCFs with small mode areas, which can execute particularly strong waveguide dispersion, the ZDW can be shifted into the visible spectral region so that anomalous dispersion is obtained in the visible wavelength region, allowing for soliton transmission. PCFs, as well as some other fiber designs, can generate two or three different ZDW. SCG (Supercontinuum Generation) can lead to particularly broad optical spectra when the pump light has a wavelength near the ZDW.

5 Applications of PCF

A. A highly nonlinearly designed PCF with four strands of air holes with different diameters can be accessed for broadband supercontinuum generation that is used in dermatology, ophthalmology, dentistry, and detection [11].
B. A PCF-in-PCF structure shows ultra-flattened negative dispersion at a large range of wavelength ranging from 1360 to 1690 nm and can be utilized for residual dispersion compensation in optical transmission [17].

C. By varying the diameters of the inner air, holes can be utilized for supercontinuum generation and gives a flat dispersion profile for mid-infrared range from 1 to 10 μm. A highly nonlinear photonic hexagonal crystal fiber with a structure of five rings can be used [13].
D. Photonic crystal fiber with central core region doped with GeO_2, a butterfly lattice structure, and fiber Bragg grating (FBG) fetched in the core can be used as an optical fiber pressure sensor [15].
E. A chalcogenide glass PCF with a square lattice and hexagonal lattice structure with the pitch of 0.2 μm can be used as dispersion compensating fibers. In comparison to silica, this fiber makes available high negative dispersion in the wavelength range 1.2–1.6 μm [19, 20].

6 Conclusion

The unprecedented properties and features of photonic crystal fiber (PCF) strands such as spot measure, novel cutoff, and scattering properties besides allowing spillage-free wave bearing in a low-record focus region has extended the use of PCF in future applications. We have included the distinctive upgrades accomplished in the field of PCF in the recent couple of decades and a part of the focal and essential properties of these new fibers, in connection with conventional strands and demonstrated a bit of their impending future applications. We have in a like manner discussed the basic guided properties of this new class of fiber. These strands can be used in multidisciplinary applications in different fields with their phenomenally arranged features. In like manner, these strands are under wide research and have an extremely splendid future in the field of optical fiber correspondence.

References

1. N. Mahnot, S. Maheshwary, R. Mehra, Photonic crystal fiber-an overview. Int. J. Sci. Eng. Res. **6**(2), 45–53 (2015)
2. L. Chaudhary, A. Jb, H. Purohit, Photonic crystal fibre : developments, properties and applications in optical fiber communication. Int. J. Res. Appl. Sci. Eng. Technol. **5**(Xi), 1828–1832 (2017)
3. J.C. Knight, T.A. Birks, P.S.J. Russell, D.M. Atkin, All-silica single-mode optical fiber with photonic crystal cladding. Opt. Lett. **21**(19), 1547–1549 (1996)
4. R. Buczynski, Photonic crystal fibers. Acta Phys. Polonics **106**(2) (2004)
5. J.M. Dudley, G. Genty, S. Coen, Super continuum generation in photonic crystal fiber. Rev. Mod. Phys. **78**(4) (2006)
6. H.F. Wei, H.W. Chen, P.G. Yan, A compact seven-core photonic crystal fiber super continuum source with 42.3 W output power. Laser Phys. Lett. **10**(4), 04 (2013)
7. P.S.J. Russell, Photonic-crystal fibers. J. Lightwave Technol. **24**(12), 4729–4749 (2006)
8. K. Kaneshima, Y. Namihira, N. Zou, H. Hoga, Y. Nagata, Numerical investigation of octagonal photonic crystal fibers with strong confinement field. IEICE Trans. Electron. (2006)

9. R. Buczynski, Photonic crystal fibers. Acta Phys. Polonica (2) (2004) (Warsaw, Poland). Prog. Electromagn. Res. **106** (2009) (PIER 99,225,244)
10. H. Ademgil, S. Haxha, Design and optimisation of photonic crystal fibres for applications in communication systems, in *Proceedings of the World Congress on Engineering* (2007)
11. A.M.R. Pinto, M. Lopez-Amo, All-fiber lasers through photonic crystal fibers, in *Nanophotonics* (Science Wise Publishing, 2013)
12. S.K. Tripathy, J.S.N. Achary, N. Muduli, G. Palai, Nonlinear rectangular photonic crystal fiber (PCF) for optical communication exclusively super continuum generation. J. Laser Opt. Photonics (2015)
13. P.St.J. Russell, R. Beravat, G.K.L. Wong, Helically twisted photonic crystal fibres. Philos. Trans. R. Soc. (2017)
14. COMSOL Multiphysics User's Guide, Version: COMSOL 4.3 (2012)
15. M.-Y. Chen, R.-J. Yu, Member, OSA, A.-P. Zhao, Senior Member, IEEE, Confinement losses and optimization in rectangular-lattice photonic-crystal fibers. J. Light Wave Technol. **23**(9), (2005)
16. L. Xiao, M.S. Demokan, W. Jin, Y. Wang, C.-L. Zhao, Fusion splicing photonic crystal fibers and conventional single-modefibers: microhole collapse effect. J. Lightwave Technol. **25** (2007)
17. K.-Y. Y., Y.-F. Chau, Y.-W. Huang, H.-Y. Yeh, D.P. Tsai, Design of high birefringence and low confinement loss photonic crystal fibers with five rings hexagonal and octagonal symmetry air-holes in fiber cladding. J. Appl. Phys. (2011)
18. J. Johny et al., Determination of photonic crystal fiber parameters with effects of nonlinearities in supercontinuum generation, in *Optical Networking Technologies and Data Securities (OPNTDS)* (2012)
19. Md. Rabiul Hasan, S. Akter, A.A. Rifat, S. Rana, S. Ali, A highly sensitive gold-coated photonic crystal fiber biosensor based on surface plasmon resonance. Opt. Appl. **XLIV**(3) (2014)
20. S.H. Keerthana, S.K. Sudheer et al., Simulation of a highly nonlinear photonic crystal fiber using COMSOL multiphysics for supercontinuum generation, in IEEE International Workshop on Optical Networking Technologies and Data Security (2014). ISBN 978-1-4799-5291-5

Multiplier-Less Architecture for 4-Tap Daubechies Wavelet Filters Using Algebraic Integers



Mohd. Rafi Lone and Najeeb-ud-Din Hakim

1 Introduction

Discrete Cosine Transformation (DCT) has been replaced in most of the applications by Discrete Wavelet Transform (DWT) because DWT provides many advantages over DCT. Some of these include embedded bitstream coding, progressive image transformation, efficient bit-plane coding, ease of compressed image manipulation, and region of interest coding. In most of the recent compression algorithms including state-of-the-art JPEG 2000 compression standard, Discrete Wavelet Transform is being used extensively. This is due to Multi-Resolution Analysis (MRA) provided by DWT, which provides full control over both time and frequency domains simultaneously. DWT also provides good energy compaction as compared to DCT. The spatial orientation of sub-bands in one level with the corresponding sub-bands at other levels of DWT has changed the trend for new compression algorithms. This helps in achieving bit-plane coding and zero-tree coding simultaneously. The recent compression standards that are extensively being researched are JPEG 2000 and Set Partitioning in Hierarchical Trees (SPIHT), and both use DWT as the transformation technique. Traditionally, DWT was implemented using convolution. Lifting-based wavelet transformation, proposed by Sweldens [1, 2] replaces convolution-based wavelet transformation [3–6] in most of the applications. The high-pass and low-pass wavelet filters are further split into smaller filters that are easily implementable with simple arithmetic operations and also provide in-place computation. The selection of a wavelet filter depends upon the specific application. JPEG 2000 and SPIHT

Mohd. Rafi Lone (✉)
National Institute of Technology, Srinagar, J&K, India
e-mail: rafimiet@gmail.com

Baba Gulam Shah Badshah University, Rajouri, J&K, India

N.-D. Hakim
National Institute of Technology, Srinagar, J&K, India

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_122

are usually based upon Cohen–Daubechies–Feauveau (9,7). Other frequently used wavelet families for image compression applications include Daubechies wavelets [7–9]. Many of the wavelets have coefficients, irrational in nature. These irrational numbers cannot be perfectly represented by integers. Some noise is injected into the system as the coefficients have to be truncated at a certain decimal value. Due to the irrational coefficients, some wavelets are seldom used although they are powerful.

The error can however be reduced by representing the irrational numbers with algebraic integers (AI) and hence delaying the error injection into the system [10]. This way, the chances of error propagation through the system are reduced. AI quantization has numerous applications in signal and image processing domains, based on wavelet as well as Discrete Cosine Transforms [11]. AI can translate irrational coefficients into vectors or arrays of integers. Many wavelet filters, especially Daubechies filters use irrational coefficients and hence, AI is applied for their implementation [12, 13].

This work provides implementations for 4-tap Daubechies wavelet. The organization of the rest of the paper is as follows: Sect. 2 describes the two methods for implementing Daub-4, Convolution and Lifting. Section 3 provides the proposed design. Section 4 throws light on FPGA implementation and results. Finally, concluding remarks are given in Sect. 5.

2 Background

Algebraic integers have been extended to wavelets by [14]. 1-D and 2-D implementation of Daubechies wavelets, Daub-4, and Daub-6 is provided in [15]. Eight-point Daubechies wavelet based on a two-level folded mapping technique is given in [16]. Reference [17] illustrates pipelined direct mapping for the implementation. Multiplier-free implementation of Daub-4 and Daub-6 is given in [18].

Daubechies-4 filter consists of irrational coefficients. Attempts have been made to minimize the error due to these coefficients, and hence increase the signal-to-noise ratio at reduced FPGA area. Since wavelet transformation has a very large range of applications, it is desired to reduce the area needed by wavelets on FPGA, thereby providing more area to the application. In this paper, we have made an effort to reduce the error propagation and hence improve signal-to-noise ratio for 1-D Daub-4 wavelet, at reduced hardware cost. We have considered three types of implementations possible for Daub-4 wavelet filters. These are Convolution-based, Lifting-based, and Proposed method, which is based on convolution and algebraic integers. The filter coefficients for low-pass and high-pass filter banks are given below:

Low pass: $1/4\sqrt{2}(1+\sqrt{3}, 3+\sqrt{3}, 3-\sqrt{3}, 1-\sqrt{3})$
High pass: $1/4\sqrt{2}(-1+\sqrt{3}, 3-\sqrt{3}, -3-\sqrt{3}, 1+\sqrt{3})$

Using convolution, both high-pass and low-pass filters are realized as shown in Fig. 1. The input, output, and coefficients are represented by x(n), y(n), and c(n), respectively. Both the high-pass and low-pass filters are implemented separately.

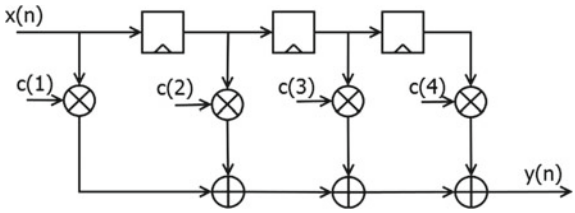


Fig. 1 4-tap FIR filter based on convolution

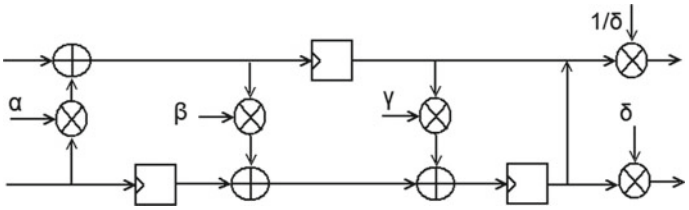


Fig. 2 Lifting-based implementation of Daub-4 wavelet

The problem with convolution-based architecture is that its computational cost is approximately double than the lifting-based method because half the transformed coefficients need to be thrown away. Second, we have four multipliers for each filter bank and these need to be implemented using floating point, which takes much more hardware than a normal multiplier.

The lifting-based architecture may vary depending on the polyphase matrix simplifications and one such implementation is given in [19]. The polyphase matrix is given by

$$P(z) = \begin{bmatrix} \delta & 0 \\ 0 & 1/\delta \end{bmatrix} \begin{bmatrix} 1 & 0 \\ z^{-1} & 1 \end{bmatrix} \begin{bmatrix} 1 & \beta z + \gamma \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \alpha & 1 \end{bmatrix} \tag{1}$$

where $\alpha = -1.7320508$, $\beta = -0.0669873$, $\gamma = 0.4330127$, and $\delta = 1.93185165$.

Figure 2 shows the implementation of the lifting-based architecture for the above polyphase matrix. Here, the number of multipliers has been reduced from eight to five, but still, we need to allocate a huge area for multipliers as they still are in the irrational form. If we use AI here, the error injection will propagate through these multiplier stages.

3 Proposed Method

The proposed method is based on convolution but takes some more considerations into account. First, it minimizes the use of these irrational multiplier coefficients. Only two coefficients are used, 3 and $\sqrt{3}$. Both of these are multiplied toward the later stage of DWT. Multiplication by $\sqrt{3}$ enters some error into the later part of the design. The design of these multipliers is shown in Fig. 3, where (>>) represents

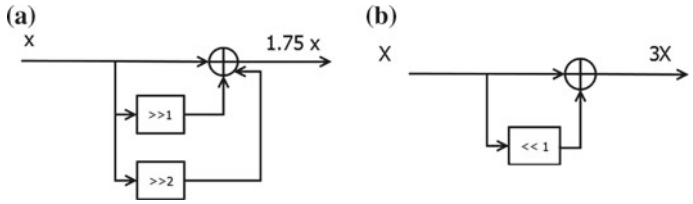


Fig. 3 Shift and add-based multipliers. **a** $\sqrt{3}$, **b** 3

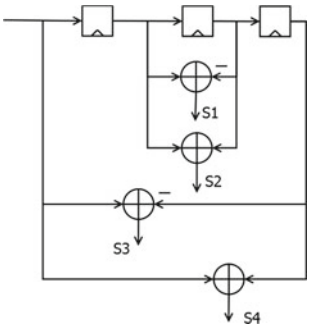


Fig. 4 Proposed architecture, common resources section for low- and high-pass filter banks

right shift and (<<) represents left shift operation. The number of bits by which the input needs to be shifted are specified immediately after the symbol. The value of $\sqrt{3}$ (1.7321) is replaced by the value 1.75, entering an error of 0.0179 in the coefficient. This means that any value will get an error of 1.79%. The error can further be decreased, but it will cost much hardware.

Using the parallel and resource sharing techniques of [20], the equations for high- and low-pass filter banks are simplified as shown below:

$$\begin{aligned} y_L =& \sqrt{3}[x(n) + x(n-1) - x(n-2) - x(n-3)) \\ &+ 3(x(n-1) + x(n-2)) + x(n) + x(n-3]; \end{aligned} \tag{2}$$

$$\begin{aligned} y_H =& \sqrt{3}[x(n) - x(n-1) - x(n-2) + x(n-3)) \\ &+ 3(x(n-1) - x(n-2)) - x(n) + x(n-3)]; \end{aligned} \tag{3}$$

where y_L and y_H are the low-pass and high-pass outputs of the wavelet and x(n-i) represents input sample delayed by i clock cycles. As the equations reflect, resources can be shared between the two filter banks efficiently. Figure 4 shows the resources which are being shared between the two filter banks. Figures 5 and 6 show the remaining resources for low-pass and high-pass filter banks. The 1-D DWT can be easily extended to yield 2-D DWT architecture of Daub-4 wavelet. $1/4\sqrt{2}$ is common to all the coefficients. For 2-D DWT architecture, it will multiply itself to make 1/32 which can be easily implemented using shifting of the final result by 5 bits to the right. As the architecture is based on convolution, half of the computed coefficients

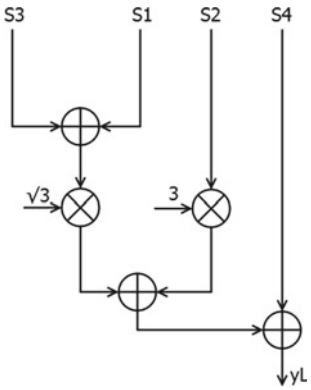


Fig. 5 Proposed architecture, low-pass section

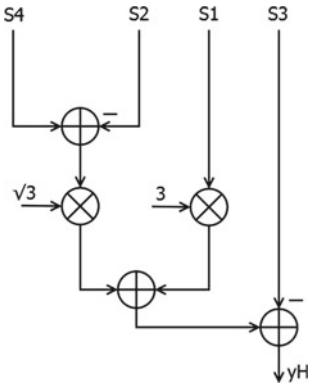


Fig. 6 Proposed architecture, high-pass section

in low-pass and high-pass filters are thrown away, to yield the output transformed image of the same size.

4 FPGA Implementation and Results

We have proposed an algebraic integer based convolution architecture in this paper. The proposed method and convolution method are synthesized on the Xilinx Virtex-6 xc6vcx240t-1ff1156 FPGA platform using ISE 14.2. Images of size 512×512, 256×256, and 128×128 have been transformed through the architectures. The input and output word lengths are set to 8 bits, while in between, the word length has been changed to 13 bits to prevent overflow and hence the quality of the images. The convolution based-architecture is designed for fixed-point arithmetic. As such, we employed 8 bits for the integer part and 6 bits for the fractional part. Also, the hardware can be easily replicated for obtaining a 2-D DWT architecture. Table 1 provides the comparison of hardware utilization for the proposed method with the convolution-based method. It is clear from Table 1 that the proposed method requires

lesser hardware than that of the convolution-based method. The convolution method has used 8 multipliers while the proposed method is multiplier-less. In addition to that, the proposed architecture has utilized less than 37% of the resources required by the convolution method. The reduction in hardware requirement is because of multiplication-less and resource-sharing ability in the proposed method. We have simulated the implementation of the proposed architecture for many image sizes. The PSNR results for the images, *Lena, Boat*, and *Mandrill* are shown in Table 2. The input to the test bench and the output from the test bench are processed in MATLAB 2012. There are not any visible distortions in the recovered images.

Finally, the proposed method is compared with many existing methods as shown in Table 3. All the architectures except the proposed architecture are based upon 2-D DWTs. References [12, 14] have only given a mathematical concept of AI for 4-tap and 6-tap Daubechies wavelet filters while [17] has given MATLAB implementation for the same. Although it has also hinted on its possible hardware implementation, the hardware implementation results have not been provided in the paper. From the table, it is clear that our proposed architecture costs much lesser than the existing architectures. Reference [18] has a better PSNR than our architecture by approximately 3 dB, but it needs approximately 10 times more resources than needed by our architecture. Also, [17] has provided better results by approximately 0.5 dB, but it has not provided any synthesis or simulation results for hardware; hence, FPGA resource comparison can not be done.

Table 1 Synthesis results for hardware utilization of the proposed architecture and convolution based-architecture

Architecture	Adders	Registers	Multipliers	Slice Registers	Slice LUTs	IOBs
Convolution based	23	3	8	28	311	41
Proposed	16	3	0	24	115	25

Table 2 PSNR value for recovered images in the proposed architecture

Size	Image	PSNR (dB)
512 × 512	Lena	62.27
	Boat	64.24
	Mandrill	64.79
256 × 256	Lena	60.77
	Boat	63.73
	Mandrill	64.04
128 × 128	Lena	59.09
	Boat	62.64
	Mandrill	63.01

Table 3 PSNR and synthesis results for hardware utilization of the existing architectures and their comparison with the proposed architecture (FP: Fixed Point)

Architecture		Slice registers	Slices	PSNR (dB)
Wahid-2004 [11]	FP	422	268	45.00
	AI	200	124	46.00
Wahid-2003 [10]		–	–	50.49
Wahid-2002 [9]		–	–	22.26
Islam-2010 [16]		115	106	46.00
Madishetty [18]		236	2482	66.82
Sangeetha [17]		–	–	64.12
Proposed method		24	115	63.76

5 Conclusions

In this paper, we have presented an approach to implement Daubechies 4-tap DWT using algebraic integers and convolution. The Daub-4 provides coefficients that are irrational and implementing them in hardware means lossy compression. So to design an architecture that reduces the loss in compression, it is important to reduce the use of irrational coefficients. Our proposed method uses only one irrational coefficient and utilizes algebraic integers to represent it with the nearest possible integer operations. A shift adder design is used to represent the multiplication with the rational and irrational coefficients. The hardware cost has been effectively reduced to less than 37% of that used with the convolution-based method. Furthermore, it is a multiplier-less architecture. The average PSNR of the recovered images is approximately 63.75 dB, hence the distortion is not visible to the human eye. Also, the architecture was verified for more than 10 images of different sizes. The results are good for smooth images as well as highly textured images such as Mandrill.

References

1. I. Daubechies, W. Sweldens, Factoring wavelet transforms into lifting schemes. J. Fourier Anal. Appl. **4**(5), 247–269 (1998)
2. W. Sweldens, The lifting scheme: a custom-design construction of biorthogonal wavelets. Appl. Comput. Harmon. Anal. **3**(15), 186–200 (1996)
3. M. Vishwanath, R. Owens, M.J. Irwin, VLSI architectures for the discrete wavelet transform. IEEE Trans. Circuits Syst. II **42**(5), 305–316 (1995)
4. K.K. Parhi, T. Nishitani, VLSI architectures for discrete wavelet transforms. IEEE Trans. VLSI Syst. **1**(6), 191–202 (1993)
5. A. Grzeszczak, M.K. Mandal, S. Panchanathan, T. Yeap, VLSI implementation of discrete wavelet transform. IEEE Trans. VLSI Syst. **4**(6), 421–433 (1996)

6. C. Chakrabarti, M. Vishwanath, Efficient realizations of the discrete and continuous wavelet transforms: from single chip implementations to mappings on SIMD array computers. IEEE Trans. Signal Process. **43**(3), 759–771 (1995)
7. J.P. Andrew, P.O. Ogunbona, F.J. Paoloni, Comparison of wavelet filters and sub-band analysis structures for still image compression, in *Proceedings of IEEE International Conference on Acoustics, Speech, Signal Processing* (1994), pp. 589–592
8. M. Rafi, Najeeb-ud-Din, A high speed architecture for lifting-based 2-D Cohen-Daubechies-Feauveau (5,3) discrete wavelet transform used in JPEG2000. Int. J. Adv. Telecommun. Electrotech. Signals Syst. **6**(1), 23–29 (2017)
9. K.A. Wahid, V.S. Dimitrov, G.A. Jullien, W. Badawy, An algebraic integer based encoding scheme for implementing Daubechies discrete wavelet transforms, in *Proceedings of Asilomar Conference on Signals, Systems and Computers*, vol. 1 (2002), pp. 967–971
10. K.A. Wahid, V.S. Dimitrov, G.A. Jullien, Error-free arithmetic for discrete wavelet transforms using algebraic integers, in *16th IEEE Symposium on Computer Arithmetic* (2003), pp. 238–244
11. K. Wahid, V. Dimitrov, G. Jullien, VLSI architectures of Daubechies wavelets for algebraic integers. J. Circuits Syst. Comput. **13**(6), 1251–1270 (2004)
12. K.A. Wahid, V.S. Dimitrov, G.A. Jullien, W. Badawy, An analysis of Daubechies discrete wavelet transform based on algebraic integer encoding scheme, in *Proceedings of 3rd International Workshop on Digital Computational Video* (2002), pp. 27–34
13. V.S. Dimitrov, G.A. Jullien, W.C. Miller, A new DCT algorithm based on encoding algebraic integers, in *IEEE International Conference on Acoustics, Speech and Signal Process*, Seattle, WA, vol. 3 (1998), pp. 1377–1380
14. K.A. Wahid, V.S. Dimitrov, G.A. Jullien, Error-free arithmetic for discrete wavelet transforms using algebraic integers, in *Proceedings of 16th IEEE Symposium on Computer Arithmetic* (2003), pp. 238–244
15. K. Wahid, V. Dimitrov, G. Jullien, VLSI architectures of Daubechies wavelets for algebraic integers. J. Circuits Syst. Comput. **13**(6), 1251–1270 (2004)
16. M.A. Islam, K.A. Wahid, Area- and power-efficient design of Daubechies wavelet transforms using folded AIQ mapping. IEEE Trans. Circuits Syst. II **57**(9), 716–720 (2010)
17. K. Sangeetha, K. Devi, VLSI architectures for the 4-tap and 6-tap 2-D Daubechies wavelet filters using pipelined direct mapping method, in *IEEE Sponsored 2nd International Conference on Innovations in Information Embedded and Communication Systems*, no. 8 (2015)
18. S.K. Madishetty, A. Madanayake, R.J. Cintra, V.S. Dimitrov, D.H. Mugler, VLSI architectures for the 4-tap and 6-tap 2-D Daubechies wavelet filters using algebraic integers. IEEE Trans. Circuits Syst. I: Regul. Pap. **60**(6) (2013)
19. C. Desmouliers, E. Oruklu, J. Saniie, Discrete wavelet transform realisation using run-time reconfiguration of field programmable gate array (FPGAs). IET Circuits Devices Syst. **5**(4), 321–328 (2011)
20. K.K. Parhi, *VLSI Digital Signal Processing Systems: Design and Implementation* (Wiley, Pvt. Ltd., New York, 1999)

Bit Representation for Candidate Itemset Generation



Carynthia Kharkongor and B. Nath

1 Introduction

Association rule mining is one of the most captivating topics in data mining. Most of the organizations and industries require mining algorithms to solve their problems. It gives them the solutions to help the growth of the organization and new strategies which improves the overall performance of their business [1]. Association rule mining was not only used in business but also in many other applications. With rapid growth and emerging technologies, mining of data has become a crucial part in association rule mining. With increasing population, the number of internet users increase day by day and the amount of data also grows exponentially. The process of analyzing and processing these data require not only time but also memory [2].

1.1 Association Rule Mining

Association rule mining is the process of finding out and extracting the rules from the dataset. The dataset contains a collection of itemsets. These itemsets may represent the number of transactions in market basket analysis, the stock market data, the traffic in network communication, the symptoms of a patient, and so on depending on the type of applications used [3, 4].

Each dataset D consists of a collection of transactions $\mathrm{T} = \{t_1, t_2, t_3, \ldots, t_n\}$. Each transaction t contains a set of items called itemsets $I = \{i_1, i_2, i_3, \ldots, i_n\}$ [2].

Some of the important terms in association rule mining are as follows:

1. kth—itemset: is a set of itemsets that contains kth number of elements.
2. Frequent itemsets: is a set of itemsets whose support is greater than or equal to the threshold value.
3. Support count: is the number of times an itemset occurs in the entire dataset.

C. Kharkongor (✉) · B. Nath
Tezpur University, Tezpur, India
e-mail: carynethia@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_123

4. Confidence: of the rule W → Y means that when an item W occurs then the items W and Y also occur in the same transaction. It is interpreted as Confidence(W,Y) = support(W U Y)/Support(W) [1].

Association rule is a rule which is elucidated in the form of W → Y where W and Y are two items. A rule is said to be true if it satisfies the minimum confidence threshold [3, 4]. There are many algorithms that have been introduced to mine the association rules [5–8]. One of the oldest algorithms is the Apriori Algorithm [9].

1.2 Apriori Algorithm

The Apriori Algorithm was the first algorithm introduced by Agarwal [4]. Since then, the algorithm has become the basis for the other association rule mining algorithms. Apriori Algorithm consists of two steps:

– The generation of frequent itemsets.
– The generation of the rules from these generated frequent itemsets.

First, the database is scanned to find the occurrences of the itemset. The scan itemsets at kth level, is called Large Itemsets L_k. The generated large itemsets will self-joined with itself to generate the next candidate itemsets. The itemset L_{k-1} is self-joined with $L_{k-1}$to generate L_k and the support count of L_k is also measured. These itemsets L_k are called candidate itemsets, C_{k-1}. The union operation of (a ∪ b) itemsets ∈ L_{k-1} is generated only when they have a common first $k-2$ itemsets.

The itemsets a[1], a[2] … a[$k-1$], b[1], b[2] … b[$k-1$] are inserted into C_k from L_{k-1} a, L_{k-1} b where a[1] = b[1], … a[$k-2$] = b[$k-2$], a[$k-1$] < b[$k-1$]. The itemsets which have support count greater than the minimum threshold support are considered as frequent itemsets.

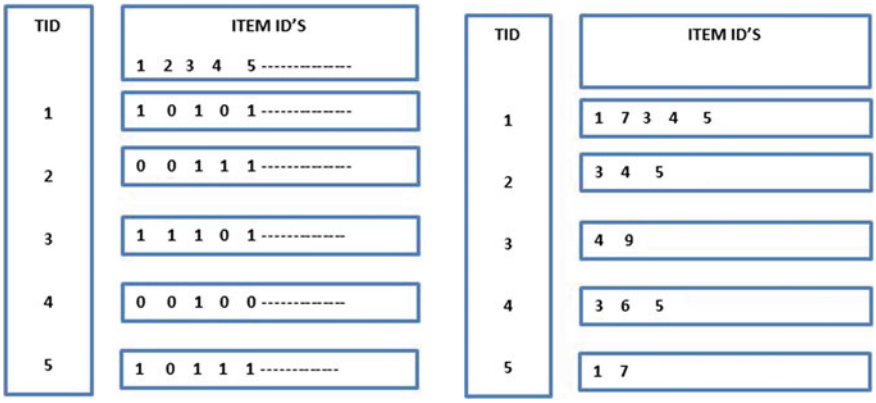
The next step is the prune step, in which the itemsets which are not frequent are deleted from C_k. The itemset $c \in C_k$ are deleted when its subset $(k-1)$ of c is not in L_{k-1}. The principle used in Apriori Algorithm is that any itemset is not frequent if its subset is also not frequent. The itemsets C_k is obtained when the itemsets $(k-1)$ are deleted when they are not in L_{k-1} [1, 2].

For example, let the itemsets L_3 be {{2 3 4}}, {{1 3 4}}, {{1 2 4}}, {{1 2 3}}, {{1 3 5}}. In the join step, the candidate itemsets C_4 generated will be {{1 2 3 4}}, {{1 3 4 5}}. However, in C_4, the itemset {{1 3 4 5}} will be deleted because the itemset {{1 4 5}} is not present in L_3 [4].

1.3 Representation of Itemsets

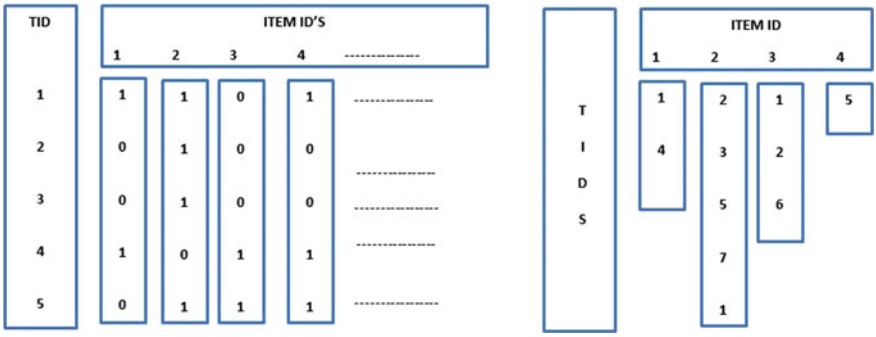
The most common representation of itemsets used in association rule mining are as follows:

– Horizontal Vector Item (HVI): The database is represented as a collection of rows with an identifier(ID) and "1" or "0" is marked when the item is present or absent respectively. The diagram is shown in Fig. 1a.
– Horizontal List (HL): The items are stored in a similar way as horizontal vector item but except that only the items ID's that occur in each transaction are stored which is shown in Fig. 1b.
– Vertical Vector Item (VVI): The database is represented in such a way that each column stores the items ID and "1" or "0" is labeled only when the items are present or absent as given in Fig. 2a.
– Vertical List (VL): The dataset is arranged into set of columns where each column represents the data items storing only those TID's in which the items occur. The representation is shown in Fig. 2b [10].



(a) Horizontal vector item (b) Horizontal list

Fig. 1 Horizontal layout



(a) Vertical vector item (b) Vertical list

Fig. 2 Vertical layout



Fig. 3 Representation of itemset = {2, 15, 19, 26, 30, 31} using linked list

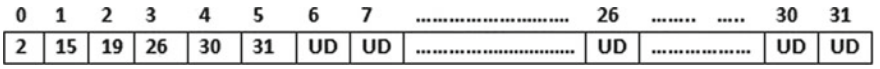


Fig. 4 Representation of itemset = {2, 15, 19, 26, 30, 31} using array

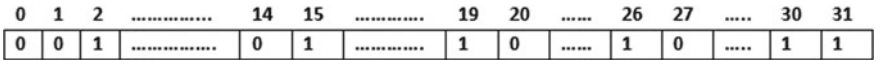


Fig. 5 Representation of itemset = {2, 15, 19, 26, 30, 31} using bitmap

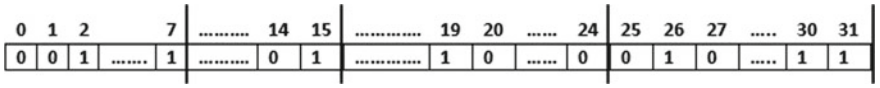


Fig. 6 Representation of itemset = {2, 15, 19, 26, 30, 31} using set

Some of the algorithms such as [10, 11] have used vertical layout for representing their items whereas [12] uses horizontal representation. Suppose the itemsets $I = \{2, 15, 19, 26, 30, 31\}$ needs to be represented in the memory, the different data structures used are as follows:

1. Linked list: There are six number of elements in the itemsets I. Using linked list representation, the space requirement will be $6 \times (2 \times \text{Integers})$ irrespective of the itemset size. With each integer consuming 4 bytes, then the space requirement will be $6 \times (2 \times 4) = 48$ bytes to store six elements. The representation using linked list is shown in Fig. 3.
2. Array: For representing the itemset I using array, the memory storage will require $32 \times$ Integers irrespective of the size of the itemset. There are only six elements that are stored in the array and the rest of the space is left unused. With integer consuming 4 bytes, the total memory space is $32 \times 4 = 128$ bytes. The representation is shown in Fig. 4.
3. Bitmap: uses only one byte for representing the items in the itemsets. The items are represented by "1" if the item is present and "0" if the item is absent on the itemset as shown in Fig. 5. The memory requirement to represent the itemset I will be 32×1 byte $= 32$ bytes irrespective of the itemset size [9, 13].
4. Set Representation: The itemsets are represented in 4-bytes long integer consuming only one bit per element. The cardinality of the set will be the maximum size of the attributes. For the itemset I, the number of bits that are represented is 32×1 bit $= 32$ bits or 4 bytes as shown in Fig. 6 [14].

The algorithms such as [12, 15] have used array for representing their itemsets. However, there are some algorithms which use a hybrid of the data structure such as array, bitmap, and prefix tree [5]. Other algorithms have used bitmap for representing the itemsets such as [16–19].

2 Problem Definition

The main challenge in data mining that persists for a considerable time is memory storage. With the new upcoming technologies, the problem of storage seems to be solved for sometime. However, with the increasing number of internet users, the amount of data generated grows at a very alarming rate. The process of retrieval of data and processing not only takes time but also huge amount of storage.

In association rule mining, storage of candidate of itemsets has become a challenge. If the database size is n, the number of candidate itemsets generated will be 2^n. If n is very large, storing the candidate itemsets in main memory will not fit. Secondary storage is required if this happens. However, time will be wasted in retrieving the itemsets to main memory. Furthermore, if the itemsets are stored using array, each element will require 4 bytes for storage. If the itemsets size is 10, the number of candidate itemsets generated is $2^{10} \approx 1024$ itemsets and the total memory requirement will be 4096 bytes (1024×4). Therefore, if we are handling the large database, space and time requirements will also increase.

This paper will represent the itemset using bit representation for candidate generation in Apriori Algorithm.

3 Bit Representation for Candidate Generation

In the paper [17], a compression scheme called the Word-Aligned Hybrid (WAH) code was introduced to compress the bitmap indices for higher cardinality attributes. In this paper, the itemsets are represented using set representation. The layout used is horizontal layout. Each item of the itemsets is represented by "1" or "0" for storing in the memory. Each item is marked by "1" when an item is present or "0" when the item is absent. The cardinality of the array will be the maximum size of the attributes in the dataset [14]. Some of the operations used in candidate generation are as follows:

1. Union: The union operation for itemsets is performed using OR bitwise operation which takes only O(1) time complexity. The union operation is shown in Fig. 7.
2. Intersection Operation: This operation is also performed using bitwise AND operation taking only O(1) time complexity. The operation is given in Fig. 8.
3. Superset Operation: This involves using AND operation to check whether the itemset L_1 is superset of another itemset L_2 or not.

4. Subset Operation: It is similar to superset operation that checks whether an itemset L_1 is a subset of L_2 or not.
5. Membership Operation: It checks whether an item c is a part of the itemset L or not.
6. Set Difference: It finds the set difference of the two itemsets using bitwise operation taking only O(1) in both time and space complexity. An example of set difference is given in Fig. 9.

3.1 Experimental Analysis

The experiment is performed using both array and set representation. First, all the set operations such as union, intersection, set difference, subset, superset, and membership operations have been tested using both array and set representation. Then, the candidate generation phase in Apriori Algorithm is represented using both array and set representation. The experiment is done using 4GB RAM 64-bit computer and the representation is tested using synthetic dataset. There are three datasets:

1. First dataset: It has 50 attributes with a size of 1000. The dataset has been tested for different set operations. The result is shown is Table 1.
2. Second dataset: It consists of 50 attributes with a total of size 2000. The set operations are tested using both the array and set representation for the second dataset. The time and space consumption is shown in Table 2.
3. Third dataset: It contains 50 attributes with a size of 3000. It has been tested for different set operations using set and array representation. The result of the performed experiments is shown in Table 3.

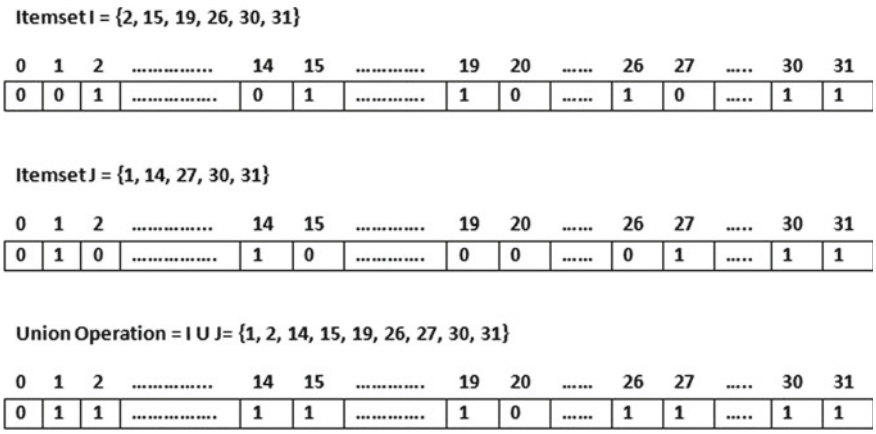


Fig. 7 Union operation using bitwise OR operation

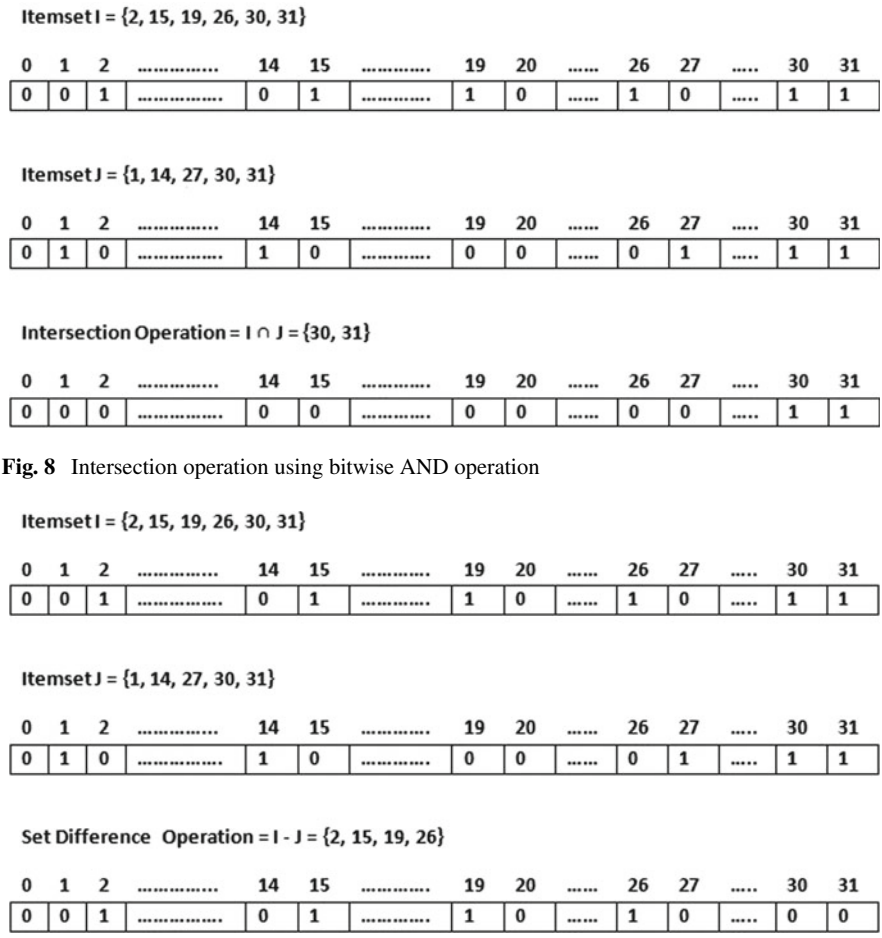


Fig. 8 Intersection operation using bitwise AND operation

Fig. 9 Set difference operation using bitwise operation

Table 1 Time and memory consumed by array and set representation with first dataset

Array	Union	Intersection	Set difference	Subset	Superset	Membership
Time (ms)	1560000	64923.4	1070000	65839.8	66002.2	72.292
Memory (kbs)	21747	21330	215649	21460	21613	21102
Bit representation	Union	Intersection	Set difference	Subset	Superset	Membership
Time (ms)	53596.9	52924.3	54209.7	49849	48711.9	50.597
Memory (kbs)	13593	13698	13617	13534	13533	13213

Table 2 Time and memory consumed by array and set representation with second dataset

Array	Union	Intersection	Set difference	Subset	Superset	Membership
Time (ms)	1922080	308341	1070000	324762	323403	166
Memory (kbs)	41246	21982	40354	22976	22821	21970
Bit representation	Union	Intersection	Set difference	Subset	Superset	Membership
Time (ms)	303045	302483	306509	276760	273995	109.311
Memory (kbs)	14504	14288	14243	14710	14883	13989

Table 3 Time and memory consumed by array and set representation with third dataset

Array	Union	Intersection	Set difference	Subset	Superset	Membership
Time (ms)	2196838	902228	2072704	972939	972364	356.708
Memory (kbs)	74458	24000	57968	24493	25658	22978
Bit representation	Union	Intersection	Set difference	Subset	Superset	Membership
Time (ms)	850639	863937	849334	797165	801369	252
Memory (kbs)	17680	16961	16861	16978	16839	15989

The array and set representation are represented for candidate generation of Apriori Algorithm. The candidate generation is tested for varying support count, $S = 1$, 2.5, and 5% and the results are shown in Tables 4, 5 and 6. As seen from the above tables, the set representation is more concise for different set operations and candidate generation. The memory requirement for different set operations such as union, intersection, set difference, subset, superset, and membership is less than array representation. Moreover, the time execution for the set operations is less in set representation than in array representation. The results from the candidate generation also show that both space and time requirement is better in set representation than array representation. Overall, the performance of set representation is more efficient for different set operations and candidate generation.

4 Conclusion

In this paper, we have seen that the set representation is more efficient than array representation. Using set representation, the performance of the mining algorithms

Table 4 Time and memory consumed by array and set representation for different set operations with first dataset

Array	s = 1%	s = 2.5%	s = 5%
Time (ms)	39275.6	38018	24610.4
Memory (kbs)	21360	21280	21070
Bit representation	s = 1%	s = 2.5%	s = 5%
Time (ms)	27668.1	25734.1	15424
Memory (kbs)	13491	13358	13308

Table 5 Time and memory consumed by array and set representation for different set operations with second dataset

Array	s = 1%	s = 2.5%	s = 5%
Time (ms)	115834	107161	87914
Memory (kbs)	210206	21174	210220
Bit representation	s = 1%	s = 2.5%	s = 5%
Time (ms)	103700	92106.8	52147
Memory (kbs)	13644	13439	13348

Table 6 Time and memory consumed by array and set representation for different set operations with third dataset

Array	s = 1%	s = 2.5%	s = 5%
Time (ms)	235874	234753	178790
Memory (kbs)	22170	21569	21268
Bit representation	s = 1%	s = 2.5%	s = 5%
Time (ms)	203060	193564	103033
Memory (kbs)	13874	13684	13483

also improves especially when we are dealing with very large databases. The set representation consumes very less space which takes only one bit for each element whereas array representation takes up 4 bytes per element. Furthermore, the time execution is less for set representation rather than array representation. Using array representation, generation of candidate itemsets may take 3–4 days. However, the execution time using set representation may take only few hours. Therefore, set representation for mining algorithms is more effective than array representation.

References

1. R. Srikant, R. Agrawal, Mining generalized association rules (1995)
2. M. Abdel-Basset, M. Mohamed, F. Smarandache, V. Chang, Neutrosophic association rule mining algorithm for big data analysis. Symmetry **10**(4), 106 (2018)
3. R. Agrawal, H. Mannila, R. Srikant, H. Toivonen, A. I. Verkamo et al., Fast discovery of association rules. Adv. Knowl. Discov. Data Min. **12**(1), 307–328 (1996)
4. R. Agrawal, T. Imieliński, A. Swami, Mining association rules between sets of items in large databases, in *ACM Sigmod Record*, vol. 22 (ACM, 1993), pp. 207–216
5. G. Grahne, J. Zhu, Efficiently using prefix-trees in mining frequent itemsets, in *FIMI*, vol. 90 (2003)
6. V. Umarani et al., A bitmap approach for closed and maximal frequent itemset mining. Int. J. Adv. Res. Comput. Sci. **3**(1) (2012)
7. T.-T. Nguyen, Mining incrementally closed item sets with constructive pattern set. Expert Syst. Appl. **100**(2018), 41–67 (2018)
8. D. Serrano, C. Antunes, Condensed representation of frequent itemsets, in *Proceedings of the 18th International Database Engineering & Applications Symposium* (ACM, 2014), pp. 168–175
9. C.-Y. Chan, Y.E. Ioannidis, Bitmap index design and evaluation, in *ACM SIGMOD Record*, vol. 27 (ACM, 1998), pp. 355–366
10. P. Shenoy, J.R. Haritsa, S. Sudarshan, G. Bhalotia, M. Bawa, D. Shah, Turbo-charging vertical mining of large databases, in *ACM Sigmod Record*, vol. 29 (ACM, 2000), pp. 22–33
11. M.J. Zaki, K. Gouda, Fast vertical mining using diffsets, in *Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (ACM, 2003), pp. 326–335
12. M. Song, S. Rajasekaran, A transaction mapping algorithm for frequent itemsets mining. IEEE Trans. Knowl. Data Eng. **18**(4), 472–481 (2006)
13. J. Ayres, J. Flannick, J. Gehrke, T. Yiu, Sequential pattern mining using a bitmap representation, in *Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (ACM, 2002), pp. 429–435
14. C. Kharkongor, B. Nath, Set representation for itemsets in association rule mining, in *2018 IEEE International Conference on Intelligent Computing and Control Systems (ICICCS)* (IEEE, 2018)
15. R. Srikant, R. Agrawal, Mining sequential patterns: generalizations and performance improvements, in *International Conference on Extending Database Technology* (Springer, 1996), pp. 1–17
16. G. Antoshenkov, Byte-aligned bitmap compression, in *Proceedings DCC'95 Data Compression Conference* (IEEE, 1995), p. 476
17. K. Wu, E.J. Otoo, A. Shoshani, Optimizing bitmap indices with efficient compression. ACM Trans. Database Syst. (TODS) **31**(1), 1–38 (2006)
18. A. Moffat, J. Zobel, Parameterised compression for sparse bitmaps, in *Proceedings of the 15th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval* (ACM, 1992), pp. 274–285
19. N. Koudas, Space efficient bitmap indexing, in *CIKM*, pp. 194–201 (2000)

PROD: A Potential Rumour Origin Detection Model Using Supervised Machine Learning



Akshi Kumar and Harshita Sharma

1 Introduction

Social media is an indispensable and ubiquitous part of the everyday life of millions of people around the world. It has been explored by researchers for sentiment analysis [1–4], sarcasm detection [5], cyberbullying detection [6, 7], expert mining [8] and rumour analytics [9–11]. Currently, social media is used as an outlet to express not only one's opinions, sentiments, support, opposition but also the 'go to' source for staying updated about current events be it elections, terrorist attacks or natural disasters. Although, government officials and common citizens have used it in critical situations to bring awareness and for propagating updates and latest news, sadly, it has also been a fertile ground for rumour breeding. Rumours, in general, spread fake information and in particular may defame someone and in extreme cases may prove to be life-threatening as well.

A rumour is defined as any piece of information put out in public without sufficient knowledge and/or evidence to support it thus putting a question on its authenticity. It may be true, false or unspecified [10]. As important it is to detect and debunk a rumour, equally important is finding the origin of the rumour post. Doing so, will not only help to catch the culprit but can hint screening agencies to proactively keep a check on users activities. This subtask of rumour analytics has been explored the least and this work aims to detect a user who can be a potential rumour origin based on certain user account and tweet content-based features. The work proposes a Potential Rumour Origin Detection (PROD) Model which labels each user based on whether he can be a potential rumour source or not defining a credibility quotient derived from

A. Kumar (✉) · H. Sharma
Department of CSE, Delhi Technological University, New Delhi 110042, India
e-mail: akshikumar@dce.ac.in

H. Sharma
e-mail: harshitasharma23@ymail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_124

the values of 8-tuple feature vector extracted based on user data and tweet content. Finally, three supervised machine learning algorithms Naïve Bayes (NB), Support Vector Machines (SVM) and Multilayer Perceptron (MLP) are used to evaluate the performance of the model.

The rest of the paper is ordered as follows: the second section summarizes the existing work, the third section expounds the proposed **P**otential **R**umour **O**rigin **D**etection (PROD) model, the fourth section presents the results. The last section concludes the work and presents some future directions.

2 Related Work

The area of rumour origin detection has been explored little. We summarize some of the most notable works. The first systematic study of finding rumour sources used the SIR model where the authors built an estimator for rumour source in trees and general graphs [12]. They use the approach of maximum likelihood. The problem of rooting out a single rumour source from a set of suspect nodes under an SI model has also been studied [13]. They build a Maximum A Posteriori (MAP) to identify the rumour source. Some authors introduce monitors in the network and build algorithms based on the approach that receive information from the network [14, 15]. Rumour sources are detected based on which nodes receive more information. In another work, researchers compare time of tweet posting to find the origin [16]. A track of rumour propagation trails and empirical investigation of the rumour teller's position using statistics has also been done [17]. The problem of single rumour source detection with multiple observations has also been addressed [18]. The authors propose a unified inference framework using the concept of union rumour centrality.

3 Potential Rumour Origin Detection (PROD) Model

Identifying potential rumour mongers is imperative to prevent further rumour spread. Moreover, this step can be very helpful in calculating the veracity of a rumour that is the final step in rumour analytics pipeline. The following Fig. 1 depicts the system architecture of the work undertaken. It consists of four primary modules: Data acquisition, feature extraction, classification and evaluation. The dataset used in this work is the PHEME dataset of rumours. It consists of labelled tweets for five breaking news events, namely, #charliehebdo (458 rumours and 1621 non-rumours), #ferguson (284 rumours and 859 non-rumours), #germanwingscrash (238 rumours and 231 non-rumours), #ottawashooting (470 rumours and 420 non-rumours) and #sydneysiege (522 rumours and 699 non-rumours). As the primary aim of this work is to identify potential rumour source, we only considered rumourous tweets from the dataset for all the five events. Preprocessing is then done to remove useless words, whitespaces and other data that might act like noise.

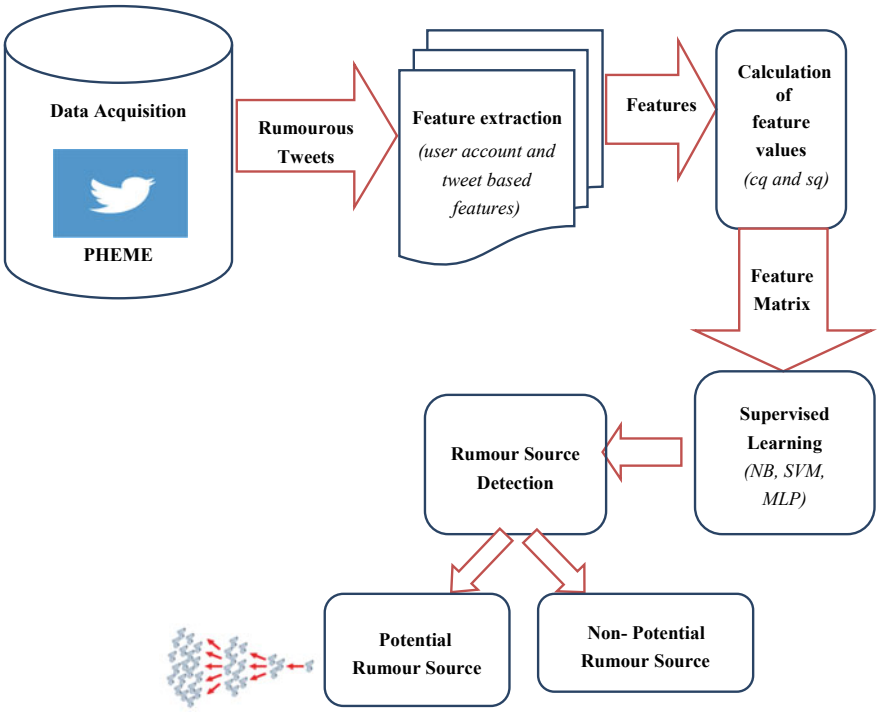


Fig. 1 Systematic flow of the model

In this work, we use meta-features derived from user account and tweet-based features to detect if a particular user is a rumour source or not. The features extracted are: Date of tweet posting, Date of account creation, Retweet status of post, Retweet count of post, Followers count of user, Status count of user, Favourite count of user, User account verification status, URL presence, and media presence. Thus, the resulting feature tuple is given as <*naiv, orig, resp, activ, sr, av, tv, media*>, where each term is assigned a weight indicating how negatively or positively each value affects the credibility quotient of the user. Each of these features used to build the PROD model is described in the following Table 1.

All these feature values are summed up to calculate the *credibility quotient* (*cq*) of a user which would range from [−10, 10]. The higher the credibility, less likely is to be a rumour source. Hence, *the source quotient* (*sq*) is calculated to be the inverse of the credibility quotient. The pseudocode for the model is given in the following Fig. 2.

The final value of the source quotient is used to calculate the *prs* (*potential rumour source*) binary value, which is also the class label for our dataset. A negative value of source quotient indicates high chance of being a rumour source. Credibility points anywhere between the values 5–10 are considered to indicate that concerned user

Table 1 Features to build the model

Feature	Description
<naiv>	It is indicative of user naivety and depends on the account age of the user. The younger the user account age, more naïve he is and hence more prone to spread rumours. Naivety has been assigned a score of −2
<orig>	It indicates whether a particular post is actually composed by the user or not. If the post is retweeted, orig is assigned a value of 0 as this user cannot be a rumour source. It has been assigned a score of 1
<resp>	It is short for responsibility and is set to 1 if the user has posted a URL in his tweet. Posting URL is considered to be indicative of genuine news and hence the user cannot be a potential rumour source. It has been assigned a score of 4
<activ>	It indicates how active a user is and is measured as the ratio of number of status updates to the no. of days since account creation. A study [19] shows that active users are more likely to spread rumours. Hence, this parameter has been assigned a score of −2
<sr>	sr or social reputation is calculated as the ratio of #favourite_count of a user to the #followers of that user and represents how popular a user is
<av>	It signifies the account verification status of a person. A verified account, usually of public figures, is less likely to be a rumour source. This factor is thus assigned a positive weight of 5. <av> in combination with <sr> has been assigned a negative score of 2 considering that an account which is not verified but has high popularity can be rumour source
<tv>	It is short for tweet viralty. A viral tweet from an unverified and popular account has very high chances of being a rumour source. Hence, this feature in combination with <av> and <sr> has been assigned a score of (−2)
<media>	Many rumours are spread, when people use images from a different contexts in situations where they would seemingly fit, hence misleading people. This factor, thus, has also been assigned a negative score of 2

cannot be a potential rumour source, while a credibility score from 1 to 4 is considered to be a rumour origin suspect. Accordingly, the final prs value is labelled as 1 or 0.

The model is trained using supervised learning wherein a number of labelled training instances are given to the ML classifier Naïve Bayes (NB), Support Vector Machine (SVM) and Multilayer Perceptron (MLP) as input. Tenfold cross-validation method is used for evaluating the performance of the various algorithms. ZeroR has been used as the baseline classifier which simply predicts the majority class for all instances.

```
Input: Rumourous tweets
For each post,
Extract tweet_created_at, account_created_at, retweeted, retweet_count, followers_count,
friends_count, status_count, verified, url, media
acc_age = tweet_created_at – account_created_at
if acc_age< 180 then naiv =1
else naiv = 0
if retweeted = False then orig =1
else orig = 0
if url = 1 then resp = 1
else resp = 0
if (statuses_count/followers_count> 4) then activ =1
else activ = 0
if friends_count/followers_count< 0.2 then sr =1
else sr = 0
if verified = true then av =1
else av =0
if retweets > 500 then tv =1
else tv = 0
if media = 1 then media =1
else media = 0
cq =[ orig * 1 + resp * 4 + activ * (-2) + av * 5 + (not(av) and sr) * (-2) + (tv and  not(av) and sr) * (-
2) + media * (-2) + naiv * (-2)]
source_quotient = 1/cq
if (source_quotient<= 0) or (source_quotient> 0.2 and source_quotient< 1), then
prs =1
else prs = 0
   end for
```

Fig. 2 Pseudocode of the proposed PROD model

4 Results

WEKA[1] tool has been used for evaluation purposes. It is an open-source software tool with a collection of machine learning algorithms and tools for classification, visualization, etc. The results for all five events in PHEME dataset have been evaluated for four performance measures: accuracy, precision, recall and F-measure [20] and are summarized in Table 2. The highest accuracy for each event is obtained by MLP for all events followed by SVM and NB. The accuracy result for MLP classifier for all events is shown as a graph in Fig. 3.

[1] https://www.cs.waikato.ac.nz/ml/weka/.

Table 2 Performance Results

Classifier	Accuracy	Precision	Recall	F-measure
Charlie Hebdo				
ZeroR	65.2838	0.653	0.653	0.784
NB	93.2314	0.937	0.932	0.933
SVM	97.5983	0.976	0.976	0.976
MLP	97.5	0.987	0.987	0.987
Ferguson				
ZeroR	82.3944	0.824	0.824	0.903
NB	87.6761	0.872	0.877	0.857
SVM	95.0704	0.951	0.951	0.951
MLP	97.9	0.97	0.97	0.97
German Wings Crash				
ZeroR	65.9664	0.660	0.660	0.795
NB	88.2353	0.890	0.882	0.884
SVM	92.437	0.924	0.924	0.924
MLP	95.7983	0.958	0.958	0.958
Ottawa Shooting				
ZeroR	70.6383	0.706	0.706	0.828
NB	95.5319	0.955	0.955	0.955
SVM	96.383	0.964	0.964	0.964
MLP	97.234	0.972	0.972	0.972
Sydney Seige				
ZeroR	68.3301	0.683	0.683	0.812
NB	93.666	0.937	0.937	0.937
SVM	95.0096	0.952	0.950	0.951
MLP	97.8887	0.979	0.979	0.979

5 Conclusion

This work proposes a PROD model for detecting users that can be probable rumour sources. An 8-tuple feature vector has been used to evaluate the credibility of each user. The likelihood of the user being a potential rumour source is the inverse of the obtained credibility quotient. The resulting value is used to label each user as a potential or non-potential rumour source. The multilayer perceptron classifier outperforms the naïve bayes and support vector machine in terms of performance accuracy, for all five events of PHEME, thus demonstrating the superior use supervised learning for potential rumour source detection model. A limitation of the proposed model is that only few features have been used to detect a potential rumour origin. The model currently does not also detect rumour source for end to end encrypted social media

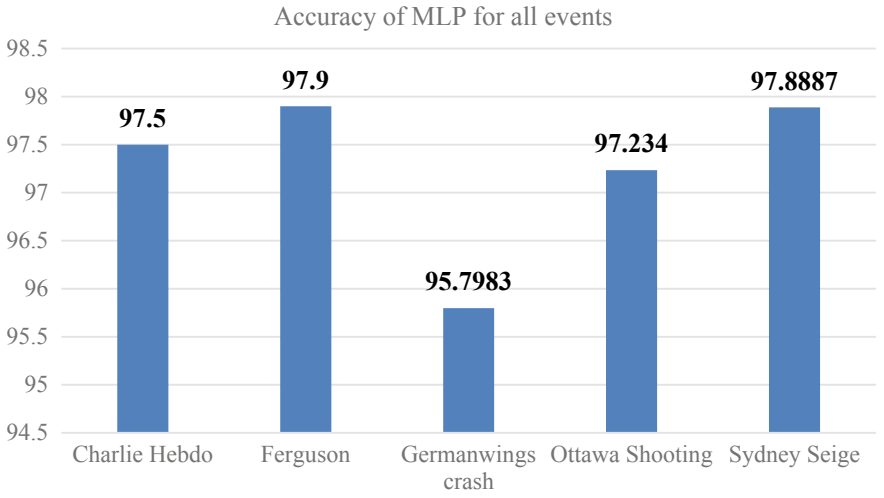


Fig. 3 Accuracy of MLP for all events

such as Whatsapp. This work can be extended to include more meta-features and using fuzzy values instead of binary values. Also, optimization on the feature set can be carried out to determine the most relevant features for rumour origin prediction.

References

1. A. Kumar, R. Khorwal, S. Chaudhary, A survey on sentiment analysis using swarm intelligence. Indian J. Sci. Technol. **9**(39), 1–7 (2016)
2. A. Kumar, G. Garg, Sentiment analysis of multimodal twitter data, in *Multimedia Tools and Applications* (2019), pp. 1–17
3. A. Kumar, G. Garg, Systematic literature review on context-based sentiment analysis in social multimedia, in *Multimedia Tools and Applications* (2019), pp. 1–32
4. A. Kumar, A. Jaiswal, Swarm intelligence based optimal feature selection for enhanced predictive sentiment accuracy on twitter, in *Multimedia Tools and Applications* (2019), 1–25
5. A. Kumar, S.R. Sangwan, A. Arora, A. Nayyar, M. Abdel-Basset, Sarcasm detection using soft attention-based bidirectional long short-term memory model with convolution network. IEEE Access (2019)
6. A. Kumar, N. Sachdeva, Cyberbullying detection on social multimedia using soft computing techniques: a meta-analysis, in *Multimedia Tools and Applications* (2019), pp. 1–38
7. A. Kumar, S. Nayak, N. Chandra, Empirical analysis of supervised machine learning techniques for cyberbullying detection, in *International Conference on Innovative Computing and Communications* (Springer, Singapore, 2019), pp. 223–230
8. A. Kumar, N. Ahmad, ComEx miner: expert mining in virtual communities. Int. J. Adv. Comput. Sci. Appl. (IJACSA) **3**(6) (2012)
9. A. Kumar, S.R. Sangwan, A. Nayyar, Rumour veracity detection on twitter using particle swarm optimized shallow classifiers, in *Multimedia Tools and Applications* (2019), pp. 1–19

10. A. Kumar, S.R. Sangwan, Rumour detection using machine learning techniques on social media, in *International Conference on Innovative Computing and Communications* (Springer, Singapore, 2019), pp. 213–221
11. A. Kumar, S.R. Sangwan, Information virality prediction using emotion quotient of tweets. Int. J. Comput. Sci. Eng. **6**(6), 642–651 (2018)
12. D. Shah, T. Zaman, Rumours in a network: who's the culprit? IEEE Trans. Inf. Theory **57**(8), 5163–5181 (2011)
13. W. Dong, W. Zhang, C.W. Tan, Rooting out the Rumour culprit from suspects, in *2013 IEEE International Symposium on Information Theory* (IEEE, 2013), pp. 2671–2675
14. E. Seo, P. Mohapatra, T. Abdelzaher, Identifying Rumours and their sources in social networks, in *Ground/Air Multisensor Interoperability, Integration, and Networking for Persistent ISR III*, vol. 8389 (International Society for Optics and Photonics, 2012), p. 83891I
15. W. Xu, H. Chen, Scalable Rumour source detection under independent cascade model in online social networks, in *2015 11th International Conference on Mobile Ad-hoc and Sensor Networks (MSN)* (IEEE, 2015), pp. 236–242
16. V.P. Sahana, A.R. Pias, R. Shastri, S. Mandloi, Automatic detection of rumoured tweets and finding its origin, in *2015 International Conference on Computing and Network Communications (CoCoNet)* (IEEE, 2015), pp. 607–612
17. D. Król, K. Wiśniewska, On Rumour source detection and Its experimental verification on twitter, in *Asian Conference on Intelligent Information and Database Systems* (Springer, Cham, 2017), pp. 110–119
18. Z. Wang, W. Dong, W. Zhang, C.W. Tan, Rumour source detection with multiple observations: fundamental limits and algorithms, in *ACM SIGMETRICS Performance Evaluation Review*, vol. 42, no. 1 (ACM, 2014), pp. 1–13
19. Hindustan times website, https://www.hindustantimes.com/tech/active-twitter-users-most-likely-to-spread-fake-news-study/story-sxrZe611IBYPxv0Pmn8hGO.html
20. M.P.S. Bhatia, A.K. Khalid, A primer on the web information retrieval paradigm. J. Theor. Appl. Inf. Technol. **4**(7) (2008)

Double-Stage Sensing Detectors for Cognitive Radio Networks



Ashish Bagwari, Jyotshana Kanti and Geetam Singh Tomar

1 Introduction

To detect LU frequency bands there are various sensing techniques available [1, 2]. In [3], the authors introduced two-stage detectors (ED and cyclo detector-2010), but this scheme was computationally more complex and had longer observation time. Then [4] worked over observation time and presented an adaptive spectrum sensing scheme (ASS-2012). However, the authors minimized sensing time but complexity was still there. Furthermore, the authors discussed the status of the LU signal in [5].

In this paper, we have improved detection performance by introducing a double-stage sensing scheme in CRN. The proposed sensing scheme works on two phases, the first phase carry ED-SAT that senses the LU signal and computes the energy (E), if $E \geq \lambda_1$, detector confirms LU band busy else free then the second phase detector detects the LU signal, λ_1 is the predefined threshold. In the second phase, two adaptive thresholds-based ED (ED-TAT) detects LU frequency band, if observed energy of ED-TAT (Z) $\geq \gamma$, shows LU frequency band is busy else free, where γ is the predefined threshold.

A. Bagwari (✉)
Department of Electronics and Communication Engineering,
Uttarakhand Technical University, Dehradun, India
e-mail: ashishbagwari@ieee.org

J. Kanti
Department of Computer Science and Engineering,
Uttarakhand Technical University, Dehradun, India
e-mail: jyotshanakanti@gmail.com

G. S. Tomar
T.H.D.C.I.H.E.T., Tehri, Uttarakhand, India
e-mail: gstomar@ieee.org

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_125

The rest of the paper is divided into five sections, Sect. 2 is system description, Sect. 3 is the proposed system model, Sect. 4 is simulation results, and finally Sect. 5 shows the conclusion.

2 System Description

If the LU signal is present then the mathematical expression to get the received signal is defined in terms of hypothesis test [1]

$$u(n) = v(n) \times h(n) + g(n), \quad H_1 \tag{1}$$

Similarly, if the LU signal is absent then the received signal becomes

$$u(n) = g(n), \quad H_0 \tag{2}$$

In Eqs. (1) and (2), H_1 and H_0 are alternate and null hypothesis, respectively, $u(n)$ is received signal, $v(n)$ is licensed users signal, $g(n)$ is additive white gaussian noise, $h(n)$ is the channel gain, and n is number of samples, i.e., $n = 1, 2, \ldots, N$.

3 The Proposed System Model

3.1 Double-Stage Sensing Detectors

Figure 1 illustrates the flow chart of double-stage sensing detectors. There are two phases, the first phase consists of ED-SAT, and the second phase has ED-TAT. In the given figure, first phase ED-SAT detects the LU signal, then calculates signal energy (E) and compares with a predefined threshold (λ_1). If $E \geq \lambda_1$, the detector confirms that the LU signal is present else the second phase detector ED-TAT will repeat the same process. Second phase detector performs the detection operation in three levels, first is above a predetermined threshold (γ_2), the second is below a predetermined threshold (γ_1), and third is between predetermined thresholds (γ_1) and (γ_2).

Now, considering all cases and their outcomes: the first case: If $E \geq \gamma_2$, where γ_2 is the predetermined threshold, then the detector confirms that the LU signal is present.

Second case: Signal energy (E) is below the predetermined threshold (γ_1), then the detector confirms that the LU signal is absent.

Third Case: E is between predetermined thresholds (γ_1) and (γ_2). This condition is known as a confused region and a phenomenon is sensing failure problem. To resolve this issue, there are two adaptive thresholds concept as discussed in Sect. 3.1.2.

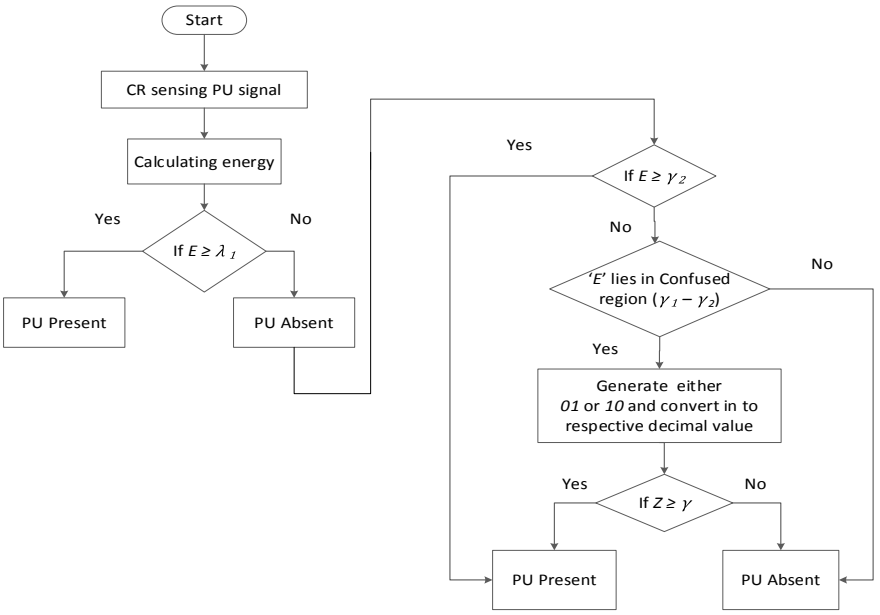


Fig. 1 Flow chart of double-stage sensing detectors

Finally, the output (Z) compares with a predetermined threshold (γ) and gives the final decision as shown in Fig. 1.

- Detection probability of double-stage sensing detectors can be defined as

$$P_D^{\text{Advanced-TPD}} = P_r\left(P_d^{\text{ED-SAT}} - P_d^{\text{ED-TAT}}\right) + P_d^{\text{ED-TAT}} \tag{3}$$

where $P_d^{\text{ED-SAT}}$ and $P_d^{\text{ED-TAT}}$ are the detection probability throughout ED-SAT and ED-TAT detector, respectively. P_r is the probability factor, i.e., $0 \leq P_r \leq 1$.

3.1.1 First Stage: ED-SAT

Figure 2 shows the block diagram of ED-SAT, the incoming LU signal is received by a band-pass filter (BPF) that passes a particular band of frequency signal to the ADC. ADC is analog to digital converter that produces digital signal to the square law device (SLD). SLD computes the incoming signal energy then integrates using integrator, and passes to the decision-making device (DMD) to take the final decision in terms of LU is present or not.

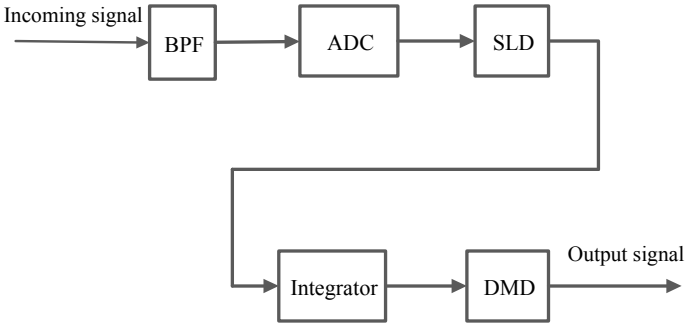


Fig. 2 Block diagram of ED-SAT

Expression of Single Adaptive Threshold

The mathematical expression of a single adaptive threshold (λ_1) can be defined as [6, 7]

$$\lambda_1 = \left[N \times \sigma_g^2 \left\{ Q^{-1}\left(\overline{P_f}\right) \times \sqrt{\frac{2}{N}} + 1 \right\} \right] \tag{4}$$

where N is a number of samples, $Q^{-1}()$ denotes inverse- Q-function, P_f is false alarm probability, and σ_g^2 is noise variance. The energy can be calculated as

$$E = \frac{1}{N} \sum_{n=1}^{N} |u(n)|^2 \tag{5}$$

Finally, the local decision of ED-SAT becomes

$$\text{ED-SAT}|_{\text{o/p}} = \begin{cases} E < \lambda_1, & H_0 \\ E \geq \lambda_1, & H_1 \end{cases} \tag{6}$$

Detection Probability for ED-SAT Detector

The final expression for detection probability can be written as [4]

$$P_d^{\text{ED-SAT}} = Q\left[\sqrt{\frac{N}{2}} \times \left(\frac{\lambda'}{N} - 1 \right) \right] \tag{7}$$

In Eq. (7), λ' is defined as $\lambda' = \frac{\lambda_1}{(\sigma_v^2 + \sigma_g^2)}$, where λ_1 is a single adaptive threshold and σ_v^2 is the LU signal variance.

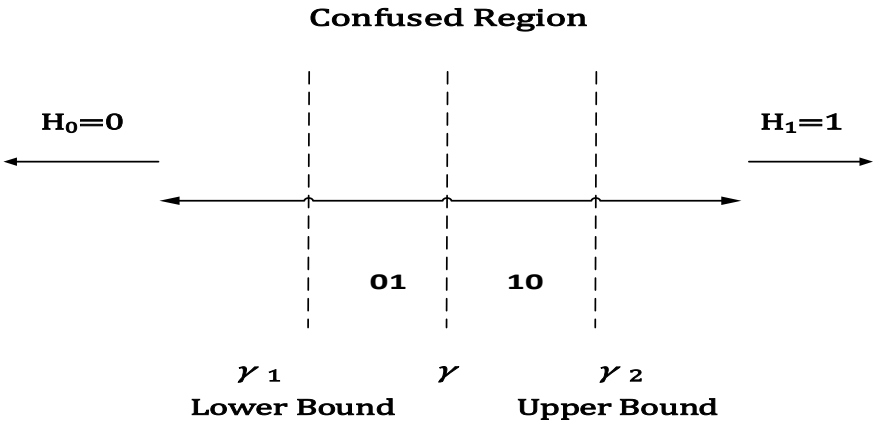


Fig. 3 The double threshold detection scheme

3.1.2 Second Stage: ED-TAT

In Fig. 3, the entire part is divided into three sections, the first section, i.e., below γ_1, has only noise and defined by H_0, the second section, i.e., above γ_2 has the LU signal and defined by H_1, and third section is between γ_1 and γ_2, the combination of noise and the LU signal known as *confused region* [7].

Therefore, the detector output can be written as

$$\text{ED-TAT}|_{\text{o/p}} = \text{UP} + \text{LP} = Z \tag{8}$$

Consider the confused region, if observed energy exists between (γ_1–γ) it shows 01 and further converts binary di-bits into decimal, i.e., 1, similarly, if observed energy exists between (γ–γ_2) shows 10 and further, its decimal value is 2. Now, the predefined threshold (γ) can be calculated as [8]

$$\gamma = \left[\left(\frac{N}{\sigma_g^2}\right) \times \left\{Q^{-1}\left(\overline{P_f}\right) \times \sqrt{\frac{2}{N}} + 1\right\}\right] \tag{9}$$

The mathematical expression for the lower threshold (γ_1) and the upper threshold (γ_2) can be found as

$$\gamma_1 = \left[\left(\frac{N}{\rho \times \sigma_g^2}\right) \times \left\{Q^{-1}\left(\overline{P_f}\right) \times \sqrt{\frac{2}{N}} + 1\right\}\right] \tag{10}$$

$$\gamma_2 = \left[\left(N \times \rho \times \sigma_g^2\right) \times \left\{Q^{-1}\left(\overline{P_f}\right) \times \sqrt{\frac{2}{N}} + 1\right\}\right] \tag{11}$$

The value of the lower part (LP) and the upper part (UP) is defined as

$$LP = \begin{cases} \text{if } \gamma_1 \leq E < \gamma, \text{ represent bit } 01 \\ \text{if } \gamma \leq E < \gamma_2, \text{ represent bit } 10 \end{cases} \tag{12}$$

$$UP = \begin{cases} \text{if } E < \gamma_1, \text{ represent bit } 0 \\ \text{if } \gamma_2 \leq E, \text{ represent bit } 1 \end{cases} \tag{13}$$

Now, combining LP and UP, the final output of ED-TAT becomes [8]

$$\text{ED-TAT}|_{\text{o/p}} = \begin{cases} (\text{UP} + \text{LP}) < \gamma, \ H_0 \\ (\text{UP} + \text{LP}) \geq \gamma, \ H_1 \end{cases} \tag{14}$$

Detection Probability for ED-TAT Detector

Now, the detection probability for ED-TAT becomes

$$P_d^{\text{ED-TAT}} = \exp\left[-\frac{\left\{ (\gamma)^{\frac{1}{a}} \right\}^2}{(1+U)} \right] \tag{15}$$

4 Numerical Results and Analysis

In our simulations, there are 1000 number of samples, P_f is set at 0.1, thresholds are $_1 = 1.7$, $\gamma_1 = 0.9$, $\gamma_2 = 1.4$ and $\gamma = 1.15$, and range of SNR varies -20 to 0 dB.

Analysis Fig. 4, double-stage sensing technique outperforms ED and cyclo detector-2010, adaptive spectrum sensing-2012, EDT-ASS-2015, and ED and ED-ADT detector-2015 by 34.4%, 33.9%, 13.4%, and 2.9% at—10 dB SNR in terms of detection probability, respectively. Under the consideration of IEEE 802.22 rules, the proposed detector is able to detect the LU signal at—12.0 dB SNR level approximately.

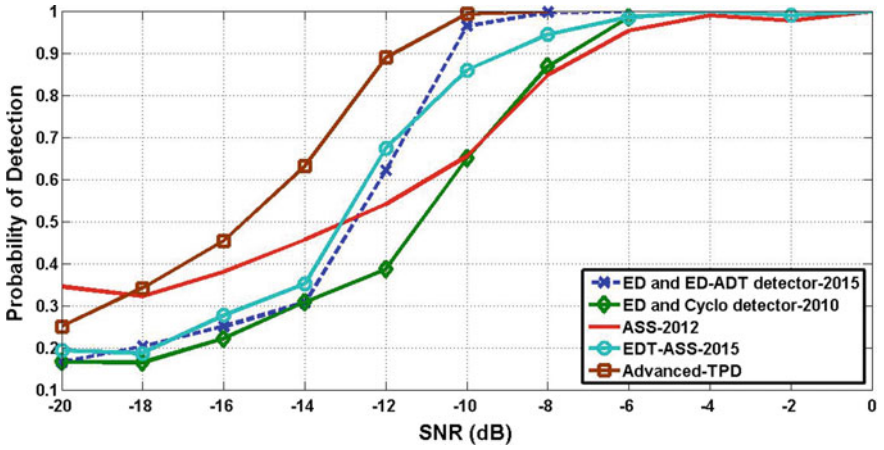


Fig. 4 Detection probability versus SNR

5 Conclusion

In this paper, a double-stage sensing technique for spectrum sensing has been proposed. This scheme enhances detection performance. Simulation results confirm that the proposed double-stage sensing scheme beats the existing sensing techniques, i.e., it outperforms the ED and cyclo detector-2010, adaptive spectrum sensing-2012, energy detection technique for adaptive spectrum sensing-2015 (EDT-ASS-2015), and ED and ED-ADT detector-2015 by 34.4%, 33.9%, 13.4%, and 2.9% at—10 dB SNR respectively.

Acknowledgements The authors wish to thank their parents for supporting and motivating this work because without their blessings and God's grace this was not possible.

References

1. A. Bagwari, B. Singh, Comparative performance evaluation of spectrum sensing techniques for cognitive radio networks, in *Fourth IEEE International Conference on Computational Intelligence and Communication Networks (CICN-2012)*, vol. 1 (2012), pp. 98–105
2. Xiaoshuang Xing, Tao Jing, Wei Cheng, Yan Huo, Xiuzhen Cheng, Spectrum prediction in cognitive radio networks. IEEE Wirel. Commun. **20**(2), 90–96 (2013)
3. S. Maleki, A. Pandharipande, G. Leus, Two-stage spectrum sensing for cognitive radios, in *IEEE Conference on Acoustics Speech and Signal Processing (ICASSP)* (2010), pp. 2946–2949
4. W. Ejaz, N. ul Hasan, H.S. Kim, SNR-based adaptive spectrum sensing for cognitive radio networks. Int. J. Innov. Comput. Inf. Control **8**(9), 6095–6105 (2012)
5. I. Sobron, P.S.R. Diniz, W.A. Martins, M. Velez, Energy detection technique for adaptive spectrum sensing. IEEE Trans. Commun. **63**(3), 617–627 (2015)

6. A. Bagwari, G.S. Tomar, Adaptive double-threshold based energy detector for spectrum sensing in cognitive radio networks. Int. J. Electron. Lett. (IJEL) **1**(1), 24–32 (2013) (Taylor & Francis Group)
7. R. Tandra, A. Sahai, SNR walls for signal detection. IEEE J. Sel. Top. Signal Process. **2**(1), 4–16 (2008)
8. A. Bagwari, J. Kanti, G.S. Tomar, A. Samarah, Reliable spectrum sensing scheme based on dual detector with double-threshold for IEEE 802.22 WRAN. J. High Speed Netw. **21**(3), 205–220 (2015) (IOS Press)

Modified Soft Combination Scheme for Cooperative Sequential Detection Considering Fast-Fading in Cognitive Radios



Mayank Sahu and Amit Baghel

1 Introduction

Today, cognitive radios (CR) emerged as a means, which enabled secondary users to enhance the spectrum utilization by opportunistically detecting spectrum holes or white spaces to mitigate the problem of spectrum scarcity. Spectrum sensing, an essential function of cognitive radio networks, detects spectrum holes or white spaces, to exploit spectrum access opportunities without causing interference to the primary user [1–4].

Energy detector (ED) is the often used spectrum sensor, as it does not require any prior information about the received signal and the noise signal. It is also easy to implement [5]. However, to make detection decision for the case of low SNR, ED requires a huge number of the samples, consequently long sensing time. Sequential detection method as proposed by Wald [6]; reduces the number of samples required by energy detector through sequential probability ratio test (SPRT).

In cooperative sensing, many CRs collaborate to mitigate the effects of shadowing and fading. Each CR senses the spectrum for availability of white spaces individually and sends the information to the fusion center that collects information from all the participating CRs and takes the detection decision. The cooperative sequential detection scheme was well analyzed by [7]; to decrease the average number of samples required by a CR, composite hypotheses employing generalized log-likelihood ratio (GLLR) are performed excellently for the case of independent and identically distributed (i.i.d.) samples received by the detector.

In practice, fast-fading becomes prominent hence it is important to study sensing of spectrum by cognitive radios considering i.n.i.d. samples, signifying fast-fading, which was studied by Cho [8]. In [9] it was exhibited that using soft combination in

M. Sahu · A. Baghel (✉)
Jabalpur Engineering College, Jabalpur 482011, MP, India
e-mail: amitbaghel77@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_126

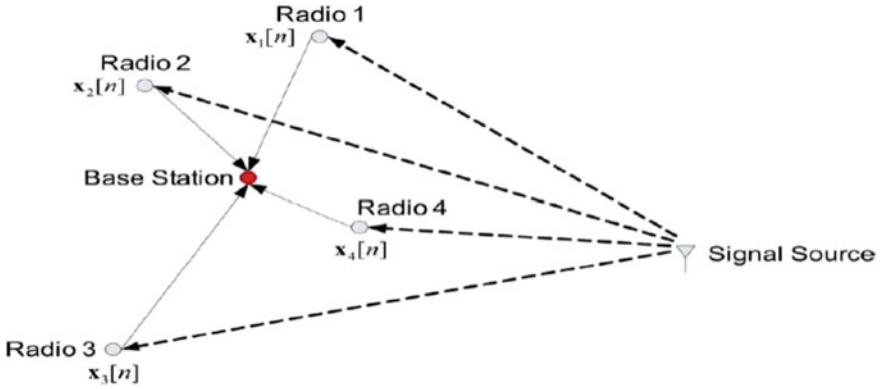


Fig. 1 Cognitive radio network with a number of cooperating cognitive radios [7]

cooperative sensing, like equal gain combination (EGC) and maximal ratio combination (MRC), improved detection performance in cooperative sensing.

Here, cooperative sequential detector is used by us with energy detector. Acquired samples of the signal and noise are dependent on the received signal variance and noise signal variance, respectively. The network consists of a number of cognitive radios that are sensing the spectrum for the presence primary user, as shown in Fig. 1. Also, we proposed a modified soft combination rule based on linearly combining, weighted local log-likelihood ratio of participating CRs for the cooperative sequential spectrum sensing when the acquired samples are independent and nonidentically distributed (i.n.i.d.) due to fast-fading. This is due to fast-fading, where channel characteristics change rapidly.

2 System Model

Here, we adopted similar system model as that of [8], where M is the number of cooperating CRs and assumed that the mth ($m = 1, 2, \ldots, M$) cognitive radio acquires nth sample as a zero-mean Gaussian random variable, where $v_0^m(n)$ is the noise signal variance and $v_1^m(n)$ is the received signal variance ($v_1^m(n) > v_0^m(n)$), i.e.,

$$H_i : x^m[n] \sim \frac{1}{\sqrt{2\pi\, v_i^m(n)}} \exp\left(-\frac{(x^m(n))^2}{2v_i^m(n)}\right) \tag{1}$$

Here the two hypotheses are as follows:

H_0 primary user is absent
H_1 primary user is present

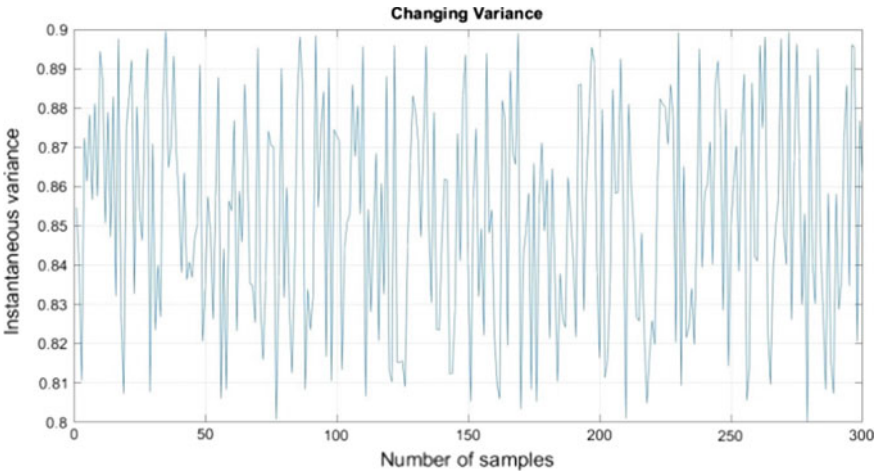


Fig. 2 Instantaneous noise variance as a uniform random variable with $L = 0.8$ and $U = 0.9$

$\nu_0^m(n)$ and $\nu_1^m(n)$ are the instantaneous variances for nth sample, under H_0 and H_1, respectively. In fast-fading environment, instantaneous variances, $\nu_0^m(n) \in V_0^m$ and $\nu_1^m(n) \in V_1^m$, are unknown and are taken as random variables for $m = 1, 2, \ldots, M$ and $n = 1, 2, \ldots, N$, and $\nu_0^m(n)$ and $\nu_1^m(n)$ have probability density functions as $p(\nu_0^m)$ and $p(\nu_1^m)$, respectively (Fig. 2).

We have taken sample spaces V_0^m and V_1^m as disjoint, i.e.,

$$V_0^m = \left\{x | L_0^m \leq x \leq U_0^m\right\} \tag{2}$$

and

$$V_1^m = \left\{y | L_1^m \leq y \leq U_1^m\right\} \tag{3}$$

The distributions of $\nu_i^m(n)$ $(m = 1, 2, \ldots, M)$ between the other CRs differ from each other but the distribution of $\nu_i^m(n)$ $(n = 1, 2, \ldots, N)$ is same for single CR for long duration of time. Here, $p_0^m(\mathbf{x}^m[n]; v_0^m(n))$ and $p_1^m(\mathbf{x}^m[n]; v_1^m(n))$ are the probability distribution functions of the acquired signal by the mth CR under H_0 and H_1, respectively.

3 Sequential Detector for Fast-Fading Scenario

3.1 *Sequential Detector with* GLLR

Here, summary of the work of Cho in [8] is given for better understanding. In the previous section, we assumed that the received signal variance and noise signal variance changes rapidly while sensing. The ideal sequential probability ratio test (SPRT) for this case is as follows:

(i) The mth ($m = 1, 2, \ldots, M$) CR acquires sample $x^m[N]$ then computes $\ln\left(\frac{p_1^m\left(x^m[N];v_1^m(N)\right)}{p_0^m\left(x^m[N];v_0^m(N)\right)}\right)$

(ii) Fusion center then updates ideal log-likelihood ratio $LLR_{\mathrm{N}}^{\mathrm{ideal}}$, sequentially as

$$LLR_{\mathrm{N}}^{\mathrm{ideal}} = LLR_{\mathrm{N-1}}^{\mathrm{ideal}} + \sum_{m=1}^{M} \ln\left(\frac{p_1^m\left(x^m[N];\, v_1^m(N)\right)}{p_0^m\left(x^m[N];\, v_0^m(N)\right)}\right) \tag{4}$$

(iii) H_0 is accepted, if $LLR_{\mathrm{N}}^{\mathrm{ideal}} \leq \lambda_0$; H_1 is accepted, if $LLR_{\mathrm{N}}^{\mathrm{ideal}} \geq \lambda_1$, where λ_0 and λ_1 are conceptual thresholds.

(iv) If not, one more sample is taken and Steps (i)–(iii) are repeated.

However, $v_i^m(n)$ ($n = 1, 2, \ldots, N$; $m = 1, 2, \ldots, M$; $i = 0, 1$) is considered as a random variable, so $LLR_{\mathrm{N}}^{\mathrm{ideal}}$ can not be calculated exactly. Thus, the cooperative sequential detection using the GLLR [8] is given after replacing $v_i^m(n)$ by their maximum likelihood estimates (MLE). Therefore we have

$$\mathrm{GLLR}_N = \sum_{n=1}^{N}\sum_{m=1}^{M} \ln\left(\frac{p_1^m\left(x^m[n];\, \tilde{v}_1^{m,(N)}\right)}{p_0^m\left(x^m[n];\, \tilde{v}_0^{m,(N)}\right)}\right) \tag{5}$$

where $\tilde{v}_0^{m,(N)}$ and $\tilde{v}_1^{m,(N)}$ are the MLE of v_0^m and v_1^m [7], given by

$$\tilde{v}_0^{m,(N)} = \arg \max_{v_0^m \in V_0^m} \sum_{n=1}^{N} \ln p_0^m\left(x^m[n];\, v_0^m\right) \tag{6}$$

$$\tilde{v}_1^{m,(N)} = \arg \max_{v_1^m \in V_1^m} \sum_{n=1}^{N} \ln p_1^m\left(x^m[n];\, v_1^m\right) \tag{7}$$

For fast-fading environments, the MLEs, $\tilde{v}_0^{m,(N)}$ and $\tilde{v}_1^{m,(N)}$, converges to the following values:

(a) Under H0, $\tilde{v}_0^{m,(N)}$ converges to $\mathbf{E}\{\tilde{v}_0^{m,(N)}\}$ and $\tilde{v}_1^{m,(N)}$ converges to L_1^m.

(b) Under H1, $\tilde{v}_0^{m,(N)}$ converges to U_0^m and $\tilde{v}_1^{m,(N)}$ converges to $\mathbf{E}\{v_1^m(n)\}$.

Table 1 Simulation Parameters

	L_0^m	U_0^m	L_1^m	U_1^m	$\tilde{v}_0^m$	$\tilde{v}_1^m$
$m = 1$	0.64	0.90	0.90	1.15	0.7465	1.0515
$m = 2$	0.75	0.87	0.90	1.18	0.7976	1.0526
$m = 3$	0.72	0.82	0.87	1.02	0.7633	0.9522
$m = 4$	0.69	0.85	0.86	0.99	0.7575	0.9386

Therefore, under H_0, the corresponding GLLR_N^m can be expressed by $\sum_{n=1}^{N} \ln\left(\frac{p_1^m\left(x^m[n]; L_1^m\right)}{p_0^m\left(x^m[n]; \boldsymbol{E}\{v_0^m\}\right)}\right)$. Similarly, under $H1$, the corresponding GLLR_N^m can be expressed by $\sum_{n=1}^{N} \ln\left(\frac{p_1^m\left(x^m[n]; \boldsymbol{E}\{v_1^m\}\right)}{p_0^m\left(x^m[n]; U_0^m\right)}\right)$.

3.2 *Log-Likelihood Ratio When* $\mathbf{P}(\mathbf{V}_0^m)$ *and* $\mathbf{P}(\mathbf{V}_1^m)$ *Are Unknown*

Since we do not use the conventional *GLLR* scheme, hence, here we define optimal log-likelihood ratio for cooperative sequential detection:

$$\widetilde{LLR}_N^m = \sum_{n=1}^{N} \ln\left(\frac{p_1^m\left(x^m[n]; \tilde{v}_1^m\right)}{p_0^m\left(x^m[n]; \tilde{v}_0^m\right)}\right) \tag{8}$$

where $\tilde{v}_0^m \in V_0^m$ and $\tilde{v}_1^m \in V_1^m$ are given in Table 1.

3.3 *Modified Soft Combination Rule*

The rule for soft combination, as given in [9] was modified by us on the hit and trial basis, first LLRs are obtained from each CR, which are then sent to a fusion center, at the fusion center a decision is made to show whether primary user is present or not by linearly combining the weighted local LLRs of each CR. The log-likelihood ratio for mth ($m = 1, 2, \ldots, M$) CR is expressed as follows:

$$\mathrm{MLLR}_n^m = \sum_{n=1}^{N}\sum_{m=1}^{M}\left[\frac{1}{2}\left(\frac{1}{\left(\sigma_{0,m}^*\right)^2} - \frac{1}{\left(\sigma_{1,m}^*\right)^2}\right)(X_m[n])^2 + \frac{1}{2}\ln\left(\frac{\left(\sigma_{0,m}^*\right)^2}{(\sigma_{0,m}^*)^2}\right)\right] \tag{9}$$

Here N is the required number of samples. The fusion center employs soft combination as the following:

$$Y = \sum_{m=1}^{M} \omega_m \text{MLLR}_n^m \tag{10}$$

where ω_m is the mth CR's weighting coefficients [10], given as

$$\omega_m = \frac{1}{\sqrt{M}} \tag{11}$$

On modifying equal gain combination (EGC), the weight for the mth CR will be as follows:

$$\omega_m = \frac{4}{\sqrt{M}} \tag{12}$$

where M is the number of cooperating CRs.

4 Simulation and Results

The performance of our proposed method is illustrated by performing simulations considering the average number of samples required for detection. Here, number of cooperating cognitive users were taken as $M = 4$. Other simulation parameters are given in Table 1. In this scenario, received signal variance and noise signal variance were assumed to have uniform probability distribution. Extensive Monte Carlo simulations for both scenarios under H_0 and H_1 were performed. Detection thresholds are dependent on the false- and missed-detection constraints, α and β which are predefined and, hence, detection thresholds were computed as given in [7]. Simulation results, under cooperative sequential detector considering fast-fading under H_0 and H_1 are shown in Fig. 3 and Fig. 4, respectively. As evident from the figures that our proposed method performed better than the LLR method as proposed by Cho in [8].

5 Conclusion

We in this paper proposed a modified soft combination rule for cooperative sequential detection in fast-fading environment. We also, showed how the cooperative sequential detector using LLR performed in terms of average number of samples and a modified soft combination method was proposed by us which is applicable to the same scenario. Our simulations results exhibited that the proposed modified method for cooperative sequential detection performs better with respect to average sample numbers in fast-fading scenario.

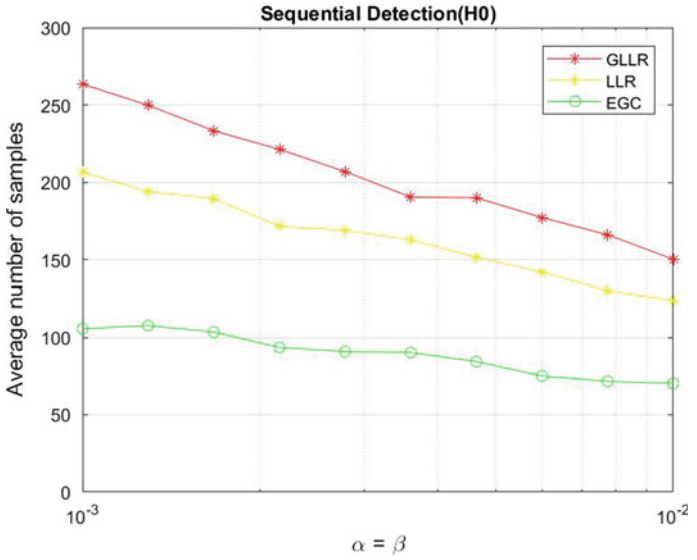


Fig. 3 Cooperative sequential detector considering fast-fading scenario under H_0, where α and β are false alarm and missed-detection constraints and are predefined

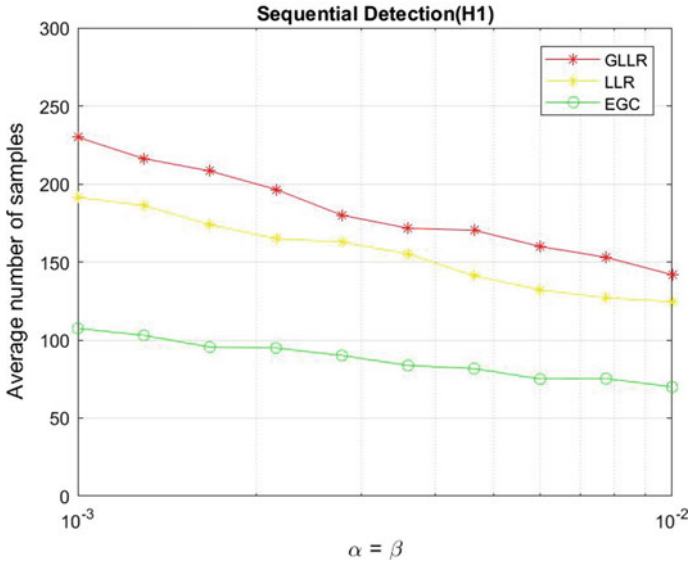


Fig. 4 Cooperative sequential detector considering fast-fading scenario under H_1, where α and β are false alarm and missed-detection constraints and are predefined

References

1. F. Akyildiz, B.F. Lo, R. Balakrishnan, Cooperative spectrum sensing in cognitive radio networks: a survey. Phys. Commun. J. **4**(1), 40–62 (2011)
2. J. Mitola III, G.Q. Maguire, Cognitive radio: making software radios more personal. IEEE Personal Commun. **6**(4), 13–18 (1999)
3. S. Haykin, Cognitive radio: brain-empowered wireless communications. IEEE J. Sel. Areas Commun. **23**(2), 201–220 (2005)
4. A. Ali, W. Hamouda, Advances on spectrum sensing for cognitive radio networks: theory and applications. IEEE Commun. Surv. Tutor. **19**(2), 1277–1304 (2016) (Secondquarter 2017)
5. D. Cabric, A. Tkachenko, R.W. Brodersen, Spectrum sensing measurements of pilot, energy, and collaborative detection, in *Proceedings of MILCOM*, pp. 1–7, Oct 2006
6. A. Wald, *Sequential Analysis* (Wiley, New York, 1947)
7. Q. Zou, S. Zheng, A.H. Sayed, Cooperative sensing via sequential detection. IEEE Trans. Signal Process. **58**(12), 6266–6283 (2010)
8. Y.J. Cho, J.H. Lee, A. Heo, D.J. Park, Smart sensing strategy for SPRT considering fast changing sample statistics in cognitive radio systems, in *Proceedings of 2012 International Conference on Future Generation Communication Technology*, no. 9 (2012), pp. 1–8
9. X. Zhang, Z. Qiu, D. Mu, A modified SPRT based cooperative spectrum sensing scheme in cognitive radio, in *Proceedings of 2010 IEEE International Conference on Signal Processing*, no. 10 (2010), pp. 1512–1515
10. J. Ma, Y. Li, Soft combination and detection for cooperative spectrum sensing in cognitive radio networks, in *Proceedings of IEEE GLOBECOM*, pp. 3139–3143, Nov 2007

Role of Chaos in Spread Spectrum Communication



Devendra Kumar, Divesh Kumar and Dheeraj Kalra

1 Introduction

In the fifth-generation communication, the spread spectrum technique is preferred for high data rate communication, in which the pseudo-noise codes are utilized to expand the bandwidth of information signal uniformly at the same transmitted power. Walsh codes, Gold Codes, and PN sequences are the famous spreading codes. PN codes are random sequences and hence nondeterministic in nature. On the other hand, the signal or codes generated on the basis of chaos theory (chaotic signals) are deterministic in nature. The chaotic signals are also limited and nonperiodic as well as highly sensitive to the initial condition [1, 8]. So the noise-like behavior of the chaotic signals makes them suitable for the generation of spreading codes.

Communication systems based on chaos are having a potential advantage over conventional pseudorandom-based systems in terms of security and synchronization [48]. Over the past few years, some more chaos-based spreading systems have been proposed such as chaotic masking, chaotic modulation, and chaos shift keying. More interestingly, the chaos-based SS techniques are well suited for analog and digital signals.

This paper meets the following objectives: first, it provides the introduction to chaos theory and its role in spread spectrum communication engineering and second, the synchronization issue [2, 3]. It will also explore the new area of developments on the basis of signal processing capability.

D. Kumar (✉) · D. Kumar · D. Kalra
Department of Electronics and Communication Engineering, GLA University, Mathura, UP, India
e-mail: devendrakumar.ec@gla.ac.in

D. Kumar
e-mail: divesh.kumar@gla.ac.in

D. Kalra
e-mail: dheeraj.kalra@gla.ac.in

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_127

2 Chaos Theory

The phenomenon of chaos is robust and universally present in many nonlinear systems with some characteristics. With reference to the chaos-based dynamics, these characteristics are (1) sensitivity toward the initial conditions, (2) broadband frequency spectrum, (3) noise-like behavior, and (4) high complexity [32, 35]. Due to these characteristics, the chaos theory can be used to generate the random-like and secure sequences for spread spectrum communication. Chaos can be used as masking or modulation of data as well as a suitable replacement of pseudorandom sequence in direct sequence spread spectrum system. Historically, Faraday's experiment in 1831 may be considered the first to express chaos scientifically. Because of its characteristics, it was not easy to use chaos in communication, but in 1990 [48], it was reported that chaos could be controlled. Later, synchronization issues were also reported in 1992, after which the researchers gradually realized the role of chaos in communication. The chaotic sequence can serve as a better replacement of pseudorandom sequences, that is why, many scientists are involved in the research of use of chaos in secure communication, specifically in spread spectrum system [11–13].

Chaos can be better understood with the help of difference equations. The logistic population model is very popular to understand chaos.

Definition 1 The function of the logistic map can be described as

$$f(x_n) = x_{n+1} = r\, x_n(1 - x_n) \tag{1}$$

where r is an important parameter to discuss and popularly known as the growth rate of population. The parameter r can be elaborated for different values.

The values of r are in the range of $0 \leq r \leq 4$.

Proposition—for the above values of r, the logistic map produces [0, 1] and inputs it to self.

One more point, i.e., $x = 1 - \frac{1}{r}$, can be considered stable, however, only for the range, i.e., $1 < r < 3$ and becomes unstable for $r > 3$.

Proposition—from the analysis of the chaotic logistic map, it can be also shown that it has two cycles for $r > 3$.

With the help of the above propositions, it can be concluded that as the value of r increases, the stability coefficients of the fixed state point decreases and at the value of $r = 3$, a stable 2-cycle is formed. So the value of $r = 3$ is popularly known as *period doubling bifurcation.* Further, if the values of r increases, then 2-cycle stability becomes unstable; however, simultaneously a stable 4-cycle started at $r = 3.4494897$.

In Fig. 1, the one-dimensional logistic map with the bifurcation parameter r is pictorially depicted and for $r < 4$, an orbit converges at the stable fixed point. In Fig. 2, a different situation is depicted; at $r = 4$, i.e., it may be considered as chaotic because the stable point becomes unstable.

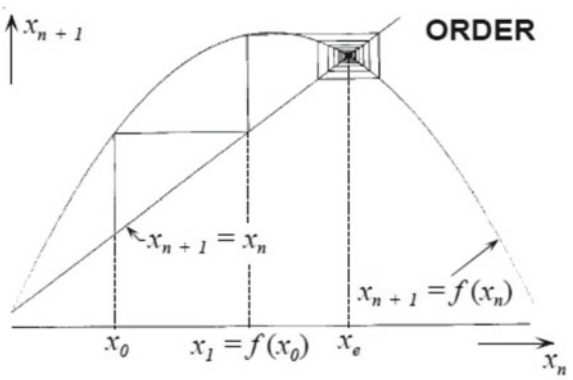


Fig. 1 One-dimensional logistic map with the bifurcation parameter *r* set below 3

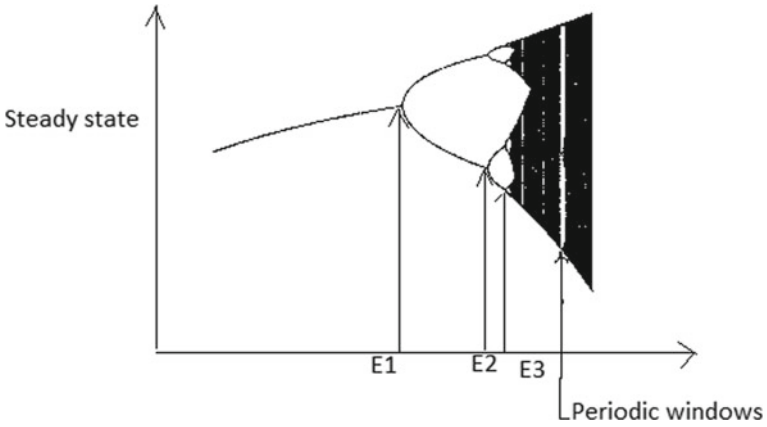


Fig. 2 Bifurcation diagram with period-doubling rate

Similarly, Fig. 2 shows the bifurcation diagram and a very popular characteristic of any nonlinear dynamical system. Now the other interesting property, which any chaotic system exhibits, is dependence on initial conditions of any nonlinear dynamical system.

Case—When the value of *r* is greater than 4, then the logistic map can be considered chaotic.

Now we may have several important observations of chaos. First, chaos can occur only in nonlinear systems and can be described with the help of differential equation or difference equations; behavior of chaos can also be studied in the following types such that chaos in periodically forced system can be named as Transient chaos; similarly, *Intermittent chaos* or *intermittency* as it is commonly called, usually occurs when a mapping exhibits a form of criticality that leads into a behavior with chaotic episodes interspersed with regular behavior at random periods.

3 Chaos-Based Communication

Why chaotic signal can be the replacement of sinusoidal carriers in communication systems?

This question has become surprisingly very interesting and holds great importance.

Some researchers raised this question in a better way; here, we can quote them one by one to find the answer to the above-said statement. When a sinusoidal signal is used to transmit information, the power spectral density concentrates on a narrow range of frequencies. Whereas chaotic signals, can occupy a large bandwidth; their autocorrelations and cross-correlation properties are also favorable. These characteristics made chaotic signals a better choice in communication systems.

Chaos-based SS systems have several properties, namely (i) difficult to interfere with any unauthorized user; (ii) information transfer is more secure than any other communication system; and (iii) resistant to jamming [19, 24, 32, 50].

Shannon et al. discussed three fundamental aspects of secret communication systems in one of his popular papers, viz., concealment, privacy, and encryption. Hiding the existence of a transmitted message is known as concealment and chaotic carrier signal can be used for concealment because of its irregular shape and a periodicity. The privacy aspect of communication is to keep the information transmitted private form other non-intended recipients. The third aspect of security is encryption, which uses "key" to encrypt and decrypt the messages.

The use of chaos in communication arises as chaotic signals satisfying the characteristics mentioned by Shannon. These signals are proved to be better in the use of spread spectrum and can be a good substitute for other spreading codes. In further sections, we intend to discuss the role of chaos in SS system.

4 Chaos-Based DS-SS (CDS-SS) Communication

As discussed, CDS-SS System has drawn much attention recently. Existing methods can be subdivided into two broad categories: one is coherent and the second is noncoherent. Here, the coherent system refers to both transmitter and receiver having the same spreading sequence. On the other hand, noncoherent systems do not use the same spreading sequence but it uses some other method to detect the information. One important point is to be noted that different signaling schemes can be used in chaos-based direct sequence spread spectrum system such as BPSK and QPSK.

Tabulation is the best way to represent the use of chaos. The below-mentioned table presents the role of chaos in different spread spectrum systems. Lau et al. suggested the use of chaotic sequence to distinguish the users. Azou et al. used the chaotic sequences in a code division multiple access for the application of underwater communication. Similarly, the whole table presents the use of chaos in different ways. Most of the researchers attempted to present chaos as a substitute for spreading

sequences. Without any doubt, this review table represents the rapid growth of chaos in communication.

Some of the applications are in a coherent category and rest are in the noncoherent one. In general, coherent communication systems are more preferred and less complex. Bandwidth is utilized more in comparison to noncoherent systems. The receiver structure is relatively simple and needs less processing [32, 33]. However, with all the advantages of coherent systems, some limitations are also present, such that these systems require the exact replica of the transmitter spreading sequence, which means that both the sequences should be properly synchronized. Furthermore, synchronization may be treated as the main constraint in coherent systems. On the other hand, noncoherent systems are not good in the utilization of bandwidth and they need extra overhead to achieve synchronization. Hence, noncoherent systems have poor throughput. Although this is a debatable issue to find the optimal choice between these two systems, here, we choose coherent systems because it can be considered optimal under some constraints.

5 System Overview

Many Authors proposed CDSS systems based on different chaotic maps. Some of them are already enlisted in Table 1; here, we first propose a general block diagram of the CDSS system and later discuss the details of each block (Fig. 3).

Here, the proposed block diagram is under the category of a coherent system, hence chaos generator blocks at transmitter and receiver are using the same chaotic sequences. Now, the first part of the system is the Data block, which produces input data sequences. In the second part, this data sequence is spread by a particular chaotic sequence. Researchers proposed many methods to generate chaotic sequences, namely logistic map, tent map, Gauss-iterated map, Gingerbread man map, Henon map, Chebyshev polynomial function of order 2 (CPF), and the second is a one-dimensional noninvertible piecewise inner map (PWL). Some of them are shown in a tabular format. Among all the maps, logistic map is more popular to illustrate the phenomenon of chaos. Once chaotic sequence is generated; it is multiplied with the data bits to produce the transmitter output. Hence irrespective of chaotic map, we can write the general output of transmitter as follows [49, 61].

$$G_t = \sum_{s=1}^{N} \gamma_i^k l_t^k \tag{2}$$

where N is the no. of users, k is the user index, and i is the bit index. Given that

$$\forall i, k, \gamma_i^k = \begin{Bmatrix} +1 \\ -1 \end{Bmatrix} \tag{3}$$

Table 1 Review of role of chaos in spread spectrum systems

System	Year	Authors	Features
DCSK	2001	Lau et al.	Differential CSK is proposed and chaos is used to distinguish the users
CD3S	2002	Azou et al.	Chaotic sequence is used in CDMA for underwater communication
CSK-DSSS	2003	Lau et al.	Combined chaos-based and conventional DSSS system is proposed
DCSK	2004	Xia et al.	Performance of the noncoherent DCSK communication over a multipath channel is evaluated with delay spread
DS/CDMA	2005	Kurian et al.	DS/CDMA system is proposed using chaotic Ikeda map as spreading sequence
DS/CDMA	2006	Assad et al.	Chaotic cequences is used for secure communication
Digital system	2007	Mooney et al.	Chaotic signal and white noise are compared
Fractional	2009	Kiani et al.	FCS method is implemented by using Kalman filter chaotic system
DCSK	2011a, b	Xu et al.	DCSK cooperative communication is proposed
DCSK	2012	Zhang et al.	Image transmission is suggested with chaos shift keying
CDS-SS	2013	Li et al.	A particle filter is used to demodulate CDS-SS
OFDM	2014	Quen et al.	CDSSS communication system based on OFDM is proposed.
DS/CDMA	2014	Zhao et al.	Hyper chaotic sequences are generated
DS/CDMA	2015	Mahmood et al.	Orthogonal chaotic vector is generated as a spreading sequence in DS/CDMA

DCSK Differential chaos shift keying
CD3S Chaotic direct sequence spread spectrum system
CSK-DSSS Chaos shift keying-Direct sequence spread spectrum
DS/CDMA Direct sequence/Code division multiple access
OFDM Orthogonal frequency division multiple access

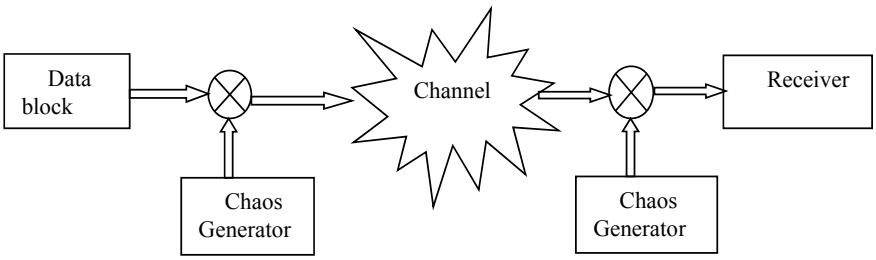


Fig. 3 General block diagram of CDS-SS system

Below, some of the popular chaotic maps are given which are popular in the generation of chaotic sequences. Simulated bifurcation diagrams are also presented of the Logistic map, Tent map, and Bernoulli's map (Fig. 4 and Table 2).

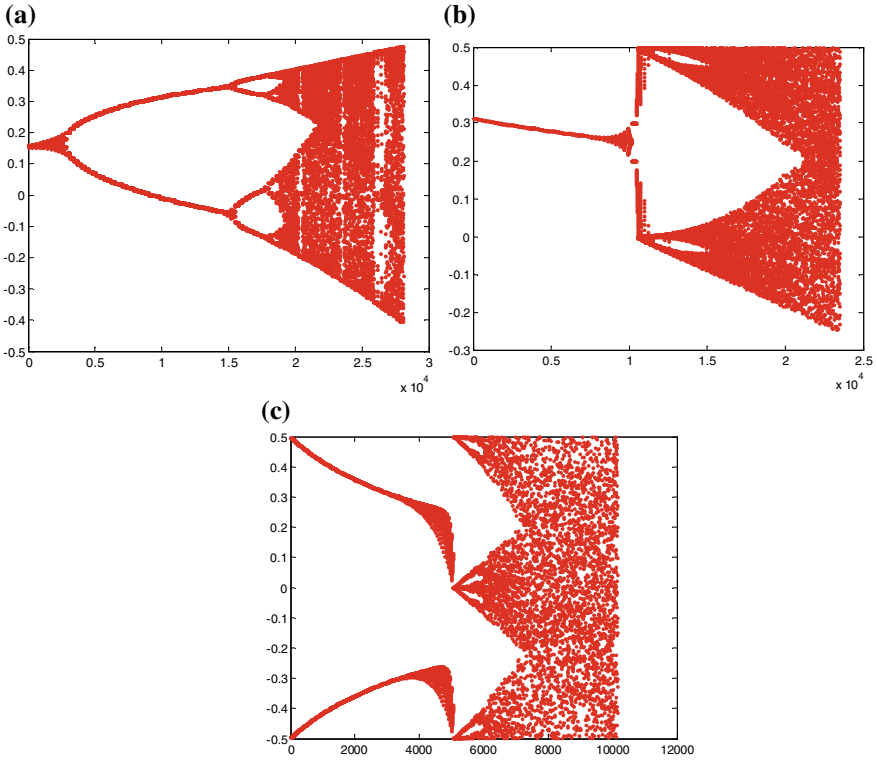


Fig. 4 **a** Bifurcation diagram of Logistic Map. **b** Bifurcation diagram of TENT map. **c** Bifurcation diagram of Bernoulli's map

Table 2 Use of different chaotic maps for the generation of chaotic sequences

Chaotic map	Time domain/dimension	Equation
Logistic map	Discrete/one-dimensional	$X_{n+1} = AX_n(1 - X_n)$
Tent map	✓/✓	$X_{n+1} = A\min(X_n, 1 - X_n)$
Binary shift	✓/✓	$X_{n+1} - 2\,X_n$
Baker map	✓/two-dimensional	$F(x, y) =$ $\left\{ \begin{array}{ll} (2x, y/2 & 0 < \mathrm{x} < 1/2 \\ ((2x - 1), (y + 1/2)) & 1/2 < \mathrm{x} < 1 \end{array} \right\}$
Henon map	✓/✓	$X_{n+1} = 1 - aX_n^2 + Y_n$ $Y_{n+1} = bX_n$

The transmitter output is then moved toward the channel which normally consists of noise, fading, interference, and propagation delay. The popular noise model is additive white Gaussian noise and it simply gets added to the transmitted signal to generate input to the receiver. This is an ideal form of noise where white represents large bandwidth and Gaussian represents a statistical distribution of the values.

The channel may also consist of reflection, diffraction, and scattering mechanism and may constitute variations in received signal strength; this phenomenon is called fading and such channels can be modeled with the help of Rayleigh, Rician, and many other popular distribution models. The Rayleigh distribution model has the following probability density function (pdf).

$$p(r) = \left\{ \frac{r}{\sigma^2} \exp\left(-\frac{r^2}{2\sigma^2} \right) \right\} \text{ for } 0 \leq r \leq \infty \tag{4}$$

where r is the rms value of the received signal before detection and σ^2 is the time average power of the received signal before detection. Similarly, Rician fading distribution can be given as

$$p(r) = \left\{ \frac{r}{\sigma^2} \exp\left(-\frac{(r^2 + A^2)}{2\sigma^2} \right) I_0\left(\frac{Ar}{\sigma^2} \right) \right\} \text{ for } (A \geq 0, r \geq 0) \tag{5}$$

The parameter A denotes the peak amplitude and $I_0(\cdot)$ is the modified Bessel function of the first kind and zero order. Diversity techniques are used to mitigate the effect of fading. RAKE receiver is also very popular in DS-CDMA to combat this effect.

Here, as we use a coherent CD-SS system, the same chaotic sequence is used at the receiving side to decode the received signal. Now this is obvious that both the chaotic sequences should generate the output at the same time and must be synchronized for the proper recovery of message signal. However, the transmitted signals take some time to reach the receiver, known as propagation delay. This delay produces time misalignment and hence the need for synchronization arises.

Up to now, nearly all literature of CD-SS has been presented by assuming perfect synchronization. Although sequence synchronization is very difficult to achieve, some of the researchers suggested moving toward the need for noncoherent systems and some prove that synchronization is perfectly achievable. However, motives of both the researches are to make a communication system better and noise free [28, 52, 53].

As we discussed that the purpose of the acquisition is to accurately estimate the time difference between the sequence at transmitter and receiver so that information can be decoded properly, accuracy is not considered to be more than one chip duration for proper de-spreading. Sequence acquisition can be achieved with the help of detection schemes and search strategies. Usually, acquisition is achieved in two steps:

- Initial code acquisition, which synchronizes the transmitted and received data within one-chip period ($\pm T_c$). This is also known as a coarse acquisition or coarse synchronization.
- Code tracking, which performs fine synchronization between $T_x \& R_x$.

6 Synchronization in DS-SS System

With the help of the abovementioned steps, it is easy to perform synchronization in spread spectrum systems. Although tracking is somewhat an easy task and is performed with the help of DLL and TDL, i.e., Delay locked loop and Tau-dither loop, these loop algorithms work continuously throughout the communication period possible phases (delays). If both the uncertainties exist, then two-dimensional searches should be initiated [23, 45, 62]. Many acquisition techniques have already been proposed for phase matching. Some of them briefly are reviewed and presented in the subsequent sections.

6.1 Serial Search

This acquisition strategy is simple and easy to implement. In this method, the circuit tries to examine all possible phases one by one (serially); that is why the technique is known as serial search. In this technique, acquisition time is relatively large so this search is also treated as slow acquisition [55] (Fig. 5).

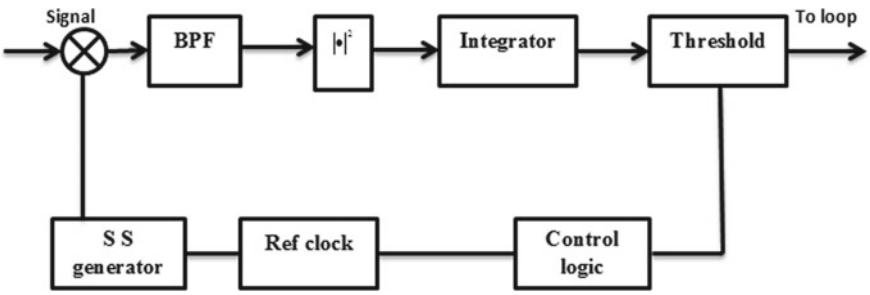


Fig. 5 Serial search scheme

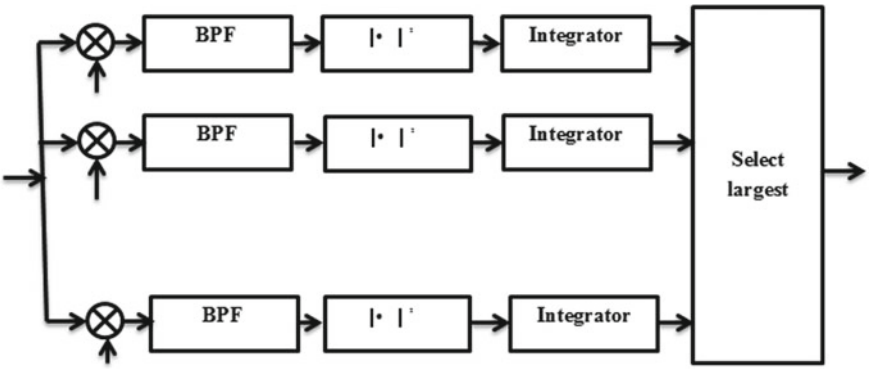


Fig. 6 Parallel search scheme

6.2 Parallel Search

In this approach, circuit test all the possible phases simultaneously, the complexity of the circuit in parallel search is high. The required acquisition time is much smaller than that of the serial search scheme [10] (Fig. 6).

7 Matched Filter Correlator

This technique is also popular for search in acquisition strategy, wherein a matched filter is employed which is matched to the spreading sequence. The impulse response of the matched filter can be written as $b * (T - t)p_T(t)$; here, $b(t)$ is the spreading sequence of period T. Neglecting the presence of noise and if the received signal is $r(t)$, then the square of the magnitude of the matched filter can be written by

$$Y(t) = \frac{1}{T^2}\left|\int_{-\infty}^{+\infty} r(\tau)h(t-\tau)\right|^2 \tag{6}$$

After putting the value of channel response in the above equation, the output can be rewritten as

$$Y = \frac{P}{T^2}\left|r_{b,b}(\zeta - (t-T)\right|^2 \tag{7}$$

By continuously observing the output of the matched filter, different postulated phases can be evaluated. Overall acquisition time requited is very small in this technique but the performance of this technique is mainly limited by frequency variation or uncertainty.

8 Sequential Search

This scheme is very popular and was designed by *Ward* for initial tracking purposes. In this technique, the present state of PN sequence is sequentially estimated with the help of a small number of input bits. The performance of sequential estimation is significantly improved and gives shorter acquisition times for improved SNR.

Other than the basic search algorithms, some more search schemes are also addressed by researchers. A brief review is presented of the search schemes which presents many recent algorithms used for initial code acquisition (Table 3).

9 Sequence Acquisition for Chaos-Based Spreading Sequences

As we know that our main area of interest is chaos-based sequences, the research on sequence acquisition for chaos-based spread spectrum is more recent and very small in volume. The research on chaos-based synchronization was first proposed by Yamada and Fujiska in 1983. As discussed, in traditional CDMA, synchronization is achieved through code acquisition and code tracking. More interestingly, orthogonality between two chaotic signals motivates researchers to investigate the role of traditional synchronization techniques, chaos-based spread spectrum communication systems [11]. Generally, in chaos-based systems, only the code acquisition process is sufficient to achieve code synchronization. Setti et al. suggested the detailed process of code acquisition for chaos-based spread spectrum system.

Many researchers proved that chaotic time series performs better than traditional spreading sequences, [49]. It has also been shown that quantized spreading codes can be generated for any number of users and outperforms classical sequences. The optimally chosen chaotic sequence can perform better in terms of capacity also. DSP processors are used for the implementation of chaos-based spread spectrum system.

Once the acquisition stage has estimated the time difference between the transmitter and receiver to within one-chip duration, the synchronization block enters the tracking mode. The aims of the tracking stage are two fold; first, it is to estimate the time difference between the transmitter and receiver to a time resolution finer than the acquisition phase. As a result, the tracking is sometimes referred to as fine synchronization. The second goal is to maintain time synchronization between the transmitter and receiver when timing jitter, which is a result of clock wander and oscillator discrepancies, is present in the system. Most of the tracking stages in the literature are feedback based and have a loop configuration. As a result, they are often called tracking loops. This is because, they use the carrier tracking phase-locked loops (PLL) theory as the basis of operation. In the continuation of the above discussion, a review table is also presented of a few papers. In this table, we investigate that some of the papers are having both acquisition and tracking. But some

Table 3 Review on initial acquisition techniques

S. no.	Title of paper	Year	Author	System	Initial acquisition
1	Robust synchronization for asynchronous multiuser chaos-based DS-CDMA	2008	Kaddoum et al.	Asynchronous multiuser chaos-based system	Serial search mode
2	A robust sequence synchronization unit for multiuser DS-CDMA chaos-based communication systems	2007	Jovic et al.	Multiuser DS-CDMA based chaotic system	NA
3	A novel acquisition technique of spread spectrum signals using two correct cells jointly	2001	Yoon	DS/SS system	Novel technique based on joint decision rule
4	A new approach to Rapid PN Code Acquisition using Iterative message-passing techniques	2005	Chugg and Zhu	DS/SS system	Iterative message passing algorithm
5	Acquisition of PN sequence in chip synchronous DS/SS system using SPRT	1994	Chawla and Sarwate	DS/SS system	SPRT
6	Code Detection and Acquisition Techniques in the DS/CDMA communication	2009	Barbary et al.	DS/CDMA	Z-search algorithm
7	Combined scheme for fast PN code acquisition	2009	Salah alagooz	DS-WB system	FSSSE Scheme

(continued)

Table 3 (continued)

S. no.	Title of paper	Year	Author	System	Initial acquisition
8	Fast Anti-jamming Timing Acquisition using Multilayer Synchronization Sequence	2013	Zhang et al.	DS/SS system	Algorithm based on BCH codes
9	A Rapid Code Acquisition Scheme in OOC-Based CDMA system	2013	Yoon et al.	OOC- CDMA	Multiple shift algorithm
10	Efficient Acquisition Algorithm for Long Pseudorandom Sequence	2013	Hsieh et al.	DS/SS system	Phase coherence acquisition algorithm
11	Adaptive Acquisition of PN sequence in Nonfading AWGN Channel	2008	Benkrinah et al.	DS/SS system	Adaptive single dwell serial search
12	Adaptive Filter Based PN Code Phase Acquisition Under Frequency Selective Rayleigh Fading Channels	2013	Lee et al.	DS/SS system	Adaptive filter-based search algorithm

researchers suggested the necessity only for the code acquisition part and not for tracking (Table 4).

10 Sequence Tracking for DS-SS Systems

After sequence acquisition schemes, sequence tracking is also an important step in the process to acquire full synchronization. On the availability of carrier phase information, sequence tracking loops can be divided into two categories, viz., coherent and noncoherent loops. For phase discrimination, tracking loops make use of a correlator. One of the popular tracking loop configurations is known as Delay Lock Loop (DLL). This tracking loop uses two independent correlators for the purpose of

Table 4 Review on initial tracking techniques

S. no.	Title of paper	Author/Year	System	Details	Tracking
1	Spread spectrum communication system with sequence synchronization unit using Chaotic Symbolic Dynamics	Kaddoum et al./2013	CBD SS	Time acquisition based on serial search technique (×)	×
2	Code synchronization algorithm based on segment correlation in spread spectrum communication	Li et al./2015	CBD SS	Code acquisition method based on segment (Delay Locked Loop)	√
3	A robust sequence synchronization unit for multiuser DS-CDMA chaos-based communication systems	Jovic et al./2007	CBD SS	PRBS signal is used as a pilot signal used for synchronization (Delay Locked Loop)	√

CBDSS Chaos-based direct sequence spread system

optimum tracking. The block diagram of DLL has also been proposed. Some of the popular DLL configurations are the early-late DLLs, the decision-directed, and the data-aided DLL.

Another popular type tracking loop configuration is early-late gates or Tau-dither Loop (TDL). In TDL, only one correlator is used for the advanced and delayed branches.

Modified and improved versions of DLL and TDL are also available in the literature, namely the double dither loop (DDL), modified code tracking loop (MCTL) and more specifically the product of the sum and difference DLL and the complex sums DLL are more suitable for the fast fading environments.

10.1 Sequence Tracking for Chaos-Based Spreading Sequences

Very small research is available in the field of the tracking of chaos-based spreading sequences. Many of them assume that the tracking of the chaos-based spreading sequence integrated with the acquisition. Recently, some of the papers suggested sequence tracking of the chaos-based spreading sequences.

11 Discussion and Conclusion

In the above discussion, a brief study on the role of chaos in the spread spectrum systems is given. Moreover, in-depth surveys of the literature for CDS-SS acquisition and tracking were also presented in the preceding subsections.

The aim of this paper was to present the literature on chaos theory and possible application of chaotic sequences at the place of PN sequences in DS-SS system. The major shortcoming of the chaos system, i.e., synchronization, is also being discussed.

References

1. H. Abdullah, A. Radhi, M. Majeed, Chaotic multiple access system based on orthogonal chaotic vector of Rossler sequence. J. Sel. Areas Telecommun. (JSAT) **4**(2), 40–45 (2014)
2. K. Abualnaja, E. Mahmoud, Analytical and numerical study of the projective synchronization of the chaotic complex nonlinear systems with uncertain parameters and its applications in secure communication. Math. Probl. Eng. 1–10 (2014) (Hindawi Publishing Corporation)
3. Z. Ali, Q. Memon, Time delay tracking for multiuser synchronization in CDMA networks. J. Netw. **8**(9), 1929–1935 (2013)
4. S. Azou, G. Burel, C. Pistre, A chaotic direct-sequence spread-spectrum system for underwater communication. IEEE Oceans **4**, 2409–2415 (2002)
5. J. Baek, J. Park, Y. Lee, S. Kim, G. Jee, J. Yang, S. Yoon, Low complexity long PN code acquisition scheme for spread spectrum systems, in *EMERGING 2011: The Third International Conference on Emerging Network Intelligence* (2011), pp. 25–28
6. S. Benkrinah, M. Barkat, M. Benslama, A. Benmeddour, R. Bekhakhecha, Adaptive acquisition of PN sequence in nonfading AWGN channel. Afr. Phys. Rev. **2**(Special Issue (Microelectronics): 0059), 120–122 (2008)
7. K. Berbra, M. Barkat, A. Anou, PN Code acquisition using smart antenna and adaptive thresholding CFAR based on ordered data variability for CDMA communications. Progr. Electromagn. Res. B **57**, 139–155 (2014)
8. R. Candido, D. Soriano, M. Silva, M. Eisencraft, Do chaos-based communication systems really transmit chaotic signals. Signal Process. **108**, 412–420. ISSN 0165-1684 (Elsevier)
9. K. Chae, S. Yoon, A rapid code acquisition scheme in OOC-based CDMA systems. Int. J. Electr. Comput. Energ. Electron. Commun. Eng. **9**(4), 407–410 (2015)
10. K. Chawla, D. Sarwate. Acquisition of PN sequence in chip synchronous DS/SS systems using a random sequence model and the SPRT. IEEE Trans. Commun. **42**(6), 2325–2334 (1994). ISSN 0090-6778
11. C. Chen, K. Yao, K. Umeno, E. Biglieri, Design of spread-spectrum sequences using chaotic dynamical systems and ergodic theory. IEEE Trans. Circuits Syst. I: Fundam. Theory Appl. **48**(9), 1110–1114 (2001). ISSN 1057-7122
12. D. Chen, X. Huang, T. Ren, Study on chaotic fault tolerant synchronization control based on adaptive observer. Sci. World J. 1–5 (2014) (Hindawi Publishing Corporation)
13. T. Chien, T. Liao, Design of secure digital communication systems using chaotic modulation, cryptography and chaotic synchronization. Chaos Solitons Fractals **24**, 241–255 (2005) (Elsevier)
14. K. Chugg, M. Zhu, A new approach to rapid PN code acquisition using iterative message passing techniques. IEEE J. Sel. Areas Commun. **23**(5), 884–897 (2005). ISSN 0733-8716
15. N. Corron, D. Hahs, A new approach to communications using chaotic signals. IEEE Trans. Circuits Syst. I: Fundam. Theory Appl. **44**(5), 373–382 (1997). ISSN 1057-7122

16. S. Elagooz, Combined scheme for fast PN code acquisition, in *13th International Conference on Aerospace Sciences & Aviation Technology, ASAT-13* (2009), pp. 1–8
17. K. El-Barbary, E. ElWaniss, A. Abd El Azziz, Code detection and acquisition techniques in the DS/CDMA communications, in *13th International Conference on Aerospace Sciences & Aviation Technology, ASAT-13* (2009), pp. 1–11
18. M. El-Dessoky, M. Yassen, Adaptive feedback control for chaos control and synchronization for new chaotic dynamical system. Math. Probl. Eng. (2012) (Hindawi Publishing Corporation)
19. Z. Elhadj, J. Sprott, Some open problems in chaos theory and dynamics. Int. J. Open Probl. Comput. Math. **4**(2), 1–10 (2011). ISSN 1998-6262
20. R. Gaudenzi, M. Luise, Decision-directed coherent delay-lock tracking loop for DS-spread-spectrum signals. IEEE Trans. Commun. **39**(5), 758–765 (1991)
21. L. Guo, Z. Sun, S. Chen, B. Wang, X. Ning, Research on time-frequency domain acquisition algorithm of parallel combinatory spread spectrum system based on FFT, in *International Conference on Computer Information Systems and Industrial Applications* (2015)
22. W. Hai, H. Jiandong, Chaotic spread-spectrum communication using discrete-time synchronization. J. China Univ. Posts Telecommun. **4**(1), 65–70
23. A. Hassani, M. Zouak, M. Mrabti, Contribution to synchronization and tracking modelisation in a CDMA receiver. J. Eng. 1–7 (2013) (Hindawi Publishing Corporation)
24. G. Heidari-Bateni, C. McGillem, A chaotic direct-sequence spread-spectrum communication system. IEEE Trans. Commun. **42**(2/3/4), 1524–1527(1994). ISSN 0090-6778
25. W. Hsieh, C. Chang, M. Kao, Efficient acquisition algorithm for long pseudorandom sequence. IEEE Trans. Aerosp. Electron. Syst. **50**(3), 1786–1797. ISSN 0018-9251
26. W. Jibrail, H. Al-Zubiady, Search strategies for the acquisition of DS spread spectrum signals. Int. J. Electron. **84**(2), 83–104 (1998). ISSN 0020-7217
27. B. Jovic, C. Unsworth, G. Sandhu, S. Berber, A robust sequence synchronization unit for multi-user DS-CDMA chaos-based communication systems. Signal Process. (2007). ISSN 0165-1684 (Elsevier)
28. G. Kaddoum, P. Chargé, D. Roviras, D. Fournier-Prunaret, Comparison of chaotic sequences in a chaos-based DS-CDMA system, in *International Symposium on Nonlinear Theory and its Applications NOLTA'07*, Vancouver, Canada (2007), pp. 212–215
29. G. Kaddoum, G. Gangan, F. Gagon, Spread spectrum communication system with sequence synchronization unit using chaotic symbolic dynamics modulation. Int. J. Bifurc. Chaos **23**(2), 1–14 (2013)
30. G. Kaddoum, D. Roviras, P. Charge, D. Fournier-Prunaret, Robust synchronization for asynchronous multi-user chaos-based DS-CDMA. Signal Process. **89**, 807–818 (2009) (Elsevier)
31. A. Kiani-B, K. Fallahi, N. Pariz, H. Leung, A chaotic secure communication scheme using fractional chaotic systems based on an extended fractional Kalman filter. Commun. Nonlinear Sci. Numer. Simul. 863–879 (2009). ISSN 1007-5704 (Elsevier)
32. G. Kolumban, M. Kennedy, L. Chua, The role of synchronization in digital communications using chaos—part ii: chaotic modulation and chaotic synchronization. IEEE Trans. Circuits Syst. I: Fundam. Theory Appl. **45**(11), 1129–1140 (1994). ISSN 1057-7122
33. P. Kumar, A. Sampath, P. Indumathi, Improving security of communication systems using CHAOS. Indian J. Sci. Technol. **4**(5), 561–565 (2011). ISSN 0974-6846
34. A. Kurian, S. Puthusserypady, S. Htut, Performance enhancement of DS/CDMA system using chaotic complex spreading sequence. IEEE Trans. Wirel. Commun. **4**(3), 984–989 (2005). ISSN 1536-1276
35. F. Lau, C. Tse, Performance of chaos-based communication systems under the influence of coexisting conventional spread-spectrum systems. IEEE Trans. Circuits Syst. I: Fundam. Theory Appl. **50**(11), 1475–1481 (2003). ISSN 1057-7122
36. A. Li, Z. Yang, R. Qi, F. Zhou, G. Han, Code synchronization algorithm based on segment correlation in spread spectrum communication. Algorithms **2015**(8), 870–894 (2015)
37. S. Li, G. Alvarez, Z. Li, W. Halang, Analog chaos-based secure communications and cryptanalysis: a brief survey, in *PhysCon'2007*, Potsdam, Germany (2007)

38. S. Li, H. Lian, Y. Zhao, A. Wu, Hyperchaotic spread spectrum sequences selection and its application in DS-CDMA system 1–8. ISSN 0021-3470
39. D. López-Mancilla, C. Cruz-Hernández, A note on chaos-based communication schemes. Revista Mexicana De Física **51**(3), 265–269 (2005)
40. E. Lorenz, Deterministic nonperiodic flow. J. Atmos. Sci. **20**, 130–141 (1963)
41. M. Mahmood, B. Nathem, Performance evaluation of DS-CDMA system based orthogonal chaotic vectors over Rayleigh fading channel. Int. J. Comput. Appl. **114**(3), 8–14 (2015). ISSN 0975-8887
42. H. Mansour, Y. Fu, A new method for generating a zero mean self-balanced orthogonal chaotic spread spectrum codes. Int. J. Hybrid Inf. Technol. **7**(3), 345–354 (2014)
43. J. Mata-Machuca, R. Martinez-Guerra, R. Aguilar-Lopez, C. Aguilar-Ibanez, A chaotic system in synchronization and secure communications. Commun. Nonlinear Sci. Numer. Simul. 1706–1713 (2012)
44. Z. Ma, T. Chen, M. Zhang, P. Kecerski, S. Dang, Literature review of spread spectrum signaling: performance, applications and implementation. J. Commun. **10**(12), 932–938 (2015)
45. G. Mazzini, R. Rovatti, G. Setti, Sequence synchronization in chaos-based DS-CDMA systems (1998), pp. 485–488
46. G. Mazzini, G. Setti, R. Rovatti, Chaotic Complex spreading sequences for asynchronous DS-CDMA—part I: system modeling and results. IEEE Trans. Circuits Syst. I: Fundam. Theory Appl. **44**(10), 937–947 (1997). ISSN 1057-7122
47. H. Ong, S. Tilahun, S. Tang, A comparative study on standard, modified and chaotic firefly algorithms. Pertanika J. Sci. Technol. **23**(2), 251–269 (2015). ISSN 0128-7680
48. L. Pecora, Carroll, Synchronization in chaotic systems. Phys. Rev. Lett. **64**(8), 821–824 (1990)
49. N. Quyen, V. Yem, T. Hoang, A chaos-based secure direct-sequence/spread-spectrum communication system. Abstr. Appl. Anal. 1–11 (2013) (Hindawi Publishing Corporation)
50. C. Silva, A. Young, Introduction to chaos-based communications and signal processing, in *2000 IEEE Aerospace Conference Reprint* (2000), pp. 1–21
51. R. Takahashi, K. Umeno, Performance evaluation of CDMA using chaotic spreading sequence with constant power in indoor power line fading channels. IEICE Trans. Fundam. **97**(A), 1619–1622 (2014)
52. R. Vali, S. Berber, S. Nguang, Accurate derivation of chaos-based acquisition performance in a fading channel. IEEE Trans. Wirel. Commun. **11**(2), 722–731 (2012). ISSN 1536-1276
53. R. Vali, S. Berber, S. Nguang, Analysis of chaos-based code tracking using chaotic correlation statistics. IEEE Trans. Circuits Syst. **59**(4), 796–805 (2012)
54. C. Wang, Y. He, J. Ma, L. Huang, parameters estimation, mixed synchronization, and antisynchronization in chaotic systems. Wiley Online Libr. 64–73 (2013)
55. R. Warsi, A. Chaturvedi, A New Adaptive Serial Search PN Code Acquisition Scheme for DS-CDMA Systems (2000), pp. 245–248
56. S. Won, L. Hanzo, *Initial and Post-Initial Acquisition in the Serial Search Based Noncoherent Multiple Transmit/Receive Antenna Aided DS-CDMA Downlink* (IEEE, 2006). ISBN 7803-9392
57. W. Xu, L. Wang, G. Chen, Performance of DCSK cooperative communication systems over multipath fading channels. IEEE Trans. Circuits Syst. **58**(1), 196–204 (2011). ISSN 1549-8328
58. W. Xu, L. Wang, G. Kolumban, A novel differential chaos shift keying modulation scheme. Int. J. Bifurc. Chaos **21**(3), 799–814 (2011)
59. R. Yamapi, G. Filatrella, M. Aziz-Alaoui, H. Enjieu Kadji, Modeling, stability, synchronization, and chaos and their applications to complex systems. Abstr. Appl. Anal. 1–2 (2014) (Hindawi Publishing Corporation)
60. L. Yang, L. Hanzo, Acquisition of m-sequences using recursive soft sequential estimation. IEEE Trans. Commun. **52**(2), 199–204 (2004). ISSN 0090-6778
61. T. Yang, A Survey of chaotic secure communication system. Int. J. Comput. Cogn. **2**(2), 81–130 (2004). ISSN 1542-5908
62. S. Yoon, I. Song, H. Kwon, C. Park, *A Novel Acquisition Technique of Spread Spectrum Signals Using Two Correct Cells Jointly* (IEEE, 2001). ISBN 7803-7227

63. T. Youssef, M. Chadli, H. Karimi, M. Zelmat, Chaos synchronization based on unknown input proportional multiple-integral fuzzy observer. Abstr. Appl. Anal. 1–11 (2013) (Hindawi Publishing Corporation)
64. J. Zhang, N. Ge, Z. Wang, S. Chen, Fast antijamming timing acquisition using multilayer synchronization sequence. IEEE Trans. Veh. Technol. **62**(7), 3497–3503 (2013). ISSN 0018-9545

Design Analysis of CG-CS LNA for Wideband Applications Using Noise Cancelation Technique



Dheeraj Kalra, Devendra Kumar and Divesh Kumar

1 Introduction

Today's requirement for a communication device is to be compact, less power consumption, good performance, [1] etc. The first block in a super heterodyne receiver is low noise amplifier; amplify the signal with low inclusion of noise. LNA parameters such as the gain, NF, wideband matching, and power consumption are to be adjusted to get its good performance. Various techniques such as source degeneration, resistive feedback, and common gate–common source (CG-CS) topology [2] are employed to improve these parameters. Different topologies have their different advantages in improving the LNA parameters. Source degeneration, connecting the inductor at the source , provides the linearity by matching the parasitic capacitance of MOSFETs but increases the size of circuit. Common gate topology has high noise figure at moderate power consumption but has better impedance matching and linearity. CG-CS topology helps in improving the matching, gain, improving the SNR, and lowering the NF. Resistive feedback topology increases the bandwidth, better impedance matching, good linearity, low NF, and lowering the gain [3]. Cascode topology is advantageous in increasing the gain, lowering noise figure, good linearity and increaseing the power consumption. So trade-off among the LNA parameters make difficult to design the circuit. Various design topologies are merged to get good performance of LNA. LNA [4] uses the current reuse CG topology and helps in reducing the power consumption of the circuit [5]. Current reuse, transistors use the same DC current to reduce the circuit's power consumption. CG topology is used to provide the good impedance matching at the input. MOSFET gain is dependent on transconductance as $A_V = g_m r_d$ where g_m, r_d are the conductance and drain resistance [6–10]. In CG topology, g_m cannot be lowered to decrease the circuit's power consumption [11–14]. In [4], the proposed circuit has not good IIP and NF.

D. Kalra (✉) · D. Kumar · D. Kumar
GLA University, Mathura, India
e-mail: dheeraj.kalra@gla.ac.in

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_128

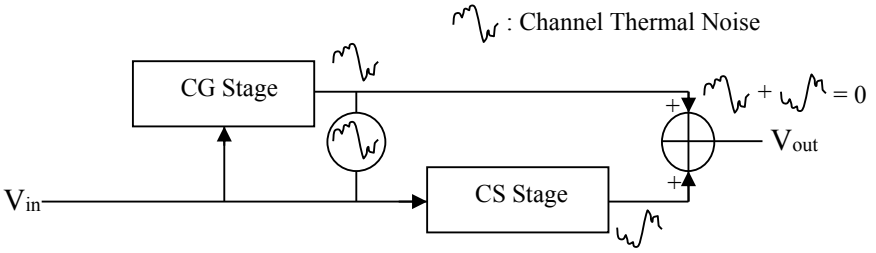


Fig. 1 Noise cancelation technique

In this paper, CG-CS source degeneration topology is implemented to get the good gain and low noise figure which has good impedance matching.

2 Circuit Analysis

Noise cancelation technique is shown in Fig. 1, thermal noise generated in the channel of CG circuits will be available at CG and CS stages. CS stage provides the 180° phase difference and then two noises are added together, so net noise at the output is zero, hence improving the noise figure.

The circuit is shown in Fig. 2, CS stage composed of two-stage amplifier, first stage consists of M_3 MOS and second-stage consists of two MOS M_4 and M_5.

Transistor M_1 is used in common gate configuration that provides good input impedance matching. Transistor M_2 acts as current source for providing the biasing to the circuits. Transistor M_3 is connected in common source configuration which provides the gain to the circuit.

The overall gain of the designed circuit is given by

$$G_V^T = G_V^{\mathrm{CS}} + G_V^{\mathrm{CG}} \tag{1}$$

where G_V^{CS} and G_V^{CG} represent the individual gain of CS and CG path.

3 Simulation Results

Designed circuit is simulated in CMOS 0.18 μm technology for the frequency range from 1 to 5 GHz. Simulation results for S21 versus RF frequency are shown in Fig. 3. Maximum achieved value for S21 is 10.756 dB at 1 GHz, in the entire frequency range 1 to 5 GHz, gain is positive.

Simulation result for S11, S22, and S12 are shown in Fig. 4. Input reflection coefficient, output reflection coefficient, and reverse transfer gain are negative throughout the entire range of RF frequency.

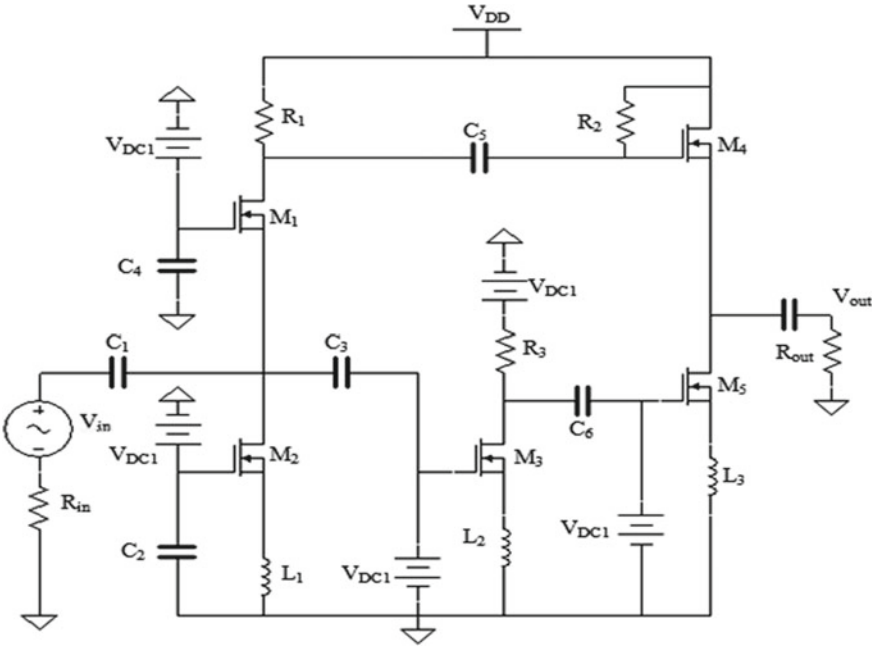


Fig. 2 Low noise amplifier circuit

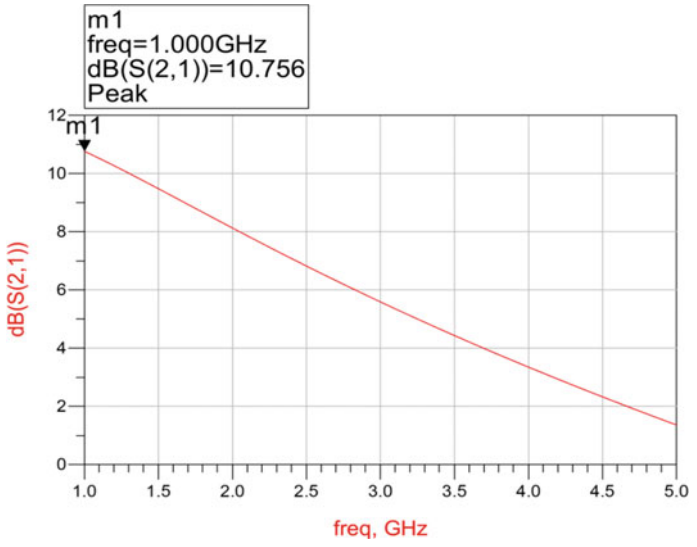


Fig. 3 Simulation for S21 versus RF frequency

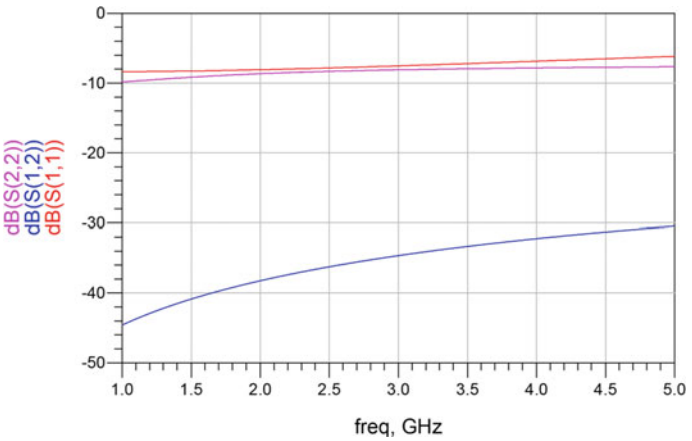


Fig. 4 Simulations for S11, S22, and S21 versus RF frequency

Simulation result for noise figure of the designed circuit is shown in Fig. 5. The minimum value of Noise Figure achieved is 2.962 dB at 1 GHz and the maximum value is 4.898 dB at 5 GHz.

The power consumption of the entire circuit is 8.111 mW which shows good in terms of battery life of the circuit.

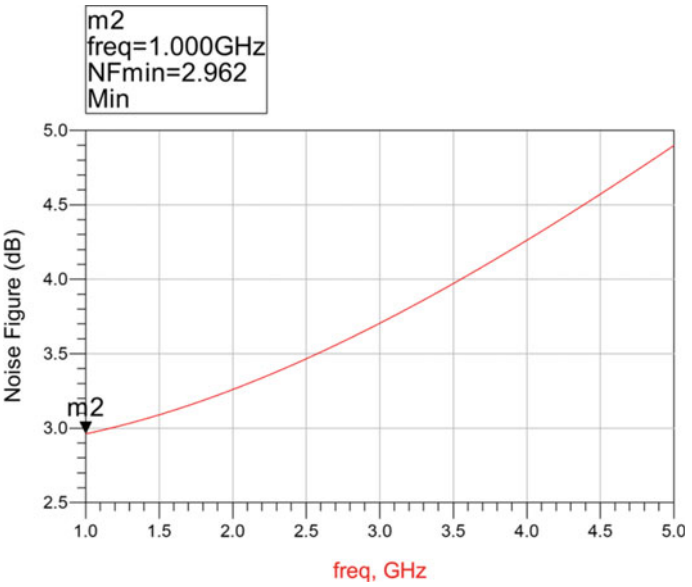


Fig. 5 Simulation for noise figure versus RF frequency

4 Conclusion

The CS and CG topologies are used to provide enough gain and linearity. The designed circuit is simulated in TSMC 0.18 μm technology. The designed circuit shows the maximum value of S21 as 10.756 dB at 1 GHz, in the entire frequency range 1–5 GHz, the gain is positive. The minimum value of Noise Figure achieved is 2.962 dB at 1 GHz and the maximum value is 4.898 dB at 5 GHz. The power consumption of the entire circuit is 8.111 mW.

References

1. A. Balankutty, S.A. Yu, Y. Feng, P.R. Kinget, A 0.6-V zero-IF/low-IF receiver with integrated fractional-N synthesizer for 2.4-GHz ISM-band applications. IEEE J. Solid-State Circ. **45**(3), 538–553 (2010)
2. H.C. Chen, T. Wang, H.W. Chiu, T.H. Kao, S.S. Lu, 0.5-V 5.6-GHz CMOS receiver subsystem. IEEE Trans. Microw. Theory Tech. **57**(2), 329–335 (2009)
3. A. Balankutty, P.R. Kinget, An ultra-low voltage, low-noise, high linearity 900-MHz receiver with digitally calibrated in-band feed-forward interferer cancellation in 65-nm CMOS. IEEE J. Solid-State Circ. **46**(10), 2268–2283 (2011)
4. A. Shameli, P. Heydari, A novel power optimization technique for ultra-low power RFICs, in *Proceedings of the 2006 International Symposium on Low Power Electronics and Design* (ACM, 2006), pp. 274–279
5. J.-F. Chang, Y.-S. Lin, 0.99 mW 3–10 GHz common-gate CMOS UWB LNA using T-match input network and self-body-bias technique. Electron. Lett. **47**(11), 658–659 (2011)
6. M. Parvizi, K. Allidina, M.N. El-Gamal, A sub-mW, ultra-low-voltage, wideband low-noise amplifier design technique. IEEE Trans Very Large Scale Integr. (VLSI) Syst. **1**(6), 1111–1122 (2015)
7. M. Parvizi, K. Allidina, M.N. El-Gamal, An ultra-low-power wideband inductorless CMOS LNA with tunable active shunt-feedback. IEEE Trans. Microw. Theory Tech. **64**(6), 1843–1853 (2016)
8. F. Bruccoleri, E.A.M. Klumperink, B. Nauta, Wide-band CMOS low-noise amplifier exploiting thermal noise canceling. IEEE J. Solid-State Circ. **39**(2), 275–282 (2004)
9. W.-H. Chen et al., A highly linear broadband CMOS LNA employing noise and distortion cancellation, in *2007 IEEE Radio Frequency Integrated Circuits (RFIC) Symposium* (IEEE, 2007)
10. J. Shim, T. Yang, J. Jeong, Design of low power CMOS ultra wide band low noise amplifier using noise canceling technique. Microelectron. J. **44**(9), 821–826 (2013)
11. S.C. Blaakmeer et al., Wideband balun-LNA with simultaneous output balancing, noise-canceling and distortion-canceling. IEEE J. Solid-State Circ. **43**(6), 1341–1350 (2008)
12. Y. Yu et al., Analysis and design of inductorless wideband low-noise amplifier with noise cancellation technique. IEEE Access **5**, 9389–9397 (2017)
13. L. Wu, H.F. Leung, H.C. Luong, Design and analysis of CMOS LNAs with transformer feedback for wideband input matching and noise cancellation. IEEE Trans. Circ. Syst. I Regul. Pap. **64**(6), 1626–1635 (2017)
14. H.-T. Chou, S.-W. Chen, H.-K. Chiou, A low-power wideband dual-feedback LNA exploiting the gate-inductive bandwidth/gain-enhancement technique, in *2013 IEEE MTT-S International Microwave Symposium Digest (MTT)* (IEEE, 2013)

Performance Evaluation of Hybrid Renewable Energy System for Supplying Electricity to an Institution and a Hospital Using HOMER



Ingudam Chitrasen Meitei, Amit Kumar Irungbam and Benjamin A. Shimray

1 Introduction

Electrical energy plays an important role in initializing a development process and further helps in maintaining the development of a country. Electrical energy acts as the main source of energy utilization purposes [1]. The challenges that we encounter with the electrical system is the generation, transmission, and distribution. The challenging thing is to set up the transmission line from the generating station to the consumer in a remote area due to high expenses. Therefore, Hybrid Renewable Energy System (HRES) is a better option for providing an economical solution to supply power for the remote areas. HRES comprises generation of power from renewable energy resources like solar, micro-hydro, wind, geothermal, tidal energy, ocean waves, biomass, and other fuel cells. In the present scenario, the total installed power capacity of renewable energy (excluding large hydro) is 71.325 GW as of June 30, 2018, which is just 20% of the total installed power capacity of India. According to the record of March 31, 2018, the installed capacity for Large Hydro was 45.29 GW, contributing around 13% of the total power; this is comparatively a low number as compared with the total availability of renewable resources in the country [2–7].

The main disadvantages of the off-grid system consisting of wind energy and solar energy are the energy variation that results in the irregular supply of electricity. This causes a problem if we require a reliable and continuous supply of power. Therefore, we need to set up a system consisting of a primary renewable resource in parallel with storage units and a standby nonrenewable source [8].

I. C. Meitei · A. K. Irungbam · B. A. Shimray (✉)
Electrical Engineering Department, National Institute of Technology Manipur, Imphal, India
e-mail: benjaminshimray@gmail.com

I. C. Meitei
e-mail: alwayschitrasen@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_129

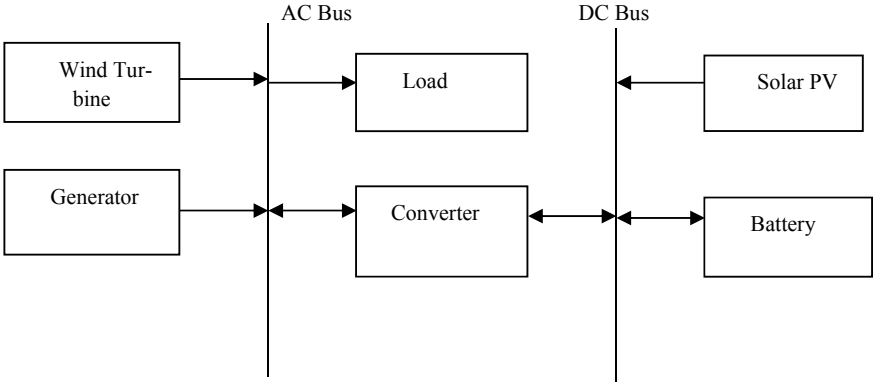


Fig. 1 Schematic diagram of the HRES model

The major advantage of an HRES is that power sources can compensate each other when one of the sources is at a lower level. On a cloudy, windy day, solar PV produces less energy and thereby becomes less while wind turbine generator produces a higher level of energy during such weather conditions. Similarly, for wind energy generator the main problem lies with the site location, where the wind energy availability is limited on a regular basis [8]. The major use of nonconventional energy decreases the price of energy for the long-term basis, and diesel generator combination provides the backup for the system in emergency cases like low renewable power availability or high loads. Based on the operation of power frequency, microgrids are classified into DC, AC, and Hybrid Microgrids [9–12]. The schematic diagram of the HRES model is given in Fig. 1.

In this paper, optimization of HRES is performed by using HOMER Energy Software. First, estimate the load consumption by N.I.T Manipur and Shija Hospital located at Langol, Manipur whose latitude 24°50′33.20″ N longitude 93°55′28.82″ E [13] (Fig. 2) is obtained for the construction of hybrid renewable energy microgrid model [5]. Then, HOMER is used by considering small wind turbines, solar PV, batteries, and diesel generator to study the optimization of the model to achieve the minimum operational cost of the system and the final result is presented [1–7, 13–15].

2 Methodology

The methodology flowchart is given in Fig. 3, which starts with load profile. Then the resource availability of the area is obtained. The required components and sensitivity cases are considered. All the collected data (electric load, resources available, components detail, sensitivity cases) are loaded in HOMER Software. Thereafter

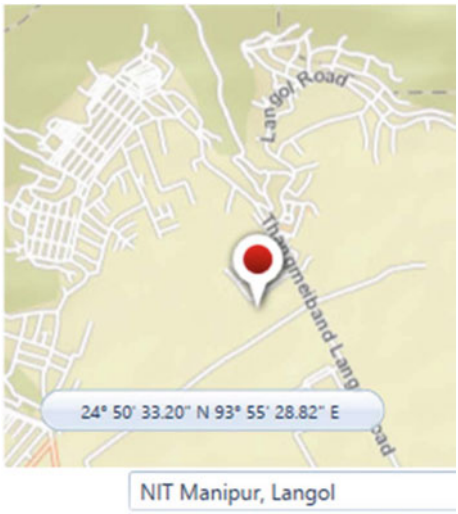


Fig. 2 A geographical view of Langol, Manipur

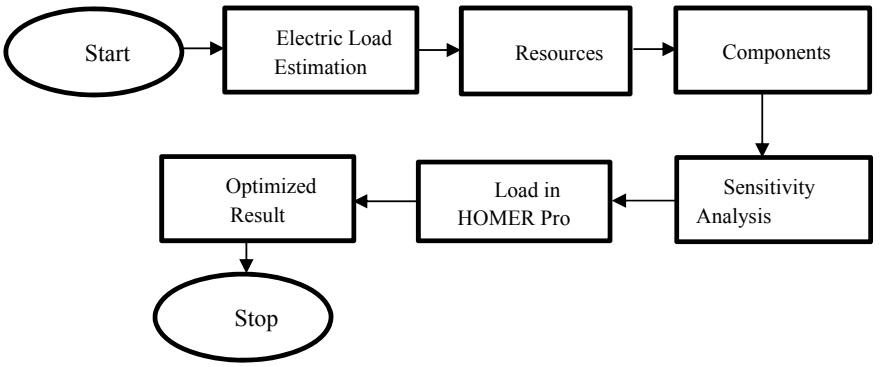


Fig. 3 Methodology flowchart

using HOMER, the simulation is performed to achieve the final optimized result. The factors which we focus on the optimization are the COE and NPC [1–7, 13–15].

3 Load Profile and Site Selection

The proposed model is used to power the National Institute of Technology Manipur and Shija Hospital which are located at the same location in Manipur. The selected area represents reputed institution and hospital of the state. The area is situated at latitude 24°50′33.20″ N longitude 93°55′28.82″ E. In our study, the electrical load requirement of the institute and hospital are obtained from Manipur State Power

Table 1 Load profile

Sl. no	Name	Daily demand (kWh/d)	Peak demand (kW)
1.	Shija hospital	2639	402.04
2.	N.I.T Manipur	1210	215.38
Total average demand		3849 kWh/d	

Distribution Company Ltd (MSPDCL). Almost all the power that we used today in Manipur are bought from the neighboring states. Manipur government spends more than Rs. 700 crore (approx) annually [4] in order to supply the demands of power of approximately 800–900 Mega Units annually [3], and most of the hospitals, institutions, factories, etc. rely on backup diesel generator in order to have uninterrupted power. Table 1 shows the load profile of the institute and hospital. Thus, the two sites are selected for optimal sizing of the HRES.

Here, Fig. 4a shows the hourly demand of Shija Hospital, in which the maximum demand can be found during 11.00–13.00 h and 19.00–21.00 h and Fig. 4b shows the hourly demand of N.I.T Manipur, in which the peak demand can be found during 08.00–17.00 h.

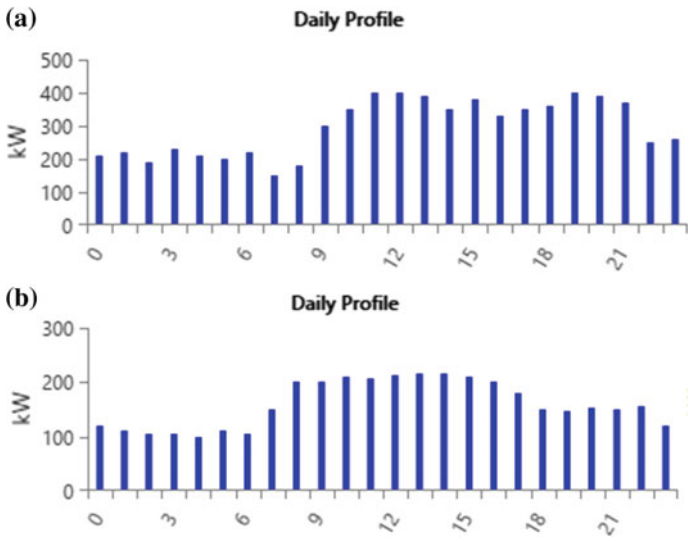


Fig. 4 The average hourly load profile of **a** Shija hospital, **b** N.I.T Manipur

4 Components and Resources Available

The following components and resources are considered for the optimization process in HOMER Software. And the diagram for HOMER design for the HRES is presented in Fig. 5.

4.1 Photovoltaic Panels

After proper research, the PV panels of loom solar 340 W Monocrystalline Panel is considered as it has a lower cost of Rs. 36,000/kW. The cost of replacement is taken as Rs. 36,000, the operation and maintenance (O and M) cost is taken as Rs. 10,000/year. Also, the derating factor is considered as 80% and the lifetime is taken as 20 years.

4.2 Small Wind Turbine

The small wind turbine is defined as the turbines which are capable to generate less than 100 kW at rated speed. The model of a generic 1 kW wind turbine is considered. The cost of capital and replacement is taken as Rs. 50,000/kW and it's O and M cost is taken as Rs. 10,000/kW/year. The lifetime is considered as 20 years.

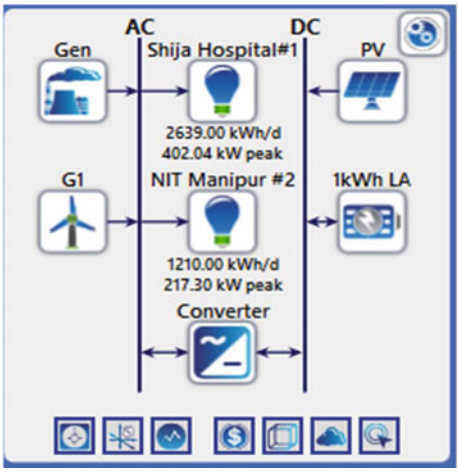


Fig. 5 Homer design for the HRES

Table 2 Properties of generic 1 kWh LA battery

Sl. no	Parameters	Values
1.	Nominal voltage	12 V
2.	Lifetime throughput	800 kWh
3.	Round trip efficiency	80%
4.	Maximum discharge current	24.33 A
5.	Maximum charging current	16.67 A

4.3 Diesel Generator

A minimum load ratio of 25% is considered. Here in this paper, a three-phase 220/440 V Kirloskar diesel generator is considered which has price of Rs. 6,000/kW. Also, the cost of replacement is taken as Rs. 6,000/kW and its O and M cost is taken as Rs. 0.5/h. The generator lifetime is considered as 44,000 h.

4.4 Battery

The considered battery is Lead Acid (LA) by Amaron HCV 620D31R Hiway 12 V/80 Ah which has a capital cost of Rs. 6,300 per battery and its replacement cost is taken as Rs. 6,300 per battery. The operation and maintenance are considered as Rs. 1,000 per battery and its lifetime is taken as 8 years. This battery matches the battery input used in HOMER Software which is 1 kWh lead acid battery (Table 2).

4.5 Converter

The converter is used for maintaining the energy flow between the AC and DC power system components. The capital and replacement cost of the considered converter is taken as Rs. 9,000. The operation and maintenance cost is considered as Rs. 500/year and its lifetime as 15 years. Also, efficiency is considered as 95%.

4.6 Solar Energy Resources

The solar radiation of the project site latitude 24°50′33.20″ N longitude 93°55′28.82″ E [13] is acquired from the NASA surface meteorology and solar energy database [6]. The acquired solar radiation data is shown in Fig. 6, in which the left axis gives the average daily solar radiation data, the right axis gives the solar radiation clearness index values. The clearness index indicates the fraction of the solar insolation that is

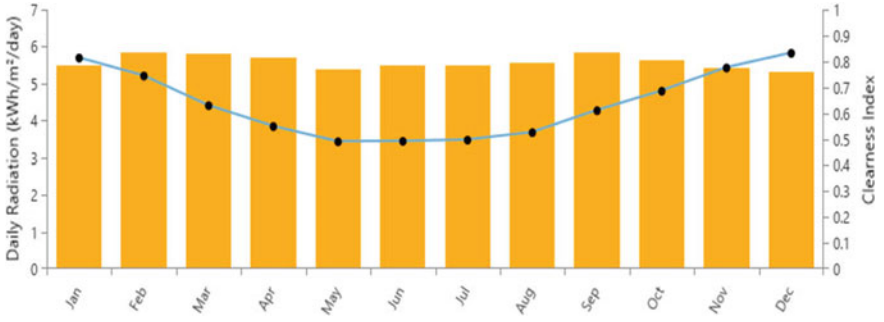


Fig. 6 Solar radiation data

transmitted through the atmosphere to strike the earth's surface; it is dimensionless which ranges between 0 and 1. The average solar radiation of the location is 5.58 kWh/m^2/day [16].

4.7 Wind Resources

The wind resource data of the location is acquired from the NASA surface meteorology and solar energy database [6]. The average wind speed from the database is shown in Fig. 7. The average wind speed of the project location is indicated on the left axis. The annual average wind speed of the location from the database is 4.2 m/s.



Fig. 7 Wind speed data

5 Results and Discussion

The HOMER Software performs a total of 1,522,542 simulation solution out of which 1,055,608 were feasible and 466,608 simulation results are not feasible due to the capacity storage constraints (Fig. 8).

After the simulation process by HOMER Software, the best optimal results are the design consisting of the components solar PV, diesel generator, and battery. Out of the total generation, the solar PV generates 83.3% of energy and the diesel generator generates 16.7% (Table 3).

The final optimization result of the simulation is given in Table 4. The optimized

Fig. 8 Calculation report

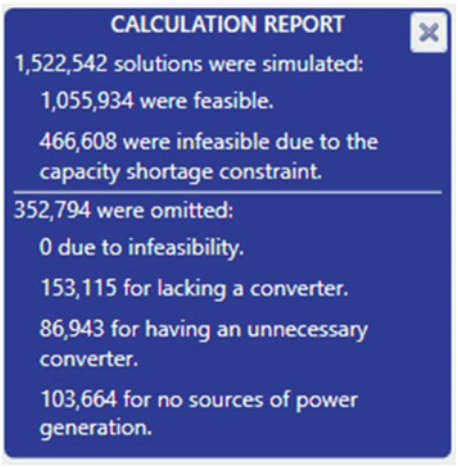


Table 3 Total energy generation in a year

Production	kWh/yr	%
Generic flat plate PV	1,420,874	83.3
Autosize genset	285,197	16.7
Total	1,706,072	100

Table 4 Optimization result

Architecture										Cost				System	
					PV (kW)	G1	Gen (kW)	1kWh LA	Converter (kW)	Dispatch	NPC (₹)	COE (₹)	Operating cost (₹/yr)	Initial capital (₹)	Ren Frac (%)	Total Fuel (L/yr)
					758		680	2,655	503	LF	₹295M	₹16.25	₹18.8M	₹52.6M	79.7	86,302
					762	2	680	2,760	490	LF	₹296M	₹16.28	₹18.7M	₹53.4M	80.4	83,199
							680	1,066	109	CC	₹361M	₹19.90	₹27.0M	₹11.8M	0	408,883
						1	680	1,083	107	CC	₹362M	₹19.90	₹27.0M	₹11.9M	0	408,494
					148		680		125	CC	₹440M	₹24.23	₹33.2M	₹10.5M	0	508,615
					170	3	680		135	CC	₹441M	₹24.27	₹33.2M	₹11.5M	0	503,298

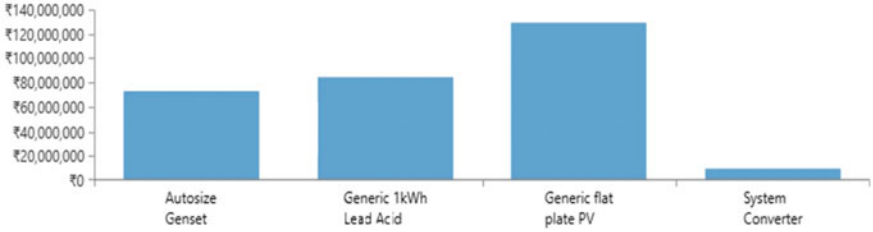


Fig. 9 Cost summary

result consists of the system with 758 kW of PV, 680 kW of diesel generator, 503 kW of converter, and 2655 numbers of 1 kWh LA battery.

The various components cost of the model is given in Fig. 9. Here, it is observed that the highest and the second highest total cost are occupied by the solar PV and the battery set, respectively.

6 Conclusion

This paper presents the simulation result to find the optimum size of the HRES components for the location of Langol, Manipur using HOMER Software. From the result, we get a clear idea of total Net Present Cost (NPC) of Rs. 295,000,000 and the Cost of Energy (COE) of Rs. 16.25 per kWh.

The important factor that can be drawn from this design of HRES is that for this particular location the usage of the wind turbine is not very advantageous as the NPC and COE increase. So, it can be concluded that the best economical solution for the supply of electrical demand of this particular location is by using the HRES components consisting of Solar PV, Battery, and Diesel Generator.

References

1. N.M. Swarnkar, R. Sharma, L. Gidwani, An application of HOMER pro in optimization of hybrid energy system for electrification of technical institute, in *2016 International Conference on Energy Efficient Technologies for Sustainability (ICEETS)*
2. https://en.wikipedia.org/wiki/Renewable_energy_in_India. Accessed 15 Jan 2019
3. http://powermin.nic.in/sites/default/files/uploads/joint_initiative_of_govt_of_india_and_manipur.pdf. Accessed 15 Jan 2019
4. Tariff Order FY 2018–19 w.r.t MSPDCL
5. P. Kumar, R. Pukale, N. Kumabhar, U. Patil, Optimal design configuration using HOMER, in *International Conference on Emerging Trends in Engineering, Science and Technology (ICETEST—2015)*; Proced. Technol. **24**, 499–504 (2016)
6. NASA surface meteorology and solar energy [Online]. https://power.larc.nasa.gov/dataaccess-viewer/

7. C. Phurailatpam, B. Rajpurohit, N. Pindoriya, Embracing microgrids: application for rural and urban India, in *10th National Conference on Indian Energy Sector* (2015)
8. K. Rout, J.K. Sahu, Various optimization techniques of hybrid renewable energy systems for power generation: a review. Int. Res. J. Eng. Technol. (IRJET) **05**(07) (2018)
9. C. Phurailatpam, B.S. Rajpurohit, L. Wang, Optimization of DC microgrid for rural applications in India, in *2016 IEEE Region 10 Conference (TENCON)*
10. C. Phurailatpam, B.S. Rajpurohit, L. Wang, Planning and optimization of autonomous DC microgrids for rural and urban applications in India. Renew. Sustain. Energy Rev. **82**(1), 194–204 (2018)
11. J.J. Justo, F. Mwasilu, J. Lee, J.-W. Jung, AC-microgrids versus DC-microgrids with distributed energy resources: a review. Renew. Sustain. Energy Rev. **24**, 387–405 (2013)
12. P. Estefanía, G.-d.-M. Asier, A. Jon, K. Iñigo, A. Iñigo Martínez de, General aspects, hierarchical controls and droop methods in microgrids: a review. Renew. Sustain. Energy Rev. **17**, 147–159 (2013)
13. H.E LLC, "HOMER", sd http://www.homerenergy.com
14. O.R. Chowdhury, H.-G. Kim, Y. Cho, C. Shin, J. Park, Optimization of a hybrid renewable energy system with HOMER. Adv. Comput. Sci. Ubiquit. Comput. 93–99
15. V. Prema, K. Vma Rao, Sizing of microgrids for Indian systems using HOMER, in *1st IEEE International Conference on Power Electronics, Intelligent Control and Energy Systems (ICPEICES-2016)*
16. N. Rana, M. Chamoli, Cost analysis of solar power generation system using homer optimization software. Int. Res. J. Eng. Technol. (IRJET) **03**(08) (2016)

Energy Scheduling of a Household with Integration of Renewable Energy Considering Different Dynamic Pricing Schemes



Bhukya Balakrishna and Sandeep Kakran

Nomenclature

A	Number of appliances in a home,
T	Scheduling time horizon (in hours),
Z_a^t	Energy (kWh) used by appliance a at hourly time interval t,
E_a	Total energy demand for complete operation of appliance a,
$Y_a^{\max}$	Maximum energy (kWh) consumption by an appliance a in an hour,
$Y_a^{\min}$	Minimum energy (kWh) consumption by an appliance a of in an hour,
$S_{a,t}$	Binary variable to represent ON/OFF time duration of appliance a on complete time horizon T (if $S_{a,t} = 1$ implies ON state, otherwise $S_{a,t} = 0$ implies OFF state),
$y_{a,t}$	Binary variable to represent ON/OFF hour of an appliance a of user,
$E_{m,t}$	Energy (kWh) consumed in hour t of the schedule m,
$E^{\max}$	Maximum energy (kWh) limit fixed by power utility in an hour,
P_t	Cost of electricity (cents/kWh) in tth hour of the day,
Q	Total electricity cost (cents) in time horizon T,
U^t	Energy demand (kWh) of user at hour t, *and*
V^t	Energy generated from rooftop PV of the consumer at tth hour.

B. Balakrishna (✉) · S. Kakran
National Institute of Technology, Kurukshetra 136119, Haryana, India
e-mail: balueee4u@gmail.com

S. Kakran
e-mail: skakran@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_130

1 Introduction

The existing power system has seen many changes in terms of demand, supply, and technology used for smooth operation from the couple of decades. Generation is much dependent on conventional energy resources, as these resources are depleting at higher rates, losses in system are high as the distance from end user to generating station is more, and carbon emission and radiation from the nuclear generating station impact on the lives of people.

To cope up with all abovementioned problems, the existing power system is shifting toward new technology, which encourages distributed generation and smart energy management known as smart grid. This technology helps in making power system cost-effective, more reliable, and more efficient [1]. As the world is heading for sustainable growth, smart technologies and energy management are need of the hour.

Smart grid uses both information technology and communication technology to get data about the behavior of power supply and consumer and to act on it in automated fashion and makes existing network as advanced energy delivery network [2]. It helps power system to be energy efficient during volatile demands by managing distributed generation.

Demand-side management (DSM) is associated with all the activities related to alteration of demand profile in time, shape, or both and also allowing the power system to be active by encouraging distributed generation. Effective application of DSM methods eliminates blackouts, decreases carbon emissions, and reduces operational costs and electricity bills. Demand response (DR) is an important part of DSM. DR is change in end customer's electricity usage with respect to per unit cost over a period of time and under these schemes incentives are given by utility to minimize usage of electricity during high prices or when system is in stress and makes power system efficient. Smart grid deployment is dependent on the design of DR programs [3].

Time-based pricing (dynamic pricing), which includes critical peak pricing (CPP), time-of-use (TOU), and real-time pricing (RTP) programs, is suitable for domestic loads [4]. Dynamic pricings have different prices in different time intervals, more the number of customer participation in the DR program more will be the success of the program.

Author in [5, 6] described about formulation of home energy management using mixed integer linear programming (MILP) with electrical vehicle and energy storage system and author in [7] presented the energy management of a household with the availability of photovoltaic energy and energy storing using neural network.

In this paper, an energy scheduler (ES) is introduced for the scheduling of different appliances and a fixed household load along with a rooftop solar PV of 2 kW. CPP, TOU, and RTP is considered for formulation of the problem. The cost is evaluated for flexible (with customer's preference) loads and fully flexible (without customer's preference) loads. MILP is used in this paper which provides a universal optimized

solution with its constructive structure. Finally, two cases are considered with customer's choice and without customer's choice for all three pricing schemes and successfully plotted the results of energy scheduling.

Rest of the paper is ordered as follows: in Sect. 2, system model is presented and pricing schemes are described. In Sect. 3, objective function is defined and solutions proposed are included. In Sect. 4, case study and results are discussed followed by conclusion in Sect. 5.

2 Energy Scheduling

2.1 System Model

Here "A" number of appliances in a house are considered and these appliances are monitored and scheduled by ES. Twenty-four hours of a day is divided into 24 intervals of 1 h each. This interval set is defined as T = {1, 2, 3 … t}. The ES manages the operating time of appliance. For each appliance, $a \in$ A of particular house energy consumption at hourly interval t is defined below:

$$Z_a \triangleq \left[Z_a^1, Z_a^2, \ldots Z_a^t\right] \tag{1}$$

where Z_a^t represents energy consumption of appliance $\in$ of user at hourly time interval t.

Let E_a is the total energy demand for complete operation of appliance "a", and hence

$$\sum_{t \in T} Z_a^t = E_a \tag{2}$$

Here, we took four different interruptible appliances with different characteristics for energy scheduling. Except these four interruptible appliances, it has been assumed that the household has other fixed load demand. The energy consumption is given by

$$Z_a^t = y_{a,t} Y_a^{\max} + \left(1 - y_{a,t}\right) Y_a^{\min} \quad \forall t \in T \tag{3}$$

where $Y_a^{\min}$ represents minimum energy (kWh) consumption by an appliance a in an hour and $Y_a^{\max}$ represents maximum energy (kWh) consumption by an appliance a in an hour, and $y_{a,t}$ represents binary variable to represent ON/OFF hour of an appliance a of user. Equation (3) gives the information about amount of load that has been consumed by user. When the appliances are ON then load of user will be maximum and when appliance is OFF, minimum energy is consumed by it. The assigned work is completed by an appliance in ON state for certain duration which is given by

$$\sum_{t \in T} (y_{a,t} * S_{a,t}) = k_a \tag{4}$$

$S_{a,t}$ is binary variable represents ON/OFF time duration of appliance $a \in$ A (if $S_{a,t} = 1$ ON time slot is selected, otherwise $S_{a,t} = 0$ implies OFF duration is selected). k_a denotes total slots for appliance to complete its assigned work.

Any household has minimum and maximum energy consumption limits indicated by $Y_a^{\min}$ and $Y_a^{\max}$ representing $\forall\, a \in$ A, and hence this constraint is represented as

$$Y_a^{\min} \leq Z_a^t * S_{a,t} \leq Y_a^{\max} \quad \forall\, S_{a,t} \tag{5}$$

The utility may fix certain amount of energy hourly as limit to reduce stress during peak hours, and this constraint is included below:

$$\Sigma_{a \in A} Z_a^t = E^{\max} \quad \forall\, t \in T \tag{6}$$

where $E^{\max}$ is maximum energy (kWh) limit fixed by the power utility in an hour.

2.2 *Pricing Schemes*

These schemes are generally decided by utility for its users. They are divided into two

- **Static Pricing Scheme and**
- **Dynamic Pricing Scheme.**

In first scheme, price per unit remains same for the day and there will be no change with respect to time. In second scheme, prices changes with respect to time according to market and power system conditions [8]. Dynamic pricing scheme includes CPP, TOU, and RTP schemes.

Real-time pricing: As the name suggests, the pricing is done on real-time usage of electricity. Pricing of this scheme is known 15 min before starting of each time period [9]. Therefore, maximum customers are required for successful implementation of RTP pricing scheme. For industrial and commercial users, this scheme is already applied, and now it is to be extended to residential loads. The prices taken in this paper are from [10].

Time-of-use: This pricing scheme is modification of flat pricing scheme. In this scheme, for different time periods, prices remain constant. The difference may be different days in a week or different hours in particular day. In order to decrease total electricity bill, customer must operate appliances in intervals of low rates. Here rate at peak load is 11.5 cents/kWh, in mid peak loads 9.2 cents/kWh, and in OFF peak loads at rate of 5.5 cents/kWh [11].

Critical peak pricing: This scheme is mostly similar to TOU in fixed pricing in different time periods. Price in CPP has change in one or two periods depending on system stress. CPP prices are announced 1 day before scheduling. CPP prices are more in event days compared to TOU and RTP programs [12].

3 Objective Function and Proposed Solution

Objective of this paper is to design an energy scheduler for a house which can manage the energy supply and operating time of appliances and incorporation of rooftop solar so that the monthly electricity bill is reduced without or less affecting consumer's priority of using the appliance, by considering all above constraints from (1) to (6). Objective function is defined below:

$$Q = \text{minimize} \sum_{t \in T} \sum_{a \in A} Z_a^t * P_t \tag{7}$$

where Q is total cost of electricity of user on time intervals T and P_t is price of electricity in tth hour in a day.

Inclusion of solar energy from rooftop leads to further reduction of electricity bill and objective function Eq. (7) is now changed as given below:

$$Q = \text{minimize} \sum_{t \in T} U^t * P_t \tag{8}$$

Subject to all the constraints from (1) to (6) and

$$U^t \geq \sum_{a \in A} Z_a^t - V^t \quad \forall\, t \in T \tag{9}$$

where V^t is energy generated from rooftop PV of the consumer at tth hour.

The objective function Eq. (8) is linear with integer values and constraints have continuous and binary values. So, this can be solved by MILP technique. Here we divided 24 h into 24 equal intervals. MATLAB and CPLEX solver of GAMS are used for problem modeling.

4 Case Study and Results

Four appliances (vacuum cleaner, dishwasher, toaster, and oven) of a household having average energy consumption per day as 1.48 kWh, 1.8 kWh, 0.55 kWh, and 4.67 kWh are considered for scheduling. As per the user's preference, these appliances are supposed to operate in the following hours as shown in Table 1.

Table 1 Operating hours of household appliances

Appliances	Vacuum cleaner	Dishwasher	Toaster	Oven
Duration of operation	9 a.m.–12 p.m./2 p.m.–8 p.m.	7 p.m.–12 a.m.	6 a.m.–8 a.m.	9 a.m.–12 p.m.

Energy scheduling under RTP with customer's preference and without customer's preference is shown in Figs. 1 and 2. It is observed that energy scheduling for customer preference takes place at 6 a.m., 9 a.m., 10 p.m., 4 p.m., 10 p.m., and 11 p.m. with energies 0.55 kWh, 3.5 kWh, 1.91 kWh, 0.74 kWh, 1.2 kWh, and 0.5 kWh, respectively. Similarly, energy scheduling without customer's preference takes place at 4 a.m., 5 a.m., and 6 a.m. with energies 0.1 kWh, 4.2 kWh, and 4.2 kWh, respectively.

According to above scheduling, electricity bill in a day with customer's choice and without customer's choice is evaluated and shown in Table 2. It can be observed that the electricity bill is less for energy scheduling without customer's preference, because the energy scheduler is free to schedule the appliances during the low-cost hours without customer's preference. It can be clearly seen in Table 2 that the electricity bill changes by 21.29 cents after scheduling the appliances with the customer's preference under RTP.

Energy scheduling for TOU scheme with and without customer's preference is shown in Figs. 3 and 4. With customer's preference, the operating hours are 6 a.m., 9 a.m., 10 a.m., 4 p.m., 10 p.m., and 11 p.m. with energies 0.55 kWh, 3.5 kWh,

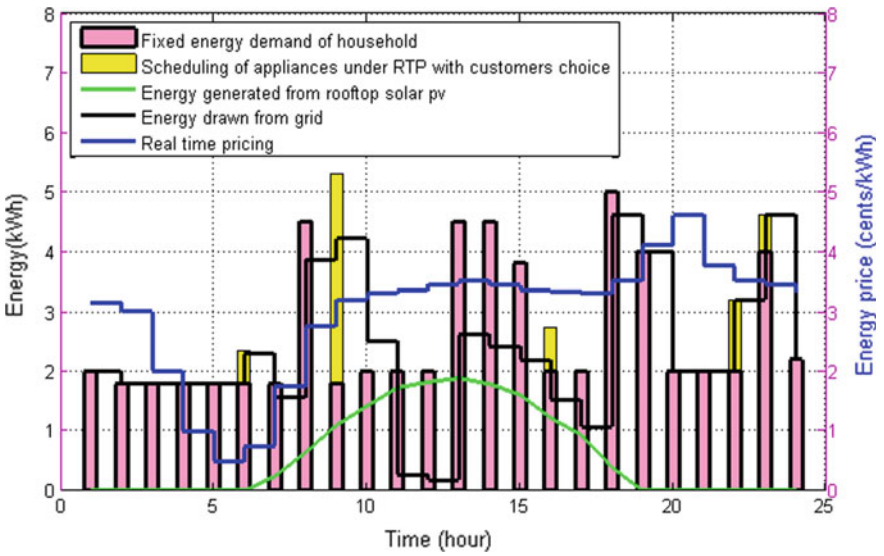


Fig. 1 Energy scheduling under RTP scheme with customer's preference

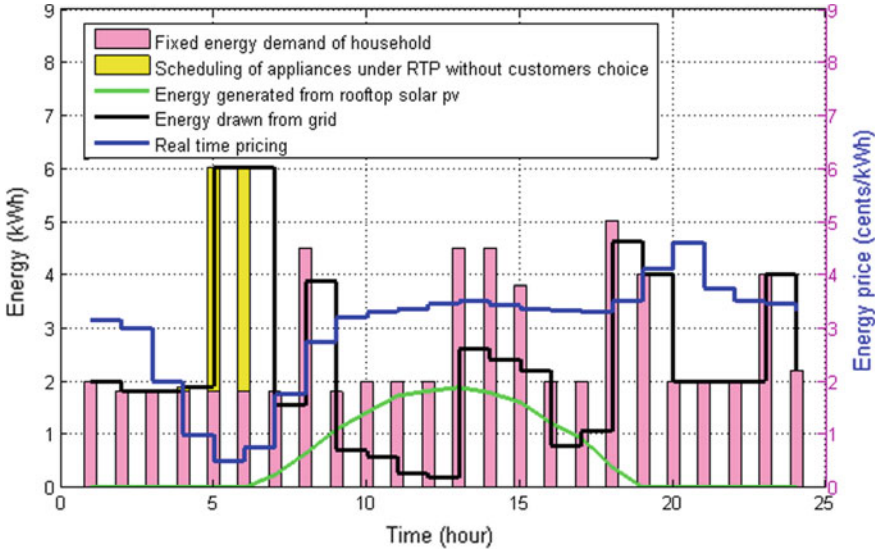


Fig. 2 Energy scheduling under RTP scheme without customer's preference

Table 2 Cost of energy consumption under RTP

	Scheduling under RTP		
Energy consumption price (cents)	With customer's preference	Without customer's preference	Change in bill amount
	173.25	151.96	21.29

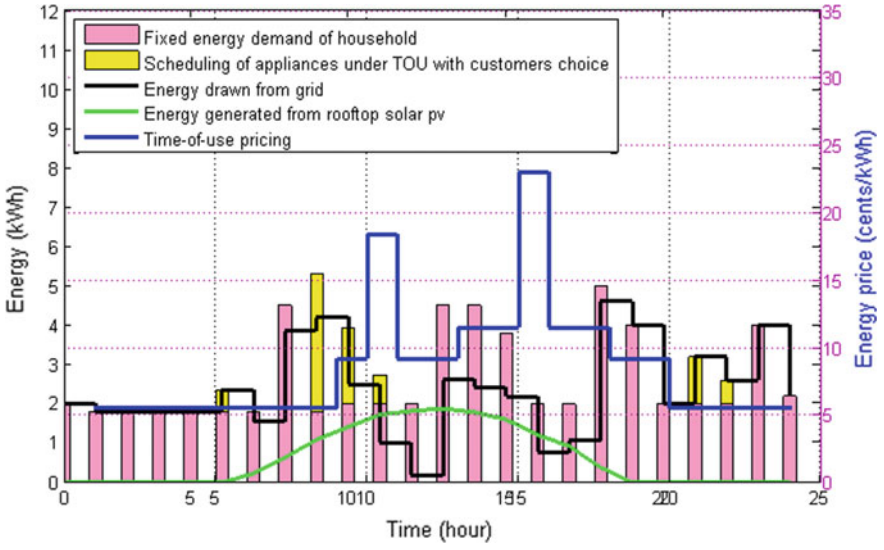


Fig. 3 Energy scheduling under TOU scheme with customer's choice

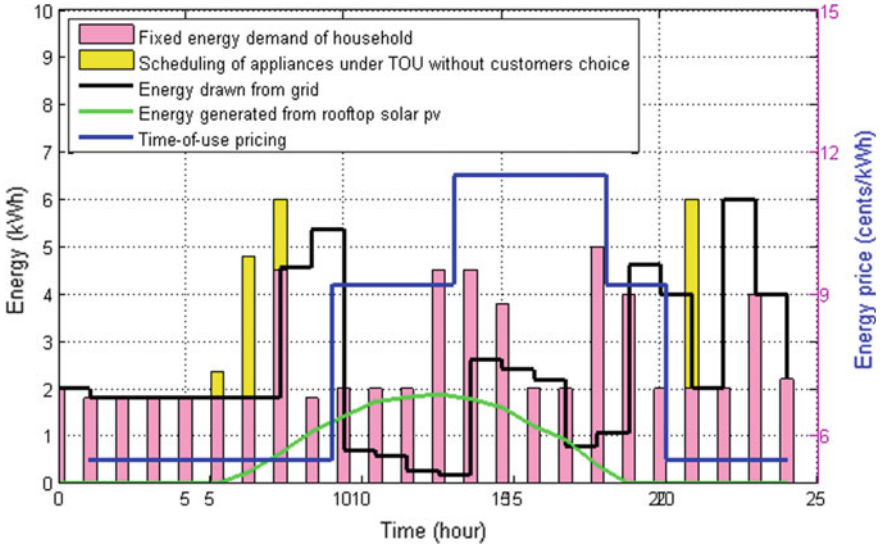


Fig. 4 Energy scheduling under TOU scheme without customer's choice

1.91 kWh, 0.74 kWh, 1.2 kWh, and 0.6 kWh, respectively. Without customer's preference, energy is scheduled at 4 a.m., 5 a.m., and 6 a.m. with energies of 0.1 kWh, 4.2 kWh, and 4.2 kWh, respectively.

The cost with customer's preference is 444.51 cents and without customer's preference is 421.75 cents. It can be seen in Table 3 that the electricity bill changes by 22.76 cents after scheduling the appliances with the customer's preference under TOU.

Energy scheduling under CPP with and without customer's preference is shown in Figs. 5 and 6. With customer's preference, the operating hours are 6 a.m., 9 a.m., 10 a.m., 11 a.m., 10 p.m., and 11 p.m. with energies 0.55 kWh, 3.5 kWh, 1.91 kWh, 0.74 kWh, 1.2 kWh, and 0.6 kWh, respectively. Without customer's preference, energy is scheduled at 7 a.m., 8 a.m., and 9 p.m. with energies of 3 kWh, 1.5 kWh, and 4 kWh, respectively, and accordingly electricity bill is evaluated.

The cost with customer's preference is 487.58 cents and without customer's preference is 452.36 cents. It can be seen in Table 4 that the electricity bill changes by 35.22 cents after scheduling the appliances with the customer's preference under CPP.

Table 3 Cost of energy consumption under TOU

	Scheduling under TOU		
Energy consumption price(cents)	With customer's preference	Without customer's preference	Change in bill amount
	444.51	421.75	22.76

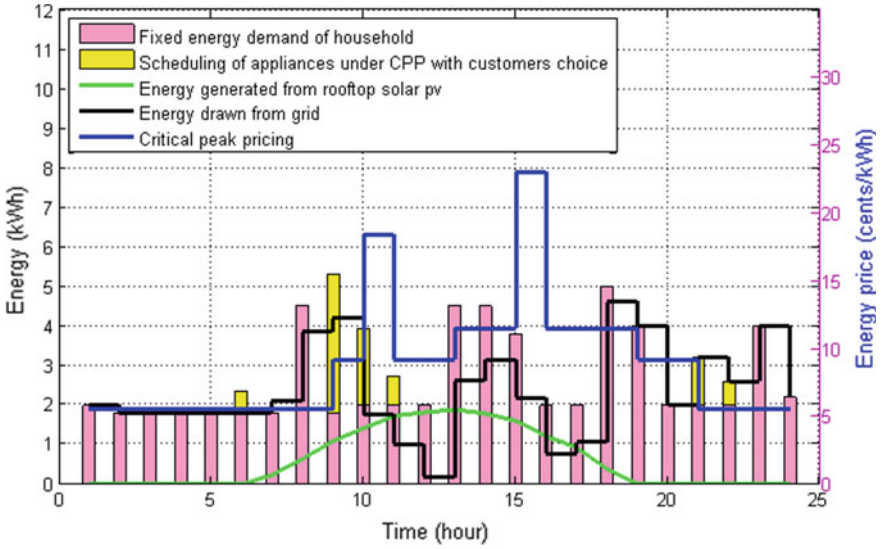


Fig. 5 Energy scheduling under CPP scheme with customer's choice

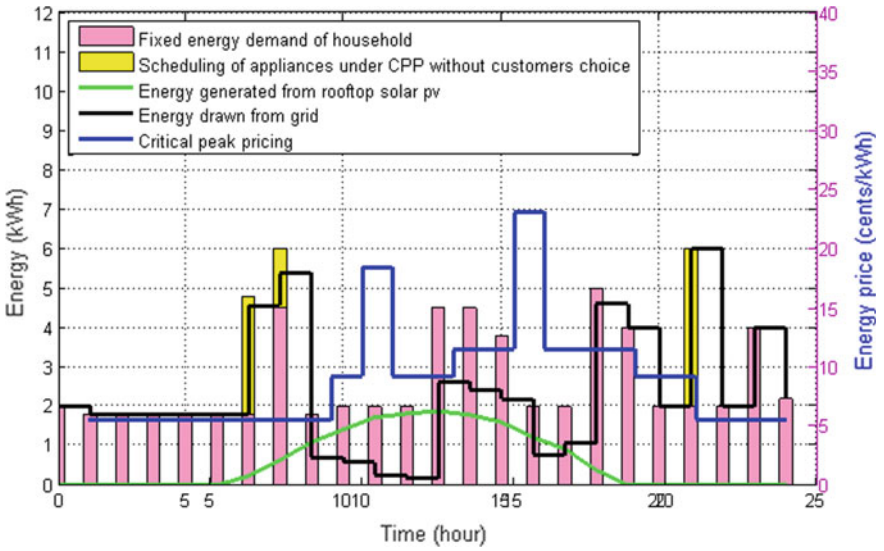


Fig. 6 Energy scheduling under CPP scheme without customer's choice

Table 4 Cost of energy consumption under CPP

	Scheduling under CPP		
Energy consumption price (cents)	With customer's preference	Without customer's preference	Change in bill amount
	487.58	452.36	35.22

5 Conclusion

In this paper, energy scheduling of fixed household load and four appliances along with distributed generation from rooftop solar PV is done. Successful implementation of different pricing schemes (RTP, TOU, and CPP) and comparison of household electricity bills are done. Proposed methodology is formulated by MILP. The main focus in this paper is comparison of different dynamic pricing schemes and adding generating capacity to the user end for mutual benefit of both the utility and customer. CPLEX solver of GAMS and MATLAB is used for scheduling. It is observed that user gets more benefits, when user shifts load from peak loads to OFF peak loads. Fuel cost of solar power is negligible or zero, and hence generation cost is also less compared to other conventional sources. It can be observed that the RTP with rooftop PV gives more benefits compared to TOU and CPP. In future, this work can be extended to multiple appliances of a house and multiple houses in a colony and solar energy can be replaced by other energy sources.

References

1. J. Gao, X. Xiao, J. Liu, W. Liang, C.L. Chen, A survey of communication/networking in Smart Grids. Future Gener. Comput. Syst. **28**(2), 391–404 (2012)
2. X. Fang, S. Misra, G. Xue, D. Yang, Smart grid-the new and improved power grid: a survey. IEEE Commun. Surv. Tuts. **14**(4), 944–980 (2012)
3. P.P. Varaiya, F.F. Wu, J.W. Bialek, Smart operation of smart grid: risk-limiting dispatch. Proc. IEEE **99**(1), 40–57 (2011)
4. N. Venkatesan, S. Jignesh, S.K. Solanki, Residential demand response model and impact on voltage profile and losses of an electric distribution network. Appl. Energy **96**, 84–91 (2012)
5. O. Erdinc, Economic impacts of small scale own generating and storage units, and electric vehicles under different demand response strategies for smart households. Appl. Energy **126**, 142–150 (2014)
6. O. Erdinc, N.G. Paterakis, T.D.P. Mendes, A.G. Bakirtzis, J.P.S. Catalao, Smart household operation considering bi-directional EV and ESS utilization by real-time pricing-based DR. IEEE Trans. Smart Grid **6**(3), 1281–1291 (2015). https://doi.org/10.1109/TSG.2014.2352650
7. E. Matallanas et al., Neural network controller for active demand-side management with PV energy in the residential sector. Appl. Energy **91**, 90–97 (2012)
8. A. Haurie, C. Andrey, The economics of electricity dynamic pricing and demand response programs. www.ordecsys.com/fr/system/files/shared/TOU-PremierRapport.pdf (2013)
9. C. Chen, S. Kishore, L.V. Snyder, An innovative RTP-based residential power scheduling scheme for smart grids, in *Proceedings of IEEE ICASSP*, Prague, Czech Republic, pp. 5956–5959, 22–27 May 2011
10. A.H. Mohsenian-Rad, A. Leon-Garcia, Optimal residential load control with price prediction in real-time electricity pricing environments. IEEE Trans. Smart Grid **1**(2), 120–133 (2010)
11. Electricity prices. http://www.hydroone.com/MyHome/MyAccount/UnderstandMyBill/Pages/ElectricityRates.aspx
12. K. Herter, Residential implementation of critical-peak pricing of electricity. Energy Policy **35**(4), 2121–2130 (2007)

Optimum Design of Photovoltaic System For a Medical Institute Using HOMER



Ingudam Chitrasen Meitei, Thounaojam Bebekananda Singh, Konjengbam Denish, Huidrom Hilengamba Meetei and Naorem Ajesh Singh

1 Introduction

Nowadays, fossil fuels dependency and environmental concerns have reached an alarming situation. The upcoming situation has made us realize the necessity of reduction in consumption of fossil products. People in the remote areas face the challenges of accessibility of the grid power. Moreover, due to insufficiency and lack of electric power supply in rural and hilly areas [1, 2], the advancement in new technologies will help us solve these problems through renewable and nonconventional energy sources and it will lead to low dependency on fossil fuels. The resources available are cost free and emission of harmful gases is very low [3]. Hybrid renewable energy system is considered to be the most suitable solution for its high efficiency.

Electricity plays a vital role in the development of small community as well as the nation. It acts as a fundamental means for advancement and growth of humankind. Major difficulties are faced regarding its production, transmission, and distribution. Many challenges are met regarding the installation of electric power system and dispatching sources for remote and hilly areas through the conventional way. Solution such as hybrid renewable energy system will help to generate electrical power from renewable energies such as wind, solar, micro hydro, biomass, etc. [4]. Comparing to conventional sources, it lacks in reliability due to its dependency on weather and high initial cost [5]. Major researches have been conducted on renewable energy technology across the globe which resulted in important development of new technologies and increase in their efficiency.

The renewable resources, i.e., wind, water, solar, geothermal, biomass, etc., are available in plenty but cannot be extracted continuously from nature and it mainly depends on various factors such as weather conditions, location, availability of

I. C. Meitei (✉) · T. B. Singh · K. Denish · H. H. Meetei · N. A. Singh
National Institute of Technology Manipur, Imphal, India
e-mail: alwayschitrasen@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_131

resources, and duration. Combination of two or more renewable resources will overcome the drawbacks and maximize the system performance and efficiency. Thus, hybrid energy system plays an important role in harnessing renewable energy [6]. System that is composed of one renewable and one conventional energy source or more than one renewable with or without conventional energy sources operates in off-grid or grid connected which is called Hybrid Renewable Energy System (HRES). Hybrid renewable energy system having two or more sources has more potential and is highly reliable than a single source [7]. The application of hybrid energy systems in remote and isolated areas is more relevant than grid-connected systems.

In this work, we designed a renewable microgrid of a particular area according to the availability of resources at minimal cost and at optimum state [8]. Here optimization and simulation are carried out using HOMER Pro software. In our system, it comprises solar PV (photovoltaic), diesel generator set, storage system, and converter [9]. The main aspect of these components is that all the components have a complementary nature and work together to meet the electricity demand.

2 Methodology

The flowchart of steps of methodology starts with load estimation or load analysis and is then followed by potential analysis of the resources, i.e., solar energy in our case. Diesel generator and battery are used during the unavailability of the resources, e.g., cloudy days. Inputs like capital cost of PV panel, battery, solar radiation value, etc. are analyzed. The simulation is performed by the HOMER software to achieve the optimized result (Fig. 1).

3 Load Estimation

In this paper, a case study is done of RIMS, Imphal, Manipur located at latitude of 24°49′ N and longitude of 93°55′ E. The load profile of this area is studied, which consists of two types of load, consisting of hospital, office, academy section, laboratory, shop, and another one consisting of quarters, hostels, and local surrounding houses. The load has been classified into two types Primary1 and Primary2 as discussed in Table 1.

Figure 2 shows the average daily load profile of Primary1 with peak demand as 287.17 kW. The daily consumption of this particular load is 1905 kWh/day as shown in Table 1.

Figure 3 shows the average daily load profile of Primary2 with peak demand as 148.53 kW. The daily consumption of this particular load is 700 kWh/day as shown in Table 1.

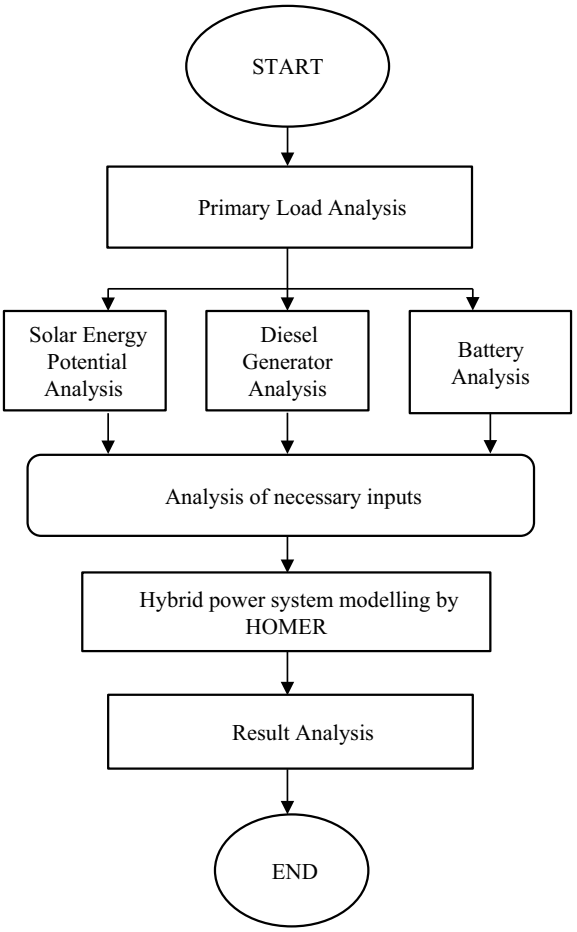


Fig. 1 Methodology flowchart

Table 1 Load classification

Serial no.	Loads	Types considered	Daily consumption (kWh/day)	Peak demand (kW)
1.	Primary1	Commercial load (hospital, office, academy section, shops, and laboratory)	1905	287.17
2.	Primary2	Residential load (quarters, hostels, and local surroundings)	700	148.53
		Total average demand	2605	



Fig. 2 Average daily Primary1 load profile (hourly) of RIMS



Fig. 3 Average daily Primary2 load profile (hourly) of RIMS

Manipur is currently facing insufficiency of power. A large amount of money is spent every year to meet its current demand (600–800 crores annually, approx.). Setting up a hybrid renewable energy system at important place with large consumption of electricity like this area can not only help improve the government economically but also increases renewable energy production of country reducing carbon emission.

4 Resource and Components

4.1 Solar Energy Resource

The solar irradiation of RIMS, Imphal located at latitude of 24°49′ N and longitude of 93°55′ E is obtained from surface meteorology and solar energy database, NASA [10] (Fig. 4).

Figure 5 gives daily solar radiation (average). The left side of Fig. 5 represents the daily solar radiation (average) data and the right side represents solar radiation clearness index value. It also gives monthly average radiation.

Average solar radiation: 4.46 kWh/m^2/day (obtained from surface meteorology and solar energy database, NASA [10]).

The solar irradiance of this particular area reaches its peak value in the month of July and drops to minimum value in the month of December.

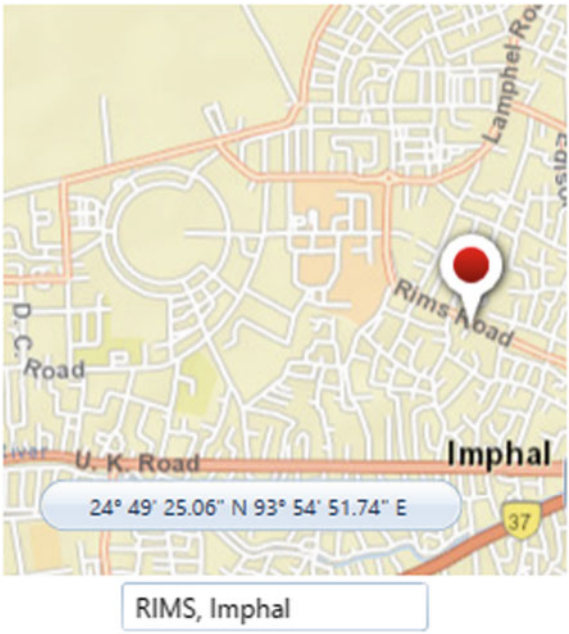


Fig. 4 Location where the systems will be set up

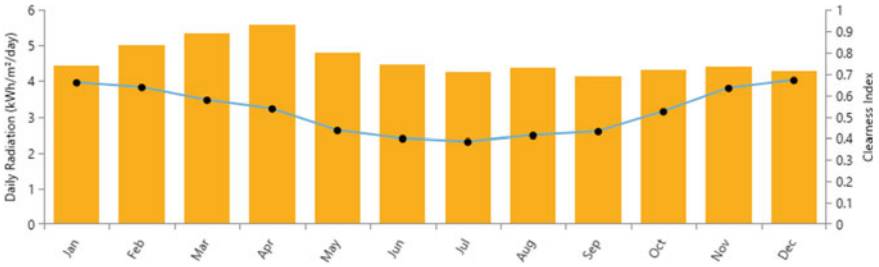


Fig. 5 Daily solar radiation of RIMS, Imphal

4.2 Solar PV Panels

According to our survey and analysis of the cost of PV panel in our location, we considered PV panel of loom solar company of size 340W in which the capital cost is Rs. 36,000 per kW, replacement cost is Rs. 36,000, and operating and maintenance cost is Rs. 10,000 per annum with a lifetime of 20 years.

4.3 Battery

We considered Hi Q battery of 12 v/85 Ah as the characteristics is the same as those given by the HOMER tool which costs Rs. 6,300, replacement cost is the same, i.e., Rs. 6,300, and its operating and maintenance cost is Rs. 1,000 per year. The battery has a lifetime of 8 years.

4.4 Converter

The size of the converter in the system is of capital cost Rs. 9,000 per kW, replacement cost is Rs. 9,000, and operating and maintenance cost is Rs. 500 per year with the lifetime of 15 years.

4.5 Diesel Generator

The diesel generator of kW is used in which the capital cost is Rs. 6,000 per kW, replacement cost is Rs. 6,000, and operating and maintenance cost is Rs. 4400 per year which will last for 44,000 h.

Table 2 gives a summary of the cost of different components such as PV panels, diesel generator, lead–acid battery, and converters used in the design. The initial cost, replacement cost, and operation cost are considered.

Table 2 Components cost details

Cost (Rupees)	Components			
	PV panels (per 1 kW)	Diesel generator (per 1 kVA)	1 kWh lead–acid battery	Converter (per kW)
Initial cost	36,000	6000	6300	9000
Replacement cost	36,000	6000	6300	9000
Operation and maintenance cost	10,000 (per year)	4400 (per year)	1000 (per year)	500 (per year)

5 Result

A schematic diagram of microgrid for RIMS is given in Fig. 6. A total of 165,855 simulations were performed by HOMER software where 123,087 were feasible and 42,429 were omitted. From the simulation result (Fig. 7), 492 kW PV panel, 480 kW generator, 2071 1 kWh LA battery, and 281 kW converter are found to be the optimal solution for the system. The cost summary for the optimized result is given in Table 3 and Fig. 8 gives the graph of it.
Production:

Generic flat plate PV: 880,517 kWh/year (79%).
Auto size generator: 23,428 kWh/year (21%).

Figure 9 gives the monthly average electric production of the optimized system.

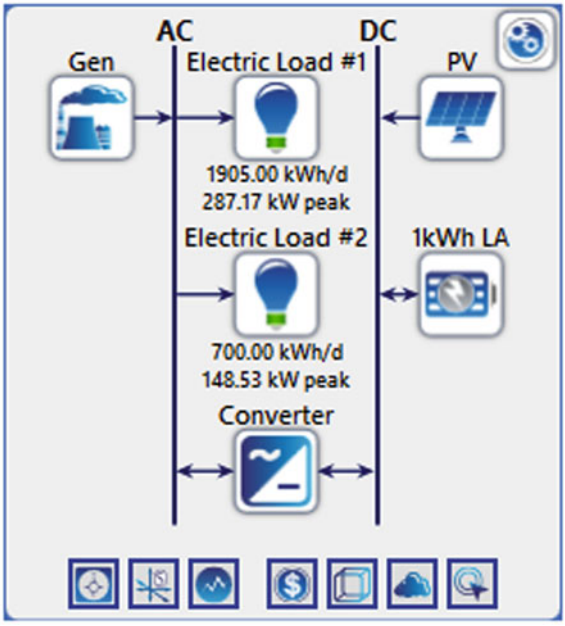


Fig. 6 Schematic diagram of microgrid for RIMS, Imphal

Architecture								Cost				System	
			PV (kW)	Gen (kW)	1kWh LA	Converter (kW)	Dispatch	NPC (₹)	COE (₹)	Operating cost (₹/yr)	Initial capital (₹)	Ren Frac (%)	Total Fuel (L/yr)
			492	480	2,071	281	LF	₹219M	₹18.95	₹14.1M	₹36.2M	74.1	70,849
				480	656	78.6	CC	₹261M	₹22.56	₹19.6M	₹7.72M	0	272,816
				480			CC	₹321M	₹27.79	₹24.6M	₹2.88M	0	367,825
			14.6	480		2.24	CC	₹323M	₹27.92	₹24.7M	₹3.43M	0	366,448
			1,792		6,611	354	CC	₹484M	₹41.88	₹29.0M	₹109M	100	0

Fig. 7 Result obtained after simulation

Table 3 Cost summary for the optimized result obtained after simulation

Component	Capital (₹)	Replacement (₹)	O and M (₹)	Fuel (₹)	Salvage (₹)	Total (₹)
Autosize genset	₹2,880,000.00	₹754,738.10	₹5,826,690.26	₹54,953,996.61	−₹643,670.17	₹63,771,754.81
Generic 1 kWh lead acid	₹13,047,300.00	₹26,654,890.59	₹26,772,886.77	₹0.00	−₹198,310.02	₹66,276,767.34
Generic flat plate PV	₹17,714,844.28	₹5,647,622.56	₹63,613,595.16	₹0.00	−₹3,182,797.51	₹83,793,264.49
System converter	₹2,524,994.24	₹1,071,288.98	₹1,813,439.16	₹0.00	−₹201,627.40	₹5,208,094.99
System	₹36,167,138.52	₹34,128,540.23	₹98,026,611.35	₹54,953,996.61	−₹4,226,405.09	₹219,049,881.63

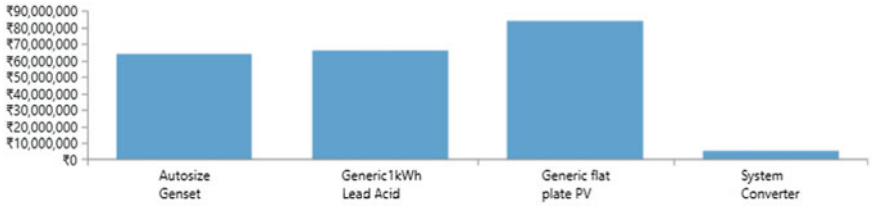


Fig. 8 Graph of the above cost summary

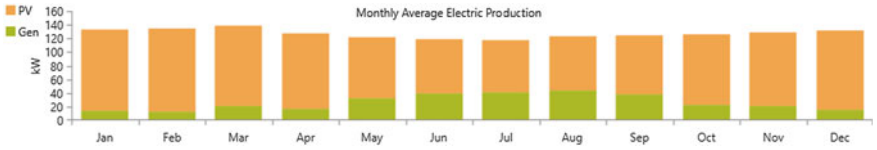


Fig. 9 Monthly average electric production of the design system

6 Conclusion

The result from the HOMER shows that PV and diesel generator with LA battery and converter were found to be a comparative economical solution.

The optimized result can provide the cost of electricity/energy at a rate of 18.95 (INR/kWh) with its initial capital cost of 36,200,000 (INR) with 74.1% of energy from RES supplying the load. The Net Present Cost (NPC) is also found to be 219,000,000 (INR) for the most optimized system design.

References

1. S. Chowdhury, S.P. Chowdhury, P. Crossley, *Microgrids and Active Distribution Networks*. Institution of Engineering and Technology (2009)
2. J.A.P. Lopes, A.G. Madureira, C.C.L.M. Moreira, A view of microgrids. Wiley Interdiscip. Rev. Energy Environ. **2**, 86–103 (2013)
3. S.N. Bhaskara, B.H. Chowdhury, Microgrids—a review of modeling, control, protection, simulation and future potential, in *2012 IEEE Power & Energy Society General Meeting*, p. 1 (2012)
4. R. Sen, S.C. Bhattacharaya, Off-grid electricity generation with renewable energy technologies in India: an application of HOMER. Renew. Energy **62**, 388–398 (2014)
5. T. Lambert, P. Gilman, P. Lilienthal, Micropower system modeling with HOMER. Integr. Altern. Sour. Energy **1**, 379–418 (2006)
6. M.A. Salam, A. Aziz, A.H.A Alwaeli, H.A. Kazem, Optimal sizing of photovoltaic systems using HOMER for Sohar, Oman. Int. J. Renew. Energy Res. **3**(2) (2013)
7. R. Srivastava, V.K. Giri, Optimization of hybrid renewable resources using HOMER. Int. J. Renew. Energy Res. **6**(1) (2016)
8. K. Rout, J.K. Sahu, Various optimization techniques of hybrid renewable energy systems for power generation: a review. Int. Res. J. Eng. Tech. (IRJET) **05**(07) (2018)

9. K. Kusakana, H.J. Vermaak, Hybrid diesel generator/renewable energy system performance modelling. Renew. Energy **67**, 97–102 (2014)
10. NASA surface meteorology and solar energy database [Online]. https://power.larc.nasa.gov/data-access-viewer/. Accessed Jan 2019

Raising Concerns on High PV Penetration and Ancillary Services: A Review



Rajesh Kumar, Aman Ganesh and Vipin Kumar

1 Introduction

With the escalating needs, desire and urge to move toward technological development, energy has become an inevitable part of the society and it could be significantly revealed through various energy reports. Reference [1] shows that the energy production overall the world in the last 26 years has increased from about 9601 Mtoe (Million tons of oil Equivalent) in 1997 to about 14,080 Mtoe in 2017 [1]. A clear increasing trend is observed in Fig. 1 based on data in [1]. Also in accordance with [2] energy needs are anticipated to expand by 55%, i.e., from 11.4 billion tons of oil equivalent to 17.7 billion considering the period from 2005 to 2030.

Further, from [3] it could be observed that the majority of the energy production in the world is due to oil, coal, and gas. This is illustrated in Fig. 2. Based on the data published in [3]. Considering Indian subcontinent escalation in energy production can be observed from [4], i.e., installed capacity in India was around 63,636 MW in 1990s and has shown a drastic escalation to about 344,002 MW in March 2018 [4]. Figure 3 gives a clear indication of increasing energy production with the ever-increasing demand for energy [4]. This is in accordance with [5] that energy consumption is anticipated to grow by 4.2% by the culmination of 2035, which would surpass major economies of the world.

Again it could be observed that the major portion of energy installed in India is contributed by Coal/Lignite [4]. This could be clearly observed in Fig. 4 [4]. Hence,

R. Kumar · V. Kumar (✉)
UIET, Maharshi Dayanand, Rohtak, Haryana, India
e-mail: vipinsaini4434@gmail.com

R. Kumar
e-mail: lncs@springer.com

A. Ganesh
Lovely Professional University, Phagwara, Punjab, India

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_132

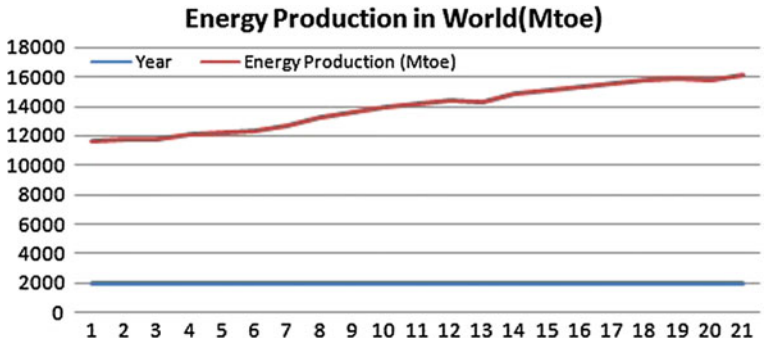


Fig. 1 Energy production in world (1997–2017)

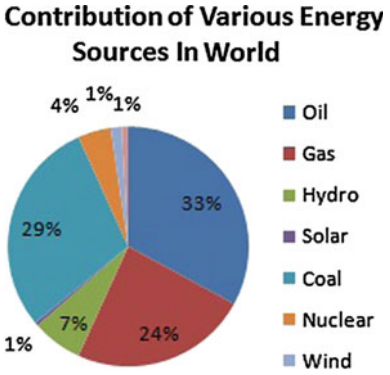


Fig. 2 Contribution of energy sources in production

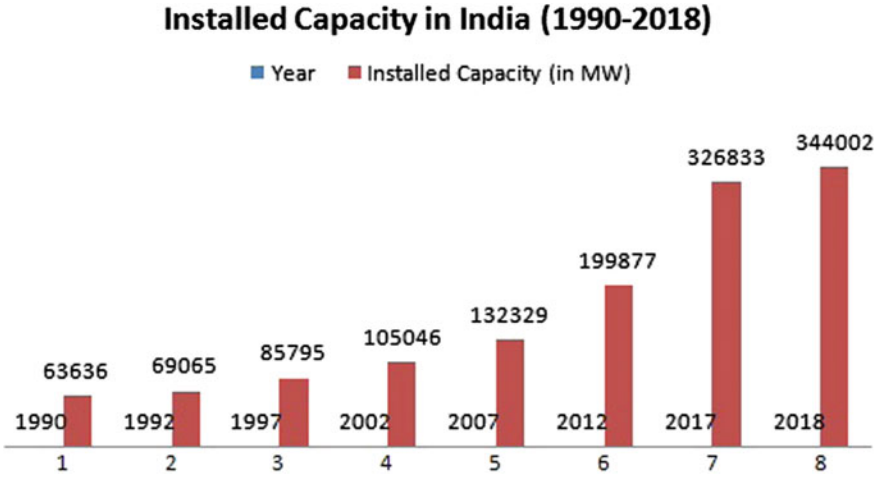


Fig. 3 Installed capacity in India (1990–2018)

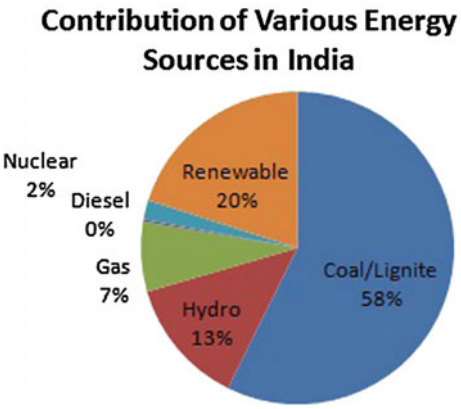


Fig. 4 Contribution of various energy sources in India

it could be scrutinized that all over the coal, gas, and oil power plants have dominance in energy production. These conventional sources emit harmful carbon dioxide and azoth gases [6]. According to [7] human health is largely burdened due to atmospheric emissions from the coal-fired power plants. The carbon dioxide emission from the Indian plant has inflated from about 1255.39 $MtCO_2$ (Metric tons of CO_2) to about 2234.27 $MtCO_2$ as shown in Fig. 5 [1].

These particulates have a diameter of less than 2.5 mm (PM 2.5) [7]. Other emissions include 2100 Ktons of sulfur dioxides, 2000 Ktons of nitrogen oxides, and 100 Ktons of volatile organic compounds which had resulted in premature deaths of about 80,000–115,000 and also many asthma cases are also reported due to intense exposure to PM 2.5 pollution levels [7]. Further, health hazard includes skin, cardiovascular, brain blood, and lung diseases [8]. However, methods for evaluation of health risk from emission of conventional power plants are not available universally [9]. The calculation of the intensity of CO_2 emission could be calculated by the methods given in [10]. With such an increase in the amount of carbon dioxide

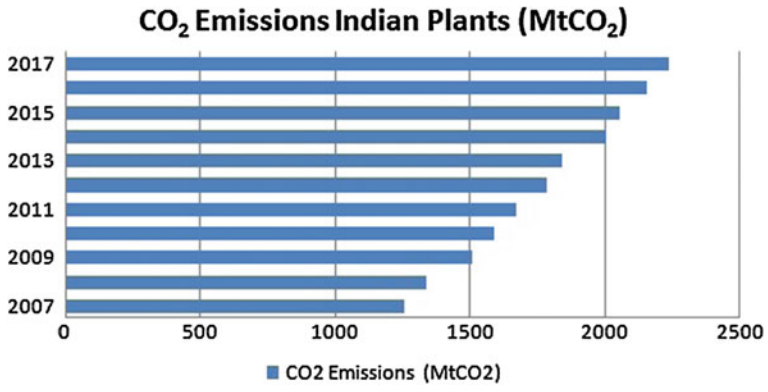


Fig. 5 Emissions by thermal plants in India

and other greenhouse emissions environmentalist over the world are raising concerns [11]. Recently, they have issued a warning regarding a point beyond which even a 0.5° rise in temperature around the globe will cause irretrievable disasters will reach within a short span of 12 years [12]. Following which the countries all around the world are gathering to find the solution to such serious threats. They have signed various treaties like the Kyoto Protocol, COP21, etc. and targeting a paradigm shift toward renewable energy sources. With these concerns, a paradigm shift toward renewable energy can be observed in Fig. 6 [1]. According to [13] renewable energy mix by 2020 is expected to be 22.5%.

Also, there is a significant increase in renewable in Indian Subcontinent which is shown in Fig. 7 [1]. The renewable energy sources include wind, small hydro,

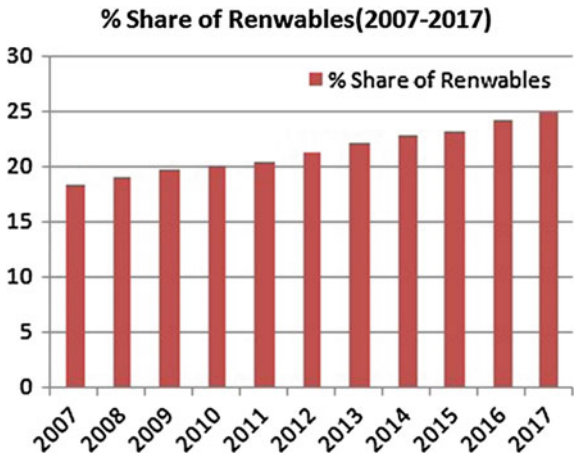


Fig. 6 Percentage share of renewable in world (2007–2017)

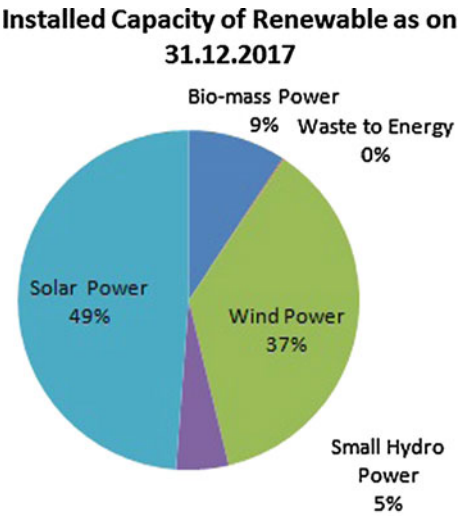


Fig. 7 Renewable energy contribution in India as on 31.12.2017

biomass power, waste to energy, and cogeneration bagasse [10]. The distribution of renewable energy resources in India is shown in Fig. 7.

The authors here have considered solar power because of the following reasons:
1. It is the cleanest source of energy,

1. Considering the geographical location of Indian subcontinent solar PV is the best suitable source.
2. Serious concern caused due to increasing emissions is zero here as generation from solar does not cause any release of harmful emissions.
3. Almost a negligible operational cost for generation of electricity.
4. Improvement in PV cell efficiencies and the significant drop in prices of PV panels.
5. The Paradigm shift of world toward cleaner sources of energy and ambitious targets set by countries toward various conferences for production of solar energy.
6. According to [14] by 2040 solar is likely to overtake wind and becoming the world's largest source of electricity by 2050.

Further sections include a brief description of solar photovoltaics and its working, its integration with the grid, contribution and policies associated with it, raising concerns due to high PV penetration, and some viewpoints on ancillary service.

2 Solar Photovoltaic: Principle and Working

2.1 PV Cell

The Rudimentary composition of PV cell includes multifarious layers, i.e., metallic contact, p-n doped semiconductor, metallic grid, antireflection coating, transparent adhesive, and glass covering, cross-sectional view of PV Cell is shown in Fig. 8 [15].

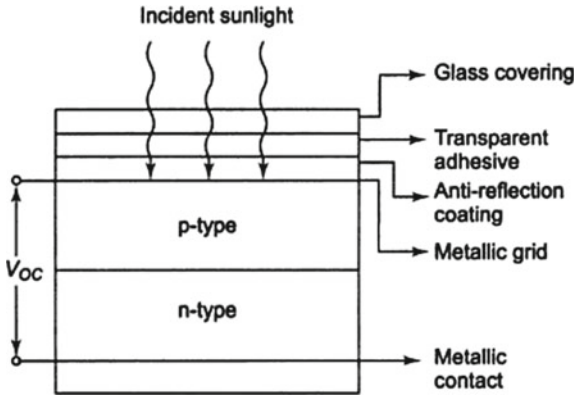


Fig. 8 Cross-sectional view of solar cell

With about 90–95% transparency of the layer of glass, the solar cell is fortified from dust and rain [15]. The transparent adhesive is used to intact the layers together, and since it is transparent in nature it does not hinder incident sunlight [15]. For ensuring maximum absorption and minimizing the reflection of sunlight antireflection layer is used [15]. The extraction of photocurrent is done with the help of a metallic grid and metallic contacts [15]. The core of the solar cell is the P-N junction diode; this is the platform wherein the electrical energy is harvested from solar energy. Photons from the sunlight energize P-N doped semiconductor and generate electron-hole pairs, which act as charge carriers and subsequently constitute photocurrent. The single unit for the production of electricity from sunlight is termed as the solar cell. The output of this unit merely 0.8 V DC (Silicon-based cell), multiple solar cells are connected to escalate the output to a significant level and is known as solar panel. The solar panel in tandem is called solar arrays. The functionality of solar cell pertains to physics of Photovoltaic effect (Fig. 9) discovered by Henry Becquerel in 1839 [16, 17]. An electron-hole pair in the P-doped semiconductor is formed due to the incident sunlight. Electrons move from p to n and exit the semiconductor from the metallic contact into the load circuit with the aid of junction potential. Meanwhile, the holes move away from the p-n junction to the metal contact on the p side where it meets the electron, which has completed its journey from the load circuit. The above-mentioned electron flow constitutes photocurrent (I_{PH}).

In the presence of a load, a potential difference develops across the terminal of the cell, generating current in the opposite direction to photocurrent termed as dark current (I_D). This gives the output current given by the Eq. (1).

$$I = I_{PH} - I_D \tag{1}$$

I_D in Eq. 1 is represented by the Shockley diode equation given by Eq. 2.

$$I_D = I_o\left(e^{\frac{qV}{nKT}} - 1\right) \tag{2}$$

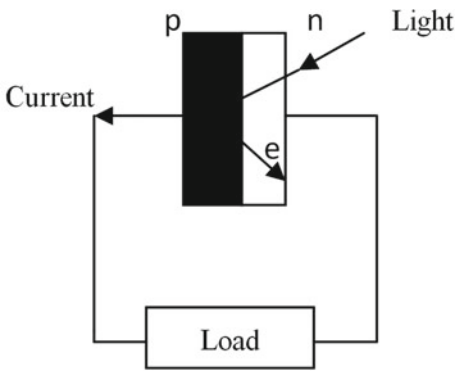


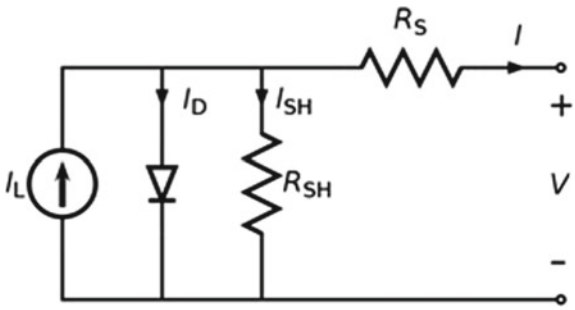
Fig. 9 Photovoltaic effect

From this equation, we equivalent circuit can be constructed as in Fig. 10, recombination losses are accounted by keeping shunt resistance, considering current source to be antiparallel with diode and metallic contacts are accounted by the series resistance [18].

Initially, solar efficiency was a major concern, it is still a concern but the efficiency has seen the growth over the period to a significant amount [19] shown in Fig. 11 due to various research and innovation in the field. With the increase in efficiency, it is also observed that over the period the prices have also fallen gradually as in Fig. 11.

From Fig. 12, it is observable the price of solar cells has fallen by over 60% from 8.82$/Watt to about $3.14/Watt [19]. Initially, solar cells have found applications in space vehicles, water heaters, and many low power applications. However, thanks to power electronics, research and innovation application of solar in the grid have also become viable, basics of which are discussed in next section.

Fig. 10 Equivalent circuit



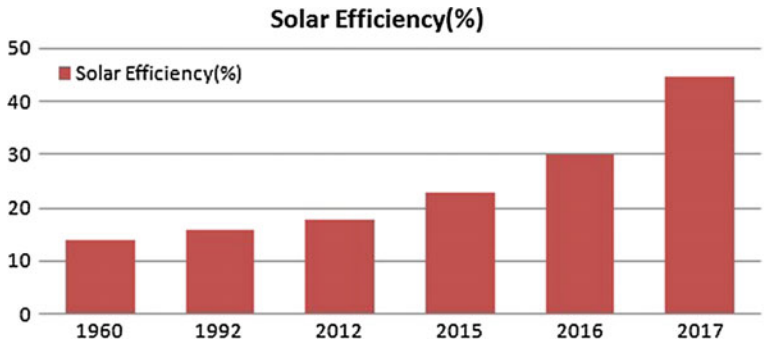


Fig. 11 Solar efficiency (1960–2017)

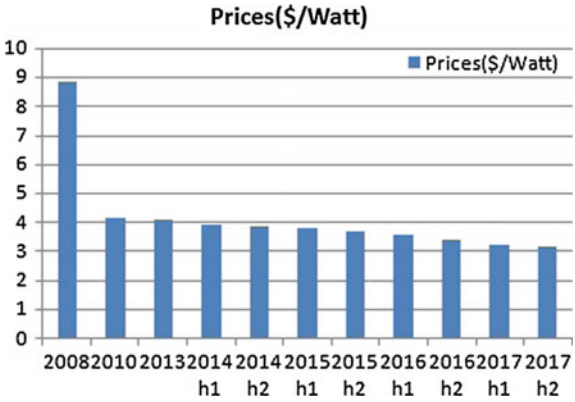


Fig. 12 Falling prices of PV cell ($/Watt)

2.2 *Grid-Connected Solar PV*

These are the power systems interconnected with the utility grid which are energized by photovoltaic panels. System composition includes PV panels, Maximum Power Point Tracking Mechanism (MPPT), inverters, units of power conditioning, and grid connection equipment. A block diagram of grid-connected interactive SPV system is shown in Fig. 13 [12].

SPV system is not compatible with the grid when directly converted to AC from DC (400 V at distribution load). Hence, first DC is escalated using DC/DC high

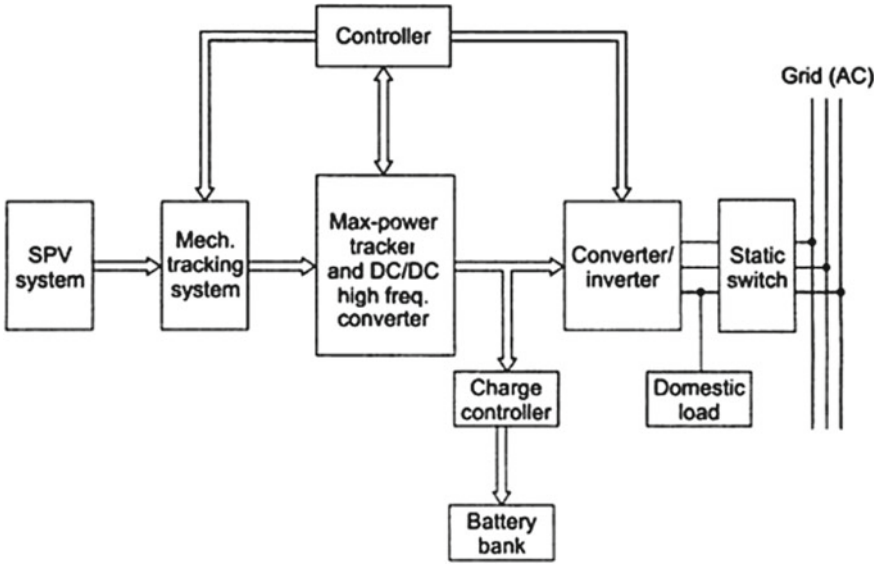


Fig. 13 Grid connected solar PV

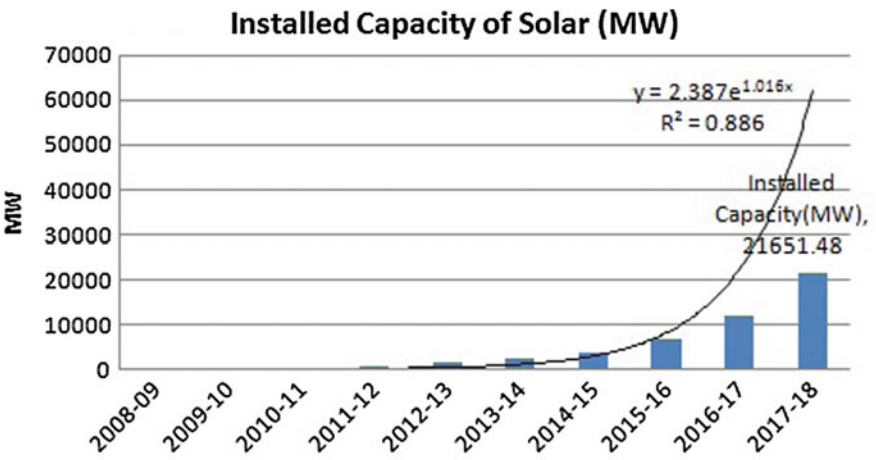


Fig. 14 Installed capacity of solar in India

frequency chopping. The combination of converter–inverter is mandatory so that power can flow in, either way, depending upon the amount of solar power availability. The battery provides back up for any grid outages occurring in the night. For bulk, solar power systems basic scheme will be similar except that it would directly feed power into the grid and no power flow in reverse direction. The increase in research and innovation in the field of solar and other factors like the increase in efficiency and decrease in price/Watt increasing trend in solar installation could be seen as in Fig. 14 [20].

Clearly, it can be observed trend is exponential since R^2 value is near to unity; hence, there anticipated the high value of the solar installation. Apart from technical factors, various other factors are responsible for increasing PV installation, which is dealt with in the next section.

2.3 Other Factors in PV Installations

The Government of India announced a target of about 175GW cumulative renewable power installed by 2022 in 2015 at Paris COP21 conference [21]. Submission of its Intended Nationally Determined Contribution (INDC) to the UNFCC by India in which following clauses are included:

1. Nonfossil-based electricity capacity to be increased to 40% by 2030 with the aid of international support.
2. India's GHG emissions intensity per unit GDP is to be reduced by 33–35% below 2005 by 2030.
3. Creation of additional carbon sinks of 2.5–3 billion tons of carbon dioxide through additional tree cover.

The first mission to be operationalized under National Action Plan on Climate Change (NAPCC) in January 2010 is National Solar Mission. It uses a three-phase approach as follows [20]:

a. Establishing policy conditions for solar technology diffusion across the country as quickly as possible to establish for acknowledging India as a global leader in solar.
b. Amendment of the initial target of 20GW grid solar plants to 100GW to be achieved by the year 2022.
c. All states are to reach 8% solar RPO in year 2022 according to revised tariff policy.

Under mission and guidelines for the development of smart cities includes 10% renewable energy as mandatory. Mandatory provision of rooftop solar or high floor area ratio has been included as amendments in building bylaws. Other incentives include tax-free solar bonds, long tenure laws, incorporating measures in Integrated Power Development Scheme (IPDS). A net of 5602.65 MW capacity has been added till 12.31.2017 during 2017–18 [21]. A cumulative achievement of various renewable energy sources until 12.31.2017 is shown in Fig. 15[21].

It is evident from Fig. 15 achievement among renewable solar has the highest share. The various commissioned and allotted projects in India are shown in Fig. 16 [21].

Many more decisions regarding ultra-mega PV power projects and development of Solar PV parks schemes, decisions by various defense institutions to plant solar on the rooftop where ever space is available, setting up of solar alliance, etc. clearly points that development of solar in India is at its peak. So, we need to rectify various concerns raised by this system. Though there are several concerns authors here have considered high PV penetration and ancillary and reviewed various works of literature concerned with it as will be discussed in the next section.

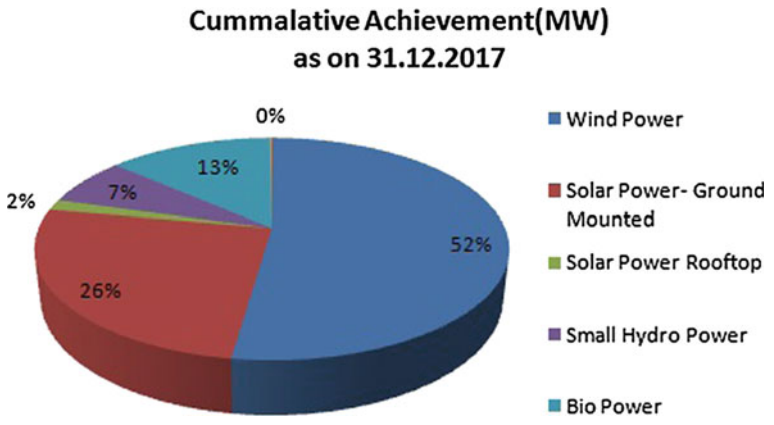


Fig. 15 Cumulative achievement of renewable energy sources

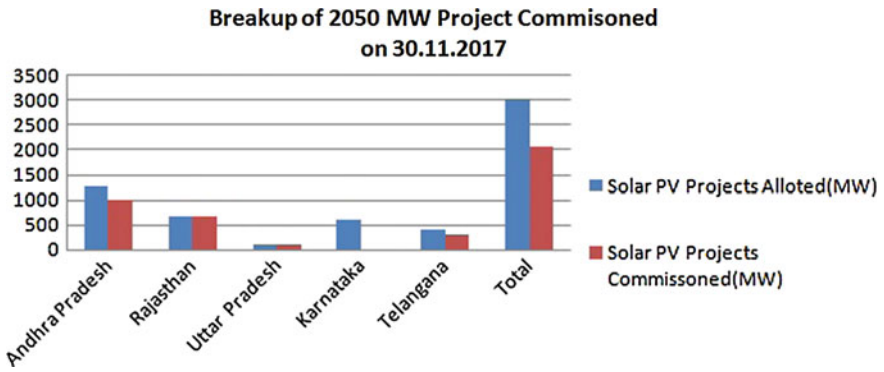


Fig. 16 Various commissioned solar projects in India

3 High PV Penetration

Increase in the PV penetration would lead to contemporary situations for consumers with regards to power supply security and power quality [22, 23]. Since the traditional power system was designed keeping in view the conventional energy sources, they constitute heavy rotating synchronous generator [24, 25]. On the contrary, PV sources hardly have the rotating parts in it; the cost of this loss of large synchronous generators is that the inertial support is lost; this inertial support helps to restore system frequency response [23–26]. Considering power system in UK integration of PV is anticipated to reduce the system inertia by 70% while considering the period 2013/14–2033/34 [27]. It was observed in [28] that increase in alteration in PV system causes nonlinear variation in frequency dynamics; also, a dip in dominant oscillation mode is observed consistently [29]. Hence, due to integration of renewable energy, the network is dominantly depended on seldom synchronous machine [30]. Increase in PV penetration also increases synchronous reactor, which in turn have negative impact on transient system stability [31]. Susceptibility of generator trip/load trip has seen an increase with the increase in the contribution of photovoltaic in the power generation [24]. With the relegation of conventional power system with wind and solar plants, power system weakens as these resources do not provide inertial response [32]. Primary frequency control reserves displacement, affecting the location of primary frequency control reserves is also raising concerns due to the increase in the PV [24]. If system infallibility or power quality is to be maintained while increasing the PV in system then there lies an inevitable need for an additional control system [33]. The probability of inverse power flow on the distribution feeder happens at a higher incessant rate with the escalating levels of solar PV [34]. With the alteration of renewable energy sources in the European continent the inertia in synchronous area diminishes [35]. This might lead to violating boundaries of voltage rise defined by ANSI C84.1 [34]. Apart from the reverse power flow as discussed in [34] two other issues related to high PV penetration include supplementary power

flows in the system from distribution link to the transmission link and grid instability due to frequency and voltage [36]. Also energy generated due to solar PV is usually done with the aid of power electronics converter which usually lacks system inertia [37]. The conventional power plant connected to the grid will be more influential while considering power system dynamics in case of lower penetration of PV (Below 50%), and would be primarily responsible for grid control [38]. On the contrary, these generators are unable to provide power system stability without the aid of other devices connected to power system [28, 38–40]. Besides PV inverter at the point of common coupling must not actively regulate the voltage according to IEEE 1547 and UL 1741 standards. Hence, null generation of reactive power from these sources [41–43]. Effects of PV system transient which are detrimental such as effects of the cloud are not mitigated to a significant amount by the work reviewed in the literature so far, hence, it becomes inevitable that power inverters provide to some extent power capability for additional voltage regulation to absorb or generate reactive power [44]. Steady-state voltage magnitudes are also affected by high PV penetration. It is evident from [45] that the detrimental or beneficial outcome on small signal stability by solar PV generation may depend on location and penetration level and dispatch of existing synchronous generator, along with the inflating size of PV solar generator and position of point of intersection from the main system transient overvoltage significantly increases [45]. The voltage problem severity caused due to high PV penetration relies on location and relative size of distributed PV generation and loads, topology of the distribution feeder and voltage regulation methods [34]. Traditional power systems pose threat to security with the increase in PV generator which is utilized from highly variable irradiance and connected overpower converter, multiple technologies for energy storage systems having different time constants, power converters are used as an interface few of them to the grid [46]. The equipments and instruments available in the system are required to work on shorter time scales, however, said variation in PV an escalation in number of operation is observed. These operations are necessary to counteract the variation and would lead to substantial reduction in the lifetime of switches and tap changers [47]. Renewable energy sources get decoupled from the ac power system that is preexisting reason being the high power converter used in PV generation. Although the generation due to PV boosts existing power capacity from the installed system, effective system inertia remains unaltered, because power converters inherited within the system creates a decoupling effect of real inertia and AC grid. The outcome of which is heavy excursions in frequency [46].

3.1 System Inertia

The opposing force to the system to vary its motion, either stationary or moving is termed as Inertia [24]. This inertia is related to kinetic energy as given in Eq. (3).

$$\textit{Kinetic Energy} = \frac{1}{2} \times \textit{Moment of Inertia} \times \textit{Angular Speed}^2 \tag{3}$$

Also, inertia constant H is termed as the ratio of Watt-seconds (MWs) at rated speed with rated power (MVA) as shown in Eq. 4 [48].

$$\boldsymbol{H} = \frac{1}{2} \times \frac{\boldsymbol{J} \times \boldsymbol{\omega}^2}{\boldsymbol{VA}} \tag{4}$$

When there is momentarily disturbance in the output, abrupt changes in the output generated can be prevented using this stored kinetic energy [48]. During generation outages, i.e., generation does not match the load; this stored kinetic energy helps in restoring power imbalance. Also from Eqs. (3) and (4), it is inevitable that increase in kinetic energy is vindicated with an increase in generator inertia. Hence, the systems with seldom inertia like PV tend to contribute negligibly to the power system stiffness [41, 48].

$$\frac{d\omega}{dt} = T_a = \frac{1}{J_g}(T_m - T_e) \tag{5}$$

Also, it is evident from Eq. (5) an increasing amount of moment of inertia results in feeble variations of the rotor angle from its synchronism [45]. In accordance with [41] with the increase up to 30% voltage at the steady-state of the system ameliorates. Whereas further PV addition there is significant decline in magnitude of steady-state voltages drop until they are culminated to the base case values. The lower inertia constant of island grids leads to high-frequency discrepancy in islands [49]. In short-term system stability, inherent rotor inertia plays a significant role [37]. The rotor speed deviation performance is improved with the aid of emulation of inertia [37]. A significant amount of duration is compelling for adjusting the power for which inertial response for current AC systems becomes crucial for frequency control [35].

3.2 Ancillary Services

The market which is vitally relied upon capacity or energy market serves as principle market. On the contrary, the ancillary service market is the fundamental term associated with security, efficient power system, infallibility, and calibration of the power system [24]. These were the activities that were previously considered to be the part of the principle market but with the integration of generation from the distribution side these services are disintegrating gradually [50]. Services that are aided to all the users present in network by the TSO are usually termed as system services, while on the contrary exactly opposite case occurs in ancillary services [51]. However, in accordance with [52] services which aid the provision of energy to support infallibility of power system is regarded to be part of ancillary services. While considering

overall system, need for voltage control emanates in ancillary service. These services are aided by resources that are proficient in providing the need (voltage control) and imbibing centralized control for directing resources to meet the requirement [53]. Reconciliation of conflict between buyers and sellers for the mutual benefit can be done by propounding market for ancillary services [54, 55]. Considering the distribution of these services, unambiguous preference for a particular resource allocation is not available. For instance, combining obligatory provisions, auctions, offers which are competitive, distinct duration length and bilateral contracts are not imbibed under a single definition or provision [56]. For fortifying supply for short-term and maintaining a balancing demand throughout the power system, frequency control ancillary services are necessitated [57]. These services should not be included as an extension to the energy market, rather distinct mechanisms must be developed for successful deployment for ensuring enactment and proper remuneration for ancillary service market [58]. General core areas in ancillary services are composed of voltage support, system restore service, and voltage support, which responds to disturbance or contingencies [24]. Also few of the regions in the world include incentivizing the provision of primary frequency response on to ancillary services [59]. While reviewing various regions in the world it would be found that system inertia rules are not enforceable in various countries until now.

a. Africa/Middle East: Frequency regulation, spinning reserves, voltage and reactive power support, black start and load shedding facilities are included in the grid code of Kenya and Nigeria. Operating reserves, black start, and unit islanding, constrained generation, reactive power supply and voltage control from units and regulation are mandated for the system operator by the South African grid code. Though South African grid code requires maintaining the frequency exceeding 49.5 Hz after consideration of contingency losses with aid of primary frequency control [24].
b. Asia: CEC India published draft grid code, according to which aiding support for active power, reactive power support, black start, etc. whereas regulation, reserve, reactive support, and voltage control, black start service, and reliability 'must run' services are incorporated in Singapore Electricity Market rules.
c. Europe: Support for voltage, frequency and system restoration is categorization of services in Europe. For maintaining grid security, infallibility, calibration of power, these services are an essential commodity.
d. North America: The United States categorized ancillary system according to US electricity regulation frequency response, spinning reserve, non-spinning reserve, replacement reserve, reactive supply, and voltage control. While in Canada Alberta Electric system Operator ancillary services are defined as satisfactory levels of service with acceptable levels of voltage and frequency support, inertia services are not mandatory [24].

4 Conclusions

The current scenario of the world toward increasing pollution, statuary warnings from the scientist and environmentalist all over the world caused serious concerns. These concerns had raised the consciousness of the authors to dig into the reality of the situation. The authors have vindicated these facts with the literature survey. This literature review guided the authors to reach a conclusion that there is an escalation of energy consumption overall the world and this energy is generated generally from resources, which have led to carbon emission in the environment. Hence there is a paradigm shift toward renewable energy is noted overall the world. This shift is aided by technical, political, and geopolitical reforms. Authors concluded PV resources as the most promising one among the renewable energy sources considering the present scenario and future investments into the projects. While considering generation with nonconventional energy sources authors raised concerns further. That is the fact that the conventional power system was designed keeping in view the conventional energy sources having a large synchronous generator attached with them. The merit of these generators is that they provide system inertia, which helps to restore frequency response during an excursion. However, while considering renewable energy source it was found that these are the system with feeble inertia. Also, ancillary service over different regions was reviewed. After reviewing the literature, the authors wish to carry out work in system designing to provide additional inertia with the help of MATLAB/PSAT. Also, authors will carry out the study of ancillary services and system optimization using linear optimization techniques.

References

1. Global Energy Statistics 2018. www.yearbook.enerdata.net
2. Reasons for increase in demand for energy. Part of geography, energy, https://www.bbc.com/bitesize/guides/zpmmmp3/revision/1
3. World Energy Resources 2016. World Energy Council
4. "Energy Statistics 2018", Ministry of Statistics and Programme Implementation, Government of India
5. ET Bureau, India's energy consumption to grow faster than major economies, 27 Jan 2017, https://economictimes.com.indiatimes.com/industry/energy/oil-gas/indias-energy-consumption-to-grow-faster-than-major-economies/articleshow/56800587.cms
6. I. Bozkurt, Energy resources and their effects on environment. WSEAS Trans. Environ. Dev. **6**(5), 327–334 (2010)
7. Sarath K. Guttikunda, Puja Jawahar, Atmospheric emissions and pollution from the coal fired thermal power plants in India, Elsevier. Atmospheric Environment **92**, 449–460 (2014)
8. M.E. Munawer, Human health and environmental impacts of coal combustion and post combustion wastes. J. Sustain. Min. **17**(2), 87–96 (2018)
9. U.C. Mishra, Environmental impact of coal industry and thermal power plants in India. J. Environ. Radioact. **72**(1–2), 35–40 (2004). ISSN 0265-931X
10. Manoj Kumar, Rajesh Kumar, Perspectives of renewable energy in 21st century in India: statistics and estimation. WASET Int. J. Energy Power Eng. **11**(6), 747–755 (2017)

11. P.F. Nelson, P. Shah, V. Strezov, B. Halliburton, J.N. Carras, Environmental impacts of coal combustion: a risk approach to assesment of emissions. Fuel **89**(4), pp. 810–816 (2010). ISSN 0016-2361
12. J. Watts, We have 12 years to limit climate change catastrophe warns UN. The Guardian, Global Environment Editor, 8 Oct 2018
13. GlobalData Energy, Renewable energy to reach 22.5% share in global power mix in 2020, 16 July 2018, https://www.power-technology.com/comment/renewable-energy-reach-22.5-share-global-power-mix-2020/
14. R.R. Kothari, K.C. Singhal, *Renewable Energy Sources and Technology* (PHI, 2007)
15. D.P. Kothari, *Modern Power System Analysis*, 4th edn. (Tata Mc Graw Hill, New Delhi, India, 2011)
16. J. Nelson, *The Physics of Solar Cell*, vol. 57 (Imperial College Press, London, 2003)
17. R.A. Messenger, J. Ventre, *Photovoltaics System Engineering*, 3rd edn (CRC Press, 2003)
18. J.L. Grey, The physics of solar cell, in *Handbook of Photovoltaic Science and Engineering* ed. by A. Luque, S. Hegedus, chap. 3 (John Wiley and Sons, Chichester, West Sussex, England, 2003), pp. 61–112
19. https://news.energysage.com/solar-panel-efficiency-cost-over-time/
20. Growth of electricity sector in India from 1947–2018. Ministry of Power, Central Electricity Authority, New Delhi, Government of India
21. Annual report, Ministry of New and Renewable Energy, 2017–18
22. H. Beck, R. Hesse, Virtual synchronous machine, in *2007 9th International Conference on Electrical Power Quality and Utilisation*, Barcelona (2007), pp. 1–6
23. P. Tielens, D. Van Hertem, Grid inertia and frequency control in power systems with high penetration of renewables, in *Young Researchers Symposium in Electrical Power Engineering,* vol. 93, no. 101, pp. 1–6 (2012)
24. C. Seneviratne, C. Ozansoy, Frequency response due to a large generator loss with the increasing penetration of wind/PV generation—a literature review, Elsiever. Renew. Sustain. Energy Rev. **57**, 659–668 (2016)
25. A. Ulbig, T.S. Borche, G. Andersson, Impact of low rotational inertia on power system stability and operation. IFAC Proc. **47**(3), 7290–7297 (2014). ISSN 1474-6670
26. M. Rezkalla, M. Marinelli, M. Pertl, K. Heussen, Trade-off analysis of virtual inertia and fast primary frequency control during frequency transients in a converter dominated network, in *2016 IEEE Innovative Smart Grid Technologies—Asia (ISGT-Asia)*, Melbourne, VIC (2016), pp. 890–895
27. M. Dreidy, H. Mokhlis, S. Mekhilef, Inertia response and frequency control techniques for renewable energy sources: A review. Renew. Sustain. Energy Rev. **69**, 144–155 (2017)
28. Ye Wang, Vera Silva, Miguel Lopez-Botet-Zulueta, Impact of high penetration of variable renewable generation on frequency dynamics in the continental Europe interconnected system. IET Renew. Power Gener. **10**(1), 10–16 (2016)
29. S. You et al., Impact of high PV penetration on the inter-area oscillations in the U.S. eastern interconnection. IEEE Access **5**, 4361–4369 (2017)
30. R. Yan, T. Kumar Saha, N. Modi, N. Masood, M. Mosadeghy, The combined effects of high penetration of wind and PV on power system frequency response. Appl. Energy **145**, 320–330 (2015)
31. M. Yagami, N. Kimura, M. Tsuchimoto, J. Tamura, Power system transient stability analysis in the case of high-penetration photovoltaics, in *2013 IEEE Grenoble Conference*, Grenoble (2013), pp. 1–6
32. G. Delille, B. Francois, G. Malarange, Dynamic frequency control support by energy storage to reduce the impact of wind and solar generation on isolated power system's inertia. IEEE Trans. Sustain. Energy **3**(4), 931–939 (2012)
33. J.W. Smith, W. Sunderman, R. Dugan, B. Seal, Smart inverter volt/var control functions for high penetration of PV on distribution systems, in *2011 IEEE/PES Power Systems Conference and Exposition*, Phoenix, AZ (2011), pp. 1–6

34. Y. Liu, J. Bebic, B. Kroposki, J. de Bedout, W. Ren, Distribution system voltage performance analysis for high-penetration PV, in *2008 IEEE Energy 2030 Conference*, Atlanta, GA (2008), pp. 1–8
35. H. Thiesen, C. Jauch, A. Gloe, Design of a system substituting today's inherent inertia in the European continental synchronous area. Energy **9**(18), 582 (2016)
36. J. von Appen, M. Braun, T. Stetz, K. Diwold, D. Geibel, Time in the sun: the challenge of high PV penetration in the German electric grid. IEEE Power Energy Mag. **11**(2), 55–64 (2013)
37. M.P.N. van Wesenbeeck, S.W.H. de Haan, P. Varela, K. Visscher, Grid tied converter with virtual kinetic storage, in *2009 IEEE Bucharest PowerTech*, Bucharest (2009), pp. 1–7
38. M. Torres, L.A.C. Lopes, Frequency control improvement in an autonomous power system: an application of virtual synchronous machines, in *8th International Conference on Power Electronics-ECCE Asia*, 30 May–3 June 2011, pp. 2188–2195
39. Q. Zhong, G. Weiss, Synchronverters: inverters that mimic synchronous generators. IEEE Trans. Ind. Electron. **58**(4), 1259–1267 (2011)
40. V. Calderaro, V. Galdi, F. Lamberti, A. Piccolo, A smart strategy for voltage control ancillary service in distribution networks. IEEE Trans. Power Syst. **30**(1), 494–502 (2015)
41. S. Efekharnejad, V. Vittal, G.T. Heydt, B. Keel, J. Loehr, Impact of increased penetration of photovoltaic generation on power systems. IEEE Trans. Power Syst. **28**(2) (2013)
42. IEEE 1547 Standard for Interconnecting Distributed Resources With Electric Power System, Oct. 2003, http://grouper.ieee.org/groups/scc21/1547/1547/index.html
43. UL 1741 Standard for Inverters, Converters, Controllers, and Interconnection System Equipment for Use with Distributed Energy Resources, https://ulstandardsinfonet.ul.com/scopes/1741.html
44. Y. Zhang, C. Mensah-Bonsu, P. Walke, S. Arora, J. Pierce, Transient Over—Voltages in High Voltage Grid Connected PV solar Interconnection. IEEE
45. H. Liu, L. Jin, D. Le, A.A. Chowdhury, Impact of high penetration of solar photovoltaic generation on power system small signal stability, in *2010 IEEE International Conference on Power System Technology*
46. F. Gonzlez-Longatt, Frequency control and inertial response schemes for the future power networks, in *Large Scale Renewable Power Generation, Green Energy and Technology* (Springer Science+Business Media Singapore, 2014)
47. K. Turitsyn, P. Šulc, S. Backhaus, M. Chertkov, Distributed control of reactive power flow in a radial distribution circuit with high photovoltaic penetration, in *IEEE PES General Meeting*, Providence, RI (2010), pp. 1–6
48. P. Kundur, *Power System Stability and Control* (Tata McGraw Hill, New Delhi, 2007)
49. G. Delille, B. François, G. Malarange, Dynamic frequency control support: a virtual inertia provided by distributed energy storage to isolated power systems, in *2010 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT Europe)*, Gothenberg (2010), pp. 1–8
50. A.G. Madureira, J.A. Peças Lopes, Ancillary services market framework for voltage control in distribution networks with microgrids. Electr. Power Syst. Res. **86**, 1–7 (2012)
51. Yann G. Rebours, Daniel S. Kirschen, Marc Trotignon, Sebastian Rossignol, A survey of frequency and voltage control ancillary services—Part I: technical feautres. IEEE Tran. Power Syst. **22**(1), 350–357 (2007)
52. E. Ela, B. Kirby, N. Navid, J.C. Smith, Effective ancillary services market designs on high wind power penetration systems, in *2012 IEEE Power and Energy Society General Meeting*, San Diego, CA (2012), pp. 1–8
53. B. Kirby, E. Hirst, *Ancillary service details: Voltage control*. United States: N. p., (1997), https://doi.org/10.2172/607488
54. Yann G. Rebours, Daniel S. Kirschen, Marc Trotignon, Sebastian Rossignol, A survey of frequency and voltage control ancillary services—Part II: economic feautres. IEEE Tran. Power Syst. **22**(1), 358–366 (2007)
55. E.G. Read, Co-optimization of energy and ancillary service markets, in *Handbook of Power Systems I. Energy Systems*, ed. by P. Pardalos, S. Rebennack, M. Pereira, N. Iliadis (Springer, Berlin, Heidelberg, 2010)

56. R. Raineri, S. Ríos, D. Schiele, Technical and economic aspects of ancillary services markets in the electric power industry: an international comparison. Energy Policy (Elseiver) **34**(13), 1540–1555 (2006)
57. J. Riesz, J. Gilmore, I. MacGill, Frequency control ancillary service market design: insights from the Australian national electricity market. Electr. J. **28**(3), 86–99 (2015). ISSN 1040-6190
58. Daniel S. Kirschen, Goran Strbac, *Fundamentals of Power System Economics* (Wiley, Ltd, 2004)
59. E. Ela, V. Gevorgian, A. Tuohy, B. Kirby, M. Milligan, M. O'Malley, Market designs for the primary frequency response ancillary service—part I: motivation and design. IEEE Trans. Power Syst. **29**(1), 421–431 (2014)

Analysis of 150 kW Grid-Connected Solar PV System Using Fuzzy Logic MPPT



Ashutosh Tiwari and Ashwani Kumar

1 Introduction

There are various sources of electricity generation in which one of the main sources is fossil fuels. Due to the depletion of fossil fuels, which is a major source of GHG emissions, electricity generation is shifting toward renewable energy sources [1]. Solar energy is the most widely used renewable energy source for electricity generation as it is free of pollution and require less maintenance. In this paper, solar-connected grid system is discussed. Photovoltaic system converts the direct sunlight into electricity [2].

PV array is the connection of a number of modules in series or parallel. The panel is connection of a number of solar cells in series or parallel [3]. A solar cell is made up of semiconductor through which sunlight passes [4]. With the change in environment conditions output power generated always changes, and hence to track the maximum power, different MPPT methods are used irrespective of these conditions [2]. Various perturb steps have been discussed to increase efficiency of MPPT [5, 6]. In a grid-connected solar PV system, electricity generated by solar system can be sent to the grid and can be taken by grid depending upon the requirement. For these systems, battery storage requirements are not needed so power losses in storage are reduced.

In solar photovoltaic system, MPPT is used for increasing the efficiency of photovoltaic system [7]. There are various MPPT techniques such as perturb and observe, incremental conductance, fuzzy logic based, and neural network MPPT methods [8]. Comparison of different MPPTs is discussed [9]. For variable step incremental conductance, MPPT is used, but it is more complex as compared to P and O MPPT and

A. Tiwari (✉) · A. Kumar
National Institute of Technology, Kurukshetra 136119, Haryana, India
e-mail: asutosh.tiwari0@gmail.com

A. Kumar
e-mail: ashwaks@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_133

for varying climate conditions fuzzy logic-based MPPT exhibits good performance as compared to others but it is more expensive due to its design complexity [10]. Fuzzy logic-based controller for stand-alone and grid-connected system is discussed [11]. The perturb and observe method is widely used because it is cheap and can be easily implemented but it oscillates around the maximum power point so it causes power losses. On comparing P and O and INC, INC exhibits better performance, but it is more complex [5].

In this paper, 150 kW grid-connected solar system has been designed with the use of MATLAB/Simulink. The proposed system consists of a solar photovoltaic array, MPPT for extracting maximum power from the solar system, boost converter for regulating and boosting the solar system output, voltage source inverter for inverting the DC output into AC for supplying to the grid, and LC filter for filtering current harmonics from the output of inverter current. Section 2 presents the detailed structure of proposed systems, Sect. 3 describes the simulation and experimental result of model, and finally Sect. 4 describes the conclusion.

2 Methodology

2.1 Modeling of PV Array

A photovoltaic array is the connection of a number of modules or panels connected in series/parallel. A PV module can be defined as the number of series and parallel connection of solar cell which, in turn, is a P-N junction that converts the sunlight into electrical energy. If the intensity of solar rays is higher, current from solar cell is more. Figure 1 shows the circuit of a solar cell.

Now from figure,

$$I = I_{ph} - I_d - I_p \tag{1}$$

For a solar cell equation,

$$I_d = I_0\left[e^{\frac{V+I_p}{nV_T}} - 1\right] \tag{2}$$

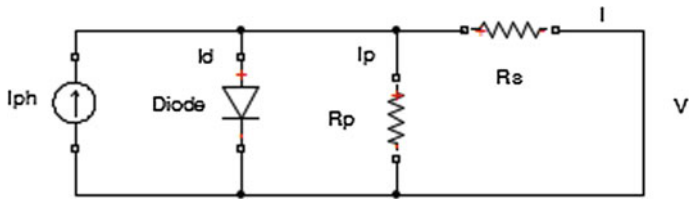


Fig. 1 Circuit of solar cell

$$I_p = \frac{V + IR_s}{R_p} \tag{3}$$

Equation (1) can now be written as

$$I = I_L - I_0\left[e^{\frac{V+I_{sh}}{nV_T}} - 1\right] - \frac{V + IR_s}{R_{sh}} \tag{4}$$

where

I_0	Reverse saturation current,
R_s and R_p	series and parallel resistances of solar cell, respectively,
n	Diode ideality factor, and
V_T	Thermal voltage.

And V_T is given by

$$V_T = \frac{KT}{q} \tag{5}$$

K = Boltzmann Constant = 1.38×10^{-23} J/K and
$q = 1.6 \times 10^{-19}$ C.

A P–V and I–V characteristic of the PV system is shown in Fig. 2.

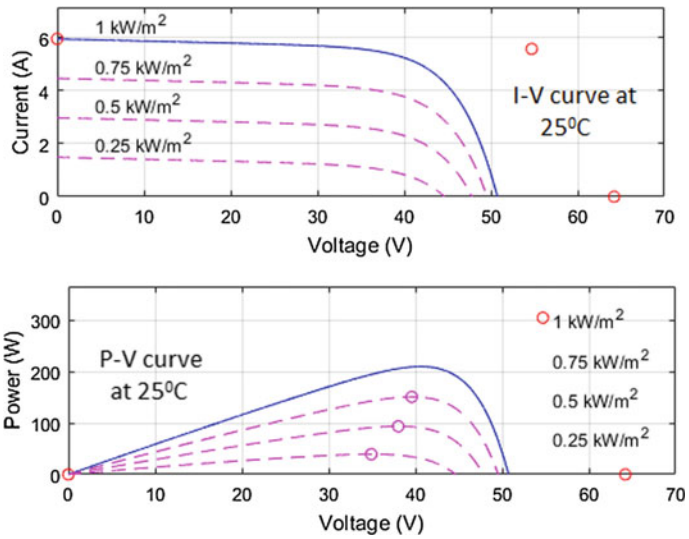


Fig. 2 P–V and I–V characteristics at 25 °C temperature

2.2 Maximum Power Point Tracking (MPPT)

For extracting the maximum power from the array, different types of MPPT are used. For varying environmental conditions, it is necessary to use MPPT. There are different types of MPPT methods available such as fuzzy logic based, neural network based, perturb and observe based, INC, etc. [12]. Among all MPPTs, fuzzy logic-based MPPT is best and it is more efficient than others but its drawback is that it is more expensive than others [13]. Perturb and observe is mostly used because of its easiness and INC type is better in performance than P and O.

Fuzzy Logic-based MPPT: As with the variation of temperature and irradiance, response of perturb- and observe-based and INC-based MPPT methods are slow, and so to improve these problems fuzzy logic-based methods are used. This method is having fast response as compared to others [14]. It also increases the stability of the system. There is no need of mathematical modeling in this case. Figures 3 and 4 are the fuzzy logic designer and membership function of fuzzy systems.

There are three stages of fuzzy logic controller: fuzzification, logic building, and defuzzification. In fuzzification *E,* i.e., $\frac{\Delta P}{\Delta V}$ is coded into linguistic variable through membership function as shown [9]. There are two inputs *E* and *CE* and the duty

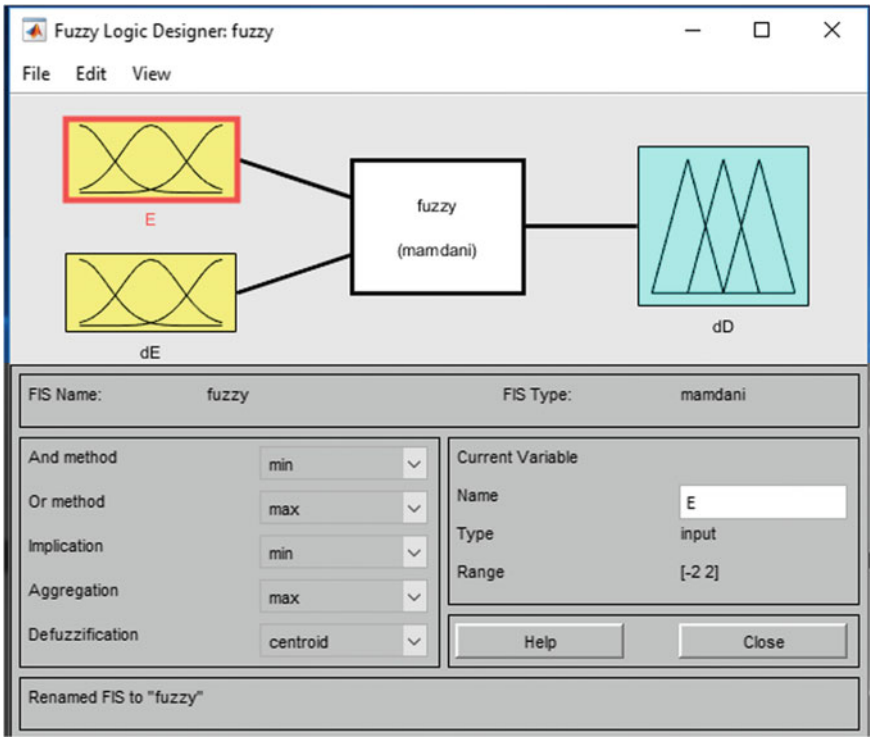


Fig. 3 Fuzzy logic designer

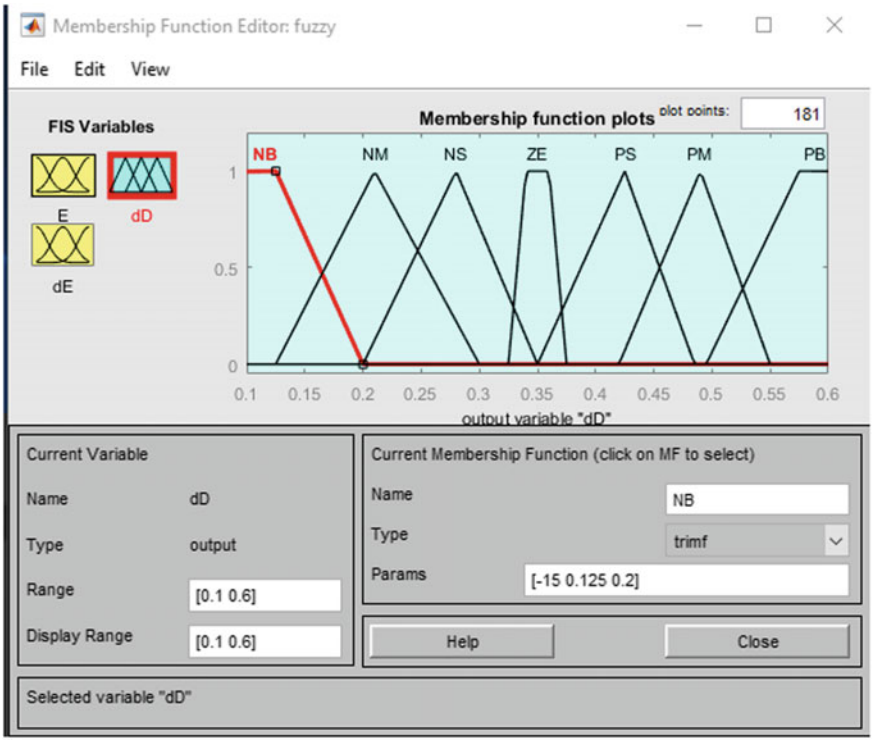


Fig. 4 Membership function of fuzzy

cycle, i.e., D is the output which is a numeric variable. E and CE can be written as

$$E = \frac{P[k] - P[k-1]}{V[k] - V[k-1]}$$
$$CE = V[k] - V[k-1] \tag{6}$$

Simulink circuit of the fuzzy logic-based MPPT is shown in Fig. 5.

2.3 Boost Converter

Boost converter is a DC converter which is used to increase the voltage level. With this converter, $Vout > Vin$ and $Iout < Iin$ [15]. Boost converter consists of IGBT and diode as shown in figure. It is also used to regulate the voltage of PV system. By adjusting the duty cycle of MPPT for different environmental conditions, we can get maximum power [16]. Simulink diagram of boost converter is shown in Fig. 6.

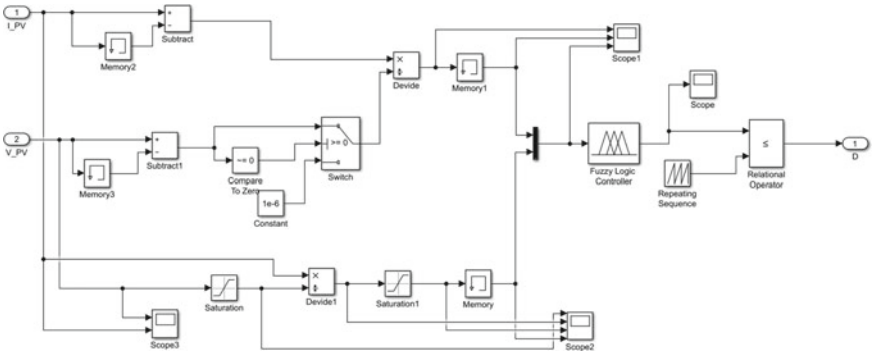


Fig. 5 Simulink diagram of fuzzy logic-based MPPT

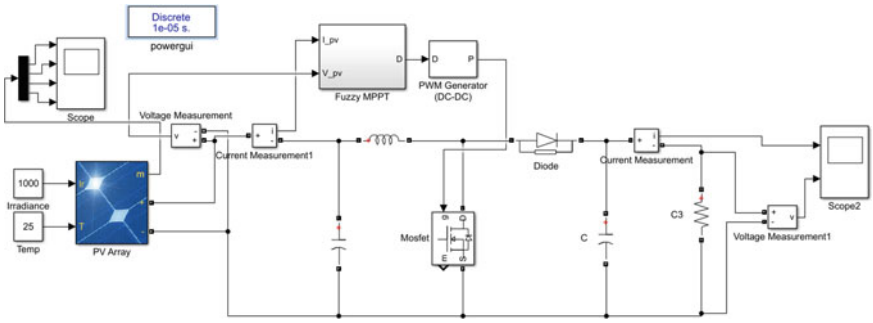


Fig. 6 Simulation model of PV array with boost converter

2.4 Three-Level Inverter

Three-phase VSC is used for converting the DC level into AC for supplying the power generated by the solar array to the grid. Simulink block of three-phase VSC is shown in Fig. 7. Two capacitors are used at the output of boost converter [17]. These capacitors provide a neutral point for the inverter. Since IGBT is having fast switching speed, it is used for high power level [18]. For controlling the inverter, HCC is used. Hysteresis current controller (HCC) provides the gate signal for voltage source inverter [19]. Six IGBTs are used in VSC for converting the DC output into AC. Hysteresis current controller provides pulse width-modulated (PWM) signal to the inverter [20]. HCC i.e. is PI controller works the basis of the current error. When load current is compared with band limit, the error signal is produced, i.e., called switching pulses for inverter.

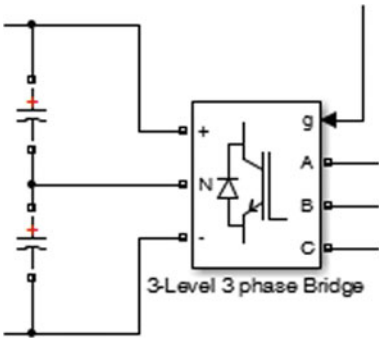


Fig. 7 Simulink block of three-phase three-level inverter

3 Simulation and Result

In this paper, 150 kW grid-connected Simulink model is designed with the use of MATLAB/Simulink software. The complete Simulink model is given in Fig. 8. The array consists of 71 parallel strings and 7 series string and each consists of 96 panels. Irradiance and temperature used for this system are 1000 W/m^2 and 25 °C. Fuzzy logic-based MPPT is used which is having better efficiency. With the help of the duty ratio, boost converter is used to increase the voltage level.

Results are compared with "Fuzzy Logic Controller based PV System Connected in Standalone and Grid Connected Mode of Operation with Variation of Load" [11]. The obtained results are matched with the reference paper, and then applied for large model.

Waveform of mean power of the PV array and output voltage of the boost converter voltage are shown in Fig. 9. Inverter input and output voltage are shown in Fig. 10. Waveform of grid voltage and current is shown in Fig. 11. Also, the output power at the bus which is supplied to the grid is shown in Fig. 12.

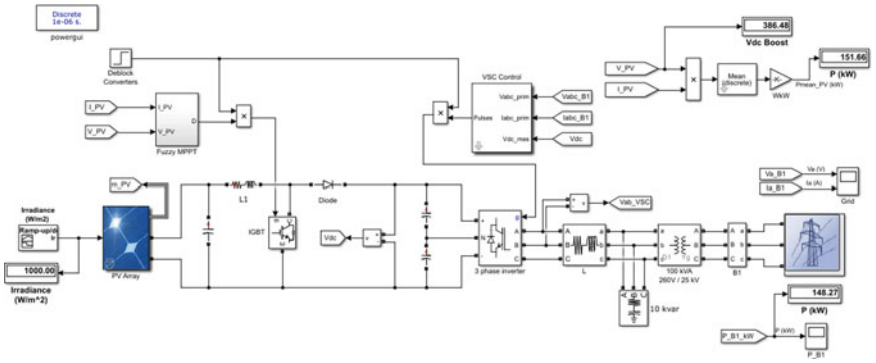


Fig. 8 Complete Simulink model

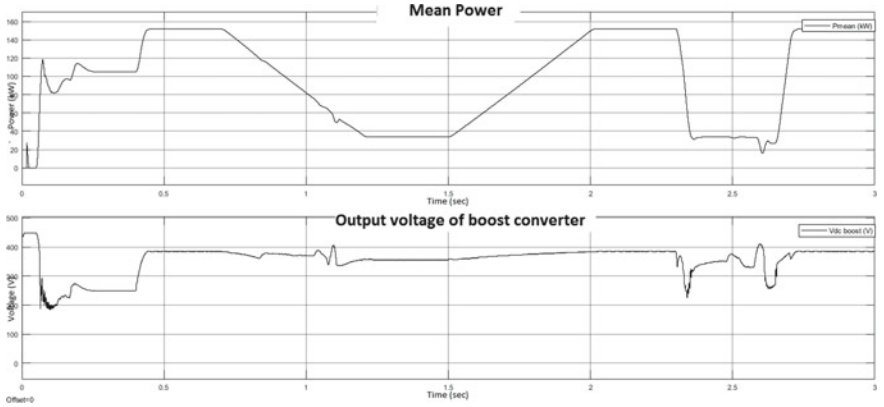


Fig. 9 Mean power of PV array and output voltage of boost converter

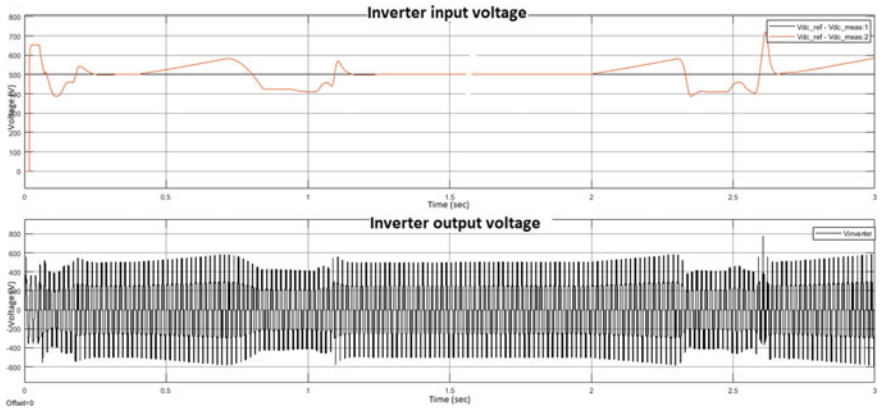


Fig. 10 DC reference voltage and output voltage of inverter

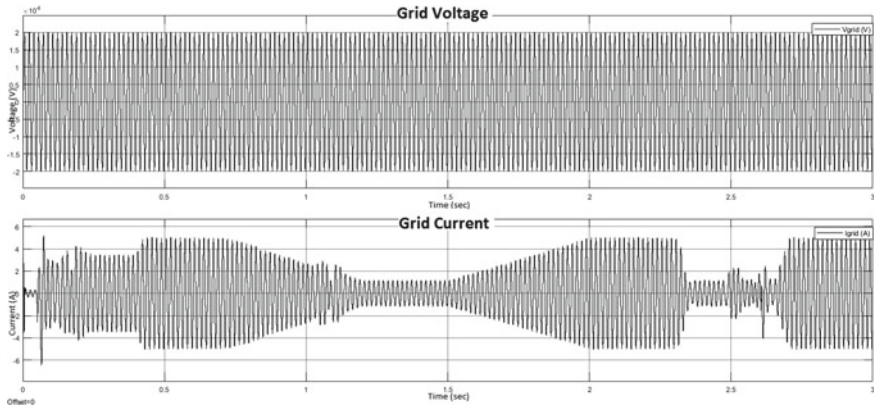


Fig. 11 Waveform of grid voltage and current

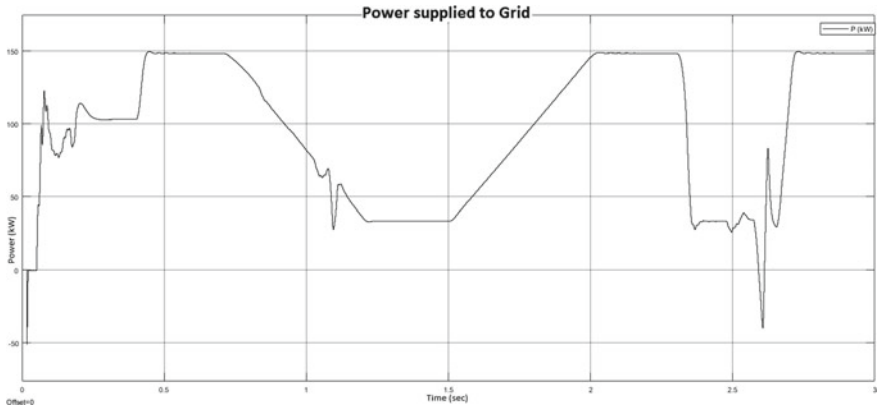


Fig. 12 Waveform of power at bus connected to grid

4 Conclusion

In this paper, 150 kW solar photovoltaic grid-connected system is designed in MATLAB/Simulink. MPPT is used here by fuzzy type which makes the system more efficient. These can easily and better perform in varying climate conditions. Boost converter is used for increasing the voltage level and also for the controlling purpose. The output of this DC converter is given to voltage source inverter for converting the DC input into AC output. AC is transferred to grid connected. Hysteresis current controller is used for controlling the pulse for voltage source inverter. A phase logic loop is used to track the phase and frequency for creating pulses for inverter. These types of controllers are easy to implement and are having better response in speed.

References

1. M.E. Ropp, S. Gonzalez, Development of a MATLAB/Simulink model of a single-phase grid-connected photovoltaic system. IEEE Trans. Energy Convers. **24**(1), 195–202 (2009)
2. M. Bharathkumar, H.V. Byregowda, Performance evaluation of 5 MW grid connected solar photovoltaic power plant established in Karnataka. Int. J. Innov. Res. Sci. Eng. Technol. **3**(6) (2014)
3. E. Isen, Modelling and simulation of hysteresis current controlled single-phase grid-connected inverter, in *Proceedings 17th International Conference on Electrical and Power Engineering*, Holland (2015)
4. R. Faranda, S. Leva, A comparative study of MPPT techniques for PV systems, in *7th WSEAS International Conference on Application of Electrical Engineering (AEE'08)*, Trondheim, Norway (2008)
5. A.K. Abdelsalam, A.M. Massoud, S. Ahmed, P.N. Enjeti, High-performance adaptive perturb and observe MPPT technique for photovoltaic-based microgrids. IEEE Trans. Power Electron. **26**(4), 1010–1021 (2011). https://doi.org/10.1109/TPEL.2011.2106221

6. N. Femia, G. Petrone, G. Spagnuolo, M. Vitelli, Optimization of perturb and observe maximum power point tracking method. IEEE Trans. Power Electron. **20**(4), 963–973 (2005). https://doi.org/10.1109/TPEL.2005.850975
7. G. Mamatha, Perturb and observe MPPT algorithm implementation for PV applications. Int. J. Comput. Sci. Inf. Technol. **6**(2), 1884–1887 (2015)
8. E. Roman et al. Intelligent PV module for grid-connected PV systems. IEEE Trans. Ind. Electron. **53**(4), 1066–1073 (2006)
9. D.P. Hohm, M.E. Ropp, Comparative study of maximum power point tracking algorithms. Prog. Photovolt: Res. Appl. **11**, 47–62 (2002). https://doi.org/10.1002/pip.459
10. Z. Ahmad, S.N. Singh, Modeling and control of grid connected photovoltaic system-a review. Int. J. Emerg. Technol. Adv. Eng. **3**(3), 2250–2459 (2013)
11. B. Krishna Naick, T.K. Chatterjee, K. Chatterjee, Fuzzy logic controller based PV system connected in standalone and grid connected mode of operation with variation of load. Int. J. Renew. Energy Res. **7**(1) (2017)
12. B.S. Kumar, K. Sudhakar, Performance evaluation of 10 MW grid connected solar photovoltaic power plant in India. Energy Rep. **1**, 184–192 (2015)
13. B. Subudhi, Senior Member, IEEE, R. Pradhan, A comparative study on MPPT techniques for PV power systems. IEEE Trans. Sustain. Energy **4**(1) (2013)
14. M. Salhi,. R. El-Bachtiri, Maximum power point tracking controller for PV systems using a PI regulator with boost DC/DC converter, ICGST-ACSE J. **8**(3) (2009)
15. J. Atiq, P.K. Soori, Modelling of a grid connected solar PV system using MATLAB/Simulink. Int. J. Simul. Syst. Sci. Technol. **17**(41) (2016)
16. N.R. Jalakanuru, Performance study of incremental inductance algorithm for PV applications. Int. J. Sci. Eng. Technol. **5** (2016)
17. J. Kasera, V. Kumar, R.R. Joshi, J.K. Maherchandani, Design of grid connected photovoltaic system employing increamental conductance MPPT algotithm. J. Electr. Eng. **12**, 172–177 (2012)
18. M. Salhi, R. El-Bachtri, Maximum power point tracker using fuzzy control for photovoltaic system. Int. J. Res. Rev. Electr. Comput. Eng. IJRRECE **1**(2) (2011)
19. F. Liu, S. Duan, F. Liu, B. Liu, Y. Kang, A variable step size INC MPPT method for PV systems. IEEE Trans. Ind. Electron. **55**(7), 2622–2628 (2008)
20. A.S Swathy, R. Archana, MPPT using modified incremental inductance for solar PV system. Int. J. Eng. Innov. Technol. **3**(2), (2013)

Clock System Architecture for Digital Circuits

Amit Saxena, Kshitij Shinghal, Rajul Misra and Alok Agarwal

1 Introduction

In digital synchronous circuits, a periodic clock is generally used to synchronize or provide timing to the circuit. Generally, a typical synchronous circuit is characterized by its setup and hold time. If the data applied at the input changes before the setup time, the output will also change accordingly to reflect the input after a finite propagation delay [1, 2]. However, if the input changes within the aperture of setup and hold times, the output data may be unpredictable. Therefore, for keeping the output predictable and make the circuit behave as desired, it becomes essential that circuit is designed in a way that it ensures that data is synchronized with clock and changes within the aperture of setup and hold time [3]. Figure 1 gives the typical example of digital circuit where all inputs I_0 to I_6 are coming from flip-flops and outputs O_9 and O_{12} are also stored in flip-flop. The arrival times at input and time required at the output are also shown. The arrival time ai at internal node i depends upon the propagation delay of the circuit.

The slack is defined as the difference between arrival time and required time. If slack is positive, it means that the circuit is fast enough and it meets timing, whereas

A. Saxena (✉) · K. Shinghal
Department of Electronics & Communication Engineering, Moradabad Institute of Technology, Moradabad, UP, India
e-mail: amitssaksena@gmail.com

K. Shinghal
e-mail: kshinghal@gmail.com

R. Misra · A. Agarwal
Department of Electrical Engineering, Moradabad Institute of Technology, Moradabad, UP, India
e-mail: rajulmisra71@gmail.com

A. Agarwal
e-mail: alokagarwal2@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_134

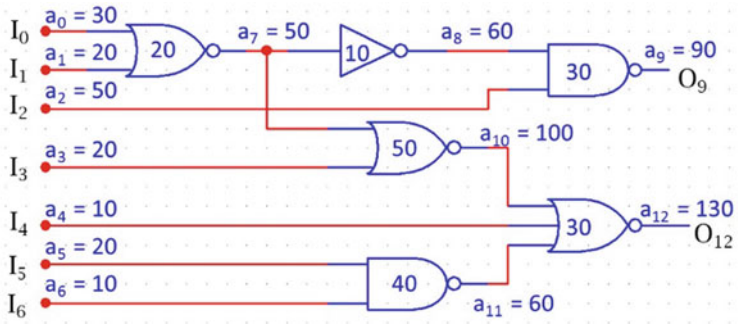


Fig. 1 Typical digital circuit with various delays

its slack is negative then it means circuit does not meet the timing requirement and not fast enough. This also means that a node is on critical path limiting the operating speed of the circuit and requires tweaking in circuit topology. The system clock plays a very important role in computation from one step to another. Ideally, the arrival time of all clocked elements should be same so that the complete circuit elements share a common time Ref. [4]. These circuit elements include latches, flip-flops, memories, registers, dynamic gates, etc.

The paper is organized as follows: the survey of the existing work in the area is carried out in Sect. 2, clock distribution network is given in Sect. 3. In Sect. 4, the proposed clock system architecture (CSA) is presented, simulation setup is presented in Sect. 5 followed by results and discussion of the proposed work in Sect. 6. Finally, the conclusions are drawn in the last section of the paper.

2 Related Work

Zarkesh-Ha et al. in their paper described a few major bottlenecks to the advancing performance of future ICs. They planned a new global interconnected design for the digital system designed and integrated on a VLSI chip consisting of signal, clock, and power-supply distribution networks [1]. Teichmann et al. in their work described about power-clock gating in adiabatic logic circuits. They mentioned that for standard CMOS, clock gating could be a standard methodology for reducing dynamic power consumption, whereas to cut back static power consumption, power gating could be a well-known methodology. They planned in their work for the primary time clock gating and power gating in adiabatic logic [4]. Wang et al. in their paper proposed two-phase sinusoidal power-clock generator. They also suggested adiabatic logic circuit adopting the two-phase sinusoidal power clocks—clocked transmission gate adiabatic logic (CTGAL) circuit [2]. Ji et al. in their paper found that employing an MCDE clock network microarchitecture, with a fine-grained adaptive

dynamic adjustment formula, will effectively decrease the microchip power by four-hundredth, compared with the initial EPIC clock network microarchitecture [3]. Strak et al. in their paper proposed that all variations of affiliation configurations within the clock generation circuits are explored to reveal potential optimum configurations [5]. Bonanno et al. in their paper proposed a technique applied on a digital filter used at intervals an ultra-low-power industrial design; comparison with alternative customary and advanced automatic clock-gating ways highlights the effectiveness of the projected technique [6]. Nagaraju et al. in their paper proposed the design of a process-variation-tolerant delay-locked loop (DLL) for use in multiphase clock generation [7]. Keller and Chakravadhanula in their paper proposed design for low-power consumption throughout practical operation in CMOS devices by gate off clocks to areas of logic not required for the present state of operation. By gating off clocks to state components that are famous to not want change, the dynamic switch current is often reduced compared with permitting state components to update once you do not care what they contain [8].

Ahuja et al. in their paper proposed clock-gating primarily based power reduction flow. System-level simulations give steering to make clock-gated RTL [9]. Teichmann et al. in their paper proposed power-clock gating to be used as a switch to disconnect the power clock from the adiabatic circuit once no operations are performed within the system. Powering down the generator, as different to the switch within the power-clock line, was conferred and according power-down schemes ensuing from various usage of the synchronization signals for powering down [10]. Teichmann et al. in their paper proposed inductance primarily based generation of the power-clock signal and they further complete gain potency figures not only on circuit, however, conjointly for the dimension of the generator driver transistors to realize highest conversion potency [11]. Houri et al. in their paper proposed the power-clock generator on the worldwide energy-performance relationship of nanoelectromechanical. They found that the outflow current of the MOSFET shift devices used among the generator constitutes a very important supply of performance degradation [12]. Zahrai et al. in their paper proposed a clock generation system. To judge the feasibleness of the clock generation approach, they conducted post-layout simulations with the interconnected ADC core [13]. It is evident from the literature review that there exists a spot within the generation and development of clock design for digital circuit style. Based on the issues identified, this paper tries to propose a solution for clock design.

3 Clocked Distribution Network

Generally, digital system requires more than one logical clock. When these clocks traverse through mismatched clock network path and environmental variations, they undergo changes, do not arrive at idle time, and become skewed physical clocks.

Clock skew can be defined as difference between ideal and actual arrival time of clocks. Figure 2 shows a circuit with two flip-flops and two clocks with zero inter-

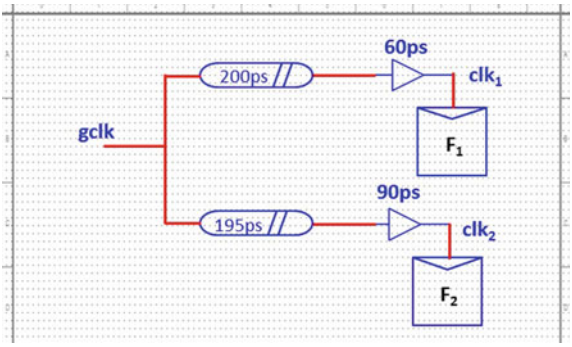


Fig. 2 Digital circuit with two flip-flops and two clocks with zero inter-arrival time

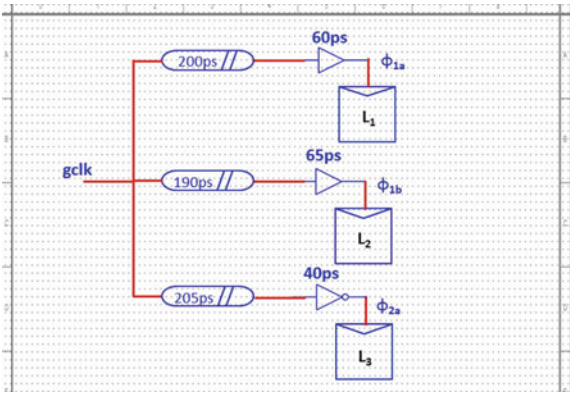


Fig. 3 Circuit with three latches with clock

arrival time but because of circuit characteristics clock 1 (clk1) arrives 25 ps before clock 2 (clk2). Therefore, the clock skew is 25 ps.

Figure 3 shows circuit with three latches with clock but it receives physical clocks with delay. We can intentionally insert these delays or clock skews to solve the setup or hold time problems for the circuit on the critical path by using a clock distribution network with clock gaters.

4 Proposed Clock System Architecture (CSA) with Clock Gaters

Figure 4 shows block diagram of the proposed clock subsystem with clock gaters for typical synchronous circuit. The circuit is designed to receive clock signal through input/output pins. This clock is then processed by a phase-locked loop (PLL) or a delay-locked loop (DLL). This clock is then distributed throughout the circuit to all the clocked elements. Clock gaters are used to receive the clock signal and modify it and distribute it accordingly. Figure 5 shows a typical clock distribution network.

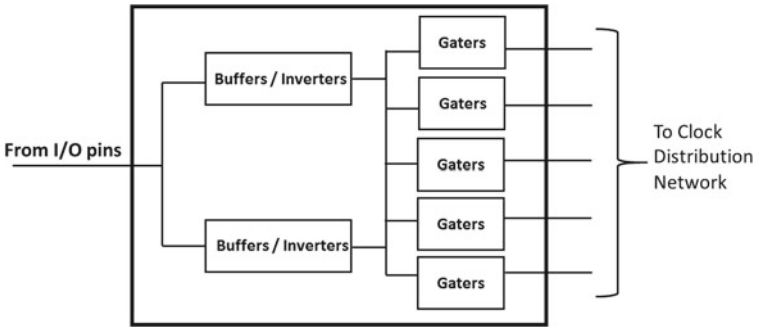


Fig. 4 Block diagram of the proposed clock subsystem with clock gaters

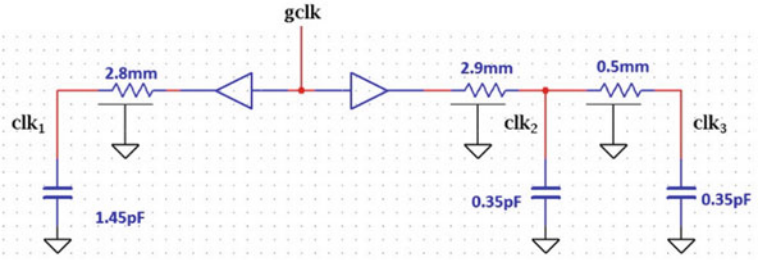


Fig. 5 Clock distribution network

Clock gaters are used to modify clock waveforms in various ways such as

1. Stop or gaining the clock to unused block in power saving modes.
2. Providing pulsed clocks.
3. For providing delayed clocks to match setup and hold times.
4. Providing stretched clocks to meet timing constraints of a node on critical path.
5. Provide nonoverlapping clocks.
6. Provide double frequency pulsed clocks.

5 Simulation Setup

For the design verification of proposed CSA, OrCad PSPICE 16.1 version was used. OrCad PSPICE is a software by Cadence used for simulation of circuits that allows users to test and verify designed circuits components on the basis of their SPICE models. The CSA proposed in this research work is simulated using conventional CMOS PSPICE models. The step-by-step verification of the circuits was done by applying voltage constraints of a standard conventional CMOS circuit. All the clock-modifying circuits shown in Fig. 6 were designed using CMOS circuits in OrCad

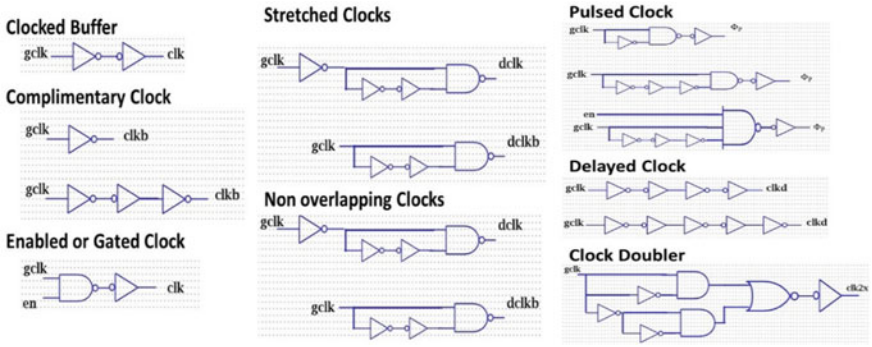


Fig. 6 Clock-modifying circuits (clock choppers/clock stretcher/clock gaters, etc.)

capture program with the integration of models given in library. Finally, the designed circuits were simulated using PSPICE A/D and the results were carefully recorded for functional verification.

6 Result and Discussion

Proposed CSA simulations and various waveforms for clock gaters were recorded. Figure 6 shows various clock-modifying circuits. Clock gaters can be used to introduce systematic delay between phases. Complimentary clock signal "clkb" is produced with three inverter while delayed clock signal "clkd" is produced with 4 or 5 inverter delays. Φ_P is pulsed clock, clh2x is clock doubler, and Φ_1 and Φ_2 are nonoverlapping clocks. Dclkb is stretched clock whereas gclk is gated clock.

The clock gater circuits were simulated and Fig. 7 shows the result. "Gclk" in Fig. 7 represents the original clock pulse generated by clock circuit, and it is fed to clock buffer circuit to generate delayed clock pulse represented by "clk". In present circuit, two buffers are considered, and more buffers can be introduced to increase the delay. Even number of buffers are introduced to get un-complimented clock.

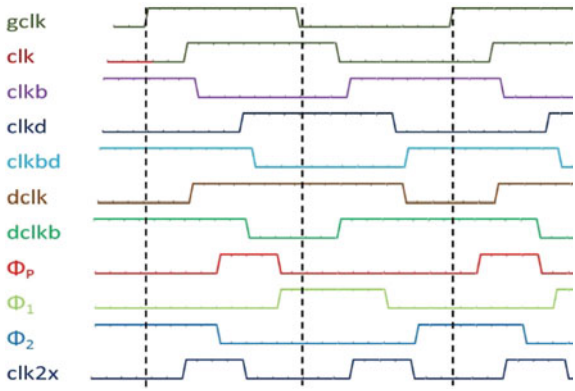


Fig. 7 Result verification waveforms of clock-modifying circuits

7 Conclusion

In this paper, a CSA for synchronous VLSI circuits was proposed. The clock architecture consisted of clock gaters, which can be successfully used for modifying the clock in different ways to meet the requirement of a typical synchronous system.

The CSA proposed in this paper gives a novel solution to the problem of clock gating for power reduction. The proposed clock system architecture was implemented using conventional CMOS circuits. This result verification waveform confirmed that proposed CSA is functioning as per specifications.

The proposed CSA suffers from the problem of synchronization and system timing. In future, a CSA architecture with clock tree can resolve the synchronization and system timing problems. However, this requires further modifications in clock system architecture.

References

1. P. Zarkesh-Ha, Power, clock, and global signal distribution, in *Interconnect Technology and Design for Gigascale Integration*, ed. by J. Davis, J.D. Meindl (Springer, Boston, MA, 2003)
2. P. Wang, J. Yu, J. Electr. (China) **24**, 225 (2007), https://doi.org/10.1007/s11767-005-0170-2
3. R. Ji, X. Zeng, L. Chen, J. Zhang, The implementation and evaluation of a low-power clock distribution network based on EPIC, in *Network and Parallel Computing. NPC 2007,* ed. by K. Li, C. Jesshope, H. Jin, J.L. Gaudiot. Lecture Notes in Computer Science, vol. 4672 (Springer, Berlin, Heidelberg, 2007)
4. P. Teichmann, J. Fischer, S. Henzler, E. Amirante, D. Schmitt-Landsiedel, Power-clock gating in adiabatic logic circuits, in *Integrated Circuit and System Design. Power and Timing Modeling, Optimization and Simulation. PATMOS 2005* ed. by V. Paliouras, J. Vounckx, D. Verkest. Lecture Notes in Computer Science, vol. 3728 (Springer, Berlin, Heidelberg, 2005)
5. A. Strak, A. Gothenberg, H. Tenhunen, Power-supply and substrate-noise-induced timing jitter in nonoverlapping clock generation circuits. IEEE Trans. Circ. Syst. I Regul. Pap. **55**(4), 1041–1054 (2008)

6. A. Bonanno, A. Bocca, A. Macii, E. Macii, M. Poncino, Data-driven clock gating for digital filters, in *Integrated Circuit and System Design. Power and Timing Modeling, Optimization and Simulation. PATMOS 2009*, ed. by J. Monteiro, R. van Leuken. Lecture Notes in Computer Science, vol. 5953 (Springer, Berlin, Heidelberg, 2010)
7. M. Nagaraju, W. Wu, C.T. Charles, Process-variation tolerant design techniques for multiphase clock generation, in *17th IEEE International Conference on Electronics, Circuits and Systems*, Athens (2010), pp. 102–105
8. B. Keller, K. Chakravadhanula, Test strategies for gated clock designs, in *Power-Aware Testing and Test Strategies for Low Power Devices*, ed. by P. Girard, N. Nicolici, X. Wen (Springer, Boston, MA, 2010)
9. S. Ahuja, A. Lakshminarayana, S.K. Shukla, System level simulation guided approach for clock-gating, in *Low Power Design with High-Level Power Estimation and Power-Aware Synthesis* (Springer, New York, NY, 2012)
10. P. Teichmann, Power-clock gating, in *Adiabatic Logic*. *Springer Series in Advanced Microelectronics*, vol. 34 (Springer, Dordrecht, 2012)
11. P. Teichmann, Generation of the power-clock, in *Adiabatic Logic*. *Springer Series in Advanced Microelectronics*, vol. 34 (Springer, Dordrecht, 2012)
12. S. Houri, G. Billiot, M. Belleville, A. Valentian, H. Fanet, Power-clock generator impact on the performance of NEM-based quasi-adiabatic logic circuits, in *Reversible Computation. RC 2015,* ed. by J. Krivine, J.B. Stefani. Lecture Notes in Computer Science, vol. 9138 (Springer, Cham, 2015)
13. S.A. Zahrai, N Le Dortz, M. Onabajo, Design of clock generation circuitry for high-speed subranging time-interleaved ADCs, in *2017 IEEE International Symposium on Circuits and Systems (ISCAS)*, Baltimore, MD (2017), pp. 1–4

TCAD Modeling and Analysis of sub-30nm Strained Channel MOSFET



Lalthanpuii Khiangte, Kuleen Kumar and Rudra Sankar Dhar

1 Introduction

Strained-silicon technology is no new concept for the improvement of device performance as also recently observed by Khiangte and Dhar [1]. Therefore, its wide range of applicability in various ways and its effects on the device operation made the research area of strained-Si technology more attractive for many scholars worldwide till date. Two types of results occur as strain is applied in the channel region of the HOI MOSFET, and these are band energy level shifting and splitting of the electronic states within the structure [2]. Therefore, the twofold valleys in the energy level shift down in contrast to the fourfold valleys at the conduction band edge due to the strained channel region, which instigate splitting of the energy bands, thereby increasing the occupancy of electrons in twofold valleys. This results in developing a twofold degenerate energy band, congregating in enhancement of electron mobility within the channel, and hence on reducing phonon scattering, it suppresses intervalley transition of electrons from lower to upper valley in the nanoregime of the MOSFET channel [3–7].

International Technology Roadmap for Semiconductors (ITRS) predicts the 40% development costs reduction in 2016 because of TCAD which confirmed the significance of TCAD tools in the global manufacturing industries [8–10]. This has given tremendous growth in design, development, and understanding of the device physics prior to manufacturing. Thus, all the possible device physics for the design structure must be described by the inclusion of appropriate rigorous models.

The development of the three-layered s-Si/s-SiGe/s-Si channel MOSFET was looked into [11, 12], which left with a question on its scalability for further miniaturization and enhancement of drive current.

L. Khiangte (✉) · K. Kumar · R. S. Dhar
Department of Electronics & Communication Engineering, National Institute of Technology Mizoram, Aizawl, India
e-mail: rdhar@uwaterloo.ca

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_135

2 Tri-Layer Channel MOSFET

2.1 Device Structure

Unlike any other strained heterostructure MOSFETs, a three-layered channel, s-Si/s-SiGe/s-Si, was developed by Khiangte and Dhar [11] with an indicative enhancement of drive current over the conventional strained-Si channel MOSFET device. The device design under investigation is shown in Fig. 1. As depicted, the channel is engineered by three layers: (i) upper strained-Si, (ii) strained-SiGe, and (iii) Lower strained-Si. The device dimensions and parameters used are tabulated in Table 1.

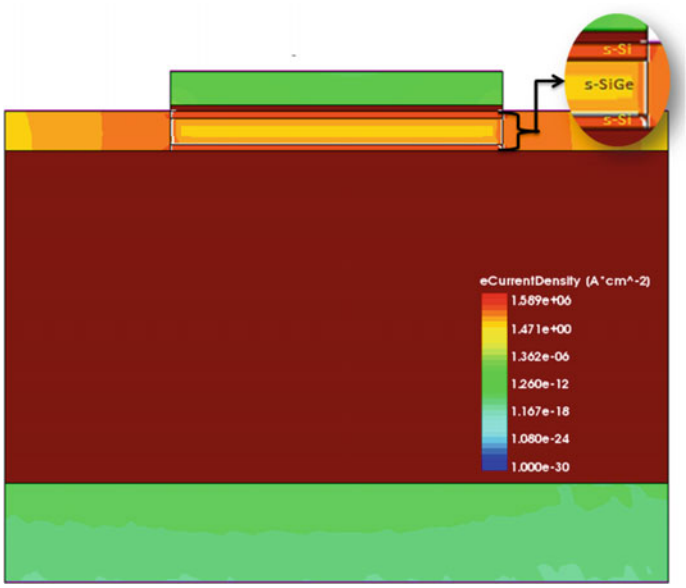


Fig. 1 Schematic of s-Si/s-SiGe/s-Si strained channel HOI MOSFET

Table 1 Device parameters

Parameters	Value
Strained-Si (t_{si})	1.5, 2 nm
Strained-SiGe (t_{sige})	3, 6 nm
Ge mole fraction	0.4
Gate-oxide	2 nm
Source/Drain doping (N_D)	10^{20} cm^{-3}
Channel doping	10^{17} cm^{-3}
Drain bias	50 mV

Both the strained-Si layers and the strained-SiGe layers are of 1.5 nm and 3 nm thickness, respectively.

2.2 Simulation Approach

The device structure and concept described earlier have been designed and simulated using Synopsis TCAD tool [13, 14]. To model such structure with inclusion of the possible physics completely is a handful task, i.e., different models are required for the band structure, mobility, quantization, effective mass, etc.

Starting with the strain-induced change of the silicon band structure, several valleys in the conduction and valence bands are needed to be considered for computing the correct dependency of the carrier concentration on the quasi-Fermi level. So, a multivalley band structure model has been included along with the Modified Local-Density Approximation (MLDA) model that calculates the confined carrier distributions that occur near semiconductor–insulator interfaces [13]. This model considers a dependency of the quantization effect on interface orientation and stress. Also, the effective mass changes for both carriers are incorporated. Additional 1D Schrodinger model for carrier confinement in the thin s-Si layer was incorporated using nonlocal meshing which encountered the threshold-voltage roll-off due to inducement of strained in the silicon layers.

3 Results and Discussion

With the reduction of channel length to 30 nm, various dimensions and parameters are precisely scaled and corrected values are induced for performance enhancement of the device and also minimize different short channel effects. The initial consideration in this novel device fall upon the selection of s-SiGe thickness when both s-Si layers are at a constant thickness value (s-Si = 1.5 nm). Thus, the effect of s-SiGe thickness on device performance has been studied for five devices as shown in Fig. 2.

As a result, increase in s-SiGe thickness lead to higher off-current in MOSFET, and hence many-fold higher current leakage is observed. Based upon these results, s-SiGe thickness of 3 nm in combination with s-Si thickness of 1.5 nm has been developed here. The calculated leakage currents (I_{off}) for both 30 and 50 nm channel length devices are illustrated in Fig. 3, which are in the acceptable range of leakage as suggested by Hu [15]. Simulated result of device characteristics is depicted in Fig. 4, and comparison with 50 nm channel length of the same device structure [12] clearly indicates the increase of quasi-ballistic carrier transport in the shorter channel region.

With the reduction of channel length scaled down to 30 nm, the electric field in the channel region increased as explicitly compared and presented in Fig. 5. Thus,

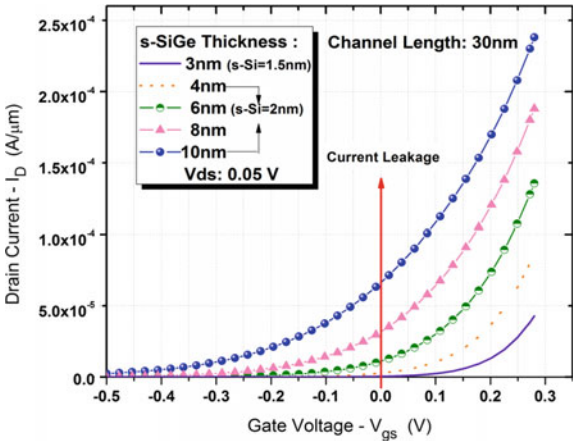


Fig. 2 HOI MOSFET with variation in s-SiGe thickness (Lg = 30 nm)

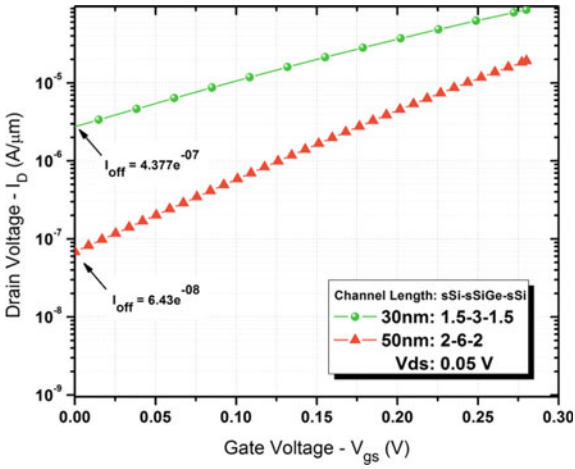


Fig. 3 Current–Voltage characteristics of 30 nm channel length HOI MOSFET indicating off-current comparison with 50 nm channel length of the same

an increase in the drift velocity of carriers is observed, which results in enrichment of drive current by ~93%.

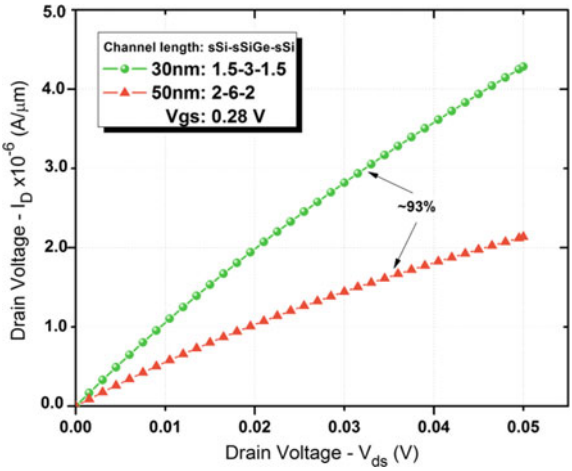


Fig. 4 Output characteristics of tri-layered channel HOI MOSFET (Lg = 30 nm)

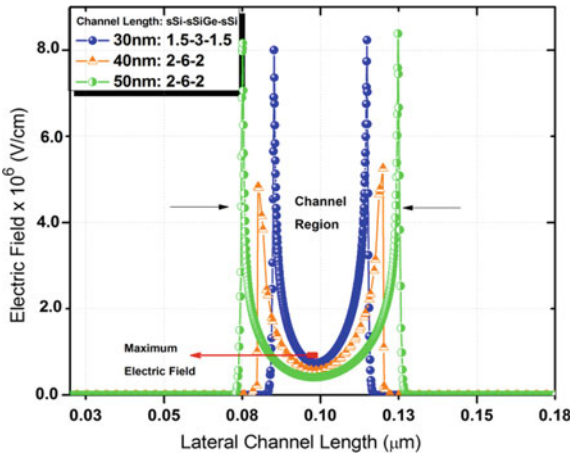


Fig. 5 Electric field along the lateral channel region of tri-layered HOI MOSFET

4 Conclusion and Outlook

Different modeling parameters for strained induced energy band structure, mobility, effective mass, and quantization were incorporated in the HOI MOSFET for the scaled dimension of 30 nm channel length using Sentaurus TCAD tool. An enhanced drive current with acceptable current leakage was modeled and developed as a result of increased drift velocity due to the increase in electric field in the channel region.

Acknowledgements The authors acknowledge NIT Mizoram and especially to the SMDP-C2SD project for providing the required amenities such as providing the workstation for simulation of the research work.

References

1. L. Khiangte, R.S. Dhar, Development of tri-layered s-Si/s-SiGe/s-Si channel heterostructure-on-insulator MOSFET for enhanced drive current. Phys. Status Solidi **255**(8), 1800034 (2018)
2. N. Kharche, M. Prada, T.B. Boykin, G. Klimeck, Valley splitting in strained silicon quantum wells modeled with 2 miscuts, step disorder, and alloy disorder. Appl. Phys. Lett. **90**(9), 92109 (2007)
3. M.L. Lee, E.A. Fitzgerald, M.T. Bulsara, M.T. Currie, A. Lochtefeld, Strained Si, SiGe, and Ge channels for high-mobility metal-oxide-semiconductor field-effect transistors. J. Appl. Phys. **97**(1), 1 (2005)
4. S.E. Thompson, G. Sun, Y.S. Choi, T. Nishida, Uniaxial-process-induced strained-Si: extending the CMOS roadmap. IEEE Trans. Electron Devices **53**(5), 1010–1020 (2006)
5. T.K. Maiti, S.S. Mahato, C.K. Maiti, Modeling of strain-engineered nanoscale MOSFETs, in *4th International Conference Nanotechnology Health Care Applications (NateHCA-07)*, Mumbai, India, D41–D45 (2007)
6. M. Willander, M.Y. Yousif, O. Nur, Nanostructure effect in Si-MOSFETs. Chalmers Univ of Technology Goeteborg (Sweden) Dept of Physics (2001)
7. M.J. Kumar, T.V. Singh, Quantum confinement effects in strained silicon mosfets. Int. J. Nanosci. **7**(2), 81–84 (2008)
8. D.Z. Pan, B. Yu, J.-R. Gao, Design for manufacturing with emerging nanolithography. IEEE Trans. Comput. Des. Integr. Circ. Syst. **32**(10), 1453–1472 (2013)
9. R. Minixhofer, TCAD as an integral part of the semiconductor manufacturing environment, in *2006 International Conference on Simulation of Semiconductor Processes and Devices*, pp. 9–16 (2006)
10. I. Lysenko, D. Zykov, S. Ishutkin, R. Meshcheryakov, The use of TCAD in technology simulation for increasing the efficiency of semiconductor manufacturing, in *AIP Conference Proceedings*, vol. 1772, no. 1, p. 60012 (2016)
11. L. Khiangte, R.S. Dhar, Development of double strained Si channel for heterostructure on insulator MOSFET, in *2017 2nd International Conference on Man and Machine Interfacing (MAMI)*, pp. 1–3 (2017)
12. L. Khiangte, R.S. Dhar, Double strained Si channel heterostructure on insulator MOSFET in sub-100 nm regime, in *2017 2nd International Conference on Man and Machine Interfacing (MAMI)*, pp. 1–3 (2017)
13. T. Sentaurus, Sdevice User Guide, ver, G-2012.06, Synopsys (2012)
14. T. Sentaurus, User Manual, Synopsys, Inc., Mt. View, CA, Version F-2011.09 (2010)
15. C. Hu, MOSFETs in ICs–scaling, leakage, and other topics, in *Modern Semiconductor Devices for Integrated Circuits* (Prentice Hall, New York, 2009)

InGaAs MOSFET for High Power Applications



Manoj Singh Adhikari, Vikalp Joshi and Raju Patel

1 Introduction

Laterally, MOS transistors are broadly used as semiconductor powerful devices in domestic to industrial (10–100 V) power applications [1–3]. For these power devices, the important physiognomies of a power MOSFET are less resistant, significant or high current, high gain, high breakdown voltage, less value of threshold potential, and low capacitance. In a simple power MOSFET, it is tough to attain the gain in every parameter simultaneously due to some parametrical effect which degrades the performance of device [2, 3]. In the past, trench-based gate MOSFET structures [3–6] on Si have been successfully demonstrated to get considerable enhancement in the semiconductor device performance. Now the Si-based device's performance is very limited, and therefore it is required to take the substitute semiconductor materials in order to achieve more improvement in the device behavior. The new material InGaAs becomes more powerful or promising semiconductors which take the place of Si in power MOS transistors [7–9]. To the best of our knowledge, the trench-gate power MOSFET structure has not been reported on InGaAs in literature. Therefore, for the first time, a power lateral trench-gate MOSFET called LTGMOS on $In_{0.53}Ga_{0.47}As$ is proposed. The simulations are done with the help of device simulator known as ATLAS [10], and the results of proposed MOS are examined and co-related with simple MOSFET.

M. S. Adhikari (✉)
SEEE Department, Lovely Professional University, Phagwara, Punjab, India
e-mail: manoj.23429@lpu.co.in

V. Joshi
HNB Garhwal University, Srinagar, Uttarakhand, India

R. Patel
ECE Department, MBM Engineering College, Jodhpur, Rajasthan, India

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_136

2 MOS Structure

The structure of the simple MOSFET is shown in Fig. 1 which is built on InGaAs material and also consists of a FIELD-PLATE which is helpful for the improvement in the off-state voltage (i.e., breakdown potential). For the device breakdown, highest (i.e., topmost) electric field is responsible. These field lines occur at the corner of the field-plate of gate region on the InGaAs material surface (as shown in Fig. 1 by "A"). Furthermore, in the simple MOSFET, the current flow on the surface is limited due to higher resistance. For the reduction of these disadvantages, we proposed the new/novel MOS as exposed in Fig. 2. In the novel MOS (N-MOS), the source contact (S) is on top and also in the middle of the device. In the N-MOS, two gates are used. These gates are connected vertically with the help of twice trenches which are placed on either sides of P-base region. The drain contacts of the N-MOS structure is taken symmetrically about the vertical axes and by applying a gate voltage, twice channels in the P-base are generated. After the input potential is applied, the current starts to flows from drain (D) terminal toward the source (S) region. Due to dual conduction path of the channels reduces the value of linear resistance. In the N-MOS structure, the RESURF effect is minimized as the trench-gate got the high breakdown voltage. In the proposed structure, reduced the value of electric field in n-drift layer (indicated in Fig. 2 by point 'B') due to RESURF effect. The MOS dimensions and the various parameters used in proposed work's simulations are specified in Table 1.

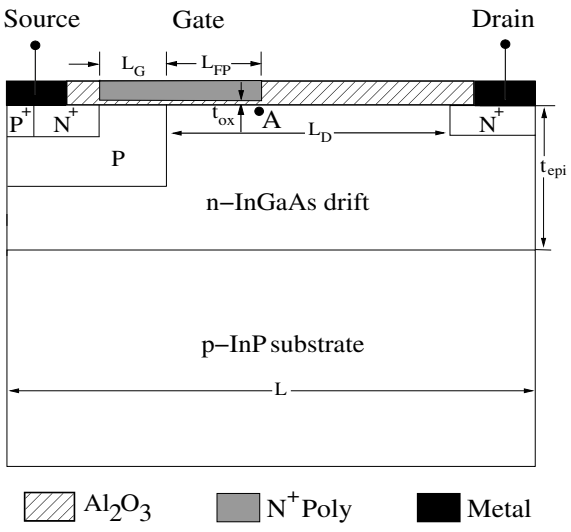


Fig. 1 Structure of the simple MOS

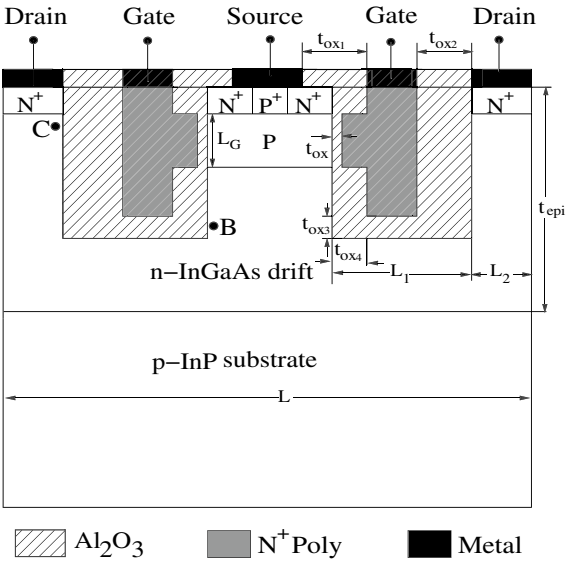


Fig. 2 Structure of the novel MOS

Table 1 Parameters of the devices

Parameter	Symbol	Conv. MOS	Proposed MOS
Gate length	L_G	0.50	0.50
Field-plate length	L_{FP}	1.0	–
Thickness (Oxide 3)	t_{ox1}	–	0.2
Thickness (Oxide 2)	t_{ox2}	–	0.5
Thickness (Oxide 1)	t_{ox3}	–	0.35
Drift epi-layer thickness	t_{epi}	0.6	1.90
Length 1: (drift region)	L_1	–	1.10
Length 2: (drift region)	L_2	–	0.5
Drift region doping	N_d	1×10^{16}	3×10^{16}
Cell pitch	L	4	4
Gate oxide thickness	t_{ox}	0.03	0.03
Drift region length	L_D	2.1	–

3 Simulation Results

The V–I characteristic of the conv. MOS and LTGMOS (proposed) devices are depicted in Fig. 3. The current of the LTGMOS is expressively high than that of the simple or conv. device. The main reason behind this is that the two channels operate in parallel mechanism in the N-MOS instead of conv. LDMOSFET. At V_{GS}= 2 V, drain current = 0.14 mA/μm of the LTGMOS and conv. devices is found to

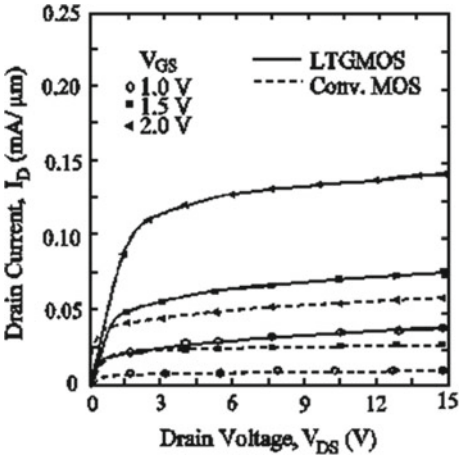


Fig. 3 V–I characteristics of the MOSFETs

be 0.06 mA/μm, the I_D of N-MOS is 2.29 times higher than the simple MOSFET. The threshold characteristics of proposed MOS and conv. MOSFET are depicted in Fig. 4. This characteristic is used to find the threshold potential or voltage of devices, which is a calculated as a straight line drawn at the point of maximum slope of the curve at input potential axis (i.e. Gate). The threshold potentials of LTGMOS and conventional devices are 0.73 and 1.0 V, respectively. LTGMOS offers 28% reduction in threshold voltage. The resistance is the ratio of V to I of drain, and at that time the device performs in the linear region. Drain characteristics are shown in Fig. 5 at gate bias of 10 V for both the devices. The resistances ($R = V_D/I_D$) of the LTGMOS and simple MOS are 31 and 52 mΩ-mm^2, respectively. In the proposed MOS, there is 41% drop in specific resistance due to its higher value of current capability.

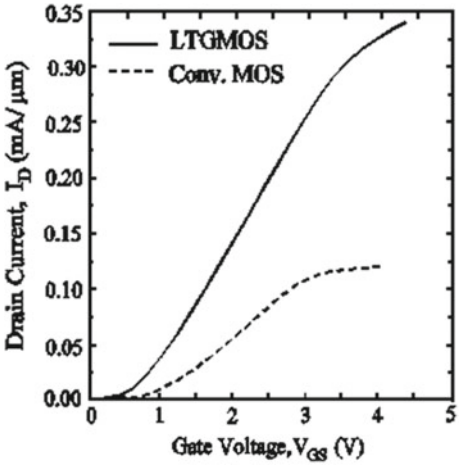


Fig. 4 Threshold characteristics of the MOSFETs

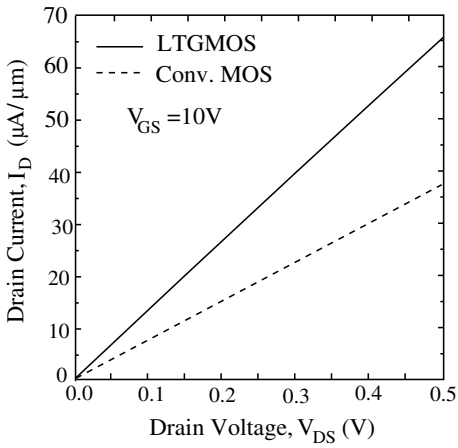


Fig. 5 Resistance lines of the conventional and the proposed MOS in linear region

Figure 6 depicted breakdown characteristics of the N-MOS and conv. MOS. The drift region doping is 3x1016 cm^{-3} for LTGMOS and 1x1016 cm^{-3} for conventional MOSFET. It is observed that off-state voltage of the LTGMOS and conv. MOS device are 82 and 41 V, respectively, which demonstrates the twice times expansion in the off-state voltage. The increment in the off-state voltage of N-MOS is due to decrement in electric field by the trench structure. In simple MOSFET, high field is responsible for low breakdown at point "A", whereas a low field is observed at points "B" and "C" in the proposed MOS.

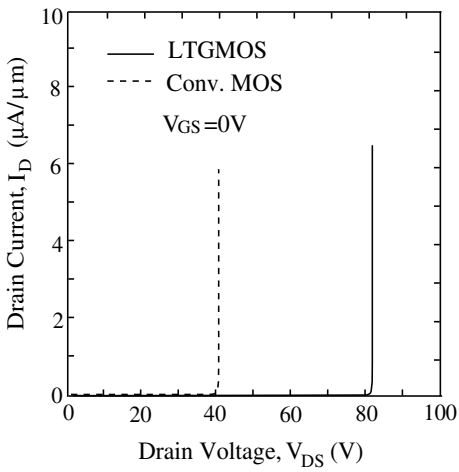


Fig. 6 Off-state characteristics of the conv. and the proposed MOSFET

4 Conclusion

An InGaAs power MOSFET having a trench-gate structure is presented. The device provides conduction of twice channels which operates in parallel to enhance the current (I_D) and causes RESURF result to improve the breakdown characteristics. Based on 2-D simulation results, this work confirmed that the proposed MOS indicates enhanced output current, i.e., 2.29 times, decrease in threshold voltage which is 28%, also reduction in resistance, i.e., 41%, and increase in breakdown voltage or off-state voltage, i.e., two times in contrast of conv. MOS.

References

1. R.P. Zingg, On the specific on-resistance of high-voltage and power devices. IEEE Trans. Electron Devices **51**, 492–499 (2004)
2. F. Schwierz, J.J. Liou, RF transistors: recent developments and roadmap toward terahertz applications. Solid State Electron. **51**, 1079–1091 (2007)
3. X. Luo, T.F. Lei, Y.G. Wang, G.L. Yao, Y.H. Jiang, K. Zhou, P. Wang, Z.Y. Zhang, J. Fan, Q. Wang, B.Z.R. Ge, Z. Li, F. Udrea, Low ON-resistance SOI dual-trench-gate MOSFET. IEEE Trans. Electron Devices **59**, 504–509 (2012)
4. Y. Guoliang, L. Xiaorong, W. Qi, J. Yongheng, W. Pei, Z. Kun, W. Lijuan, Z. Bo, L. Zhaoji, Novel SOI double-gate MOSFET with a P-type buried layer. J. Semicond. **33**, 054006-1–054006-4 (2012)
5. I. Corts, P.F. Martnez, D. Flores, S. Hidalgo, J. Rebollo, The thin SOI TGLDMOS transistor: a suitable power structure for low voltage applications. Semicond. Sci. Technol. **22**, 1183–1188 (2007)
6. Y. Singh, M. Punetha, A lateral trench dual gate power MOSFET on thin SOI for improved performance. ECS J. Solid State Sci. Technol. **2**(7), 113–117 (2013)
7. M.S. Lundstrom, On the mobility versus drain current relation for a nanoscale MOSFET. IEEE Electron Device Lett. **22**, 293–295 (2011)
8. Y. Xuan, Y. Wu, P. Ye, High-performance inversion-type enhancement-mode InGaAs MOSFET with maximum drain current exceeding 1 A/mm. IEEE Electron Device Lett. **29**, 294–296 (2008)
9. M.S. Adhikari, Y. Singh, Performance enhancement of InGaAs MOSFET using trench technology, in *IEEE International Conference on Signal Processing and Communication (ICSC)*, pp. 309–311, March 2015
10. ATLAS Users manual: device simulation software, Silvaco Int., Santa Clara, CA (2010)

Low Power Efficient $Si_{0.7}Ge_{0.3}$ Pocket Junction-Less DGTFET with Sensing Ability for Bio-species



Suman Lata Tripathi and Shekhar Verma

1 Introduction

Scaling trend as per Moor's law leads toward thermal limits on conventional CMOS circuit with addition of more transistors per unit area of chip [1]. To obtain considerable ON- and OFF-state currents, threshold voltage should be scaled in proportion with supply voltage that results in higher value of OFF-state current [2]. Also, the increase in leakage current increases the thermal stress on subthreshold parameters. A tunneling-field-effect transistor (TFET) is less affected by the conditions imposed due to thermal limits and may perform better at low VDD. A positive drain bias makes TFET reverse biased, with the band-to-band tunneling current modulated by gate voltage under this condition where charges in channel come mainly from the drain side rather than the source, for narrow bandgap TFET [3]. Here, band-to-band tunneling in TFET follows charge-plasma concept depending on work function of metallic gate contact in source/drain region and silicon body thickness less than the Debye length [4].

$$\text{Debye length is governed by expression: } L_D = \sqrt{\frac{V_T \in_{\text{Si}}}{qN}} \tag{1}$$

where N represents carrier concentration, $\in_{\text{Si}}$ represents dielectric constant of silicon, and V_T is thermal voltage.

Narrow bandgap III–V materials have potentials to optimize thin body TFET performance [5, 6] with improvement in tunneling properties. Quantum model with atomistic simulations was used to characterize the subthreshold performance of GaSb/InAs/InGaAs TFET (heterojunction TFET) [7, 8]. Tunnel FET emerged as

S. L. Tripathi (✉) · S. Verma
Lovely Professional University, Punjab, India
e-mail: tri.suman78@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_137

potential candidate in digital and memory designs in low power regime below 20 nm technology node [9]. The tunneling-field-effect transistor with short gate structure was modeled with dielectrically modulated bio-sensing features [10–13] with real-time application of emerging TFET. Heterojunction TFET with inclusion of narrow bandgap material pocket region decreases tunneling distance to enhance tunneling current in on-state condition [14, 15]. A thin dielectric region incorporated in the middle of the channel reduces OFF current (I_{on}/I_{off} ratio by 10^5) [16]. The major drawback of TFET is longer channel length to get equivalent ON current in comparison to corresponding MOSFET. To limit SCE with smaller dimensions, there is requirement of steep junctions at the source/channel and channel/drain junctions that control over the effective channel length and parasitic resistances. The multiple-gate junction-less transistors evaluated to rule out the requirement of steep junctions below 20 nm technology [17]. In addition, metal gate with higher work function and high-K dielectric material as oxide region further increases I_{on}/I_{off} current ratio. The ambipolar behavior of TFET had been exploited through several researchers to enhance bio-sensing capability with different structures [18]. Several, FET-based biomarkers were proposed which have great potential in betterment of disease diagnosis and provide new therapeutic procedures [19–21]. A CMOS-compatible SiNW FET was presented for real-time, label-free ultrasensitive prostate-specific antigen (PSA) sensor to diagnose prostate cancer [11]. The dielectrically modulated tunnel FET (DMTFET) was reported with higher sensitivity at lower subthreshold current compared with their dielectrically modulated FET counterpart [13, 14, 22]. In this paper, a novel JLDGTFET is proposed with narrow bandgap material as pocket region and elaborates the Kane's band-to-band tunneling model used in device simulations on TCAD tool for various parametric analyses. The proposed JLDGTFET provides barrier due to source–drain channel depletion and p-type pocket region that results in suppression of transistor parasitic as well as low subthreshold leakage. Using 2D/3D simulations, it is found that I_{on}/I_{off} ratio 10^9 for gate length 15 nm is higher than any other designs proposed earlier for similar dimensions. Such pocket $Si_{0.7}Ge_{0.3}$ JLDGTFET is further examined with cavity regions which are capable to sense atmospheric changes in terms of change in drain current and threshold voltage.

2 Device Model and Structures

2-D visual TCAD (Cogenda) has been used to implement the new junction-less double gate TFET(JLDGTFET). The Kane's band-to-band tunneling model [23] is used in the device simulations. The source/drain is kept at equal doping level of 1–20 cm^{-3} with dimension of 10 nm. The channel region is considered as intrinsic with dimension 20 nm. Figure 1 indicates the novel 2D JLDGTFET structure with gate height $H_g = 4$ nm and length $L_g = 15$ nm. Different oxide materials are used such as SiO_2 and HfO_2 and performance is compared. To increase band-to-band tunneling property, a pocket region of $Si_{0.7}Ge_{0.3}$ of 5 nm thickness is included near the source region under influence of gate. The proposed pocket JLDGTFET incorporated with $Si_{0.7}Ge_{0.3}$ as pocket region (doping 1–20 cm^{-3}) to suppress off-state current and to

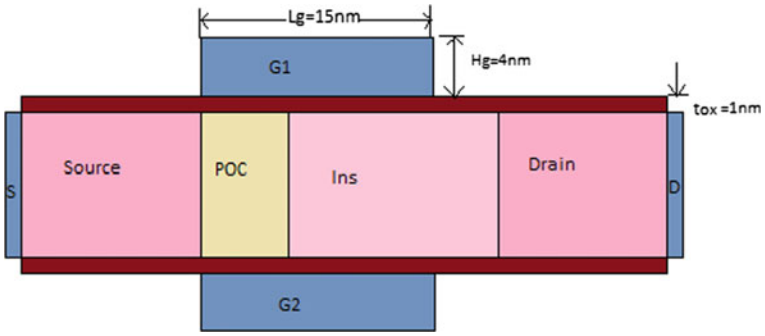


Fig. 1 2D structure of JLDGTFET with SiGe pocket region

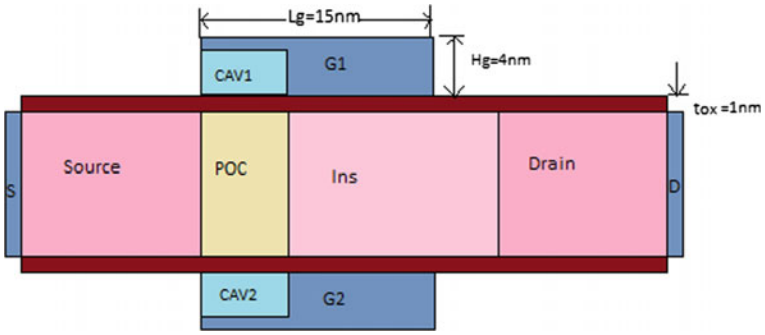


Fig. 2 2D view of JLDGTFET with bio-sensing cavity regions

achieve higher I_{on}/I_{off} current ratio. To achieve high on-state current, a metal gate with high value of work function Pt (5.7 eV) is used with oxide region made with high-K dielectric material HfO_2 (25) replacing SiO_2 (3.9) under gate and performance compared with Al/SiO_2 interface. Figure 2 shows 2-D view of JLDGTFET with bio-sensing cavity region of width 5 nm and depth 3 nm under top and bottom gate contacts.

3 Results and Discussions

First, simulation is performed for junction-based DGTFET and then performance is compared with junction-less DGTFET (JLDGTFET) in linear and saturation region of operation. The performance of proposed JLDGTFET is also compared with corresponding DGMOSFET in terms of ON/OFF-state performance and subthreshold parameters. Proposed JLDGTFET has very low OFF-state current due to barrier provided by p-type pocket region resulting in I_{on}/I_{off} current ratio up to the order

of ~10^9 (L_g = 15 nm) that is higher in comparison to TDJLT [16] varying between ~10^2 and ~10^7 with L_g variation between 10 and 20 nm. JLDGTFET performance is compared with DGTFET and DGMOSFET with similar dimension (L_g = 15 nm). Figure 3 depicts that JLDGTFET has sharp drain current variations with respect to gate voltage in subthreshold region in comparison to DGTFET and DGMOSFET in both linear (V_{ds}= 0.1) and saturation (V_{ds}= 1 V) regions.

Figure 4 concludes that JLDGTFET with Pt as gate contact and HfO_2 oxide material has better I_{on}/I_{off} current ratio of 10^9 in comparison to Al/SiO_2 of 10^7. Similar

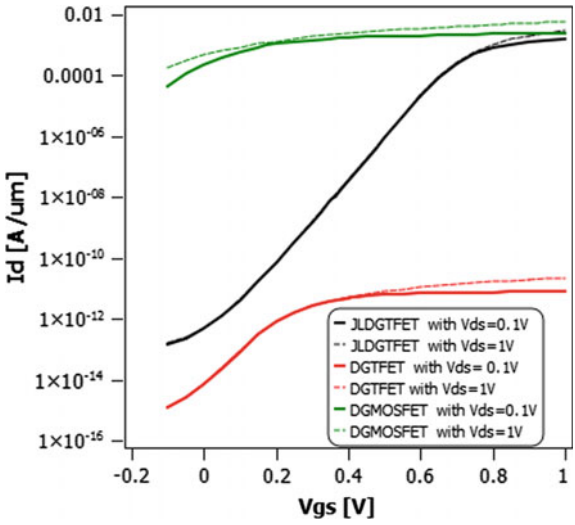


Fig. 3 I_d versus V_{gs} of pocket JLDGTFET in comparison with DGMOSFET

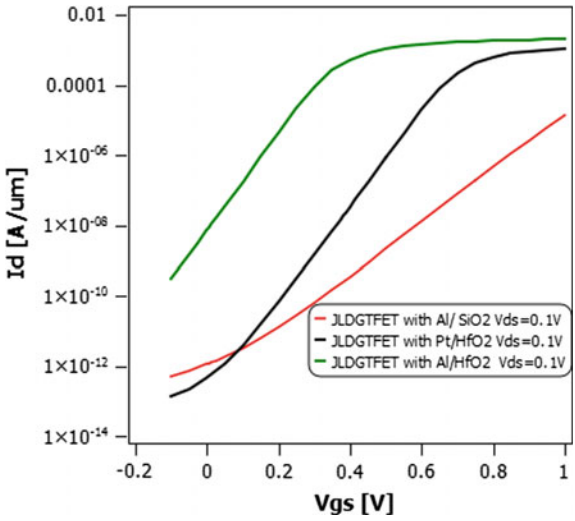


Fig. 4 I_d versus V_{gs} of pocket JLDGTFET with different gate contacts/oxide regions

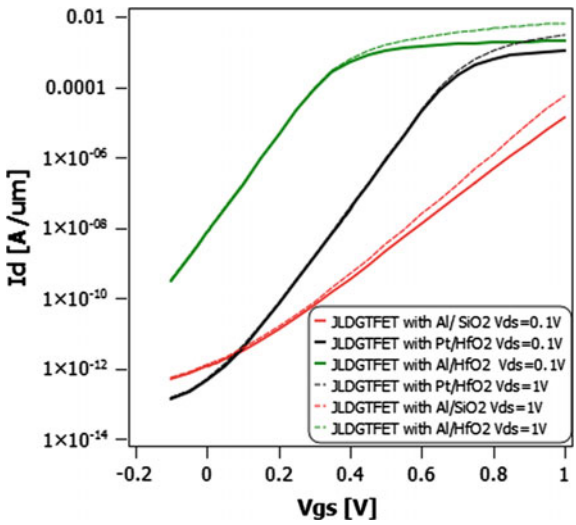


Fig. 5 I_d versus V_{gs} of pocket JLDGTFET for $V_{ds} = 0.1$ V and $V_{ds} = 1$ V

comparison has been done in Fig. 5 between all designs and proposed JLDGTFET in linear ($V_{ds} = 0.1$) as well as in saturation region ($V_{ds} = 1$ V) which also supports that JLDGTFET with Pt/HfO_2 gate/oxide interface has optimum performance in both linear and saturation regions. The Id versus V_{ds} characteristic of JLDGTFET with Al/SiO_2 as gate/oxide interface is shown in Fig. 6 matching with ideal behavior. Similar I_d versus V_{ds} of pocket JLDGTFET is observed in Fig. 7 for Pt/HfO_2 as gate/oxide interface. Here drain current shows constant nature after a particular value

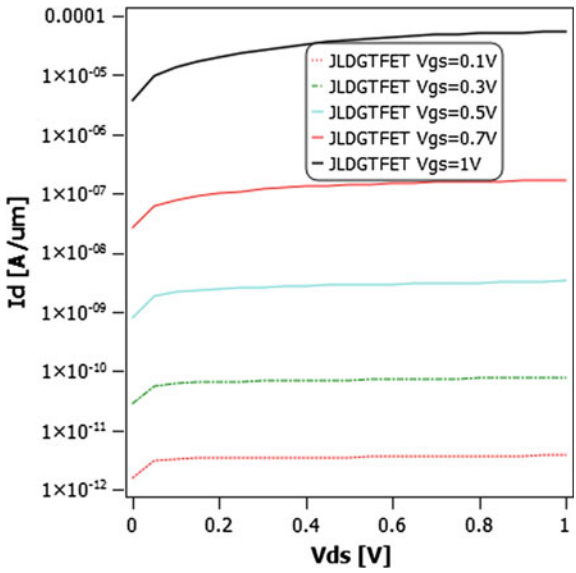


Fig. 6 I_d versus V_{ds} of pocket JLDGTFET for Al/SiO_2

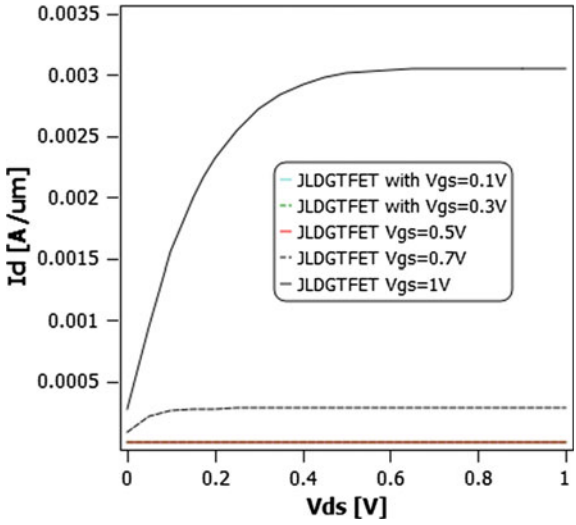


Fig. 7 I_d versus V_{ds} of pocket JLDGTFET for Pt/HfO_2

of V_{ds} for fixed value of gate voltage and increases with increase in gate voltage. This shows that proposed JLDGTFET has ideal nature in all cutoff, linear, and saturation region, and therefore can be easily implemented for analog and digital applications.

Figure 8 shows the application of JLDGTFET with inclusion of cavity region under gate to sense the changes in biomolecules present in atmosphere. A sharp change is observed with the changes in dielectric constant of material present in cavity which concludes that the device is able to sense the very minute changes in atmospheric conditions. This bio-sensing ability has application also in other

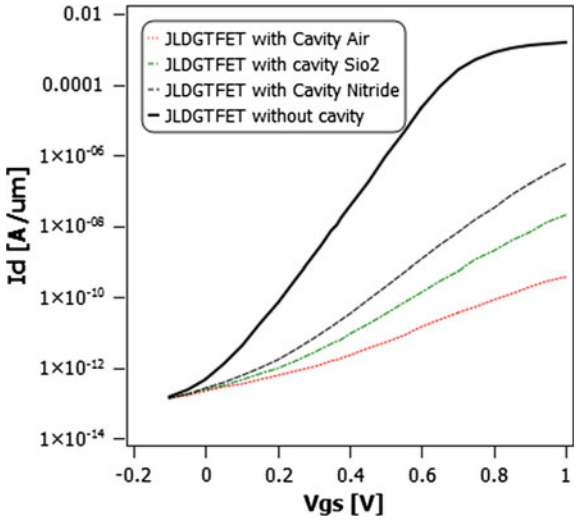


Fig. 8 I_d versus V_{gs} of pocket JLDGTFET with cavity regions

Table 1 Performance comparison of JLDGTFET with different cavity regions

Device type	SS (mV/decade)	I_{on}/I_{off}
JLDGTFET with cavity air	271	1749
JLDGTFET with cavity SiO_2	173	1E+05
JLDGTFET with cavity nitride	128	3E+06
JLDGTFET without cavity	70.3	1E+09

health monitoring systems by measuring changes in electrical parameter such as drain current, electric field, and potential variations due to the changes in bio-species present. Since the device structure proposed is up to nanolevel transistor, it will become easier to implant it anywhere in portable or wearable device for smart health monitoring system.

Table 1 shows comparison of JLDGTFET for different gate contact materials and oxides under gate in which JLDGTFET with Pt/HfO_2 gives I_{on}/I_{off} current ratio up to 10^9 in comparison to 10^7 of Al/SiO_2 as gate-oxide region and 10^5 of existing TDJLT [17] with gate length (L_g) of 15 nm. This shows that cavity region with nitride has more ON/OFF current ratio and lower threshold voltage in comparison to air and SiO_2 as cavity region material. The subthreshold slope (SS) and drain-induced barrier lowering (DIBL) are also calculated for different JLDGTFET structures. SS gives the variation of the drain current as function of the gate voltage in weak inversion region. SS can be expressed as follows:

$$\mathrm{SS} = \frac{\mathrm{d}V_{gs}}{\mathrm{d}(\log I_d)} (\mathrm{mV/decade}) \tag{2}$$

The DIBL is measured as change in threshold voltage (V_{th}) with respect to change in V_{ds} ranges 0.1–1.0 V. So for measurement of DIBL, threshold voltage of MOSFET for two different V_{ds} is measured and DIBL is calculated by the following expression:

$$\mathrm{DIBL} = \frac{\Delta V_{th}}{\Delta V_{ds}} = \frac{V_{th1} - V_{th2}}{V_{ds1} - V_{ds2}} \mathrm{mV/V} \tag{3}$$

Table 2 describes the subthreshold slope and DIBL variations comparing different DGTFET structures. It shows that the proposed pocket JLDGTFET has subthreshold slope of 70.3 mV/V and DIBL of 3.5 mV/V. Table 3 shows the variation in I_{on}/I_{off} current ratio of the proposed structure in comparison to other existing TFET structures

Table 2 Subthreshold performance comparison of JLDGTFET with existing transistors

Device type	SS (mV/decade)	DIBL (mV/V)
DGMOSFET	106.01	5.04
DGTFET	308	9.11
JLDGTFET (Al/SiO_2)	110	5.8
JLDGTFET (Pt/HfO_2)	70.3	3.5

Table 3 Comparison of JLDGTFET with other TFET for cavity region dielectric constant ($K = 1$)

Device type	L_g (nm)	I_{on}/I_{off}
Ref. [17]	42	1.00E+03
Ref. [18]	53	1.00E+03
JLDGTFET	15	1.00E+09

with cavity region ($K = 1$). The proposed JLDGTFET with channel length 15 nm has better I_{on}/I_{off} ratio up to ~10^9 in comparison to other existing TFET structures. It also shows steep subthreshold characteristics with SS and DIBL under limit.

The proposed pocket JLDGTFET with bio-sensing cavity region can be analyzed to exploit the ambipolar nature of tunneling transistors that is helpful to measure accurate positive and negative effects of bio-species. Such inherent sensing cavity region in proposed JLDGTFET will be better fitted for health monitoring portable systems.

4 Conclusion

The proposed JLDGTFET shows steep subthreshold slope of 70.3 mV/decade and very low DIBL of 3.5 mV/V with good I_{on}/I_{off} ratio that is suitable for any low-power high-speed operations. Since the device is designed for 15 nm channel length, it will be suitable for small portable systems. Similar JLDGTFET with bio-sensing cavity region shows very sharp curve of drain current with respect to the variation in dielectric constant that changes with change in atmospheric conditions. The cavity region shows a very sharp change in I_d versus V_{gs} characteristics with different dielectric constant materials present in cavity regions in comparison to existing literature. This work can be extended in future to measure PH value anywhere inside human body that is helpful in cancer treatment.

References

1. R.G. Dreslinski, M. Wieckowski, D. Blaauw, D. Sylvester, T. Mudge, Near-threshold computing: reclaiming Moore's law through energy efficient integrated circuits. Proc. IEEE **98**, 253–266 (2010)
2. S.L. Tripathi, R. Mishra, R.A. Mishra, Characteristic comparison of connected DG FINFET, TG FINFET and Independent Gate FINFET on 32 nm technology (IEEE, ICPCES, 2012)
3. D.H. Morris, U.E. Avci, R. Rios, I.A. Young, Design of low voltage tunneling-FET logic circuits considering asymmetric conduction characteristics. IEEE J. Emerg. Select. Top. Circ. Syst. **4**, 380–388 (2014)
4. C. Sahu, J. Singh, Charge-plasma based process variation immune junctionless transistor. IEEE Electron Device Lett. **35**, 411–413 (2014)

5. K. Tomioka, M. Yoshimura, T. Fukui, Steep-slope tunnel field effect transistors using III–V nanowire/Si heterojunction. in *Proceedings of the VLSI Technology (VLSIT) Symposium*, Honolulu, HI, USA, vol. 104 (2012), pp. 47–48
6. B. Ganjipour, J. Wallentin, M.T. Borgström, L. Samuelson, C. Thelander, Tunnel field-effect transistors based on InP-GaAs heterostructure nanowires. ACS Nano **6**, 3109–3113 (2012)
7. U.E. Avci et al., Understanding the feasibility of scaled III–V TFET for logic by bridging atomistic simulations and experimental results. in *Proceeding VLSI Technology (VLSIT) Symposium*, Honolulu, HI, USA (2012), pp. 183–184
8. U.E. Avci, D.H. Morris, I.A. Young, Tunnel field-effect transistors: prospects and challenges. J. Electron Device Soc. IEEE, 88–95 (2015)
9. T. Nirschl, P.-F. Wang, W. Hansch, D. Schmitt-Landsiedel, The tunneling field effect transistor (TFET): the temperature dependence, the simulation model, and its application. IEEE (2004)
10. J.-Y. Kim et al., An underlap channel-embedded field-effect transistor for biosensor application in watery and dry environment. IEEE Trans. Nanotechnol. **11**, 390–394 (2012)
11. S. Kanungo, S. Chattopadhyay et al., Comparative performance analysis of the dielectrically modulated full-gate and short-gate tunnel FET-based biosensor. IEEE Trans. Electron Devices **62**, 994–1001 (2015)
12. S. Kanungo, S. Chattopadhyay, P.S. Gupta, K. Sinha, H. Rahaman, Study and analysis of the effects of SiGe source and pocket-doped channel on sensing performance of dielectrically. IEEE Trans. Electron Devices **63**, 2589–2596 (2016)
13. D. Singh, S. Pandey, K. Nigam, D. Sharma, D.S. Yadav, P. Kondekar, A charge-plasma-based dielectric-modulated junctionless TFET for biosensor label-free detection. IEEE Trans. Electron Devices **64**, 271–278 (2017)
14. A. Vandooren, D. Leonelli, R. Rooyackers, A. Hikavyy, K. Devriendt, M. Demand et al., Analysis of trap-assisted tunneling in vertical Si homo-junction and SiGe hetero-junction tunnel-FETs. Solid State Electron **83**, 50–55 (2013)
15. W. Li, H. Liu, S. Wang, S. Chen, Z. Yang, Design of high performance Si/SiGe heterojunction tunneling FETs with a T-shaped gate. Nanoscale Res. Lett. (2017)
16. A. Lahgere, M.J. Kumar, A tunnel dielectric-based junctionless transistor with reduced parasitic BJT action. IEEE Trans. Electron Devices **64**, 3470–3475 (2017)
17. R. Rios et al., Comparison of junctionless and conventional trigate transistors with L_g down to 26 nm. IEEE Electron Device Lett. **32**, 1170–1172 (2011)
18. R. Narang, M. Saxena, R.S. Gupta, M. Gupta, Assessment of ambipolar behavior of a tunnel FET and influence of structural modifications. J. Semicond. Technol. Sci. **12**, 482–491 (2012)
19. W.C.S. Cho, Potentially useful biomarkers for the diagnosis, treatment and prognosis of lung cancer. Biomed. Pharmacother. **2007**(61), 515–519 (2007)
20. S.S. Cheng, S. Hideshima, S. Kuroiwa, T. Nakanishi, T. Osaka, Label-free detection of tumor markers using field effect transistor (FET)-based biosensors for lung cancer diagnosis. Sens. Actuators B Chem. **212**, 329–334 (2015)
21. A. Gao, L. Na, P. Dai, C. Fan, Y. Wang, T. Li, Direct ultrasensitive electrical detection of prostate cancer biomarkers with CMOS-compatible n- and p-type silicon nanowire sensor arrays. Nanoscale **6**, 13036–13042 (2014)
22. S.L. Tripathi, R. Patel, V.K. Agrawal, Low leakage pocket junction-less DGTFET with bio sensing cavity region. Turk. J. Electr. Eng. Comput. Sci., 1–11 (2018)
23. K-H. Kao, S.V. Anne, G.V. William, S. Bart, G. Guido, D.M. Kristin, Direct and indirect band-to-band tunneling in germanium-based TFETs. IEEE Trans. Electron. Devices **59**(2), 292–301 (2012)

Three-Layered Channel with Strained Si/SiGe/Si HOI MOSFET



Lalthanpuii Khiangte and Rudra Sankar Dhar

1 Introduction

An extensive study and research has been carried out in the area of strained technology in the past to comprehend its ability for enhancement of carrier mobility and in turn the drive current for MOSFETs [1]. Works in the past advocated constructive stress techniques in n-channel devices, for instance, (i) total strained substrate, (ii) memorized stress, and (iii) tensile over layer [2–9] and numerous structures formulated with strained device concept such as Si on SiGe settled as (a) strained silicon (s-Si) on relaxed SiGe, (b) s-Si on strained SiGe (s-SiGe) dual layer channel, (c) s-Si directly on oxide, and (d) heterostructure (combination of Si and SiGe) on insulator (HOI) [10]. The induced strain does not conceal on elongated span of reliability aspects and has been validated by a minimal influence on the superiority of the gate oxide [11]. Due to reduction in bandgap with escalation of strain in the channel, a major concern arises as threshold voltage (V_{th}) reduces with strained Si. On this account, the use of higher Ge mole fraction benefits in holding while preserving the V_{th} realistically large as detailed and suggested by the investigations for s-Si layer thickness of ~3 nm [12, 13]. Hence, $Si_{1-x}Ge_x$ forms a better option for the dual-channel MOSFETs.

This paper is focused on the development of a novel double strained Si channel heterostructure-on-insulator MOSFET featuring a supplementary bottom layer of s-Si in the channel region which sandwiches $Si_{1-x}Ge_x$ between making a distinguished structure than the conventional single strained Si SOI nMOSFET, ensuing in improved current. A comprehensive study and investigation of three-layered s-Si channel nMOSFET in nanoregime has been performed, where the transconductance and drain current enhancements are compared for $L_g = 100$ nm and $L_g = 50$ nm

L. Khiangte · R. S. Dhar (✉)
Department of Electronics and Communication Engineering, National Institute of Technology Mizoram, Aizawl, India
e-mail: rdhar@uwaterloo.ca

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_138

in the devices. Behavior of the effect of overshoot velocity has also been examined for $L_g = 50$ nm in the HOI MOSFET, which elucidated carriers approaching the quasi-ballistic transport.

2 MOSFET Device Structure

Incubating the benefits of SOI MOSFET with the ability to enhance carrier mobility by strained Si technology [14–19], doubly due to the additional s-Si layer discriminates this novel device structure and its superior characteristics to the conventional device (single s-Si MOSFET). Figure 1 exemplifies the structure with two layers of s-Si, which are introduced in the channel of the nMOSFET, where SiGe layer is inserted in the middle, and hence hewing a s-SiGe layer. This fresh structure of the device therefore comprises three layers in the channel creating the heterostructure-on-insulator (HOI) nMOSFET. To fabricate this HOI nMOSFET, each layer is grown individually such as strained Si is grown on strained $Si_{1-x}Ge_x$ and transported to the

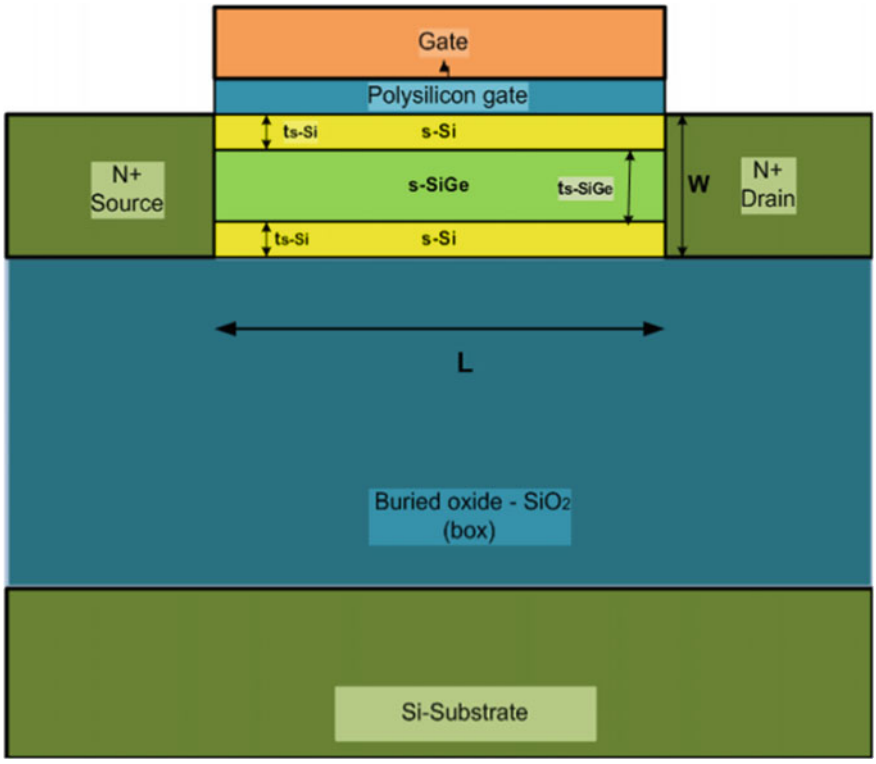


Fig. 1 Schematic of three-layered s-Si with s-SiGe HOI nMOSFET

Table 1 List of parameters for nMOSFET device

Parameters	Values
Length of the channel (L)	50 nm, 100 nm
Mole fraction for Ge	0.4
Thickness of s-Si ($t_{s\text{-}Si}$)	2 nm
Gate oxide thickness	2 nm
Doping concentration source/drain (N_D)	10^{19} cm^{-3}
Doping of channel (N_A)	10^{16} cm^{-3}
Drain bias (V_{DS})	50 mV

insulator conserving the unique strain state of the full system. So as to not confine the holes entirely in the channel, ultrathin s-Si layers of ~2 nm thick are developed as surface layers on either side of the SiGe layer. This results in condensed effect of higher Ge content on the s-Si layer and with this advancement these two s-Si layers formed the quantum well for carrier confinement. With parameters listed in Table 1, two devices have been designed: Device-A—100 nm and Device-B—50 nm gate length double s-Si channel along with the conventional device of s-Si are specified [12].

3 Results and Discussion

Employing the various device parameters as tabulated in Table 1, development and simulation of the device structure shown in Fig. 1 by 2D SYNOPSIS Sentaurus TCAD tool [20, 21] are carried out. Strain persuaded bandgap thinning phenomenon and transformation in electron mobility were modeled, respectively, by uniting the multivalley model for both the band structure and electron mobility [21]. The Schrodinger model of 1D that incorporates the simultaneous solution mechanism of both Schrodinger and Poisson's equations was accounted to visualize quantization effects in the devices produced between the dielectric interfaces and the s-Si layer by the insertion of the nonlocal-mesh. ~49% enhancement in drain current was observed for Device-A at $V_{ds} = 50$ mV as shown in Fig. 2 over the conventional device which is essentially due to the additional two strained Si layers on either side of SiGe in the channel region, a favorable complaint for the increment of electrons with enhanced mobility.

Trace of carrier confinement in the s-Si layer was obtained as illustrated in Fig. 3 due to the quantization effect. Thus, energy bandgap increased balancing of the threshold voltage roll-off due to strained layers. A clear insight of velocity overshoot [22] being more prominent in reduced channel length is depicted in Fig. 4.

With the reduction of channel length and increase in lateral electric field in Device-B, the velocity overshoot condition is acquired, i.e., nonequivalence of momentum relaxation time and energy relaxation time, resulting in low scattering and carriers do

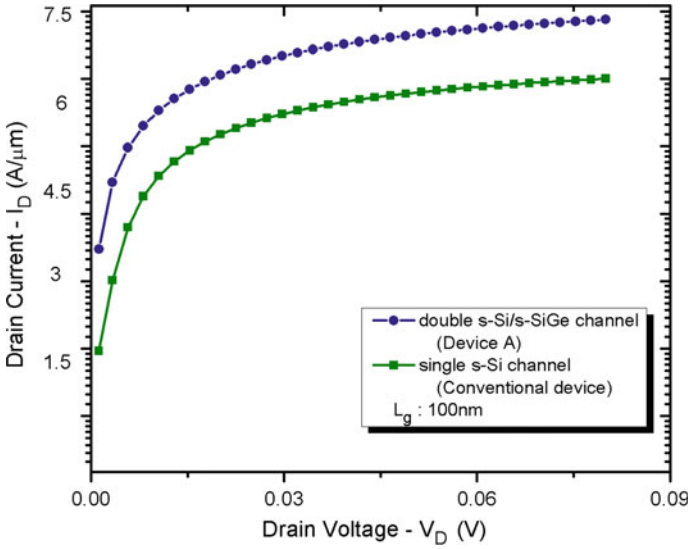


Fig. 2 $I_D - V_D$ characteristics comparison for $L_g = 100$ nm depicting drive current enhancement of 49% for HOI MOSFET over conventional MOSFET device

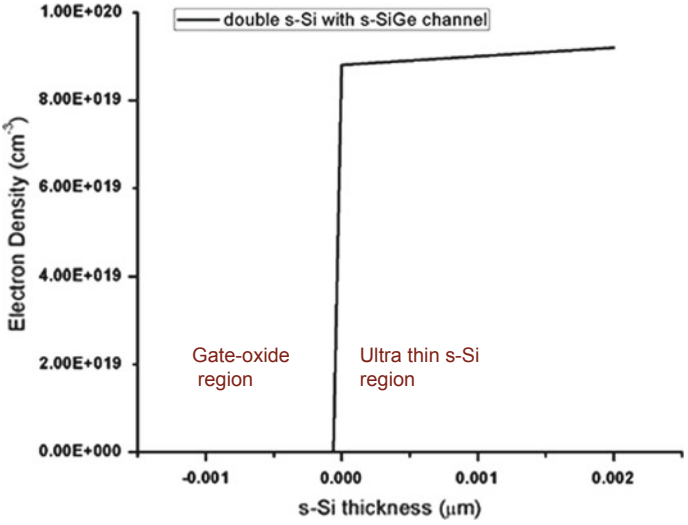


Fig. 3 Carrier confinement in 2 nm thin s-Si layers of Device-A

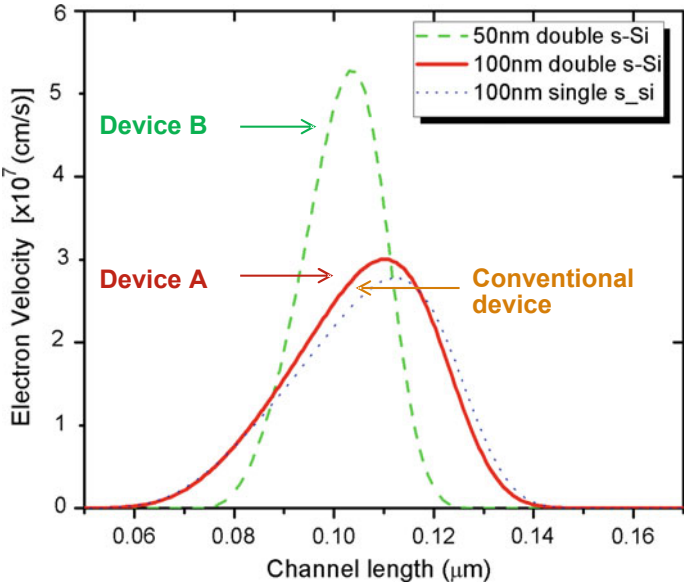


Fig. 4 A comparison of velocity of electron along the length of the channel is presented between Device-A, Device-B, and conventional MOSFET device

not have time to heat up. The electron velocity becomes greater than the saturation velocity. Hence, carrier transport approaches quasi-ballistic nature resulting in drain current enhancement ($\Delta I_D = (I_D{}^{50\text{ nm double s-Si}} - I_D{}^{100\text{ nm double s-Si}})/I_D{}^{100\text{ nm double s-Si}}$) of ~41% for Device-B ($L_g = 50$ nm) in comparison to Device-A (100 nm channel length) as shown in Fig. 5.

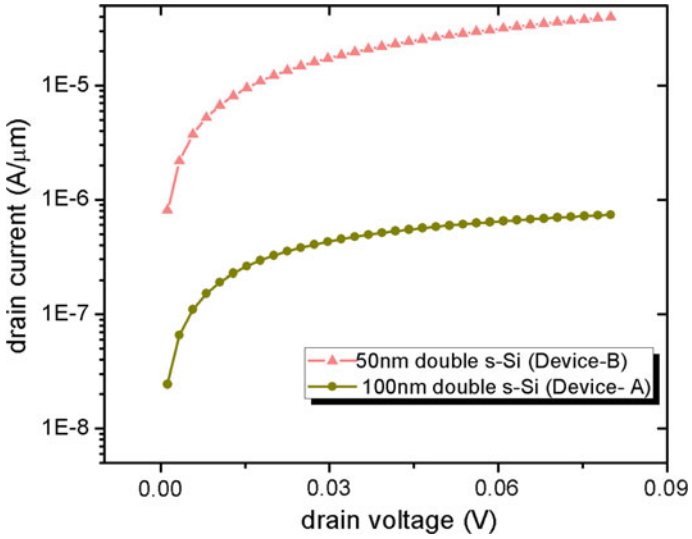


Fig. 5 A drive current enhancement of 41.3% is observed from the $I_D - V_D$ characteristics for $L_g = 50$ nm over $L_g = 100$ nm in the HOI MOSFET devices

Distinct view about the short channel effects is exemplified by drawing a comparison as shown in Fig. 6, where the transconductance (gm_{max}) is varied as a function of DIBL for all the three devices (Device-A, Device-B, and Conventional Device). For the 50 nm channel length HOI MOSFET, a threshold voltage roll-off of ~25% has been acquired in comparison to the 100 nm two s-Si nMOSFET device with a greater DIBL. An outstanding upsurge in transconductance is accomplished containing overshoot velocity [22] that is swapped between transconductance and Drain-Induced Barrier Lowering (DIBL) as detected due to the SHEs in the HOI nMOSFET devices.

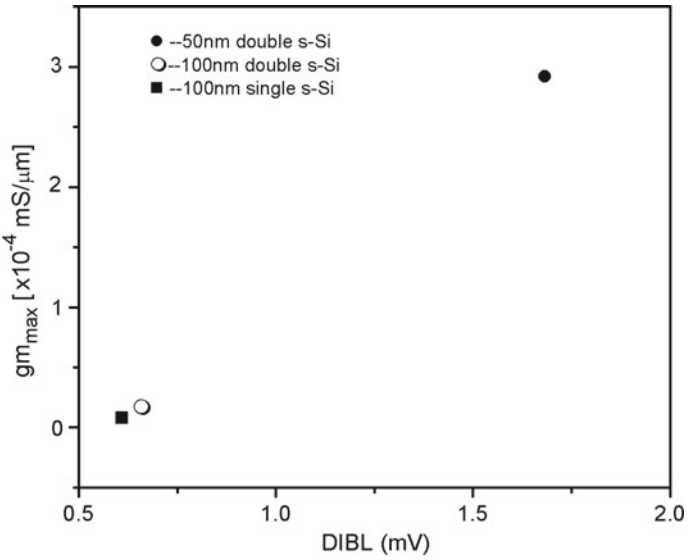


Fig. 6 Transconductance (gm_{max}) as a function of DIBL for Device-A, Device-B, and conventional device

4 Conclusion

The two strained Si layers of heterostructure-on-insulator nMOSFET have been developed and analyzed for a detailed study in the channel region. With intrusion of additional s-Si layer in the MOSFET channel, an eloquent enrichment of 49.3% drive current has been attained for the device with $L_g = 100$ nm in comparison to the conventional single-layered s-Si device. Velocity overshoot has been perceived in the sub 100 nm device (Device-B) resulting in improved drive current along with high gm_{max} attained in the short channel device with $L_g = 50$ nm. Amplified transconductance was observed in this short channel nMOSFET due to the occurrence of ballistic transport of carriers within the channel region, which is a consequence of velocity overshoot in the HOI MOSFET.

Acknowledgements The authors thank the members of Department of ECE, NIT Mizoram for the facilities and support provided throughout the research and, in particular, the SMDP-C2SD project for allowing to use the workstation.

References

1. M. Reiche, Strained silicon devices solid state phenomena. Solid State Phenom. **156**, 61–68 (2010)
2. K. Rim, Fabrication and mobility characteristics of ultra-thin strained Si directly on insulator (SSDOI) MOSFETs. IEDM Tech. Dig. **49**, 49–52 (2003)
3. A. Wei, Integration challenges for advanced process strained CMOS on biaxially-strained SOI (SSOI) substrates. ECS Trans. **6**(1), 15–22 (2007)
4. C. Auth, A. Cappellani, J.S Chun, 45 nm high-k + metal gate strain-enhanced transistors. in *Symposium on VLSI Technology Digest*, 2008, pp. 128–129
5. A. Shimzu, Local mechanical-stress control (LMC): a new technique for CMOS-performance enhancement. in *IEDM Technology Digest*, 2001, pp. 433–436
6. H.S. Yang, Dual stress liner for high performance sub-45 nm gate length SOI CMOS manufacturing. in *IEDM Technology Digest*, 2004, pp. 1075–1077
7. K. Ota, Novel locally strained channel technique for high performance 55 nm CMOS. in *IEDM Technology Digest*, 2002, pp. 27–30
8. C.H. Chen, Stress memorization technique (SMT) by selectively strained-nitride capping for sub-65 nm high performance strained Si device application. in *Symposium on VLSI Technology Digest*, 2004, pp. 56–57
9. S. Gannavaram, N. Pesovic, M.C. Ozturk, Low temperature (<800 °C) recessed junction selective silicon-germanium source/ drain technology for sub-70 nm CMOS. in *IEDM Technology Digest*, 2000, pp. 437–440
10. D.A. Antoniadis, I. Aberg, C.N. Chleirigh, O.M. Nayfeh, A. Khakifirooz, J.L. Hoyt, Continuous MOSFET performance increase with device scaling: the role of strain and channel material innovations. IBM J. Res. Dev. **50**(4/5), 363–376 (2006)
11. C. Claeys, E. Simoen, S. Put, G. Giusi, F. Crupi, Impact strain engineering on gate stack quality and reliability. Solid State Electron. **52**(8), 115–1126 (2008)
12. M.J. Kumar, T.V. Singh, Quantum confinement effect in strained silicon MOSFET. Int. J. Nanosci. **7**(2), 81–84 (2008)
13. S. Flachowsky, Understanding strain-induced drive current enhancement in strained-silicon n-MOSFET and p-MOSFET. IEEE Trans. Electron Devices **57**(6), 1343–1354 (2010)
14. K. Rim, J. Welser, J.L. Hoyt, J.F. Gibbons, Enhanced hole mobilities in surface-channel strained Si p-MOSFETs. in *IEDM Technology Digest*, 1995, pp. 517–520
15. P.K. Tiwari, G.K. Saramekala, S. Dubey, A.K. Mukkhopadhyay, Analytical model for subthreshold current and subthreshold swing of short channel double material gate MOSFETs with strained-silicon channel on silicon-germanium substrates. J. Semicond. **35**(10), (2014)
16. D.J. Paul, Si/SiGe heterostructures: from material and physics to devices and circuits. Semicond. Sci. Technol. **19**(10), 75–108 (2004)
17. S.H. Olsen, High performance nMOSFETs using a novel strained Si/SiGe CMOS architecture. IEEE Trans. Electron Device **50**(9), 1961–1969 (2003)
18. S.H. Olsen, K.S.K. Kwa, L.S. Driscoll, S. Chattopadhyay, and A.G. O'Neil, Design, fabrication and characterization of strained Si/SiGe MOS transistors. IEE Proc. Circuits Devices Syst. **151**(5), (2004)
19. S.H. Olsen et al., Optimization of alloy composition for high performance strained-Si-SiGe n-channel MOSFET. IEEE Trans. Electron Devices **51**(7), 1156–1162 (2004)
20. T. Sentaurus, Sdevice User Guide, ver, G-2012.06, Synopsys, (2012)
21. T. Sentaurus, User manual, Synopsys, Inc., Mt. View, CA, Version F-2011.09, (2010)
22. H. Batwani, M. Gaur, M.J. Kumar, Analytical drain current model for nanoscale strained-Si/SiGe MOSFETs. COMPEL: Int. J. Comput. Math. Electr. Electron. Eng. **28**(2), (2009)

Electric Field Modeling and Critiques of Dual-Halo Dual-Dielectric Triple-Material Surrounding-Gate MOSFET



Prashant Kumar, Neeraj Gupta, Rashmi Gupta and Ganesh Gupta

1 Introduction

The advancements in technology demand for the small size, high speed, and low power dissipation devices. To achieve all these features, the dimension of the MOSFET has been scaled down aggressively. But short channel effects become barrier for further downscaling and degrade the device performance. In view of difficulties of the planar MOSFET technology to get the acceptable gate control over the channel, multiple gate-based devices are better technology option for further shrinking the size of the planar MOSFET. This development in MOSFET technology has proven to be an excellent option for futuristic devices. But this advancement in technology brings new challenges for the device engineers. The design of such advanced MOSFET can further enhance the performance by reducing the SCEs. Currently, diverse gate engineering structures have been examined by Long et al. [1]. One such device structure such as the triple-material gate has been investigated by Dhanaselvam et al. [2]. In this TMG structure, three dissimilar metal gate electrodes are utilized with disparate work function. This disparity produces a step potential profile and a crest field in the channel. It boosts up the carrier transportation efficiency and, therefore, overcomes the SCEs. A narrow interfacial oxide layer is used along with high-k dielectrics to decrease the density of interface trap charges. Therefore, it minimizes gate leakage and increases the electric field inside the channel [3, 4]. Incorporating halo implant further improves the device performance [5]. Hence, the SCEs can be further minimized by proposing a novel structure which amalgamates the benefits of TM-SG MOSFET and halo. This is also called as DH-DD-TM-CGAA MOSFET.

P. Kumar
J.C. Bose University of Science and Technology, Faridabad, India
e-mail: pk.vlsi@gmail.com

N. Gupta · R. Gupta (✉) · G. Gupta
Amity School of Engineering and Technology, Amity University Haryana, Gurugram, India
e-mail: goyal.rashmi18@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_139

2 Analytical Model

Figure 1 shows the schematic diagram of DH-DD-TM-SG MOSFET structure with gate terminal that consists of three metals with different work functions for surrounding-gate MOSFET: the gate material 1 with the work function $\Phi_{m1} = 4.8$ eV (Au), gate material 2 with $\Phi_{m2} = 4.6$ eV (Mo), and gate material 3 with $\Phi_{m3} = 4.4$ eV (Ti), respectively.

$t_{SiO_2} = 1$ nm and $t_{HfO_2} = 4$ nm are the thickness of the inner and outer oxide layers, respectively. The length of channel region and the diameter of the device are more than 30 nm and 10 nm, respectively. So for the recent analysis, quantum effects are ignored [6]. The simulated structure of the present device used in the TCAD Silvaco is depicted in Fig. 2. The lengths L_1 and $L_5 - L_4$ are halo doped with concentration N_{dh}, while the rest of the region are doped with acceptor doping concentration N_{ak}, assuming that N_{dh} is greater than N_{ak}.

The cylindrical coordinate system comprises a radial direction r, an angular component θ in the plane of radial direction, and a vertical direction z, as mentioned in

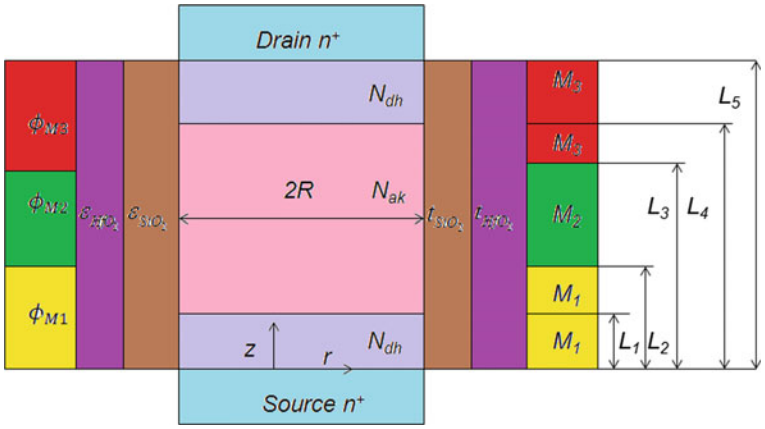


Fig. 1 Cross-sectional schematic of dual-halo dual-dielectric triple-material surrounding-gate MOSFET

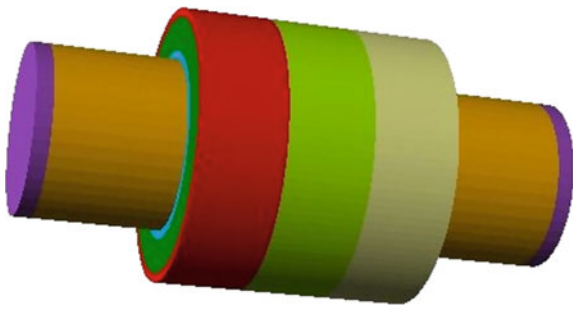


Fig. 2 Simulated structure of the DH-DD-TM-SG MOSFET used in the analysis

Fig. 1. The symmetry of the device ensures that surface potential and the field varies with r and z, but not with θ. So, 2D analysis is needed. The potential distribution in the channel can be expressed by the Poisson's equation and can be written as

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial[\phi_k(r,z)]}{\partial r}\right)+\frac{\partial^2[\phi_k(r,z)]}{\partial z^2}=\frac{qN_{\mathrm{ak}}}{\varepsilon_{\mathrm{Si}}};\quad (L_{k-1}\le z\le L_k) \tag{1}$$

where $k = 1, 2, 3, 4, 5$ and $N_{a1} = N_{a5} = N_{\mathrm{dh}}$, $N_{a2} = N_{a3} = N_{a4} = N_{\mathrm{ak}}$.

The potential distribution is approximated using the parabolic profile for DH-DD-TM-CGAA MOSFET in the radial direction and is given by [7]

$$\phi(r,z)=\zeta_0(z)+\zeta_1(z)r+\zeta_2(z)r^2 \tag{2}$$

where the constants $\zeta_0(z)$, $\zeta_1(z)$ and $\zeta_2(z)$ can be attained by substituting conditions at the boundary given in [8].

Surface potential is expressed as

$$\frac{d^2\phi_s(z)}{dz^2}-\phi_s(z)\left(\frac{2C_{\mathrm{oxeffdh}}}{\varepsilon_{\mathrm{Si}}R}\right)+(v_{\mathrm{gs}}-v_{\mathrm{fbk}})\left(\frac{2C_{\mathrm{oxeffdh}}}{\varepsilon_{\mathrm{Si}}R}\right)=\frac{qN_{\mathrm{ak}}}{\varepsilon_{\mathrm{Si}}} \tag{3}$$

The solution of second-order differential Eq. (3) through the complementary and the particular integral functions is given as

$$\phi_{\mathrm{Si}}(z)=\alpha_k e^{(\kappa z)}+\beta_k e^{(-\kappa z)}-\frac{\chi_k}{\kappa^2}\quad l_{k-1}\le \mathrm{z}\le l_k \tag{4}$$

where α_k and β_k are arbitrary constants, which are calculated with the help of continuity conditions for the surface potential distribution (ϕ) and the field distribution (E) at the interfaces of different metal gates [9].

The electric field distribution is obtained by differentiating surface potential w.r.t. vertical direction z. The electric field profile finds out the electron transport velocity through the channel.

$$E_{k(z)}=-\frac{d\phi_{\mathrm{sk}}(z)}{dz}=\kappa\left(\beta_k e^{(-\kappa z)}-\alpha_k e^{(\kappa z)}\right) \tag{5}$$

$$E(z)=\begin{cases}E_1(z);\ 0\le z\le L_1\\ E_2(z);\ L_1\le z\le L_2\\ E_3(z);\ L_2\le z\le L_3\\ E_4(z);\ L_3\le z\le L_4\\ E_5(z);\ L_4\le z\le L_5\end{cases} \tag{6}$$

3 Results and Discussion

3.1 Electric Field

The electric field of DH-DD-TM-CGAA MOSFETs has been studied with variation in device parameters such as drain bias, radius of silicon pillar, and oxide thickness. The electric field profile within the channel was simulated using ATLAS [10]. The correctness of the DH-DD-TM-SG MOSFET model was verified by ATLAS 2D device simulator and compared with the analytical model results.

3.1.1 Drain to Source Bias Variation

The deviation of electric field with channel lengths at various drain voltages of V_{ds} 0.1, 0.5, and 1 V is plotted in Fig. 3. The increase in drain bias has little effect in source field, but it raises the drain field. At higher drain bias, there is no significant change in the potential under M_1 for the present device, as compared to TM-SG. So, M_1 is known as the control gate, M_2 is known as first screen gate, and M_3 is known as a second screen gate. The reduction in electric field at the drain end is considered as a reduction in hot carrier effects [11].

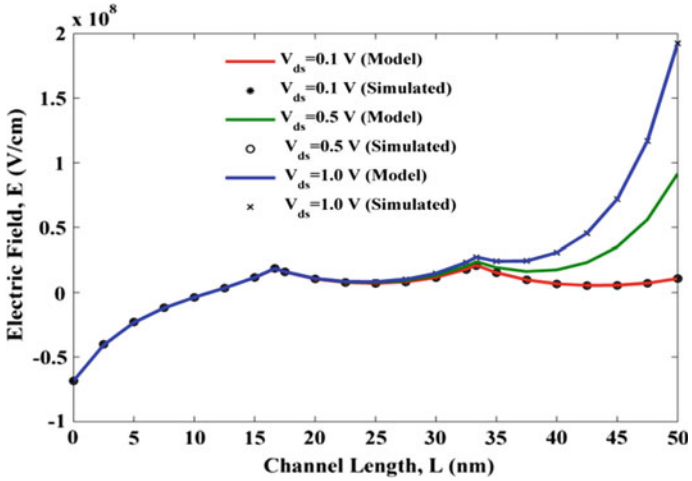


Fig. 3 Deviation in electric field with channel length at various drain voltages

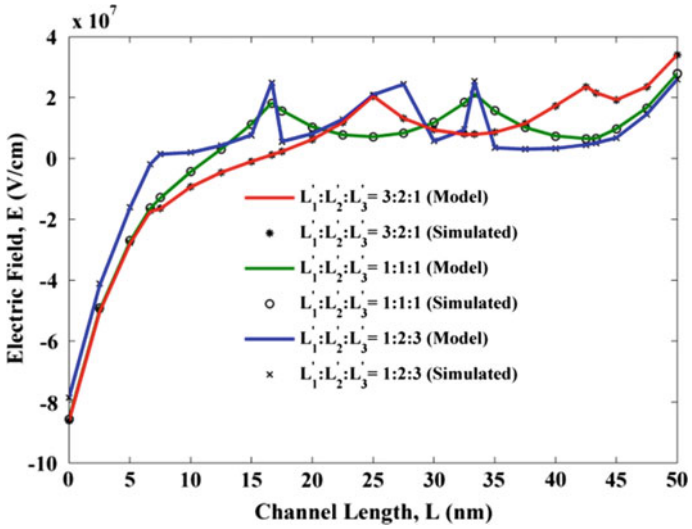


Fig. 4 Deviation in electric field with channel length at various gate length ratios

3.1.2 Gate Length Ratio Variation

The deviation in electric field with channel lengths at different combinations of gate length ratios, by keeping the total gate length constant is depicted in Fig. 4.

It is obvious from the figure that as the gate length ratio changes, then the corresponding deviation in the peak electric field. As the control gate length diminishes, the position of peak electric field is shifted toward the source side. This causes enhancement in carrier drift velocity and boosts up the carrier speed. The analytical results are well calibrated with simulated results using TCAD Silvaco.

3.1.3 Radius of Silicon Pillar Variation

The deviation in electric field with channel lengths at various silicon thicknesses is depicted in Fig. 5.

The electric field is plotted at a radius of silicon pillar 10, 20, and 30 nm, respectively. As the radius of silicon pillar/thickness of the silicon film decreases, then increase in an electric field is observed in devices. The peak of electric field is shifted toward the source side due to the difference in metal work function at the interface. This causes reduction in DIBL and SCEs. The DIBL reduced by 6.9% as compared to existing structure at $R = 10$ nm. The model results are verified using TCAD Silvaco.

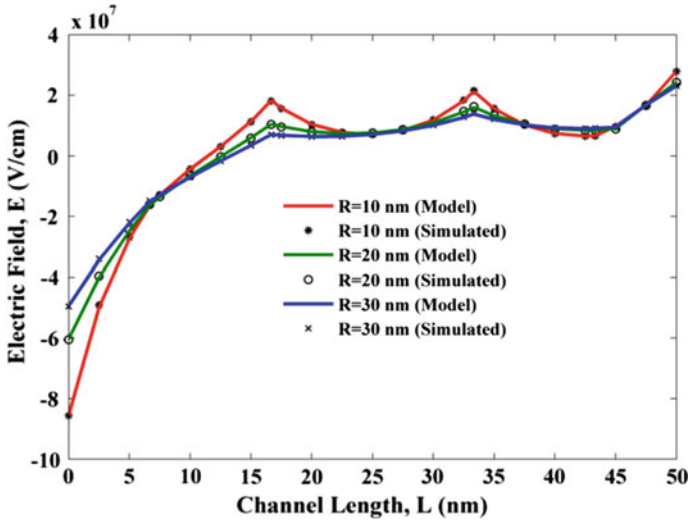


Fig. 5 Deviation in electric field with channel length at various radii of silicon pillar

3.1.4 Oxide-Thickness Variation

The deviation in electric field with channel lengths at various oxide thicknesses is depicted in Fig. 6. The electric field is plotted at oxide thickness 2, 4, and 6 nm, respectively. It is clear from the figure that lower values of oxide thickness produce

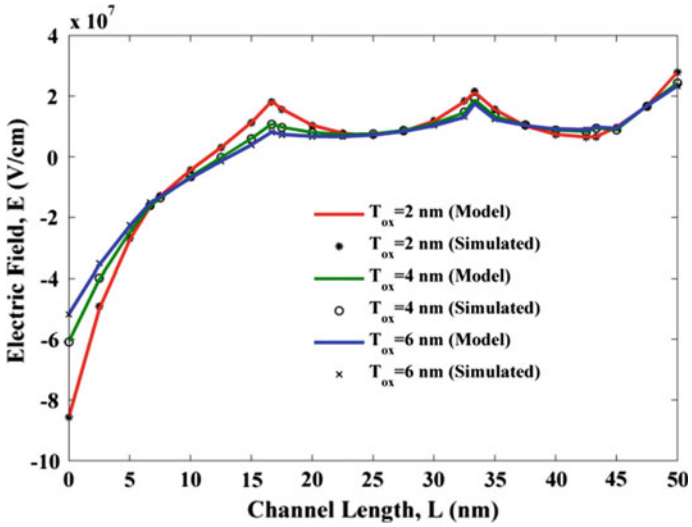


Fig. 6 Deviation in electric field with channel length at various oxide thicknesses

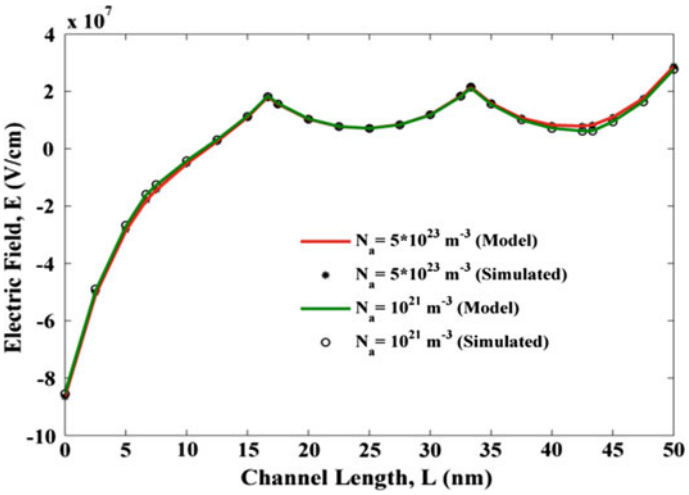


Fig. 7 Deviation in electric field with channel length at various channel doping concentrations

the highest peak of electric field. The reduce oxide thickness also produces the large electric field which increases the average electric field. The analytical results are well matched with the simulated result using TCAD Silvaco.

3.1.5 Channel Doping Variation

The deviation in electric field with channel length at various doping concentrations is plotted in Fig. 7.

The plot shows that changing the doping concentration in the channel increases the barrier height and reduces the electric field. The increase in doping concentration reduces the electric field which is mainly due to an increase in barrier height, which restricts the flow of charge carriers through the channels. This can be interpreted as reduction in hot carrier effects (HCEs). The analytical results are in close agreement with the simulated results.

4 Comparison

The device parameters used during the simulation are listed in Table 1.

Electric Field

Figure 3.19 shows the electric field distribution of DH-DD-TM-SG, SH-TM-SG, and TM-SG MOSFET structure.

Table 1 Summary of the device parameters used for simulation

Parameters	DH-DD-TM-SG	SH-TM-SG	TM-SG
S/D doping (/m^3)	1×10^{26}	1×10^{26}	1×10^{26}
Channel doping N_{ak} (/m^3)	1×10^{23}	1×10^{23}	1×10^{23}
Channel halo doping N_{dh} (/m^3)	1×10^{24}	1×10^{24}	–
Oxide thickness (nm)	1.78	1.78	1.78
Silicon thickness (nm)	10	10	10
Source/drain length (nm)	15	15	15
Permittivity of SiO_2	3.9	3.9	3.9
Permittivity of HfO_2	20	–	–

The drain side having lower metal gate M_3 work function region points to a valuable diminution in electric field. The carriers in the channel region experience fall in the electric field near to drain region which causes reduction in HCEs. Therefore, it improves the overall transport efficiency. From Fig. 8, it is noticed that an additional

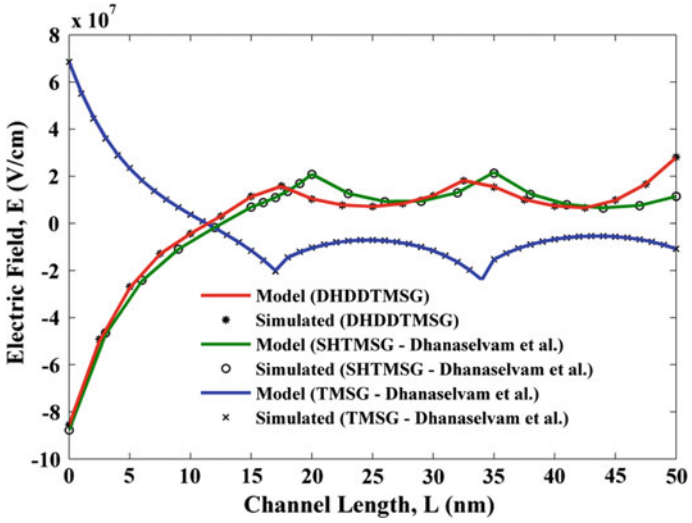


Fig. 8 Electric field distributions with varying channel lengths for TM-SG, SH-TM-SG, and DH-DD-TM-SG devices

peak of field is observed over the halo barrier, which further enhances the characteristics of the novel device. In contrary to single halo structure, there are more peaks in dual-halo structure, thereby improving immunity against SCEs. The further increment in halo doping concentration reforms the beginning peak. The second peak of the gate metal improves with further rise in work function, hence speeds up the carriers in the channel. The analytical results are close to simulated results which validate the model [12].

5 Conclusion

In this paper, a 2D analytical model of electric field has been developed. The electric field has been observed with various device parameters like oxide thickness, silicon thickness, gate length ratios, and doping concentration. Using gate engineering structure along with halo implant shows further improvement in minimization of SCEs. The step potential profiles subdue the SCEs. Thus, it is manifested that the combination of triple metal gates along with dual dielectric and halo implant architecture leads to further enhance in short channel behavior. Thus, it further ensures improvement in the carrier transport efficiency. The DH-DD-TM-SG MOSFET shows superior performance over TM-SG device for lowering HCEs and SCEs. The results obtained from TCAD Silvaco are well matched with the analytical results obtained from MATLAB.

References

1. W. Long, Dual-material gate (DMG) Field effect transistor. IEEE Trans. Electron Devices **46**(5), 865–870 (1999)
2. P.S. Dhanaselvam, G.C.V. Chakaravarthi, R.P. Ramesh, N.B. Balamurugan, A 2D analytical modeling of single halo triple material surrounding gate MOSFET. J. Electr. Eng. Technol. **9**(4), 1355–1359 (2004)
3. A. Sharama, A. Jain, Y. Pratap, R.S. Gupta, Effect of high K and vacuum dielectrics as gate stack on junctionless cylindrical surrounding gate MOSFET. Solid State Electron. **123**, 26–32 (2016)
4. J.P. Charles, D. Nirmal, P. Prajoon, J. Ajayan, Implementation of nanoscale circuits using dual metal gate engineered nanowire MOSFET with high-k dielectrics for low power application. Physica E **83**, 95–100 (2016)
5. A. Sarkar, D. Swapnadip, C.K. Sarkar, Asymmetric halo and symmetric single-halo dual-material gate and double-halo dual-material gate n-MOSFETs characteristic parameter modeling. Int. J. Numer. Model. Electron. Networks Devices Fields **26**, 41–55 (2012)
6. Y. Naveh, K.K. Likharev, Modeling of 10 nm scale ballistic MOSFETs. IEEE Electron Device Lett. **21**(5), 242–244 (2005)
7. K.K. Young, Short channel effect in fully depleted SOI MOSFETs. IEEE Trans. Electron Devices **36**(2), 399–402 (1989)
8. N. Gupta, J.K.B. Patel, A.K. Raghav, Modeling and analysis of threshold voltage for Dual-Halo Dual-Dielectric Triple-Material Surrounding-Gate MOSFET. Int. J. Pure Appl. Math. **118**(18), 3759–3771 (2018)

9. N. Gupta, J.K.B. Patel, A.K. Raghav, Performance and a new 2-D analytical modeling of a Dual-Halo Dual-Dielectric Triple-Material Surrounding-Gate-All-Around (DH-DD-TM-SGAA) MOSFET. J. Eng. Sci. Technol. **13**(11), (2018)
10. SILVACO International ATLAS User's Manual Device Simulation Software (2017)
11. E. Goel, S. Kumar, K. Singh, B. Singh, M. Kumar, S. Jit, 2-D analytical modeling of threshold voltage for Graded-Channel Dual-Material Double-Gate MOSFETs. IEEE Trans. Electron Devices **63**, 966–973 (2016)
12. P.K. Tiwari, S. Dubey, M. Singh, S. Jit, A two-dimensional analytical model for threshold voltage of short-channel triple-material double-gate metal-oxide-semiconductor field-effect transistors. J. Appl. Phys. **108**, (2010)

Modeling and Logic Synthesis of Multifunctional and Universal 3 × 3 Reversible Gate for Nanoscale Applications



Naira Nafees, Insha Manzoor, Majid Irfan Baba, Soha Maqbool Bhat, Vishal Puri and Suhaib Ahmed

1 Introduction to Reversible Logic

Reversible logic is an eminent technology in low-power calculations. It is an alternative to design digital circuits with almost zero power dissipation [1]. In reversible circuits, there is a one-to-one mapping between its input and output, i.e., number of outputs is equal to the number of inputs of the gates. This results in a unique set of output vector for each set of input vector. This one-to-one mapping is called bijective property [2, 3]. According to Landauer, for irreversible calculations, each bit loss generates KBTln2 Joules of heat energy [4]. Bennett supported this principle by proving that zero power dissipation can be obtained only if the circuit contains reversible gates [5, 6]. Reversible logic calculations will lead to improved energy efficiency and this will essentially increase the computation speed of various applications. In reversible circuits, there is no loss of information [7]. Hence, this design approach, i.e., reversible logic is of more importance in designing combinational as well as sequential circuits. There is no possibility of fan-out in reversible logic and feedback systems are not permitted. However, fan-out in reversible circuits is obtained with the help of additional gates. Some of the desirable features of a reversible logic basic circuits are [3]:

- Least number of input constants.
- Less reversible gates.
- Less garbage outputs.

Reversible logic has extensive applications in optical and DNA computation, low-power design, nanotechnology, quantum computing, etc. Reversible logic will bring down the circuit component sizes to atomic size, and hence, the devices will be much

N. Nafees · I. Manzoor · M. I. Baba · S. M. Bhat · V. Puri · S. Ahmed (✉)
Department of Electronics and Communication Engineering, Baba Ghulam Shah Badshah University, Rajouri, India
e-mail: sabatt@outlook.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_140

Table 1 Common reversible gates

Gate	Gate size	P	Q	R	S
Feyman [18]	2 × 2	A	A ⊕ B	–	–
Double Feyman [19]	3 × 3	A	A ⊕ B	A ⊕ C	–
Toffoli [20]	3 × 3	A	B	AB ⊕ C	–
Fredkin [21]	3 × 3	A	A′B ⊕ AC	A′C ⊕ AB	–
RG-QCA [11]	3 × 3	A ⊕ B	(A ⊕ B) C ⊕ AB	(B ⊕ C)′ A ⊕ B′C	–
Peres [22]	3 × 3	A	A ⊕ B	AB ⊕ C	–
HNG [23]	4 × 4	A	B	A ⊕ B ⊕ C	(A ⊕ B) C ⊕ AB ⊕ D
HNFG [24]	4 × 4	A	A ⊕ C	B	B ⊕ D
MRLG [25]	4 × 4	A	AB ⊕ A′C	B ⊕ AC	B ⊕ AC ⊕ D
RAM [26]	4 × 4	A	A ⊕ B	A ⊕ B ⊕ C	A ⊕ B ⊕ C ⊕ D

more easily carried [8, 9]. Some of the common reversible gates are given in Table 1 [10].

2 Proposed Gate

In this work, a novel 3 × 3 NNG reversible logic gate and its equations, shown in Fig. 1, are proposed for use in the design of various optimal combinational circuits like the adder circuits. The three-input and three-output vectors IV and OV of the proposed gate are given as

$$\text{IV} = (\text{A}, \text{B}, \text{C}); \text{OV} = (\text{P}, \text{Q}, \text{R}) \tag{1}$$

$$\text{P} = \text{C}, \text{Q} = \text{AB} + \text{BC}' + \text{A}'\text{B}'\text{C}', \text{R} = \text{AB} + \text{AC}' + \text{A}'\text{B}'\text{C}. \tag{2}$$

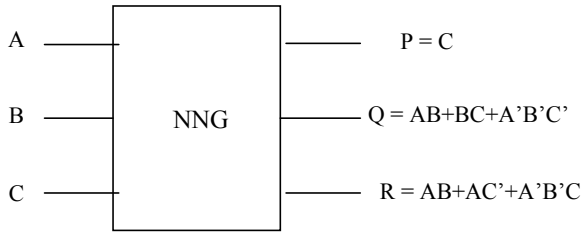


Fig. 1 Block diagram representation of proposed NNG gate

Table 2 Truth table of proposed 3 × 3 NNG gate

Input encoding	Inputs			Outputs			Output encoding
	A	B	C	P	Q	R	
0	0	0	0	0	1	0	2
1	0	0	1	1	0	1	5
2	0	1	0	0	0	0	0
3	0	1	1	1	1	0	6
4	1	**0**	0	0	0	1	1
5	1	0	1	1	0	0	4
6	1	1	0	0	1	1	3
7	1	1	1	1	1	1	7

It is observed from the truth table of the proposed NNG gate, presented in Table 2, that it is reversible in nature as it has one-to-one mapping of input and output.

3 Proposed Gate as Universal Structure

The ability to derive or implement various Boolean operations leads to the design of an efficient reversible gate and if the reversible gate is able to implement all the seven logic gates, then the gate can be called as a universal in nature. In order to verify the universality of the proposed 3 × 3 NNG reversible gate, various input combinations are considered as follows. If the NNG gate is to be used as an inverter i.e. NOT gate, then the desired input whose compliment is to be achieved is applied at input A and a logic zero is applied to input A and B. If the input B is set to a logic one an OR operation can be performed for the inputs A and C. If the input B is set to logic zero NOR function is performed between the inputs A and C at output Q of the gate. Now, if the compliments of A and C inputs are applied and logic zero is applied to input B, then AND operation is performed for inputs A and C and if complimentary inputs are applied at the A and C inputs along with logic one at input B then, a NAND function can be realized for inputs A and C. Further, logic zero is given at input B and inputs A and C are kept as it is to obtain exclusive OR operation between inputs A and C from output R of the gate and logic zero is given to input C while complimentary inputs are given at both inputs A and B to perform exclusive NOR operation for the same configuration. The universality of NNG Gate for all the seven gates is shown in Table 3.

Table 3 Proposed gate as universal structure

S. no.	Logic gate	NNG function
1	NOT	NNG(A,0,0)
2	AND	NNG(A′,0,C′)
3	NAND	NNG(A′,1,C′)
4	OR	NNG(A,1,C)
5	NOR	NNG(A,0,C)
6	EX-OR	NNG(A,0,C)
7	EX-NOR	NNG(A′,B′,0)

4 Implementation of 13 Standard Functions Using Proposed Gate

The effectiveness of NNG gate is evaluated by implementing the 13 standard logic functions [11]. These can be realized using the proposed 3 × 3 NNG reversible gate, either by manipulating the inputs or by using more than one NNG gate in the implementation. The block diagram representations of the 13 Standard Boolean Function Implemented using the proposed 3 × 3 NNG reversible gate is shown in Fig. 2.

5 Performance Comparison

The suitability of the proposed 3 × 3 NNG gate is verified by comparing the total number of gates required for the implementation of 13 functions, as shown in Table 4, whereas the percentage improvement in terms of the gates required is shown in Table 5.

It is evident that the proposed 3 × 3 NNG gate is highly area and cost-efficient as it requires minimum gates to implement all 13 Boolean functions thereby making it suitable for low-power nano-applications.

6 Future Work

Implementation of reversible circuits requires technologies which are quantum-mechanically reversible, the criteria which CMOS technology does not satisfy. Hence, there is the need for an alternate technology to implement reversible logic. One such technology which can be used is Quantum-dot Cellular Automata which involves the design of circuits using quantum-dot nanostructures [11, 12]. A QCA cell, which is the basic component of QCA, consists of four quantum dots or potential well, in which two electrons are localized, as shown in Fig. 3.

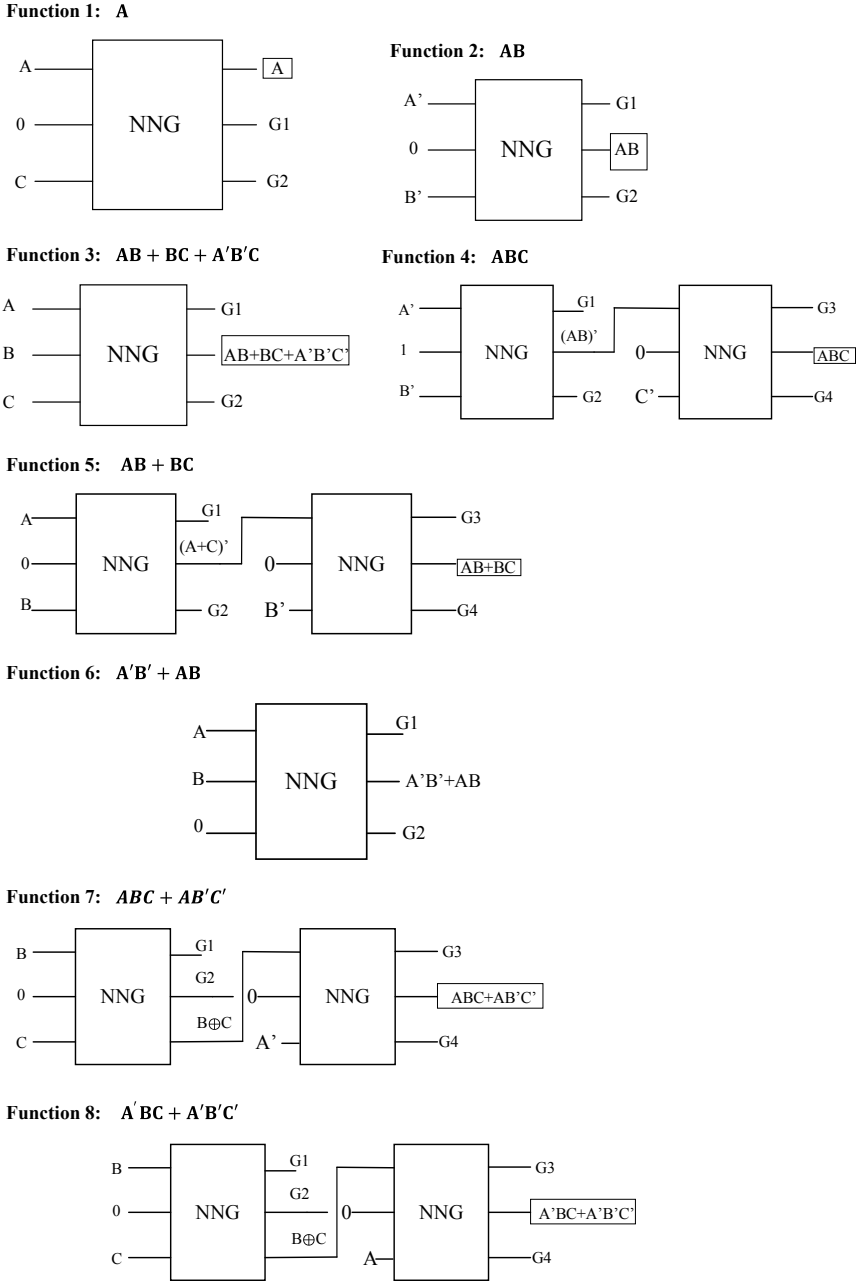
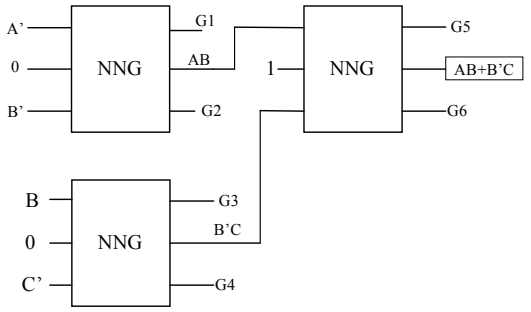
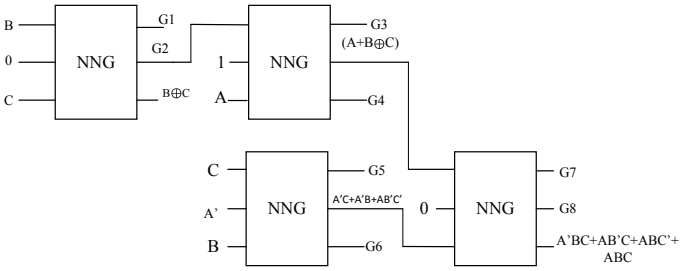


Fig. 2 Block diagram representation of implementation of 13 standard functions

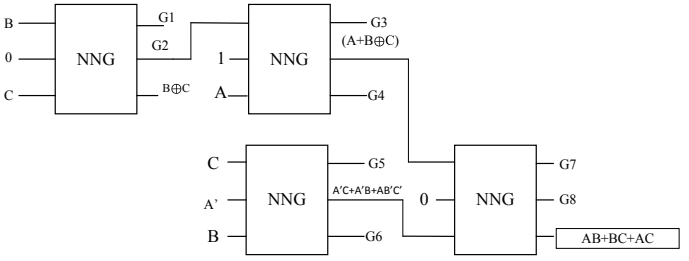
Function 9: $AB + B'C$



Function 10: $A'BC + AB'C + ABC' + ABC$



Function 11: $AB + BC + AC$



Function 12: $ABC' + A'B'C' + A'BC$

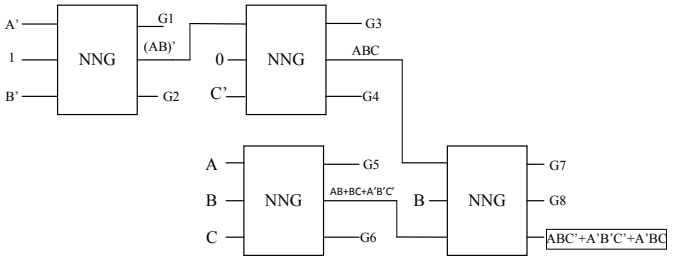


Fig. 2 (continued)

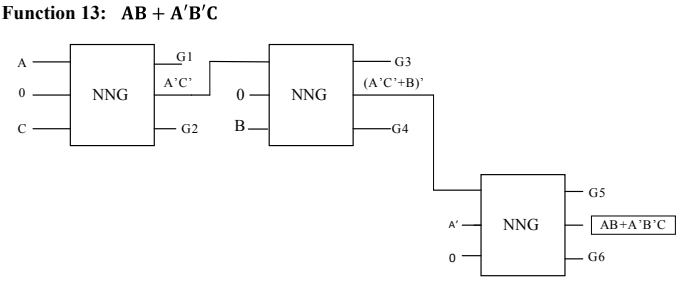


Fig. 2 (continued)

Table 4 Performance comparison of proposed gate

S. no.	CQCA [27]	Toffoli [20]	Fredkin [21]	RG-QCA [11]	Proposed NNG gate
ABC	2	2	2	2	2
AB	1	1	1	1	1
AB + BC + CA	1	5	4	1	4
A	1	1	2	1	1
A′BC + A′B′C′	3	**3**	5	2	2
AB + A′B′	4	2	2	5	1
ABC′ + A′BC + A′B′C′	6	6	1	5	4
A′BC + AB′C + ABC′ + ABC	3	2	1	1	4
ABC + AB′C′	6	5	4	5	2
AB + B′C	3	3	3	3	3
AB + BC	2	2	2	2	2
AB + A′B′C	5	5	4	5	3
AB + BC + A′B′C′	6	5	5	5	1
Total gates used	43	42	41	34	30

Table 5 Percentage of improvement in proposed gate

Reversible gate	Gate count	Gate count in proposed NNG gate	% Age improvement in proposed gate
CQCA [27]	43	30	30.23
Toffoli [20]	42		28.57
Fredkin [21]	41		26.82
RG-QCA [11]	34		11.76

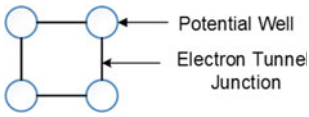


Fig. 3 Basic concept of a QCA cell

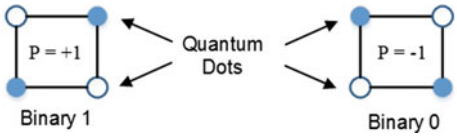


Fig. 4 Binary 1 and binary 0 representations of a QCA cell along with their polarizations

These dots are separated by tunnel junctions and it works on the concept of Coulombic repulsion and quantum mechanical tunneling which leads to two polarization states, viz., binary "1" and binary "0" as shown in Fig. 4 [13, 14].

QCA circuits have clock signals consisting of four zones, viz., switch, hold, release, and relax [15–17], which are required for its proper functioning and computing the delay in the circuit. In the future, the proposed 3 × 3 NNG gate can be implemented in QCA along with the design of various combinational and sequential circuits using the NNG gate such as half adders, half subtractors, full adders, etc.

7 Conclusion

A new 3 × 3 NNG reversible gate was proposed which was then verified to be a universal structure in nature. Further, the multifunctionality of the proposed gate was verified by implementing the 13 standard Boolean functions. Based on the performance comparison, it is observed that the proposed NNG gate is cost-efficient and achieves improvement up to 30.23% in terms of the number of gates required to implement the 13 functions.

References

1. K. Das, D. De, Characterization, test and logic synthesis of novel conservative and reversible logic gates for QCA. Int. J. Nanosci. **9**, 201–214 (2010)
2. A.N. Bahar, S. Waheed, N. Hossain, A new approach of presenting reversible logic gates in nanoscale. SpringerPlus **4**(153), 1–7 (2015)
3. P. Biswas, N. Gupta, N. Patidar, Basic reversible logic gates and its QCA implementation. Int. J. Eng. Res. Appl. **4**, 12 (2014)
4. R. Landauer, Irreversibility and heat generation in the computing process. IBM J. Res. Dev. **5**, 183–191 (1961)
5. A new design and simulation of reversible gates in quantum-dot cellular automata technology, https://arxiv.org/abs/1803.11017. Last accessed 18 Feb 2019
6. C.H. Bennett, Logical reversibility of computation. IBM J. Res. Dev. **17**, 525–532 (1973)

7. H. Thapliyal, N. Ranganathan, Reversible logic-gates concurrently testable latches for molecular QCA. IEEE Trans. Nanotechnol. **9**, 62–69 (2010)
8. M.A. AlShafi, M.S. Islam, A.N. Bahar, A review on reversible logic gates and its QCA implementations. Int. J. Comput. Appl. **128**, 27–34 (2015)
9. P.R. Yelekar, S.S. Chiwande, Introduction to reversible logic gates and its applications. in *2nd International Conference on Information and Communication Technology*, pp. 5–9 (2011)
10. S.A. Hassan, P. Sharam, Low power quantum gates for the implementation of reversible memory elements using quantum dot cellular automata. in *Proceedings of 3rd International Conference on Computing for Sustainable Global Development 2016* (IEEE 2016), pp. 2727–2732
11. B. Bilal, S. Ahmed, V. Kakkar, Modular adder designs using optimal reversible and fault tolerant gates in field-coupled QCA nanocomputing. Int. J. Theor. Phys. **57**(5), 1356–1375 (2018)
12. F. Ahmad, S. Ahmed, V. Kakkar, G.M. Bhat, A.N. Bahar, S. Wani, Modular design of ultra-efficient reversible full adder-subtractor in QCA with power dissipation analysis. Int. J. Theor. Phys. **57**(9), 2863–2880 (2018)
13. B. Bilal, S. Ahmed, V. Kakkar, Quantum dot cellular automata: a new paradigm for digital design. Int. J. Nanoelectron. Mater. **11**(1), 87–98 (2018)
14. B. Bilal, S. Ahmed, V. Kakkar, An insight into beyond CMOS next generation computing using quantum-dot cellular automata nanotechnology. Int. J. Eng. Manuf. **8**(1), 25–37 (2018)
15. B. Bilal, S. Ahmed, V. Kakkar, QCA based efficient toffoli gate design and implementation for nanotechnology applications. Int. J. Eng. Technol. **9**(3s), 84–92 (2017)
16. B. Bilal, S. Ahmed, V. Kakkar, Optimal realization of universality of peres gate using explicit interaction of cells in quantum dot cellular automata nanotechnology. Int. J. Intell. Syst. Appl. **9**(6), 75–84 (2017)
17. B. Bilal, S. Ahmed, V. Kakkar, Multifunction reversible logic gate: logic synthesis and design implementation in QCA. in *Proceedings of International Conference on Computing Communication and Automation* (IEEE 2017), pp. 1385–1390
18. R. Feyman, Quantum mechanical computers. Found. Phys. **16**(6), 507–531 (1986)
19. B. Parhami, Fault tolerant reversible circuits. in *Proceedings of 40th Asilomar Conference on Signals Systems and Computers* (2006). pp. 1–4
20. T. Toffoli, Reversible computing. in *Proceedings of International Colloquium on Automata, Languages and Programming*, (Springer, Berlin, Heidelberg 1980). pp. 632–644
21. E. Fredkin, T. Toffoli, Conservative logic. Int. J. Theor. Phys. **21**, 219–253 (1982)
22. A. Peres, Reversible logic and quantum computers. Phys. Rev. **A32**, 3266–3276 (1985)
23. A.K. Biswas, M.M. Hasan, A.R. Chowdhury, H.M.H. Babu, Efficient approaches for designing reversible binary coded decimal adders. Microelectron. J. **39**, 1693–1703 (2008)
24. M. Haghparast, K.A. Navi, Novel reversible BCD adder for nanotechnology based systems. Am. J. Appl. Sci. **5**, 282–288 (2008)
25. V. Kumar, D. Dhawan, Design of reversible adder subtractor using multifunction reversible logic gate (MRLG). Int. J. Adv. Comput. Electron. Eng. **2**, 5–11 (2016)
26. H. Rangaraju, A.B. Suresh, K. Muralidhara, Design and optimization of reversible multiplier circuit. Int. J. Comput. Appl. **52**(10), 44–50 (2012)
27. H. Thapliyal, N. Ranganathan, Conservative QCA gate (CQCA) for designing concurrently testable molecular QCA circuits. in *Proceedings of 22nd International Conference on VLSI Design* (2009), pp. 511–516

Logic Design and Modeling of an Ultraefficient 3 × 3 Reversible Gate for Nanoscale Applications



Insha Manzoor, Naira Nafees, Majid Irfan Baba, Soha Maqbool Bhat, Vishal Puri and Suhaib Ahmed

1 Introduction to Reversible Logic

Reversible logic is an eminent technology in low-power calculations and an alternative to design digital circuits with almost zero power dissipation [1]. In reversible circuits there is a one-to-one mapping between its input and output, i.e., number of outputs is equal to the number of inputs of the gates. This results in a unique set of output vector for each set of input vector. This one-to-one mapping is called bijective property [2, 3]. According to R. Landauer, for irreversible calculations, each bit loss generates KBTln2 Joules of heat energy [4]. Bennett supported this principle by proving that zero power dissipation can be obtained only if the circuit contains reversible gates [5, 6]. Reversible logic calculations will lead to improved energy efficiency and this will essentially increase the computation speed of various applications. In reversible circuits, there is no loss of information [7]. Hence, this design approach, i.e., reversible logic, is of more importance in designing combinational as well as sequential circuits. There is no possibility of fan-out in reversible logic and feedback systems are not permitted. However, fan-out in reversible circuits is obtained with the help of additional gates. Some of the desirable features of a reversible logic basic circuits are [3]:

- Least number of input constants.
- Less reversible gates and garbage outputs.

Reversible logic has extensive applications in optical and DNA computation, low-power design, nanotechnology, quantum computing, etc. Reversible logic can bring down the circuit component sizes to atomic size, and hence, the devices will be

I. Manzoor · N. Nafees · M. Irfan Baba · S. Maqbool Bhat · V. Puri · S. Ahmed (✉)
Department of Electronics and Communication Engineering, Baba Ghulam Shah Badshah University, Rajouri, India
e-mail: sabatt@outlook.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_141

Table 1 Common reversible gates

Gate	Gate size	P	Q	R	S
Feyman [17]	2 × 2	A	A ⊕ B	–	–
Double Feyman [18]	3 × 3	A	A ⊕ B	A ⊕ C	–
Toffoli [19]	3 × 3	A	B	AB ⊕ C	–
Fredkin [20]	3 × 3	A	A′B ⊕ AC	A′C ⊕ AB	–
RG-QCA [11]	3 × 3	A ⊕ B	(A ⊕ B) C ⊕ AB	(B ⊕ C)′ A ⊕ B'C	–
Peres [21]	3 × 3	A	A ⊕ B	AB ⊕ C	–
HNG [22]	4 × 4	A	B	A ⊕ B ⊕ C	(A ⊕ B) C ⊕ AB ⊕ D
HNFG [23]	4 × 4	A	A ⊕ C	B	B ⊕ D
MRLG [24]	4 × 4	A	AB ⊕ A′C	B ⊕ AC	B ⊕ AC ⊕ D
RAM [25]	4 × 4	A	A ⊕ B	A ⊕ B ⊕ C	A ⊕ B ⊕ C ⊕ D

much more easily carried [8, 9]. Some of the common reversible gates are presented in Table 1 [10].

2 Proposed Gate

In this paper, a novel 3 × 3 IMGG reversible logic gate and its equations, shown in Fig. 1, are proposed for use in the design of various optimal combinational circuits like the adder circuits. The three-input and three-output vectors IV and OV of the proposed gate are given as

$$\mathrm{IV} = (\mathrm{A}, \mathrm{B}, \mathrm{C}); \mathrm{OV} = (\mathrm{P}, \mathrm{Q}, \mathrm{R}) \tag{1}$$

$$\mathrm{P} = \mathrm{B}, \mathrm{Q} = \mathrm{A}'\mathrm{C} + \mathrm{A}\left(\mathrm{BC} + \mathrm{B}'\mathrm{C}'\right), \mathrm{R} = \mathrm{A}' \tag{2}$$

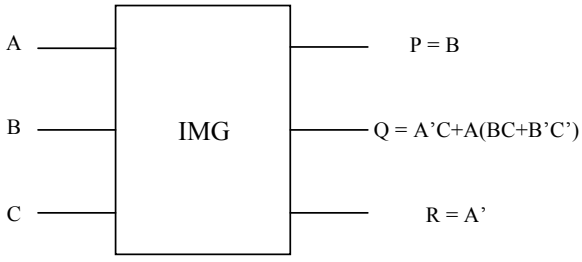


Fig. 1 Block diagram representation of proposed IMG gate

Table 2 Truth table of proposed 3 × 3 IMG gate

Input encoding	Inputs			Outputs			Output encoding
	A	B	C	P	Q	R	
0	0	0	0	0	0	1	1
1	0	0	1	0	1	1	3
2	0	1	0	1	0	1	5
3	0	1	1	1	1	1	7
4	1	**0**	0	0	1	0	2
5	1	0	1	0	0	0	0
6	1	1	0	1	0	0	4
7	1	1	1	1	1	0	6

It is observed from the truth table of the proposed IMG gate, presented in Table 2, that it is reversible in nature as it has one-to-one mapping of input and output.

3 Proposed Gate as Universal Structure

The ability to derive or implement various Boolean operations leads to the design of an efficient reversible gate and if the reversible gate is able to implement all the seven logic gates, then the gate can be called as a universal in nature. In order to verify the universality of the proposed 3 × 3 IMG reversible gate, various input combinations are considered as follows. If the IMG is to be used as an inverter, i.e., NOT gate then the desired input whose compliment is to be achieved is applied at input A, a logic 0 is applied to input B and a logic 1 is applied at input C. In order to perform AND operation for two inputs A and B, the first input is kept as it is at input A, second input is set to the complemented value of B at the input B and input C is set to logic 0. Now, if the first input is applied at input A, complemented second input is applied at input B and logic 1 is applied at input C, NAND operation can be realized for the two inputs. To obtain OR operation, input A is set to the complimented value of the first input, input B is set to the second input, and input C is set to logic 1. A NOR operation can be realized, if input A is set to the complimented value of the first input, input B is set to the second input, and input C is set to logic 0. The exclusive OR operation for inputs A and C is obtained by applying logic 0 at input B and inputs A and C are kept as it is. To perform exclusive NOR operation, the first input is applied to the input A, logic 0 is applied at input B, and the complemented value of the second input is applied at input C. The universality of IMG gate for all the seven gates is shown in Table 3.

Table 3 Proposed gate as universal structure

S. no.	Logic gate	IMG function
1	NOT	IMG(A,0,1)
2	AND	IMG(A,B′,0)
3	NAND	IMG(A,B′,1)
4	OR	IMG(B′,C,1)
5	NOR	IMG(B′,C,0)
6	EX-OR	IMG(A,0,C)
7	EX-NOR	IMG(A,0,C′)

4 Implementation of 13 Standard Functions Using Proposed Gate

The effectiveness of the proposed 3 × 3 IMG gate is evaluated by implementing the 13 standard logic functions [11]. These can be realized using the proposed 3 × 3 IMG reversible gate, either by manipulating the inputs or by using more than one IMG gate in the implementation. The block diagram representations of the 13 standard Boolean functions implemented using the proposed IMG gate are shown in Fig. 2.

5 Performance Comparison

The suitability of proposed 3 × 3 IMG gate is verified by comparing the total number of gates required for the implementation of 13 functions, as shown in Table 4, whereas the percentage improvement in terms of the gates required is shown in Table 5.

It is evident that the proposed 3 × 3 IMG gate is highly area and cost-efficient as it requires minimum gates to implement all 13 Boolean functions, thereby making it suitable for low-power nano-applications.

6 Future Work

Implementation of reversible circuits requires technologies which are quantum-mechanically reversible, the criteria which CMOS technology does not satisfy. Hence, there is a need for an alternate technology to implement reversible logic. One such technology which can be used is quantum-dot cellular automata which involve the design of circuits using quantum-dot nanostructures [11, 12]. A QCA cell, which is the basic component of QCA, consists of four quantum dots or potential well, in which two electrons are localized, as shown in Fig. 3.

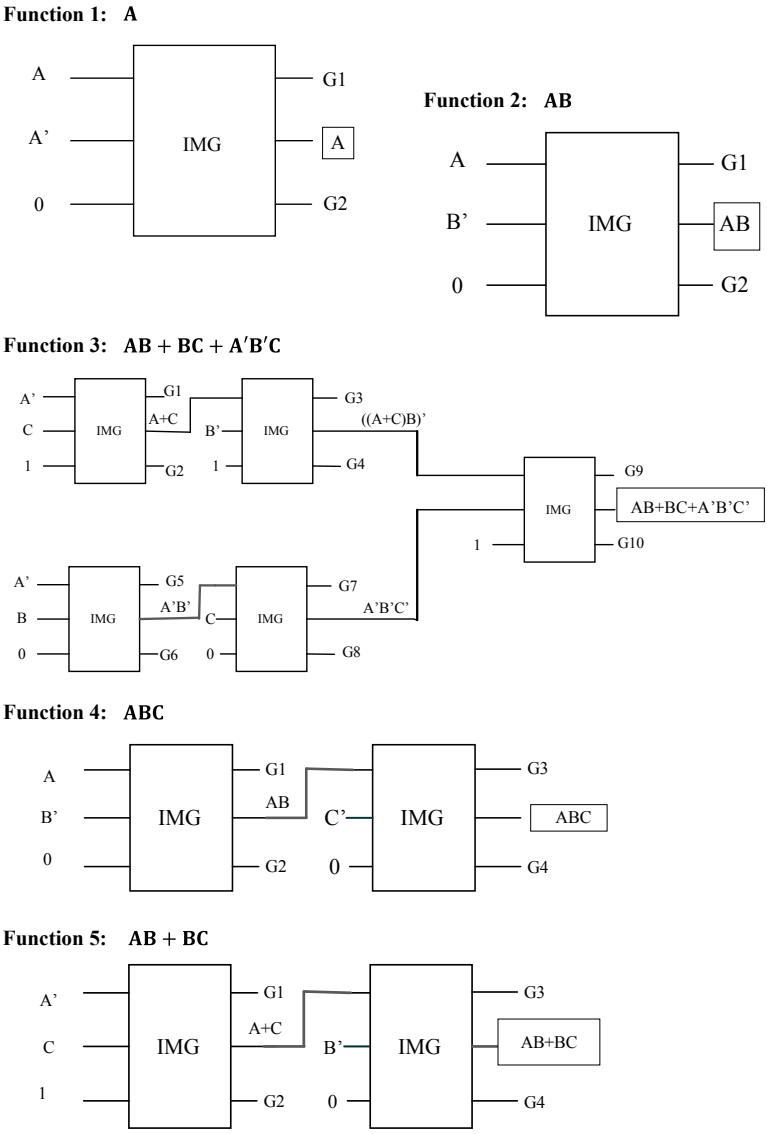
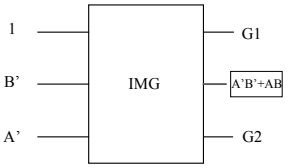
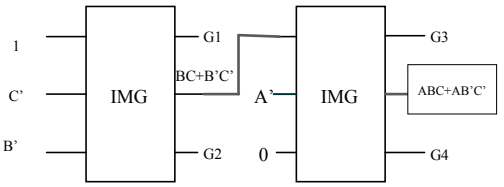


Fig. 2 Block diagram representation of implementation of 13 standard functions

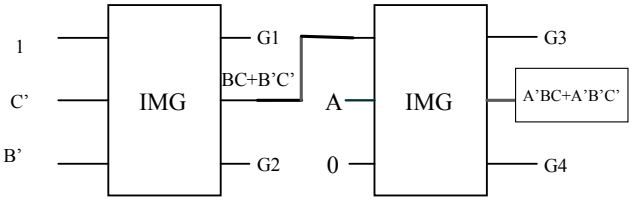
Function 6: $A'B' + AB$



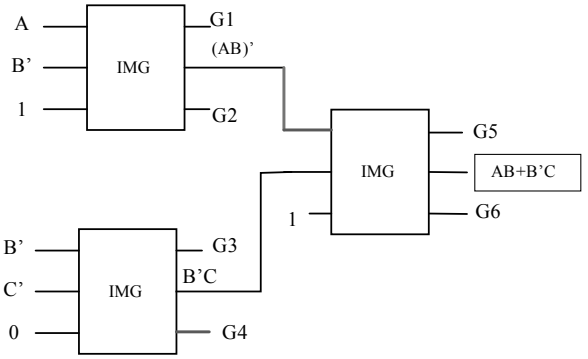
Function 7: $ABC + AB'C'$



Function 8: $A'BC + A'B'C'$



Function 9: $AB + B'C$



Function 10: $A'BC + AB'C + ABC' + ABC$

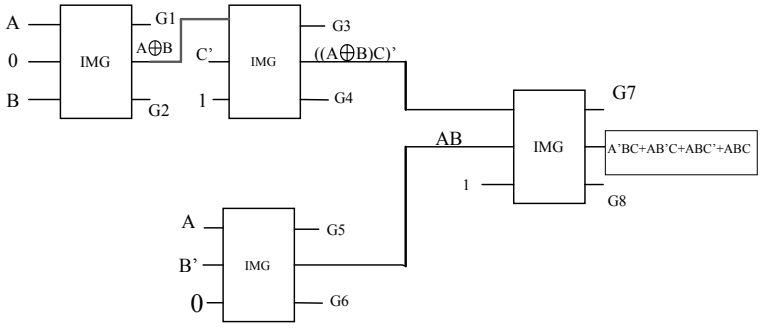
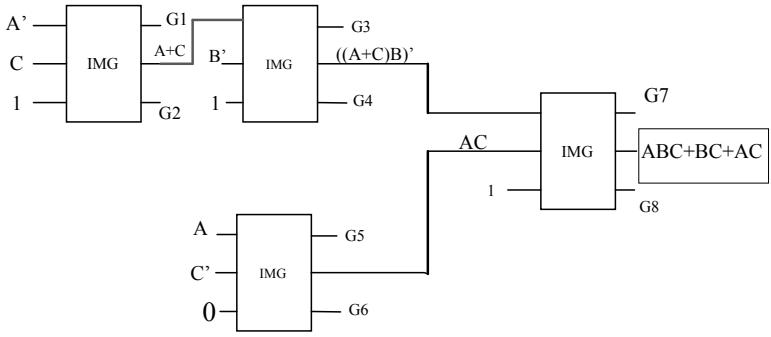
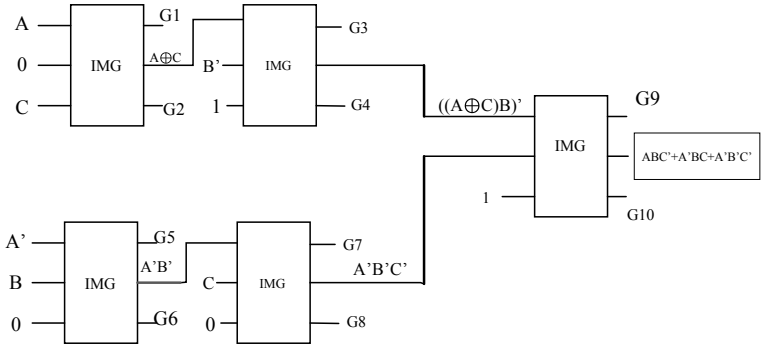


Fig. 2 (continued)

Function 11: AB + BC+AC



Function 12: $ABC' + A'B'C' + A'BC$



Function 13: $AB + A'B'C$

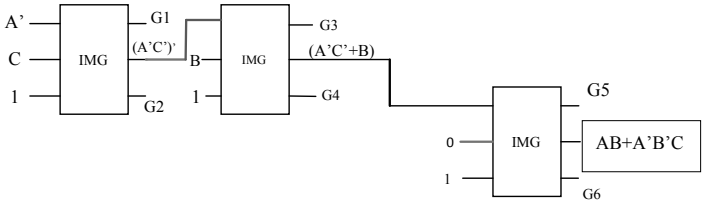


Fig. 2 (continued)

These dots are separated by tunnel junctions and it works on the concept of coulombic repulsion and quantum mechanical tunneling which leads to two polarization states, viz., binary "1" and binary "0" as shown in Fig. 4 [13, 14].

QCA circuits have clock signals consisting of four zones, viz., switch, hold, release, and relax [15, 16] which are required for its proper functioning and computing the delay in the circuit. In future, the proposed 3 × 3 NNG gate can be implemented in QCA along with the design of various combinational and sequential circuits using the NNG gate such as half adders, half subtractors, full adders, etc.

Table 4 Performance comparison of proposed gate

S. no.	CQCA [26]	Toffoli [19]	Fredkin [20]	URLG [27]	Proposed IMG gate
ABC	2	2	2	3	2
AB	1	1	1	1	1
AB + BC + CA	1	5	4	1	4
A	1	1	2	1	1
A′BC + A′B′C′	3	**3**	5	5	2
AB + A′B′	4	2	2	5	1
ABC′ + A′BC + A′B′C′	6	6	1	6	5
A′BC + AB′C + ABC′ + ABC	3	2	1	7	4
ABC + AB′C′	6	5	4	4	2
AB + B′C	3	3	3	4	3
AB + BC	2	2	2	3	2
AB + A′B′C	5	5	4	4	3
AB + BC + A′B′C′	6	5	5	5	5
Total gates used	43	42	41	49	35

Table 5 Percentage of improvement in proposed gate

Reversible gate	Gate count	Gate count in proposed IMG gate	% Age improvement in proposed gate
CQCA [26]	43	35	18.60
Toffoli [19]	42		16.66
Fredkin [20]	41		14.63
URLG [27]	49		28.57

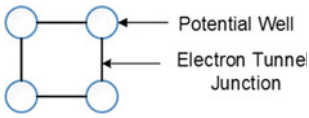


Fig. 3 Basic concept of a QCA cell

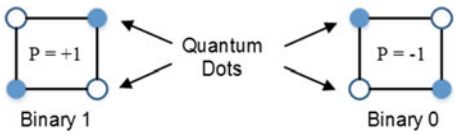


Fig. 4 Binary 1 and binary 0 representations of a QCA cell along with their polarizations

7 Conclusion

A new 3 × 3 IMG reversible gate was proposed, which was then verified to be a universal structure in nature. Further, the multifunctionality of the proposed gate was verified by implementing the 13 standard Boolean functions. Based on the performance comparison, it is observed that the proposed IMG gate is cost-efficient and achieves improvement up to 28.57% in terms of number of gates required to implement the 13 functions.

References

1. K. Das, D. De, Characterization, test and logic synthesis of novel conservative and reversible logic gates for QCA. Int. J. Nanosci. **9**, 201–214 (2010)
2. A.N. Bahar, S. Waheed, N. Hossain, A new approach of presenting reversible logic gates in nanoscale. SpringerPlus **4**(153), 1–7 (2015)
3. P. Biswas, N. Gupta, N. Patidar, Basic reversible logic gates and its QCA implementation. Int. J. Eng. Res. Appl. **4**, 12 (2014)
4. R. Landauer, Irreversibility and heat generation in the computing process. IBM J. Res. Dev. **5**, 183–191 (1961)
5. A new design and simulation of reversible gates in quantum-dot cellular automata technology (2019), https://arxiv.org/abs/1803.11017. Accessed 18 Feb 2019
6. C.H. Bennett, Logical reversibility of computation. IBM J. Res. Dev. **17**, 525–532 (1973)
7. H. Thapliyal, N. Ranganathan, Reversible logic-gates concurrently testable latches for molecular QCA. IEEE Trans. Nanotechnol. **9**, 62–69 (2010)
8. M.A. AlShafi, M.S. Islam, A.N. Bahar, A review on reversible logic gates and its QCA implementations. Int. J. Comput. Appl. **128**, 27–34 (2015)
9. P.R. Yelekar, S.S. Chiwande, Introduction to reversible logic gates and its applications. in *2nd International Conference on Information and Communication Technology* (2011), pp. 5–9
10. S.A. Hassan, P. Sharam, Low power quantum gates for the implementation of reversible memory elements using quantum dot cellular automata. in: *Proceedings of 3rd International Conference on Computing for Sustainable Global Development 2016* (IEEE 2016), pp. 2727–2732
11. B. Ahmed, S. Kakkar, Modular adder designs using optimal reversible and fault tolerant gates in field-coupled QCA nanocomputing. Int. J. Theor. Phys. **57**(5), 1356–1375 (2018)
12. F. Ahmad, S. Ahmed, V. Kakkar, G.M. Bhat, A.N. Bahar, S. Wani, Modular design of ultra-efficient reversible full adder-subtractor in QCA with power dissipation analysis. Int. J. Theor. Phys. **57**(9), 2863–2880 (2018)
13. B. Bilal, S. Ahmed, V. Kakkar, Quantum dot cellular automata: a new paradigm for digital design. Int. J. Nanoelectron. Mater. **11**(1), 87–98 (2018)
14. B. Bilal, S. Ahmed, V. Kakkar, An insight into beyond CMOS next generation computing using quantum-dot cellular automata nanotechnology. Int. J. Eng. Manuf. **8**(1), 25–37 (2018)
15. B. Bilal, S. Ahmed, V. Kakkar, QCA based efficient toffoli gate design and implementation for nanotechnology applications. Int. J. Eng. Technol. **9**(3s), 84–92 (2017)
16. B. Bilal, S. Ahmed, V. Kakkar, Optimal realization of universality of peres gate using explicit interaction of cells in quantum dot cellular automata nanotechnology. Int. J. Intell. Syst. Appl. **9**(6), 75–84 (2017)
17. R. Feyman, Quantum mechanical computers. Found. Phys. **16**(6), 507–531 (1986)
18. B. Parhami, Fault tolerant reversible circuits. in *Proceedings of 40th Asilomar Conference on Signals Systems and Computers* (2006), pp. 1–4

19. T. Toffoli, Reversible Computing. in *Proceedings of International Colloquium on Automata, Languages and Programming* (Springer, Berlin, Heidelberg 1980), pp. 632–644
20. E. Fredkin, T. Toffoli, Conservative logic. Int. J. Theor. Phys. **21**, 219–253 (1982)
21. A. Peres, Reversible logic and quantum computers. Phys. Rev. **A32**, 3266–3276 (1985)
22. A.K. Biswas, M.M. Hasan, A.R. Chowdhury, H.M.H. Babu, Efficient approaches for designing reversible binary coded decimal adders. Microelectron. J. **39**, 1693–1703 (2008)
23. M. Haghparast, K.A. Navi, Novel reversible BCD adder for nanotechnology based systems. Am. J. Appl. Sci. **5**, 282–288 (2008)
24. V. Kumar, D. Dhawan, Design of reversible adder subtractor using multifunction reversible logic gate (MRLG). Int. J. Adv. Comput. Electron. Eng. **2**, 5–11 (2016)
25. H. Rangaraju, A.B. Suresh, K. Muralidhara, Design and optimization of reversible multiplier circuit. Int. J. Comput. Appl. **52**(10), 44–50 (2012)
26. H. Thapliyal, N. Ranganathan, Conservative QCA gate (CQCA) for designing concurrently testable molecular QCA circuits. in *Proceedings of 22nd International Conference on VLSI Design* (2009), pp. 511–516
27. D. Kunal, D. Debashis, D. Mallika, Competent universal reversible logic gate design for quantum dot cellular automata. WSEAS Trans. Circ. Syst. **11**(12), 401–411 (2012)

A Novel Reversible DSG Gate and Its Quantum Implementation



Shaveta Thakral and Dipali Bansal

1 Introduction

Landauer in 1961 [1] pointed out that irreversible computing leads to energy loss. This energy loss is proportional to information (bit) loss and irreversible computation was declared dissipative. The mathematical relation between entropy and thermodynamics proved his point. Bennett in 1973 [2] provided solution for dissipative irreversible computing by replacing it with lossless reversible computing. He said that energy loss for each lost bit can be avoided by mapping output lines in the same way as that of input lines and making the system reversible. If the reversible system is allowed to run backward, then input must be obtained from the output and randomness is totally avoided to make system lossless where there is no information loss and no energy loss as well. Many reversible gates have been reported in the literature like Fredkin [3], Toffoli, FTRA, PPPG [4], HNG, F2PG, and WG which are multifunctional and are utilized to perform arithmetic and logical operations. Researchers have proposed various ALU designs [5–10] based on these gates. Some of the gates are investigated with their quantum equivalent and justified quantum cost whereas some are investigated without justifying their quantum cost. Investigating quantum cost needs a methodology to draw quantum implementation. There are two libraries, NCT and NCV, which are most widely used to draw quantum implementation of reversible logic based gates and their associated circuitry. The NCT library consists of N(NOT), C(Controlled NOT/Feynman), and T(Toffoli/Controlled Controlled NOT) gates, and the NCV library consists of N(NOT), C(Controlled NOT),V(Controlled V or Controlled V+) gates. The NOT gate is a 1*1 gate which is the same for conventional

S. Thakral (✉) · D. Bansal
Department of Electronics and Communication (FET),
Manav Rachna International Institute of Research and Studies, Faridabad, India
e-mail: shaveta.fet@mriu.edu.in

D. Bansal
e-mail: dipali.fet@mriu.edu.in

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_142

logic and reversible logic. It has quantum cost of unity. If input is A and output is P, then P will be the complement of A. CNOT gate is based on one control line and another target line. It can be considered as group of two input variables, A, B, and corresponding two output variables, P, Q. If A on control line is 0, then B on target line is passed as it is, otherwise signal B is inverted. There are two more controlled gates (two variable- based) which are popular in the literature as a part of the NCV library, i.e., controlled V and V+ gates. Controlled V and V+ gates along with controlled NOT gate make integrated qubit gates. A total eight types of integrated quantum bit gates can be constructed and their quantum cost is also considered unity. The different arrangements of integrated qubits in terms of their quantum implementation are shown in Fig. 1.

A brief insight into other popular gates which are most widely used to make any reversible logic based circuit is given in Table 1. Research on optimization metrics of reversible logic based circuit revolves around optimizing ancillary inputs

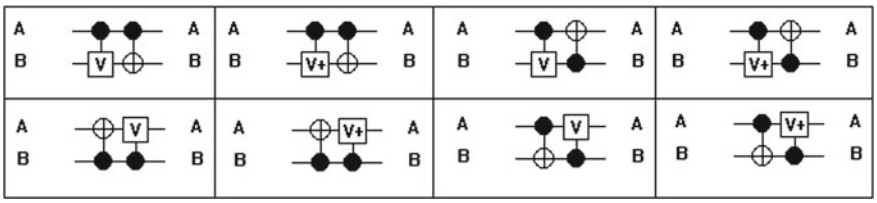
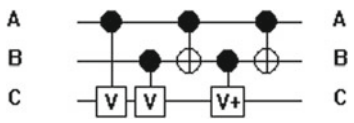
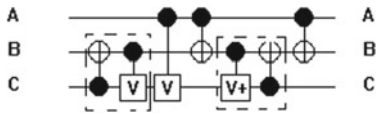
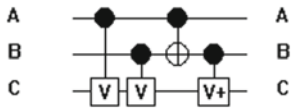
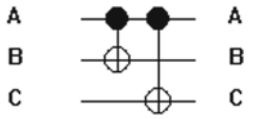


Fig. 1 Quantum implementation of eight integrated qubit gates

Table 1 Brief insight about popular reversible logic based gates

Gate	Logic	QC	Quantum diagram
Toffoli/CCNOT	P = A Q = B R = AB ⊕ C	5	
Fredkin/CSWAP	P = A $Q = \overline{A}B \oplus AC$ $R = \overline{A}C \oplus AB$	5	
Peres/NTG	P = A Q = A ⊕ B R = AB ⊕ C	4	
Double Feynman	P = A Q = A ⊕ B R = A ⊕ C	2	

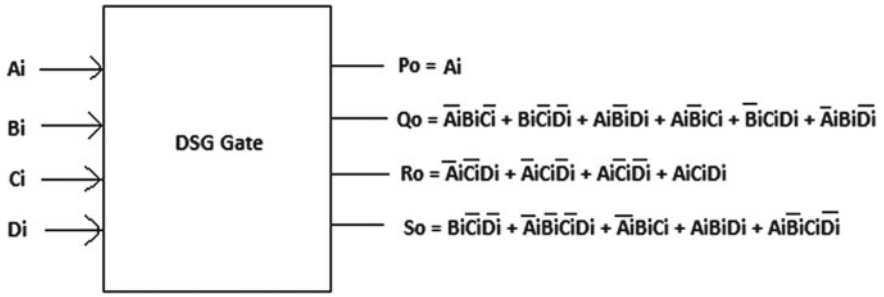


Fig. 2 Proposed novel reversible the DSG gate

(reducing them), garbage outputs (reducing them), and quantum cost. The reduction of quantum cost is the most challenging problem and needs thorough knowledge about calculation. Since every reversible gate can be synthesized in the form of a Toffoli network, and the Toffoli network can further be decomposed to primitive gates taken via the NCV gate library, the quantum cost of a reversible gate can be easily calculated by counting the total number of unity quantum cost based mentioned in the NCV library. The quantum cost for any reversible logic circuit should be as low as possible. The RCViewer+ tool provides conversion from the NCT circuit to the NCV circuits and also provides quantum cost calculator to justify the proposed quantum cost of the newly investigated gate. Section 2 briefs about the proposed novel reversible DSG gate. Section 3 provides its quantum implementation and quantum cost calculation. Section 4 provides the application of the proposed gate followed by the conclusion in the last section.

2 Proposed Novel Reversible DSG Gate

The proposed novel reversible DSG gate is a 4*4 gate that can function as a universal reversible logic based gate. The input vector of the DSG gate is Ai, Bi, Ci, and Di, and its output vector is Po, Qo, Ro, and So. The DSG gate under operation is shown in Fig. 2. And its truth table is shown in Table 2.

3 Quantum Implementation of DSG Gate

The quantum implementation of the DSG gate is done via a basic algorithm and converted to a Toffoli network based on the NCT library. The algorithm to build a Toffoli network is shown in Table 3 and the synthesized Toffoli network based on the NCT library is shown in Fig. 3.

Table 2 Truth table of the DSG gate

Ai	Bi	Ci	Di	Po	Qo	Ro	So
0	0	0	0	0	0	0	0
0	0	0	1	0	0	1	1
0	0	1	0	0	0	1	0
0	0	1	1	0	1	0	0
0	1	0	0	0	1	0	1
0	1	0	1	0	1	1	0
0	1	1	0	0	1	1	1
0	1	1	1	0	0	0	1
1	0	0	0	1	0	1	0
1	0	0	1	1	1	0	0
1	0	1	0	1	1	0	1
1	0	1	1	1	1	1	0
1	1	0	0	1	1	1	1
1	1	0	1	1	0	0	1
1	1	1	0	1	0	0	0
1	1	1	1	1	0	1	1

Then it is converted to the NCV library using the RCViewer+ tool as shown in Fig. 4. Thereafter, template matching, moving rule, and deletion rules are applied to optimize quantum implementation of the DSG gate. Then the quantum cost calculator of the RCViewer+ tool is used to calculate its quantum cost and it is obtained as 10. The quantum cost calculation of the DSG gate using the RCViewer+ tool is shown in Fig. 5. One integrated qubit is obtained in the optimized quantum implementation, and optimized cost of the DSG gate is obtained as 9. Quantum implementation with the highlighted integrated qubit is shown in Fig. 6.

4 Application of DSG Gate

The DSG gate is operated under different combinations of inputs and it is able to perform 12 logical operations like Xi, XiYi′, NOR, XNOR, OR, Xi′Yi, AND, XOR, NAND, Xi + Yi′, Xi′YiCin′ + XiYi′Cin, and Xi′Yi′Cin + XiYiCin′. It is also able to perform four arithmetic operations like Xi plus Yi plus Cin (full adder operation), Xi plus Yi (half-adder operation), Xi minus Yi minus Bin (full subtractor operation), and Xi minus Yi (half subtractor operation) where Xi and Yi are operands on which various arithmetic and logical operations can be performed, and Cin, Bin are initial carry and initial borrow, respectively. Cout and Bout are the output carry obtained after addition and the output borrow obtained after subtraction. The notation "+" is used for logical OR operation, "plus" is used for arithmetic addition, and "minus" is

Table 3 Algorithm to build Toffoli network for the DSG gate

AiBiCiDi	$Ai^0Bi^0Ci^0Di^0$	$Ai^1Bi^1Ci^1$ Di^1	$Ai^2Bi^2Ci^2$ Di^2	$Ai^3Bi^3Ci^3$ Di^3	$Ai^4Bi^4Ci^4$ Di^4	$Ai^5Bi^5Ci^5$ Di^5	$Ai^6Bi^6Ci^6$ Di^6
0000	0000	0000	0000	0000	0000	0000	0000
0001	0011	0001	0001	0001	0001	0001	0001
0010	0010	0010	0010	0010	0010	0010	0010
0011	0100	0100	0110	0111	0011	0011	0011
0100	0101	0111	0101	0100	0100	0100	0100
0101	0110	0110	0100	0101	0101	0101	0101
0110	0111	0101	0111	0110	0110	0110	0110
0111	0001	0011	0011	0011	0111	0111	0111
1000	1010	1010	1010	1010	1010	1000	1000
1001	1100	1100	1110	1111	1011	1001	1001
1010	1101	1111	1101	1100	1100	1110	1010
1011	1110	1110	1100	1101	1101	1111	1011
1100	1111	1101	1111	1110	1110	1100	1100
1101	1001	1011	1011	1011	1111	1101	1101
1110	1000	1000	1000	1000	1000	1010	1110
1111	1011	1001	1001	1001	1001	1011	1111
Toffoli gate		T2 (Di:Ci)	T2 (Bi: Ci)	T2 (Bi:Di)	T3 (Ci,Di:Bi)	T2 (Ai:Ci)	T3 (Ai,Ci:Bi)

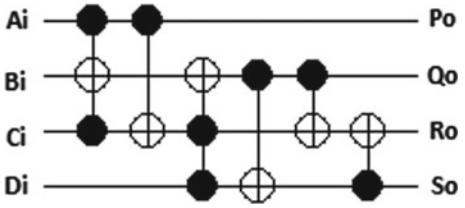


Fig. 3 Toffoli network of the DSG gate

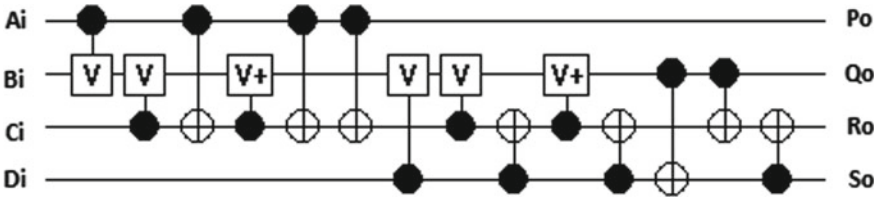


Fig. 4 Toffoli network conversion to NCV network

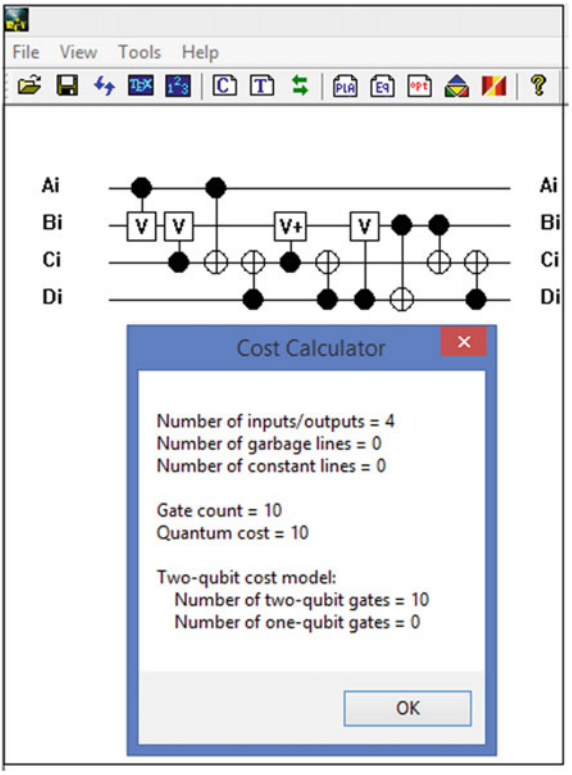


Fig. 5 Quantum cost calculation of the DSG gate using RCViewer+ tool

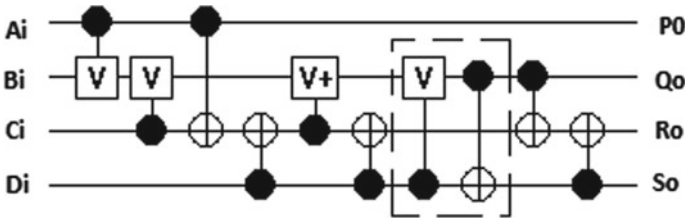


Fig. 6 Optimized quantum implementation of the DSG gate

Table 4 Operations performed by the DSG gate

Ai	Bi	Ci	Di	Po	Qo	Ro	So
Xi	1	Yi	1	Xi	Xi NOR Yi	Xi XNOR Yi	Xi OR Yi
Xi	Yi′	Yi′	1	Xi	Xi AND Yi	Xi XOR Yi	Xi NAND Yi
Xi	0	Yi	Cin	Xi	Cout	Xi plus Yi plus Cin	Xi′Yi′Cin + XiYiCin′
Xi	Yi	Yi	Bin	Xi	Xi′YiCin′ + Xi Yi′Cin	Xi minus Yi minus Bin	Bout
Xi	0	Yi	0	Xi	Cout	Xi plus Yi	Xi AND Yi
Xi	0	Yi′	0	Xi	XiYi′	Xi XNOR Yi	XiYi′
Xi	0	Yi′	1	Xi	Xi + Yi′	Xi minus Yi	Bout/Xi′Yi

used for arithmetic subtraction. The various operations performed by the DSG gate are shown in Table 4.

5 Conclusion

A novel reversible DSG gate has been investigated and its quantum implementation is done using the RCViewer+ tool to calculate its quantum cost. A single gate is able to perform 16 operations and is suitable as a universal reversible logic gate with a quantum cost of 9. There are abundant areas in which this research work on the DSG gate can be embedded to synthesize and optimize arithmetic and logical unit as well as other complex reversible logic based structures. The future scope of this research is to investigate low complexity and high functionality based ALU design using the proposed DSG gate.

References

1. R. Landauer, Irreversibility and heat generation in the computing process. IBM J. Res. Dev. **5**, 183–191 (1961)
2. C. Bennett, Logical reversibility of computation. IBM J. Res. Dev. **17**, 525–532 (1973)
3. E. Fredkin, T. Toffoli, Conservative logic. Int. J. Theor. Phys. **21**, 219–253 (1982)
4. Murthy M. Krishna, Design of efficient adder circuits using proposed parity preserving gate (PPPG). Int. J. VLSI Des. Commun. Syst. **3**, 83–939 (2012)
5. R. Saligram, S.S. Hegde, S.A. Kulkarni, H.R. Bhagyalakshmi, M.K. Venkatesha, Design of parity preserving logic based fault tolerant reversible arithmetic logic unit. Int. J. VLSI Des. Commun.Syst. **4**, 53–68 (2013)
6. Z. Guan, W. Li, W. Ding, Y. Hang, L. Ni, An arithmetic logic unit design based on reversible logic gates. in *IEEE Pacific Rim Conference on Communications, Computers and Signal Processing* (IEEE 2011), pp. 925–931
7. P. Moallem, M. Ehsanpour, A. Bolhasani, M. Montazeri, Optimized reversible arithmetic logic units. J. Electron. **31**, 394–405 (2014)
8. S. Thakral, D. Bansal, Novel reversible DS gate for reversible logic synthesis. Int. J. Modern Educ. Comput. Sci. **8**, 20–26 (2016)
9. S. Thakral, D. Bansal, Fault tolerant ALU using parity preserving reversible logic gates. Int. J. Modern Educ. Comput. Sci. **8**, 51–58 (2016)
10. M. Thomsen, R. Glück, H. Axelsen, Reversible arithmetic logic unit for quantum arithmetic. J. Phys. A: Math. Theor. **43**, 382002 (2010)

Performance Analysis of Various Embedded Linux Firmwares for ARM Architecture Based IoT Devices



Mahendra Swain, Rajesh Singh, Md. Farukh Hashmi and Anita Gehlot

1 Introduction

Linux has been known for its advantages like open source, platform independent, much secure environment, and development oriented. The Raspberry Pi board used in this prototype is significantly cost effective in comparison to other available Linux-operated boards with optimum specifications and have better community support. Popularity of Raspberry Pi estimates that to date, 5 million Raspberry Pi boards have been sold as per the survey was done in January 2015. The software platform used to design this prototype is Ubuntu 14.04 (host) and programming language used is Python 2.7. The use of Python programming language makes it more user-friendly for voice communication, GUI development, and ease to understand.

The Yocto project has been used to provide an Embedded LINUX platform to the prototype such that the configurability at both the kernel level and package level could be maintained without compromising the necessary performance. Operating system (OS) footprint and execution time for a "sample program" have been tested across available Embedded Operating systems such as Raspbian OS, PiLFS, and

M. Swain (✉) · R. Singh · A. Gehlot
Lovely Professional University, Punjab, India
e-mail: er.mahendraswain@gmail.com

R. Singh
e-mail: srajssssece@gmail.com

A. Gehlot
e-mail: eranita5@gmail.com

M. Swain
Model Institute of Engineering and Technology, Kot Bhalwal, Jammu, India

Md. F. Hashmi
National Institute of Technology, Warangal, India
e-mail: mdfarukh@nitw.ac.in

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_143

Yocto-based Embedded Linux and results produced are promising for the Yocto-based Embedded Linux Operating Systems. This Embedded Linux based device gives the prototype to attain flexibility of having minimum memory footprint and optimum speed and multitasking behavior on the same platform which makes voice calling possible along with possible inclusions of games or music or web browsing features, etc. [1, 2].

2 Various Embedded Linux Os for Raspberry Pi

2.1 Raspbian OS

Raspbian OS is a Debian-based OS, developed for Raspberry Pi. It contains all the packages required for the Raspberry Pi basic operations. Raspbian is basically designed to provide the desktop environment to Raspberry Pi. The initial build of Raspbian contains over 35,000 packages, which are optimized for the best performance on the Raspberry Pi board, which was completed in June 2012. Packages can be customized in Raspbian OS as per our requirement by removing unwanted packages from the build image, which is further explained under a methodology [3].

Debian supports two ARM architectures. The armhf (Acorn RISC Machine Hardware Float) and armel (Acorn RISC Machine Emulation Library) are two different architecture supported by the ARM family. The armel architecture supports the popular ARMv4 instruction set and it handles floating-point computation in the compatibility mode with processors without floating-point units which make it a slow performer in comparison to armhf architecture. In contrast, armhf architecture supports modern ARMv7 instruction set and provides direct hardware floating-point support but lacks the compatibility with older architectures [3, 4].

2.2 Yocto Project

The Yocto project has been used in this paper for building custom Embedded Linux OS for Raspberry Pi. To design customized embedded firmware, Yocto is an open-source platform where we can design embedded Linux firmwares for all hardware architectures [5] (Fig. 1).

2.3 PiLFS

Raspberry Pi can be operated on a customizable Linux kernel. Linux From Scratch on the Raspberry Pi is one of those which provides benefits in terms of customizable

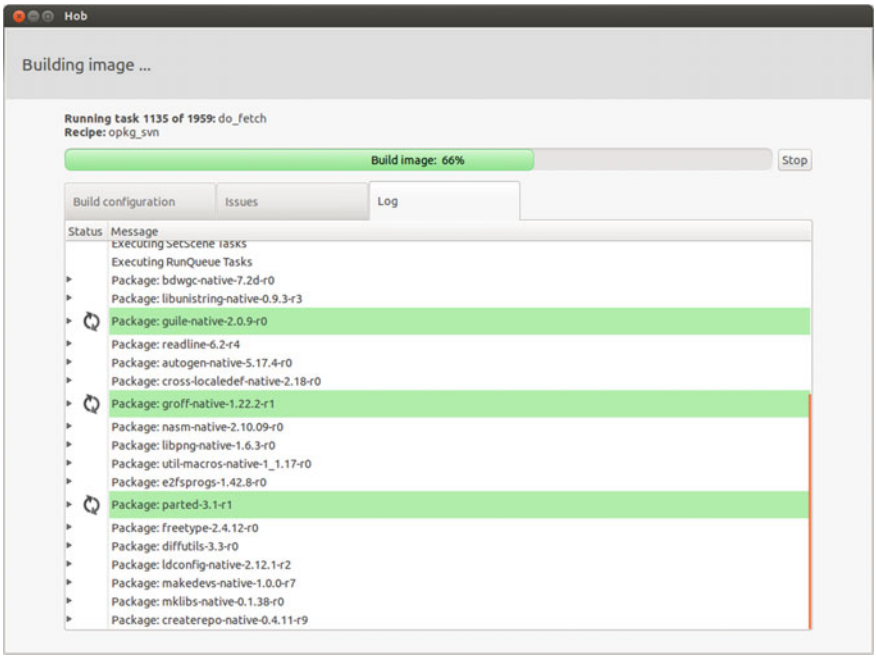


Fig. 1 Image 1: Customized embedded firmware building on a host computer for Raspberry Pi

desktop image which is used in desktop computers. To build bare essential systems where we can get complete control over software. It is free distribution-specific patches and quirks. It is curious who would like to learn working of GNU/Linux distribution [3, 4, 6].

Dissimilar Group of Data

An algorithm can be defined as a procedure to generate an output for a given set of inputs using loops and decisions as needed. Algorithm plays two vital roles in the data processing endeavor. First, they are able to determine the phenomenon of profiling which is conducted by controlling the profiling process in the form of a procedure. Algorithms may be designed to assist in capture, validation, and verification of data. Second, the algorithms can help to determine the internal relationship, trendsm and invisible patters between dissimilar groups of data. We show how algorithms find a natural home at the very heart of the profiling process and how such machine learning can actually be used to execute the task [1]. We show how algorithms find a natural home at the very heart of the profiling process and how such machine learning can actually be used to execute the task [1] (Fig. 2).

In this paper, we have considered three popular algorithms, i.e., Bubble sort, Binary search, and Merge sort. Bubble sort is a simple sorting algorithm that repeatedly passes through the sorting list in a step-by-step manner, compares each pair of adjacent items, and does the swapping operation if they are not placed in the correct

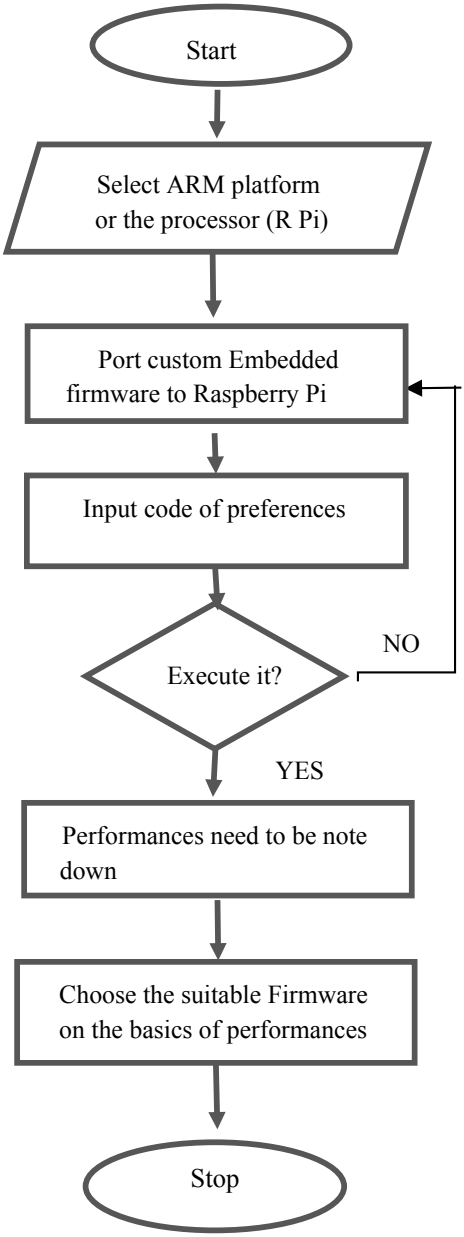


Fig. 2 Flowchart 1: Methodology of evaluation of an embedded firmware for suitable ARM architecture

order. The pass through the list is repeated until no swaps are needed, which indicates that the list is sorted. Its worst-case complexity is O($n2$), where n is the no. of items. Binary search algorithm identifies the location of the target value within the sorted array. This algorithm compares the target value to the center element of the array; if they are unequal, the half in which the target element could not be kept is eliminated and the search continues on the remaining half until it is successful. Its worst-case complexity is O (log n). Merge sort is based on the divide and conquer algorithm where it divides the input array into two halves, calls itself for the two halves, and then merges the two sorted halves. Its worst-case complexity is O (n log n).

In this paper, each of the algorithms is written in C language. During the process of writing these algorithms, we have not considered the user to provide input and output such that any unnecessary time consumption due to human intervention could be avoided. Instead, we have a predefined data set in the programs and the result is stored back in some defined array.

Each of these algorithms consists of numbers of executable sub-functions which are used to determine the functional parameters those which are responsible for utilization of CPU cycles during execution of the task. The parameters which have been taken to compare the performance of an OS are system time, real time, processing time, etc.

Performance Table

(a) Program size: 1 K

The performance table shows the user time and system time on three different languages with respect to three different embedded firmwares on Raspberry Pi. Three basics algorithms are considered for this performance analysis Binary search, Bubble sort, and Merge sort. The comparison shows that as far as the Python language is considered, the user time and system time performance are better in the customized one, i.e., Yocto. So it is preferable over PiLFS and Raspbian (Figs. 3, 4, 5, and 6, and Tables 1, 2, 3, 4, 5, 6, and 7).

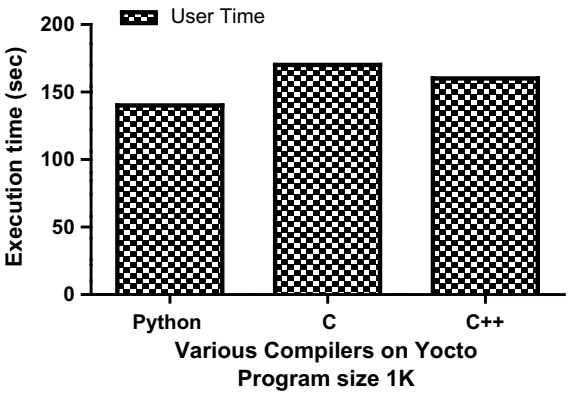


Fig. 3 Graph 1: The user time performance on various compilers for 1 K program code size

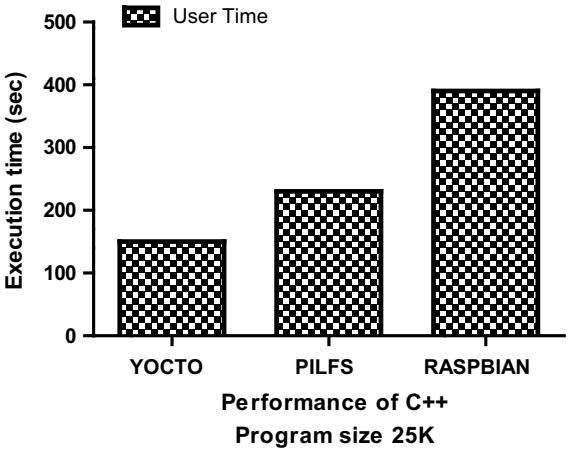


Fig. 4 Graph 2: Performance in 25 KB size of code

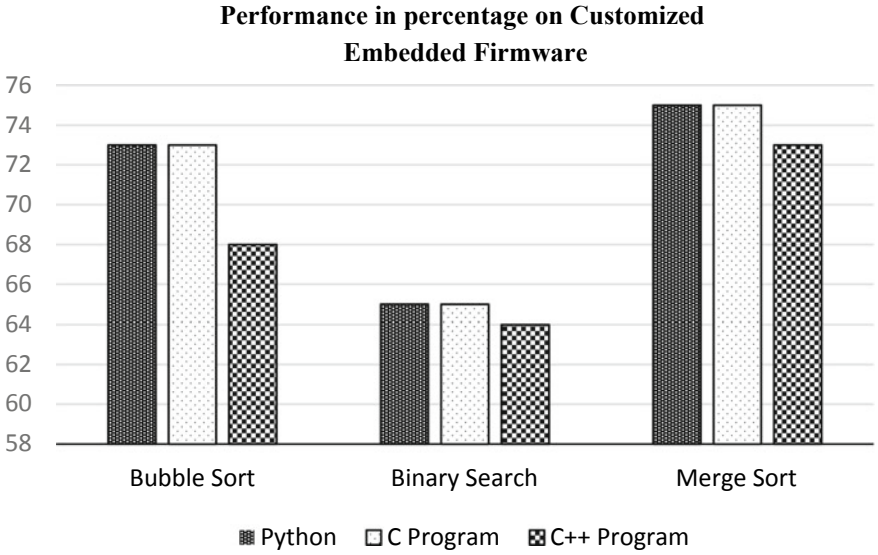


Fig. 5 Graph 3: The overall performance

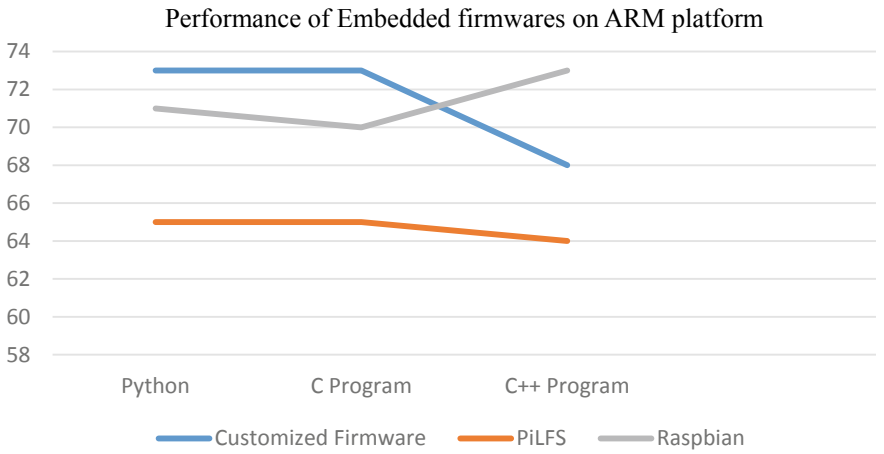


Fig. 6 Comparison of programming languages with respect to their performance on hardware platform

Table 1 Performance of 1 K size program on various compilers with respect to Embedded Linux OS

	User time in ms			System time in ms		
	Yocto	PiLFS	Raspbian	Yocto	PiLFS	Raspbian
Python	0.140	0.160	0.190	0.030	0.040	0.050
C	0.170	0.160	0.180	0.040	0.030	0.045
C++	0.160	0.100	0.150	0.025	0.025	0.035

Table 2 Performance of 10 K size program on various compilers with respect to Embedded Linux OS

	User time in ms			System time in ms		
	Yocto	PiLFS	Raspbian	Yocto	PiLFS	Raspbian
Python	0.180	0.190	0.220	0.080	0.080	0.090
C	0.190	0.180	0.190	0.070	0.070	0.065
C++	0.160	0.170	0.180	0.065	0.055	0.045

Table 3 The performance analysis over program size of 25 K. Here, for user time, Yocto performance is better whereas, for system time, Raspbian performs better over PiLFS

	User time in ms			System time in ms		
	Yocto	PiLFS	Raspbian	Yocto	PiLFS	Raspbian
Python	0.310	0.360	0.410	0.110	0.120	0.170
C	0.290	0.310	0.420	0.140	0.130	0.145
C++	0.150	0.230	0.390	0.120	0.115	0.130

Table 4 Performance of 100 K size program on various compilers with respect to Embedded Linux OS

	User time in ms			System time in ms		
	Yocto	PiLFS	Raspbian	Yocto	PiLFS	Raspbian
Python	0.910	0.960	0.910	0.410	0.420	0.970
C	0.890	0.810	0.820	0.540	0.630	0.745
C++	0.650	0.730	0.790	0.380	0.515	0.530

Table 5 Using Python on Raspbian

Evaluation parameters	Bubble sort	Binary search	Merge sort
CPU cycles used	2.2×10^{18}	3.2×10^{18}	2.5×10^{18}
Context switch time in ms	1245	1324	1367
Task clock cycle	3456	3589	3678
Cache hit time in ms	976	976	945
Overall performance in percentage	75	69	79

Table 6 Using Python on PiLFS

Evaluation parameters	Bubble sort	Binary search	Merge sort
CPU cycles used	2.1×10^{18}	2.9×10^{18}	2.3×10^{18}
Context switch time in ms	1189	1201	1235
Task clock cycle	3879	3987	3794
Cache hit time in ms	1125	1192	1232
Overall performance in percentage	84	73	65

Table 7 Using Python (Yocto)

Evaluation parameters	Bubble sort	Binary search	Merge sort
CPU cycles used	3.2×10^{18}	4.5×10^{18}	2.7×10^{18}
Context switch time in ms	1239	1287	1342
Task clock cycle	4232	3954	3875
Cache hit time in ms	1345	1356	1189
Overall performance in percentage	73	65	75

3 Conclusion

Customized embedded firmware facilitates execution speed on ARM architecture as compared to other embedded firmwares. From the result and discussion, it has been concluded that although there are various embedded firmwares that are present for Raspberry Pi, Yocto is preferable in terms of execution speed, compatibility, reliability, and small memory footprint. On the other hand, it enhances overall performance facilitates to select the suitable code needed to be embedded in IoT devices. From the above results, it has been concluded that the customized embedded firmwares are more preferable. The other parameters can be optimized in future applications. Customized embedded firmware has been developed and ported into ARM architecture in order to evaluate the performance. IoT devices embedded with customized firmware are preferable over the existing firmware.

References

1. N.N. Deniz, J.O. Chelotti, J.R. Galli, A.M. Planisich, M.J. Larripa, H.L. Rufiner, L.L. Giovanini, Embedded system for real-time monitoring of foraging behavior of grazing cattle using acoustic signals. Comput. Electron. Agric. **138**, 167–174 (2017)
2. M. Swain, Design of embedded linux based voice calling device. Int. J. Appl. Eng. Res. **10**(15), 35095–35102 (2015). ISSN 0973-4562
3. R.I. Pereira, I.M. Dupont, P.C. Carvalho, S.C. Jucá, IoT embedded linux system based on Raspberry Pi applied to real-time cloud monitoring of a decentralized photovoltaic plant. Measurement **114**, 286–297 (2018)
4. K.J. Singh, D.S. Kapoor, Create your own internet of things: a survey of IoT platforms. IEEE Consum. Electron. Mag. **6**(2), 57–68 (2017)
5. N. Asokan, T. Nyman, N. Rattanavipanon, A.R. Sadeghi, G. Tsudik, ASSURED: architecture for secure software update of realistic embedded devices. IEEE Trans. Comput. Aided Des. Integr. Circ. Syst. **37**(11), 2290–2300 (2018)
6. M. Shakeri, M. Shayestegan, H. Abunima, S.S. Reza, M. Akhtaruzzaman, A.R.M. Alamoud, N. Amin, An intelligent system architecture in home energy management systems (HEMS) for efficient demand response in smart grid. Energy Build. **138**, 154–164 (2017)
7. D. Harrison, U.S. Patent No. 9,686,596. (U.S. Patent and Trademark Office. Yocto kernel development manual, Washington, DC, 2017)
8. N. Asokan, T. Nyman, N. Rattanavipanon, A.R. Sadeghi, G. Tsudik, ASSURED: architecture for secure software update of realistic embedded devices. IEEE Trans. Comput. Aided Des. Integr. Circuits Syst. **37**(11), 2290–2300 (2018)
9. L.M. Gallegos-Canales, A. Favela-Contreras, A. Ávila, S.O. Martínez-Chapa, Embedded software implementation of the SISO adaptive predictive control algorithm. IEEE Trans. Industr. Electron. **64**(9), 7229–7238 (2017)
10. R. Zhu, B. Zhang, J. Mao, Q. Zhang, Y.A. Tan, A methodology for determining the image base of arm-based industrial control system firmware. Int. J. Crit. Infrastruct. Prot. **16**, 26–35 (2017)
11. M. Schulz, D. Wegemer, M. Hollick, The Nexmon firmware analysis and modification framework: empowering researchers to enhance Wi-Fi devices. Comput. Commun. (2018)
12. J.B. Hou, T. Li, C. Chang, Research for vulnerability detection of embedded system firmware. Procedia Comput. Sci. **107**, 814–818 (2017)

13. G. Oyarzun, R. Borrell, A. Gorobets, F. Mantovani, A. Oliva, Efficient CFD code implementation for the ARM-based Mont-Blanc architecture. Future Gener. Comput. Syst. **79**, 786–796 (2018)
14. A. Beghi, F. Marcuzzi, P. Martin, F. Tinazzi, M. Zigliotto, Virtual prototyping of embedded control software in mechatronic systems: a case study. Mechatronics **43**, 99–111 (2017)

Landslides Detection in Prone Hilly Areas Using Raspberry Pi



Prabin Kumar Das, Rajesh Singh, Anita Gehlot, Km. Vaishnavi Gupta and Arun Singh

1 Review of Literature

This system can be defined as Telemetric Sensing Node Network System (TSNNS). This system further consists of three more systems Local Sensing Node Network System (LSN NS), Host System (HS), and Cloud System (CS). The host system collects data from LSNNS for the detection of the landslide in remote areas. On the other side, HS also manages TSNNS to get the information about the prone area. By 3-D acceleration sensor, landslides can be determined in the wide area [1]. In the hilly Himalayan regions, with the use of the remote sensing–GIS techniques, landslide distribution, landslide hazard zonation, land use/land cover, drainage order can be determined. The integrated layer formed by this technique can be useful in determining the cost of route development and its maintenance. Dijkstra's algorithm, a different technique has been applied for the determination of the least cost computing routes for maintenance destroyed due to landslides [2]. This system is able to recognize the landslides occurring in the areas through the sensors used in this device. This system is used to determine the required information regarding landslide-prone areas. This data collected is basically about the occurring events before landslides. This device has been made for providing a real-time map for the landslide-prone areas. This device has been in such a manner so that it can cover the whole desired area more prone to landslide. Further, this area has been divided into local zones each consisting of smart sensors which are able to determine the sliding and spatial movement of soil particles. With the help of the algorithm one can decide the no. and size of local zones in an area. All the data has been collected to the web server using HSDPA modem for further processing which will further able to inform and give a check for civil protection [3]. This device proposed a system called LITMUS which collects data in both physical and social form and then combines the data by

P. K. Das (✉) · R. Singh · A. Gehlot · Km. V. Gupta · A. Singh
Lovely Professional University, Jalandhar, Punjab, India
e-mail: Prabindas786@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_144

filtering the information from those which are based on spatiotemporal features. It also demonstrates the live information about the landslide detection. This system was also able to detect those prone areas which were not recorded by the USGS. However, it is a prototype tool for research methods [4]. By seeing the main causes of landslide which were high altitude, deep rock backfilling, and earthquake in the Panzhihua city which is located in the mountainous regions in the province of china, the system has proposed displacement sensing technology and Global System for Mobile Communications (GSM) technology for the achievement of real-time information of unattended cracks in the ground and other causes. This system consists of five extensometers with measuring ranges of 200, 450 and 700 mm which was proved to reliable and efficient after experiments [5]. This system is based on an alerting system alerting the disaster management as well as the general population, which was implemented with use of an accelerometer which was able to detect any kind of change in the movement of the ground whether it is horizontal or angular movement. It consists of small nodes which consist of sensors that collect the data of any kind of ground movement and further connected to one hub known as Renesas GR-Kaede processor. This data is further sent to the server through ZigBee, which has been designed through XAMPP and PHP which is given as an alert to the android based application installed in mobile phones, thus helps the user to sends the data to the disaster management in case of any query or any kind of help [6]. This system has a network of wireless sensors able to detect landslide mitigation. This system consists of 50 geological sensors and 20 wireless sensor nodes which has the capability of collecting data values based on rain, moisture, pressure, and movement, along with other geological, hydrological and soil properties. This system is a three-step landslide warning system: early, intermediate, and imminent which is able to give novel solutions against difficult challenges [7]. The project named SLEWS project describes how to first collect all the data and then how to process it and further end up with its analysis. The system collects the data through wireless sensors. The required data obtained as a result of this wireless network of sensors on the mobile phones of the user which seems to be much cheaper and easy to deploy. Further sensor fusions and identification of malfunction becomes easier by using this method. The data, thus, obtained after processing by the network can be easily accessible by the user accordingly [8]. In this study, a very prone area for landslides, i.e., Myanmar which is very likely to experience landslides frequently due to its very unstable geographical conditions, slopes, and heavy monsoon rains, has been studies very deeply and geographical information tools (GIS) have been used to study the chances of landslides at various new prone locations and evolve some effective management systems. It has been aimed that proper management programs and accurate technical controls can decrease the number and effect of landslides to a great extent in upcoming future [9]. This paper describes a severe problem of Rio de Janeiro State mountain ranges which is due to being in a prone area for various natural disasters like landslides and flooding which causes a lot of damage to health as well as wealth of that area. From studies, it was found that although these disasters are cursed by nature, landscape quality degradation adds significantly for increasing risk of disasters. It is expected that these extreme events that caused disasters have

a high chance of increment in days to come. Various approaches like Ecosystem-Based Adaption (EBA) and Ecosystem-based Disaster Risk Reduction (Eco-DDR) are implemented for monitoring and restoring ecosystem for improving ability to cope up with natural disasters and reduce their adverse effect on human society [10]. For air quality monitoring the Air Quality Index (AQI) is important. It is an index prepared by government institutions to categories risk of air pollution. Lower AQI value represents pleasant air quality. Below is the AQI index value adopted by India. This proposal provides the AQI index value adopted by India along with the various algorithm and hardware descriptions to read the various sensors for control and monitor applications [11–13]. The paper discussed the use of WSN and cloud server to monitor the physical parameters. The XBee and ESP 8266 devices are the platforms to achieve the desired output [14–19].

2 Introduction

Landslide has become one of the major problems amongst natural disasters in most of the countries around the world. These disaster's elimination is not in our hands, i.e., it is not possible anyway to have control over the landslides, however its adverse effect can be reduced, if the concerned authorities will always be having proper information about environmental conditions like the current state of the earth's vibration level. This disaster alert system has been specially introduced to notify people in early stage about the possible threat. So that safety precautions can be taken to avoid any mishappening or reduce the effect on human well-being. In this system, Raspberry Pi has been used for monitoring landslide. The sensor used here to monitor earth vibration is MEMS accelerometer and for live human count PIR sensor or a camera module has been used. The outputs of sensors are given to Raspberry Pi and a buzzer is being used as the warning element when level of earth vibrations crosses a set max threshold and live human detected during that moment. The pressure sensor will be used to measure pressure during landslides. During any event of landslide, the video of whole incident will be captured by the camera module and will be uploaded to the cloud. All the real-time sensed values will be uploaded to thing speak for real-time monitoring and analysis.

In this proposal, we have basically proposed a system which is able to sense some particular elements which are the main cause of landslides in hilly areas. All these sensors are installed into nodes and these nodes are deployed at the areas prone to landslides for real-time monitoring process. These sensed data are logged on to a dedicated server wirelessly with the help of Wi-Fi Enable Module Raspberry Pi Zero W for further analysis and monitoring of environmental. Due to all these advantages, this proposed idea proves very effective in saving a lot of human health and wealth very reliably.

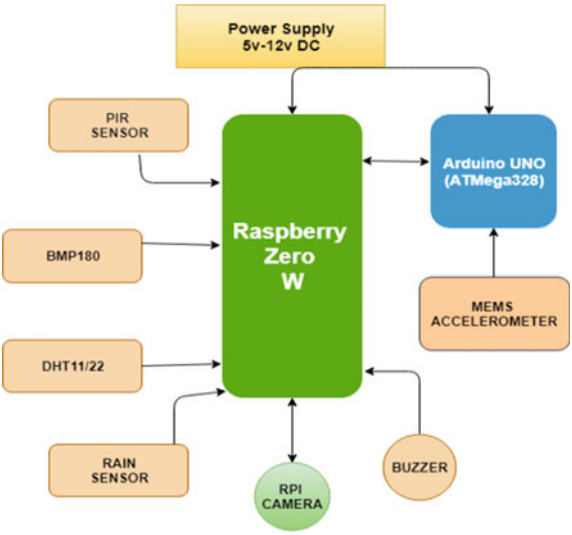


Fig. 1 Block diagram

3 Block Diagram

3.1 Connections

Figure 1 shows the block diagram of the system, which is having sensors like BMp180, MEMS accelerometer, DHT11, etc., and controlling platform like Arduino and Raspberry Pi.

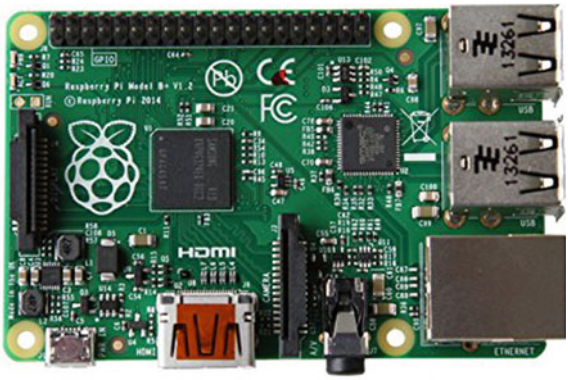
3.2 Raspberry Pi

Raspberry Pi is a tiny and affordable computer that can be used to learn programming. It was originally intended to help spur interest in computing among small school-aged children. This runs totally on open-source software and is very much compatible to be interfaced with any kind of device (Fig. 2).

3.3 Raspberry Pi Camera

Pi Camera is basically a custom add-on onto Raspberry Pi. It can be attached to the Raspberry Pi board through any of the two sockets present on the upper surface of the board. It uses the dedicated CSI interface, which was specially made for cameras interfacing with Pi board. Its bus has the ability to attain extremely high data rates

Fig. 2 Raspberry Pi



and also it is a high-quality 8 megapixel Sony IMX219 featuring a fixed focus lens (Figs. 3 and 4).

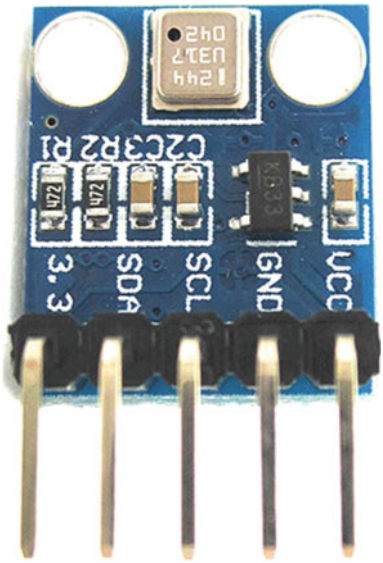
Fig. 3 Pi Camera



Fig. 4 PIR sensor



Fig. 5 BMP180



3.4 PIR Sensor

It is an electronic sensor which works by sensing the IR radiations from the objects in its line of vision. Mostly used for motion detector applications. It can sense the movement of any human or animal bodies at a particular area, thus very effectively used for anti-burglary applications.

3.5 BMP180

BMP180 is one of the sensors of BMP XXX series shows in Fig. 5. They are all designed to measure barometric pressure or atmospheric pressure. BMP180 is a high precision sensor designed for consumer applications. Barometric pressure is nothing but weight of air applied on everything. For temperature compensated pressure reading it has a temperature sensor also.

3.6 MEMS Accelerometer

It is an automated device used for the measurement of the forces of acceleration. Further, it can be described as a measurer of the acceleration of the reference body in reference to another body in a fixed reference frame and shows in Fig. 6.

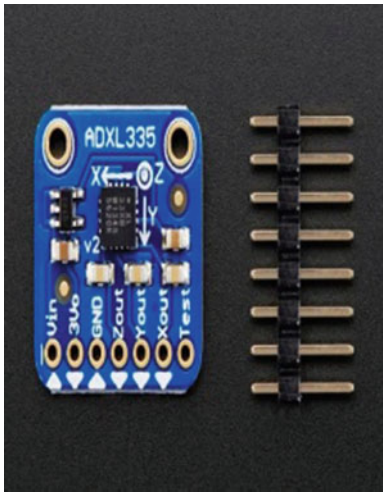


Fig. 6 MEMS accelerometer

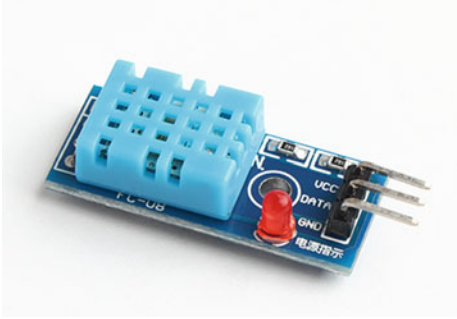


Fig. 7 DHT11

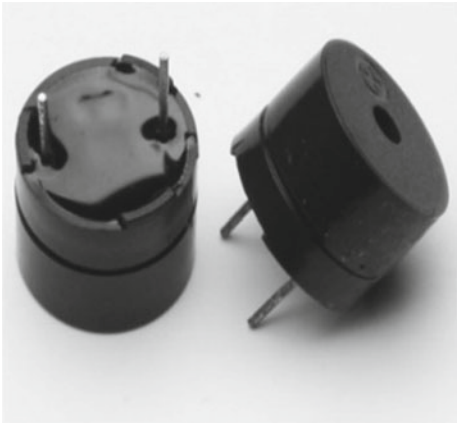
3.7 DHT11/22

It is a sensing device which measures the temperature and humidity of the respective area. It senses the parameters and gives the output in the form of analog data. It can be effectively used for getting information about temperature and humidity of any place with some precision as shown in Fig. 7.

3.8 Buzzer

A buzzer or beeper is a device basically used for alert or alarming applications. These are of different types according to their field of application, i.e., mechanical, electromechanical, or piezoelectric types as shown in Fig. 8.

Fig. 8 Buzzer



4 Circuit Diagram

In Fig. 9 of the proposed device, all the sensors are connected with Raspberry Pi Zero W. The microcontroller collects the data sensed by the sensors and log those into a dedicated server. In our project, there are some analog sensors like MEMS accelerometer and rain sensor so we have used Arduino because there is no facility of analog pins on the raspberry. Here the Arduino plays the role of a gate node for the sensor and the Raspberry. The camera is used when the parameter of the sensors detects an abrupt rise in values like pressure, temperature, vibrations, and rain for capturing the data in the form of images and video so that we can analyse

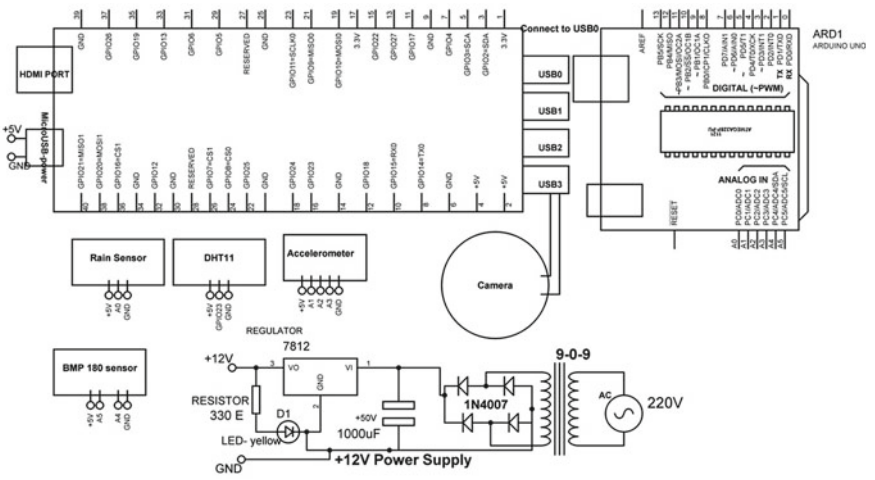


Fig. 9 Circuit diagram

the condition during a disaster at any time. These all collectively allow us to monitor the real-time data and also capture the visual values from different areas of disaster.

5 Software Development

Once the system is set to ON state, all the deployed sensor begins sensing the environmental parameters in a real-time manner and sending them on to the microcontroller. The microcontroller analyzes the values and starts logging the real-time values into a dedicated server.

This device along with its deployed sensors measure or records the real-time values of the elevation profiles, slopes, and different weather parameters of the targeted area to analyze and make a prediction of landslides occurrence by combining all information of the GIS (Geographic Information System) over its defined algorithm. Whenever the values received by the microcontroller shows an abrupt rise in specific parameter, it takes that event as a proximity condition and instantly triggers an alert information onto the server through as well as to the concerned authorities without any delay. In the case of any disaster prediction, it activates its camera module for capturing real-time images and videos of that place. The process is depicted in Fig. 10.

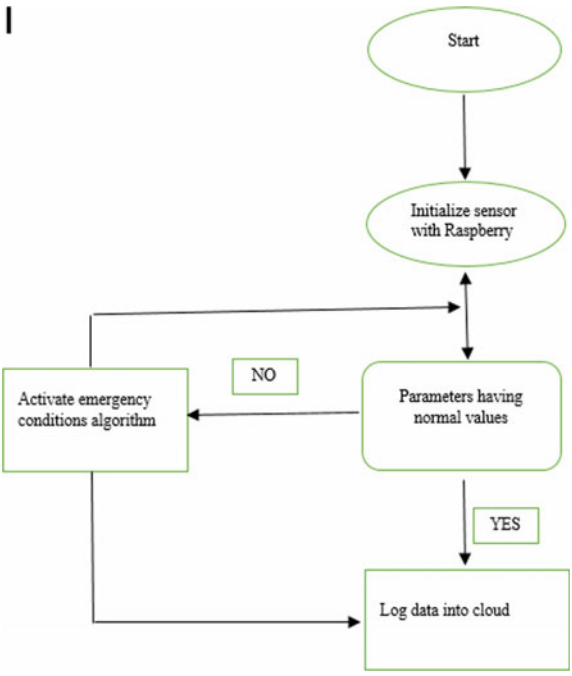


Fig. 10 Software development/algorithm

6 Results and Discussion

To get a proper information about its reliability and effectiveness in a practical situation, a prototype of this device has been designed by using sensors like DHT11, PIR, MEMS accelerometer, BMP180 pressure sensor, RPI camera, and a microcontroller. Very positive results were obtained from the working of that prototype in every tested situation.

Firstly, all these sensors are deployed at a place resembling a proximity condition of a landslide-prone area. These sensors are then connected to a microcontroller. A Raspberry Pi Zero W has been used as a microcontroller for taking all the sensed values of environmental parameters from these sensors, analyze it, and then log those values into a dedicated server. Based upon the working algorithm whenever the microcontroller analyzes the values received from the sensor and observes an abrupt or abnormal rise in any environmental parameters more than a set max threshold, it takes this event as a proximity condition for landslide and instantly activates its alarm features through the buzzer and also sends an alert warning onto the server instantly as well as to the concerned authorities. It also activates its RPI camera for real-time visual capturing of that place. A picture of that prototype has been depicted below (Figs. 11, 12, 13, 14, and 15).

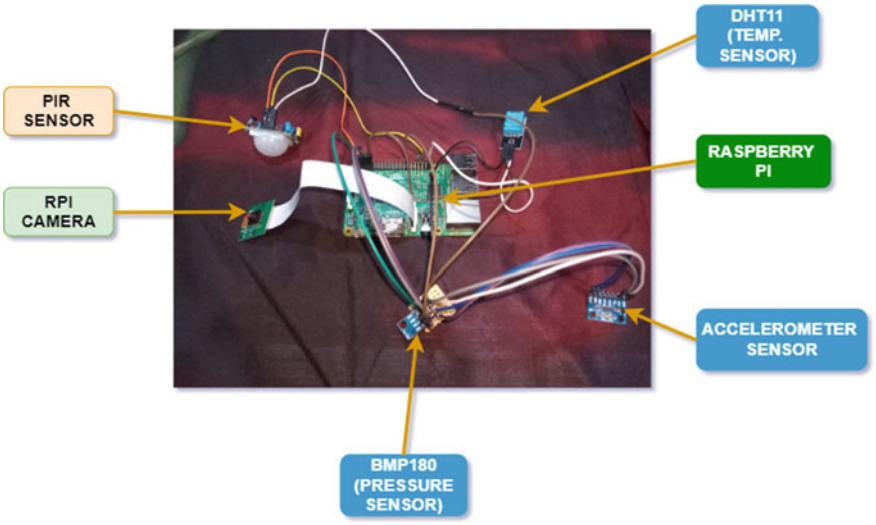


Fig. 11 All the necessary sensors are connected to Pi and used as a prototype to test the various parameters of effectiveness of this system

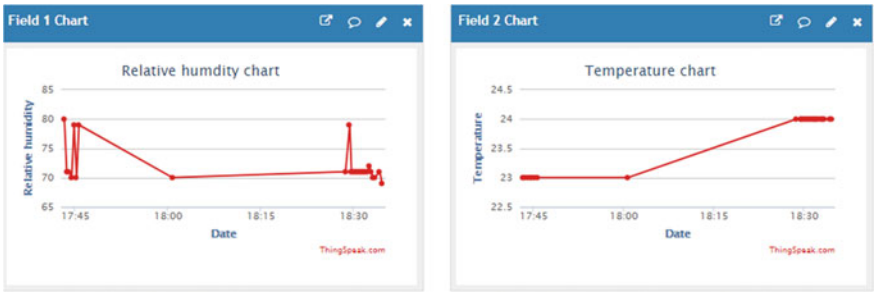


Fig. 12 The values of temperature and humidity that are logged into server during prototype testing of DHT11 in Thingspeak cloud server

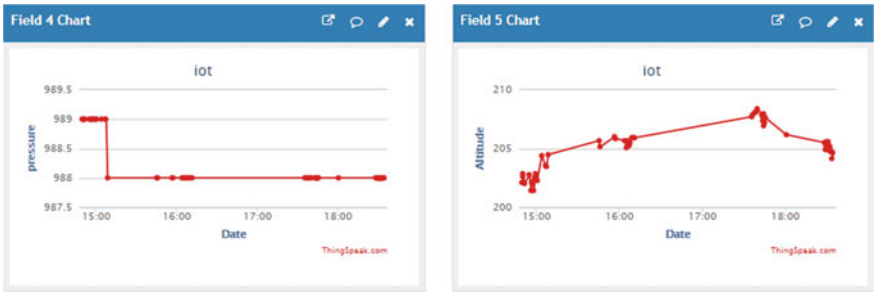


Fig. 13 The values of pressure and altitude that are logged into server during prototype testing of BMP180 in Thingspeak cloud server

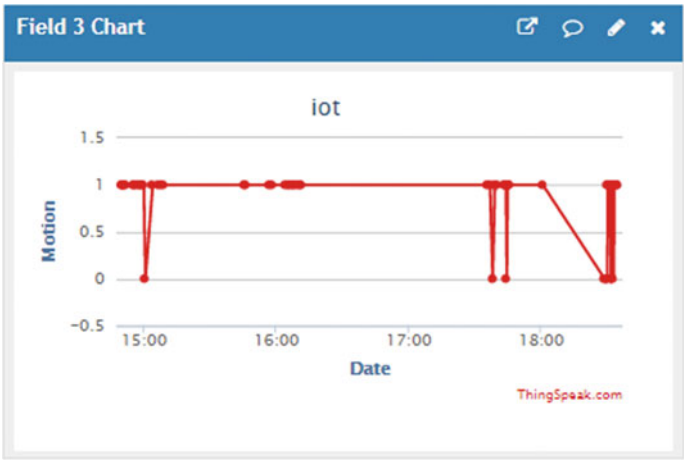


Fig. 14 The data of motion detection that are logged into server during prototype testing of PIR sensor in Thingspeak cloud server

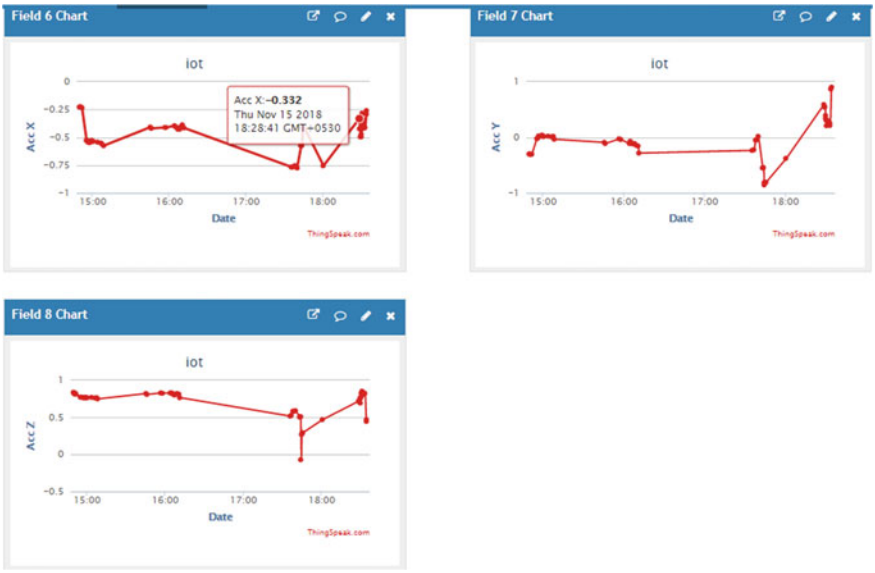


Fig. 15 The data of all three axes (x, y, z) that are logged into server during prototype testing of MEMS in Thingspeak cloud server

7 Conclusion

Various mechanisms of science are studied thoroughly and been implemented in it, especially the laws and theories of Geographical Information Tools (GIS) were implemented on microcontroller and its working mechanism/algorithm for perfect working and precise output information provision. For the purpose of detecting the analog sensors value, an Arduino interfaced with Raspberry Pi through pyfirmata protocol can be very effectively used. This system also has a scope of improvement for extending its effectiveness, which can be achieved by implementing various aspects of machine learning and artificial intelligence through a specific algorithm in a basic ambition to increase its precise decision-making capability up to a great extent.

Natural disasters are such events that cannot be eliminated by human force but their devastating effect upon the human health and wealth also cannot be ignored. This present proposal of our device can be effectively used for logging real-time values of the elevation profiles, slopes, and different weather parameters of the targeted area for early prediction, instant warning facility, and for a reliable real-time monitoring of landslide-prone areas all over the globe. Due to implementation of data logging into cloud, it enables the data acquisition and analysis at any point; as a result, it makes the system effective, efficient, as well as very reliable.

References

1. V. Van Khoa et al., Detection of landslide disaster by telemetric sensing node network system. in *2017 11th Asian Control Conference (ASCC)* (IEEE 2017)
2. A.K. Saha et al., GIS-based route planning in landslide-prone areas. Int. J. Geogr. Inf. Sci. **19**(10), 1149–1175 (2005)
3. R. Morello, C. De Capua, M. Lugarà, The design of a sensor network based on IoT technology for landslide hazard assessment. in *Proc. Imeko TC19 Symp. Environ. Instrum. Meas. Protecting Environ.* (Climate Changes Pollution Control 2013)
4. A. Musaev, D. Wang, P. Calton, LITMUS: a multi-service composition system for landslide detection. IEEE Trans. Serv. Comput. **8**(5), 715–726 (2015)
5. H. Wang et al., Panzhihua airport landslide (Oct. 3rd 2009) and an emergency monitoring and warning system based on the internet of things. J. Mt. Sci. **10**(5), 873–884 (2013)
6. S. Kapoor, H. Pahuja, B. Singh, Real time monitoring & alert system for landslide. in *2016 2nd International Conference on Contemporary Computing and Informatics (IC3I)* (IEEE 2016)
7. M.V. Ramesh, Design, development, and deployment of a wireless sensor network for detection of landslides. Ad Hoc Netw. **13**, 2–18 (2014)
8. C. Arnhardt et al., Sensor based Landslide Early Warning System-SLEWS. Development of a geoservice infrastructure as basis for early warning systems for landslides by integration of real-time sensors. Geotechnol. Sci. Rep. **10**, 75–88 (2007)
9. M.M. Mon et al., Analysis of disaster response during landslide disaster in Hakha, Chin State of Myanmar. J. Disaster Res. **13**(1), 99–115 (2018)
10. W. Lange et al., Ecosystem-based approaches for disaster risk reduction and climate change adaptation in Rio de Janeiro state. in *Strategies and Tools for a Sustainable Rural Rio de Janeiro* (Springer, Cham, 2019), pp. 345–359
11. R. Singh, R. Samkaria, A. Gehlot, N. Kumar, K. Rawat, A. De, A. Rehman, Algorithm to read various sensors to detect the hazardous parameters in industry. Int. J. Eng. Technol. (UAE) **7**(2), (2018)
12. A. Gehlot, R. Singh, R. Samkaria, S. Choudhury, A. De, Kamlesh, Air quality and water quality monitoring using XBee and internet of things. Int. J. Eng. Technol. (UAE) **7**(2), (2018)
13. N.K. Singh, A. Singh, R. Singh, A. Gehlot, Design and development of air quality management devices with sensors and web of things. Int. J. Eng. Technol. (UAE) **7**(2), (2018
14. R. Singh et al., Application of iCloud and Wireless Sensor Network in environmental parameter analysis. Int. J. Sens. Wirel Commun. Control **7**(3), 170–177 (2017)
15. S. Choudhury, P. Kuchhal, R. Singh, ZigBee and Bluetooth network based sensory data acquisition system. Procedia Comput. Sci **48**, 367–372 (2015)
16. G.N.L.R. Teja et al., Land Slide detection and monitoring system using wireless sensor networks (WSN). in *2014 IEEE International Advance Computing Conference (IACC)* (IEEE 2014)
17. R. Singh, S. Mishra, P. Joshi, Wireless pressure monitoring in wireless sensor network using 2.4 GHz transceiver module. in *ICCCT 2011, IEEE Conference on Computer and Communication Technology at MNNIT Allahabad held on September*, vol. 17, (2011)
18. R. Singh et al., *Arduino-based embedded systems: interfacing, simulation, and LabVIEW GUI* (CRC Press 2017)
19. R. Singh et al., *Arduino Meets MATLAB: Interfacing, Programs and Simulink*, (Bentham Science Publishers 2018)

Image and Video Capturing for Proper Hand Sanitation Surveillance in Hospitals Using Euphony—A Raspberry Pi and Arduino-Based Device



Navjot Rathour, Rajesh Singh and Anita Gehlot

1 Introduction

The system is designed to ensure the compliance of proper hygiene in various fields like food and catering services, healthcare and other industries where health hazards are directly related to maintenance of hygiene in work environment. In this system, an electronic system will be installed in restrooms or may be other areas where hand sanitation is required. This system is designed by using RFIDs (radio frequency identification unit) that will communicate with other electronic units by using RF waves. A central system is used to store the data captured in real time to manage the review process and to provide individual's specific data. Both checking handwashing using RF (radiofrequency) innovation and the utilization of portable RF units as operators of exchange of data starting with one RF unit then onto the next, just as to a PC information base, are outstanding exercises in earlier workmanship. The mix of said two exercises, together with the observing of the hand scouring systems as called for in the preferred embodiment to monitor handwashing, be that as it may, has never been referenced in earlier workmanship.

N. Rathour (✉) · R. Singh · A. Gehlot
Lovely Professional University, Phagwara 144411, Punjab, India
e-mail: er.rathour@gmail.com

R. Singh
e-mail: srajssssece@gmail.com

A. Gehlot
e-mail: eranita5@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_145

2 Review of Literature

In current scenario innovation human–machine collaboration is developing popular and machine needs to comprehend human signals and feelings. A real-time facial emotion recognition system has been proposed by P. Suja and Shikha Tripathi. They proposed the method by using three basic steps, i.e. Haar cascade, features extraction using active shape model (ASM). The classification is done by AdaBoost classifier to classify five basic emotions, i.e. emotions anger, disgust, happiness, neutral and surprise. They proposed a system with 94% accuracy, when it comes to implementation using Raspberry Pi II to recognize facial emotions in real time [1]. Similarly, Muhammad Sajjad et al. have presented a system for suspect distinguishing proof to inactively empower security in keen urban communities. They utilized a computationally economical Viola-Jones calculation algorithm for face detection on Raspberry Pi and ORB operation for feature extraction as well as recognition and classification of those extracted features from facial images only. Finally, the extracted features are uploaded on cloud for identification of suspects [2]. Another novel technique is implemented by Shyam Narayan Patel and V. Prakash, i.e. a wheelchair which can be controlled by the movement of eye, which is very helpful for disabled people. This system consists of a camera module that will capture the eye movement and after tracking the position of pupil of eye sends signal to motor of wheelchair. To make framework financially savvy for checking, a Raspberry Pi board permitted to get to the framework without showing unit [3].

Security issues are an essential piece of regular day to day existence. Home security is vital in the current scenario. Najmurrokhman et al. published their work in the same field and described a prototype facial recognition-based smart room access control system. A Raspberry Pi-based facial recognition system is designed to authenticate and authorize the person who can enter in room. They employed Haar cascade classifier in a picture preparing of client's face to render the face identification with high exactness [4]. Till today confront identification is a consuming point for the specialists. In the zones like computerized media, wise UI, clever visual observation and intelligent diversions. Different clamours and obscuring impacts confront pictures caught progressively. Steven Lawrence Fernandes and G. Bala Josemin proposed a Raspberry Pi-based single-board computer for face detection that can work well even in the presence of Gaussian noise, salt and pepper noise, motion blur and Gaussian blur [5]. Detection of emotional state of patient is vital for preliminary assessment of the patients in e-healthcare. Musaed Alhussein proposed an efficient facial emotion recognition system that used Weber local descriptors (WLD). Two publicly available databases were used in the experiment out of which 99.28% accuracy has been achieved in the case of Cohn–Kanade database [6].

The explosive development of sensor streams additionally prompts another worldview of edge registering. In the reconnaissance framework, edge-based mechanization is critical to get a quick reaction for quick information examination among associated gadgets. Hwin Dol Park et al. have proposed an automated surveillance system to improve to enhance power and knowledge [7]. Mrutyunjaya Sahani has

also proposed the usage and organization of remote control framework and availability into a home domain for verified individuals as it were. This system will use ZigBee based image processing technique and PCA for recognition purpose. Clients can screen guests and control the entryway bolt-on dynamic web pages improved with JavaScript and HTML. This framework finds a wide application in zones where physical nearness is beyond the realm of imagination constantly. The whole control framework is fabricated utilizing ARM1176JZF-S microcontroller and tried for real use in home condition [8]. Face recognition and following have wide applications, for instance, law implementation, gaming, picture look, showcasing and so on. B. Dahal et al. have proposed a system that can track the face using skin colour [9]. Another system has been proposed by Virginia Menezes et al. that will detect the moving objects in surveillance area. The detected objects are captured and streamed online by an MPEG streamer that can be viewed by authorized person on the go [10]. Shivam Anand et al. have proposed a smart stick that can detect human face using Pi Camera and Raspberry board. This stick can detect the obstacles and human faces and then alert the person by generating a vibration and audio in a headphone, and 98% accuracy has been achieved by using OpenCV python and Haar cascade classifier [11].

A system is proposed that can monitor a highly secured region continuously to recognize the change in the previously captured frame. The captured videos are continuously saved in the cloud. The framework contains database of face examples of nearby speculates which is contrasted and the face example of the individual setting off the framework, and picture handling is done continuously to effectively distinguish the identified face; the framework additionally continues following the face all through the area regardless of whether the individual moves out of the edge by a camera mounted over a servo engine. For better observation, face tracking has been done continuously that can track and identify the person even if the person moves out of the camera frame, and control of the camera via remote access through IoT has also been presented [12]. Abraham et al. proposed a system to communicate the notice on the signage show framework dependent on the statistic highlights like sexual orientation of the spectator. Ongoing PC vision calculations are connected to give a spectator explicit promotion communicate on the showcase framework. Real-time system has been proposed based on the observation of observer and advertisement [13].

A computer-based hand sanitation system has been proposed that can electronically monitor the sanitation frequency [14]. A smart home automation system has been proposed by Archit Aggarwal et al. that used NodeRed and visual programming through nodes to perform a certain task. MQTT communication protocol has been used for communication of different devices [15]. A XBee and microcontroller 128-based system for nuclear reactor surveillance of employees has been proposed by Sanjay Singh et al. that can detect the movement of the person with the help of camera. To interface a movement sensor and a camera with XBee and microcontroller 128, so we can utilize the hubs for having an eye on workers or in atomic reactor for observation. The movement sensor will identify any development and camera will distinguish the individual. Till any improvement the camera will be in

dead position and as it gets any development area it winds up unique looking toward the advancement and send movements through remote interface to the expert position [16].

3 Block Diagram and Hardware Description

Figure 1 shows the block diagram of the hand sanitation surveillance system which contains Arduino UNO, PIR sensor, touchpad, Raspberry Pi, Pi Camera, relay and buzzer. Raspberry Pi is the main unit that will process the data after gathering from all the sensors. Touchpad is used to take the input from user, and PIR sensor is used to detect the presence of the person. The setup will be arranged in such a way that whenever an intruder is detected the sensor will send the data to Arduino which will further send the data to Raspberry Pi. Raspberry Pi on the other side will be used to upload the sensor values on cloud. Pi Camera is used to capture the face and send the recorded data on cloud. The facial images captured by Pi Camera will be available all the time on cloud, so they can be accessed at any point of time for further use.

The hardware required to develop the system are 14,000 mAh USB power bank Raspberry Pi III (Model B+), Pi case, 3D printed tripod mount, 8 GB SanDisk SD card, 32 GB Kingston Flash Drive, camera module and 8 mm metal washer.

Figure 2 represents the step-by-step algorithm to detect face, and Fig. 4 explains the complete workflow of the algorithm designed for face detection this algorithm. The initial step is to capture the image, and for that purpose, Pi Camera is used. After capturing the image, the most important task is to detect the face from the image and remove the unwanted information. This task is performed by using Haar feature. Pre-processing is one of the most important steps in face detection as lot of noise and unwanted information presents in the detected image. To remove the background and to get a clear facial image, Sobel filter is used that will give noise free sharp image

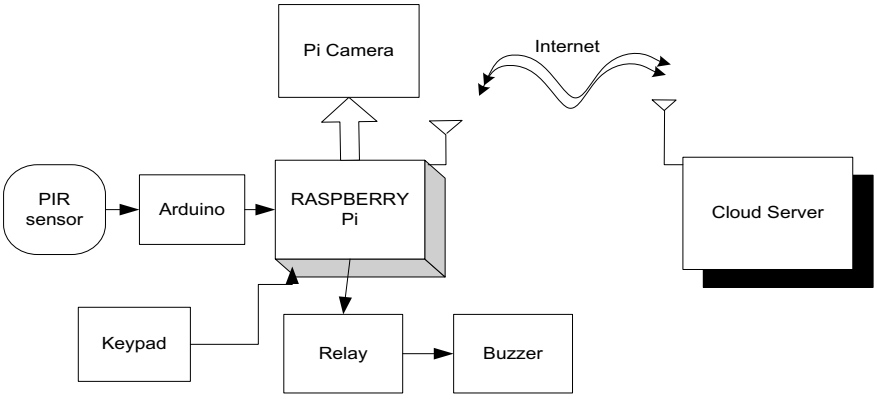


Fig. 1 Block diagram of the system

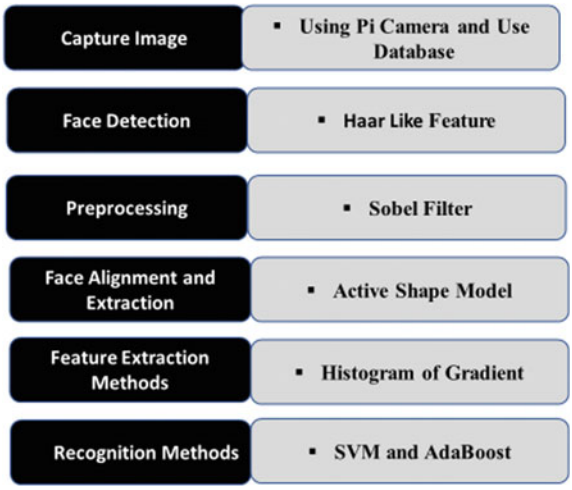


Fig. 2 Face detection and classification algorithm

of face. The next challenge is to align the facial image that is carried out with ASM (active shape model). ASM will extract the aligned face for recognition.

The programmed outward appearance investigation framework begins with determination of database pictures. Different database is accessible for detection of face in real time using appearance investigation framework. The general way to deal with programmed outward appearance investigation comprises three stages, i.e. confront identification and following, include extraction and articulation classification/acknowledgment. Classification or recognition is carried out by SVM (support vector machine) and AdaBoost. Initially, a lot of facial and non-facial images are required to train the classifier. After training, the system features should be extracted. The entire process was quite difficult and time-consuming before the existence of OpenCV. But with the presence of OpenCV, the entire process becomes quite easy as OpenCV because trainer and detector are by default present in OpenCV. Facial images can be taken from recorded and saved videos frame by frame. For that, Pi Camera is used, and the python programming with NumPy library can be used. For testing, the camera we will import NumPy library in OpenCV 2

```
import numpy as np
import cv2
```

For capturing the video of desired resolution, we can give capture command

```
cap = cv2.VideoCapture(0)
cap.set(3,640) # set Width
cap.set(4,480) # set Height
```

For displaying the captured frame imshow command can be used

```
cv2.imshow('frame', frame)
cv2.imshow('gray', gray)
```

Face detection process is basically classified in three tightly coupled steps:

- **Face Recognition**: Viola-Jones algorithm has been used to detect the facial images.
- **Feature Extraction**: After the successful detection of face, feature selection is one of the most important steps. Various feature extraction techniques available out of those Haar-like features are used to extract various features.
- **Classification**: The last step of automatic facial expression analysis system is classification. SVMs (support vector machines) and AdaBoost are the most widely used classifier and hence used.

Face detection and classification is an important part of Euphony as this will help to capture the images and stream the video from real world. As Euphony can stream live videos and images of the service providers on cloud so it can act as a base for future investigation. The process of face detection starts with the creation of database. A database of 13 images has been created. Figure 3 shows the loading of database images for real-time face detection using python on R-Pi with OpenCV.

After completion of loading the database images, the next step is to start capturing the faces in real time and start streaming the video; this step is explained in Figs. 4 and 5. Real-time images are classified after matching the images with database. Faces that are not part of the database images are classified as unknown as shown in Figs. 5 and 6.

Euphony will capture the complete image of service providers by capturing the hands also. That can be captured in real time after identifying the person with the help of database images via face detection. This video of service provider will be streamed live and will be saved on cloud also so that monitoring authorities can identify the

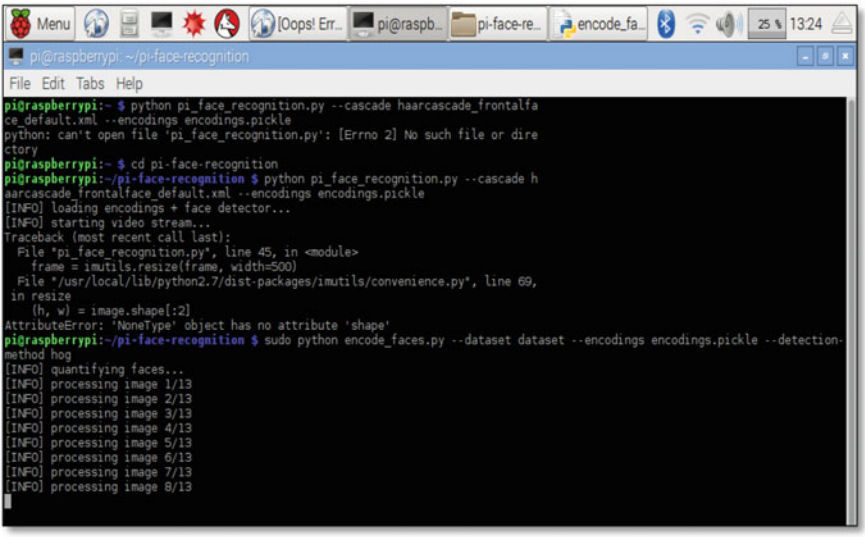


Fig. 3 Loading of database in process for face detection

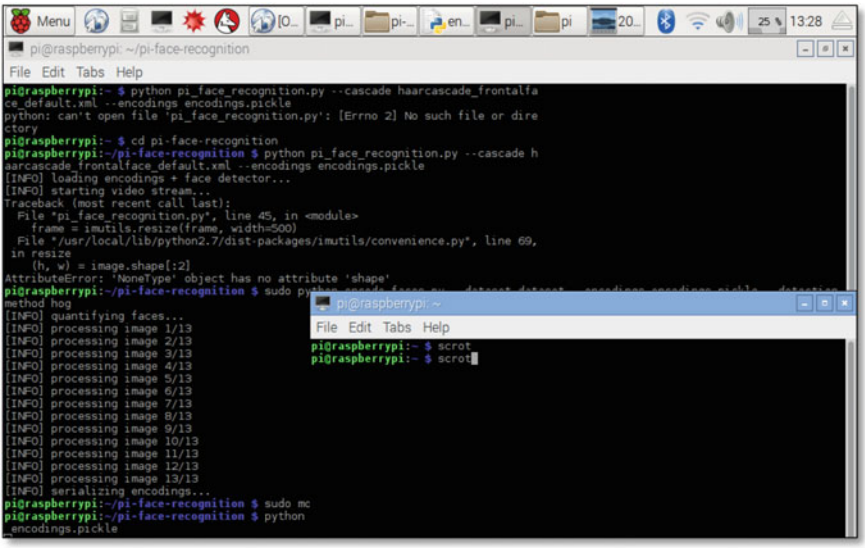
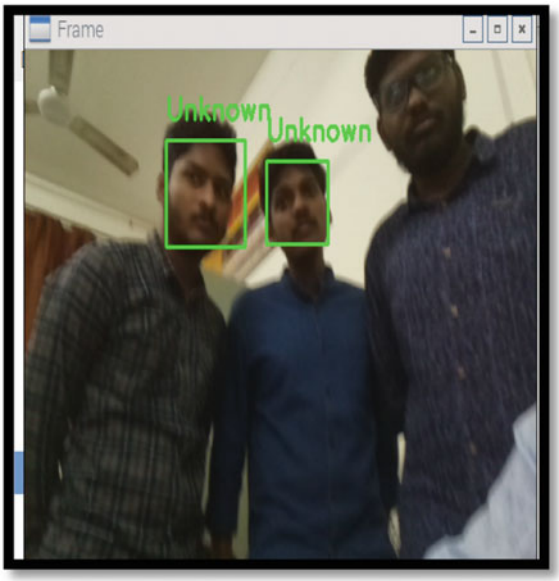


Fig. 4 Loading of database completed for face detection

Fig. 5 Face detection in real time



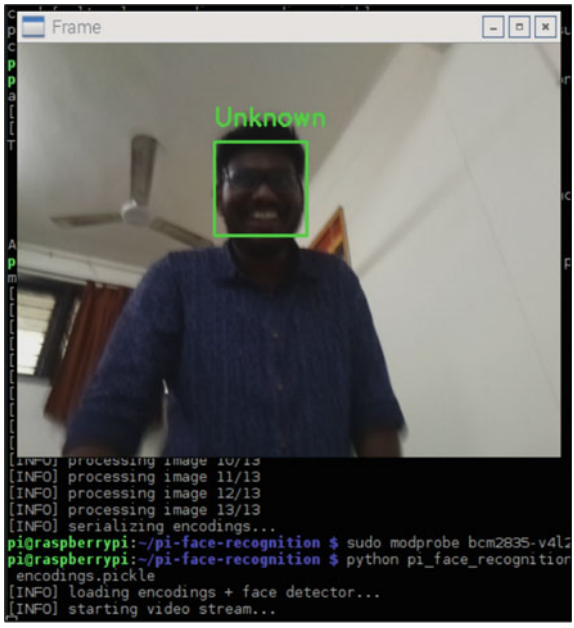


Fig. 6 Classification in real time

unhygienic conditions of service providers that will help them take a quick action to ensure hygienic conditions. Figure 7 shows the configuration of system for live streaming and Fig. 8 shows the real time image of the service provider to ensure the hygienic condition of hands.

To stream the live video, the uv4l driver needs to be installed, and for that, certain commands need to be run on terminal.

Then to start the streaming server use the commands and after that find the local IP Address by using IP Scanner Application like angry IP scanner. After that, we need to configure the router. Raspberry Pi and Arduino are connected via USB. Here we are using PIR sensor and it is connected to 4th pin of Arduino. Touchpad is attached to Arduino of 2, 3, 6, 7 pins. Buzzer with relay is connected to fifth pin of Arduino. +12 V Power supply is given to relay. Figure 9 shows the circuit diagram of the system.

4 Software Description and Algorithm

The main objective to capture the video is:

- Easy and quick to configure.
- Basic Python script that is easy to follow.
- Ability to define number and duration of video files to store.
- Minimize hardware items.

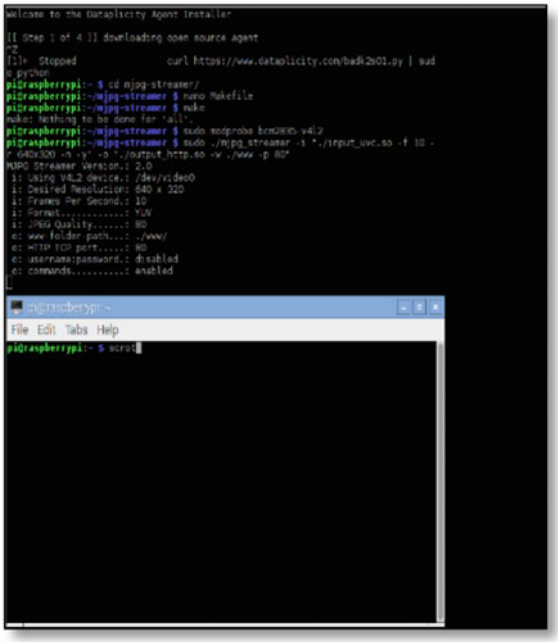


Fig. 7 Monitoring of hand sanitation in real time

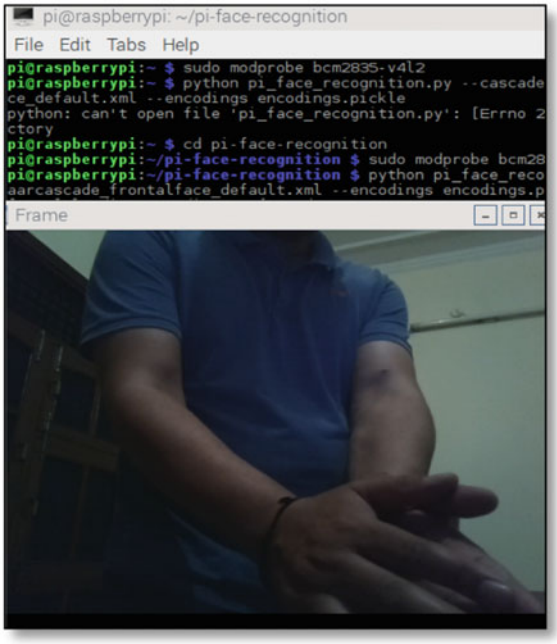


Fig. 8 Real-time streaming of hand sanitation

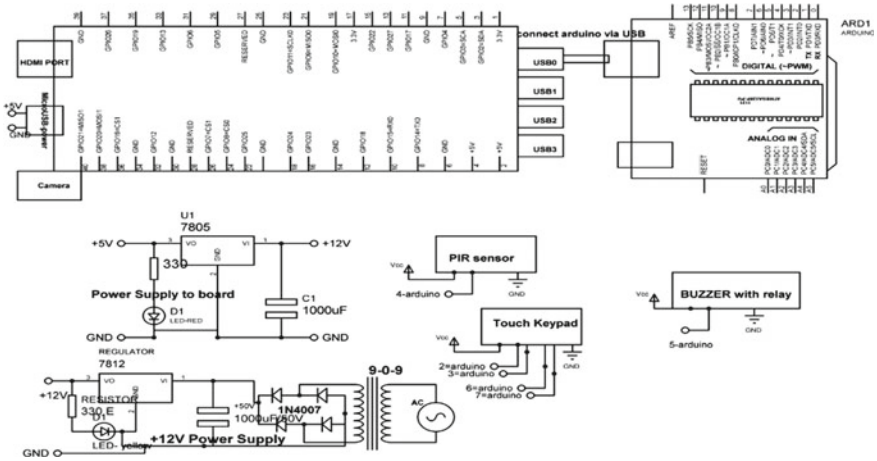


Fig. 9 Circuit diagram of the system

- Minimize wiring, soldering and construction time.
- Provide a method to stop recording and correctly shutdown Pi.
- Provide a quick method to attach to a tripod.
- Provide a method to review footage on a PC.

Figure 10 shows the complete workflow of Euphony starting from identification person near or at hand wash area where the system is installed and ensuring the hygienic conditions after monitoring the live streaming of hand sanitation activity in real time. Euphony will store the complete video of handwashing activity on cloud with name of the service provider that can be monitored anytime. Figure 11 shows the hardware setup of Euphony. This setup can be installed at any place where we want to monitor people to ensure hygienic conditions.

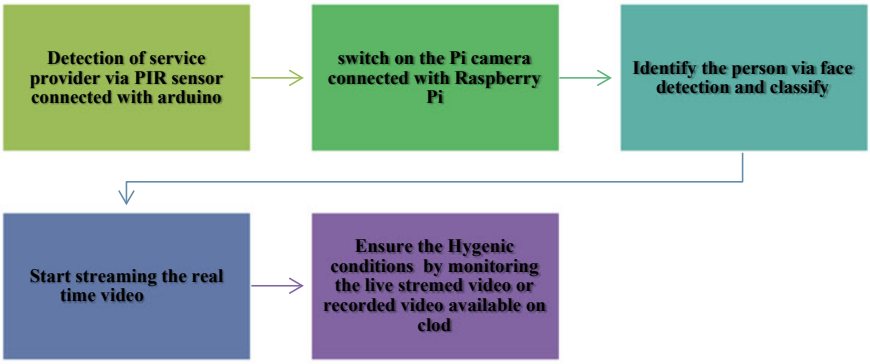


Fig. 10 Workflow of Euphony

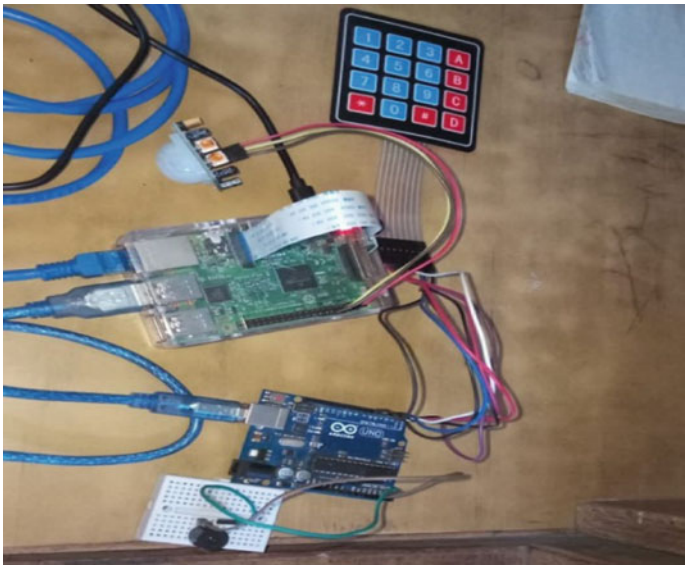


Fig. 11 Hardware setup of Euphony

5 Result and Discussion

A unique device called 'Euphony' is proposed that will monitor and inform the hand sanitation in Hospitals or other areas where hygiene is most important. This system will help to capture the person like nanny's or caretakers who comes in direct contact with patients in hospitals, to monitor the hand sanitation conditions and send alarming situations as well. This will be one of the unique systems that will ensure the hygienic conditions in hospitals.

6 Conclusion

Raspberry Pi is a remarkable device that can be efficiently used to design small applications with greater efficiency. Various sensors are used here for monitoring and sensing purposes. Pi Camera attached to system makes this device unique of its kind because we can save the captured videos on cloud for future predictions and references. In this paper, we have proposed an algorithm that will detect face as well as facial emotions. But practically, we have used the faces only. In the future, we can use those detected expressions of the person to conclude the mental or physical health conditions of caretakers in the hospitals.

Acknowledgements I would like to thank the management of Lovely Professional University for providing us the facility to conduct the experiment. All the figures are original and captured in Electronics Engineering Research Lab of Lovely Professional University.

References

1. P. Suja, S. Tripathi, Real-time emotion recognition from facial images using Raspberry Pi II, in *2016 3rd International Conference on Signal Processing and Integrated Networks (SPIN)* (IEEE, 2016), pp. 666–670
2. M. Sajjad, M. Nasir, K. Muhammad, S. Khan, Z. Jan, A.K. Sangaiah, M. Elhoseny, S.W. Baik, Raspberry Pi assisted face recognition framework for enhanced law-enforcement services in smart cities. Future Gener. Comput. Syst. (2017)
3. S.N. Patel, V. Prakash, Autonomous camera-based eye-controlled wheelchair system using raspberry-pi, in *2015 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS)* (IEEE, 2015), pp. 1–6
4. A. Najmurrokhman, K. Kusnandar, A.B. Krama, E.C. Djamal, R. Rahim, Development of a secured room access system based on face recognition using Raspberry Pi and Android based smartphone, in *MATEC Web of Conferences*, vol. 197 (EDP Sciences, 2018), p. 11008
5. S.L. Fernandes, J.G. Bala, Low power affordable and efficient face detection in the presence of various noises and blurring effects on a single-board computer, in *Emerging ICT for Bridging the Future-Proceedings of the 49th Annual Convention of the Computer Society of India (CSI)*, vol. 1 (Springer, Cham, 2015), pp. 119–127
6. M. Alhussein, Automatic facial emotion recognition using weber local descriptor for e-Healthcare system. Cluster Comput. **19**(1), 99–108 (2016)
7. H.D. Park, O.-G. Min, Y.-J. Lee, Scalable architecture for an automated surveillance system using edge computing. J. Supercomput. **73**(3), 926–939 (2017)
8. M. Sahani, C. Nanda, A.K. Sahu, B. Pattnaik, Web-based online embedded door access control and home security system based on face recognition, in *2015 International Conference on Circuits, Power and Computing Technologies (ICCPCT-2015)* (IEEE, 2015), pp. 1–6
9. B. Dahal, A. Alsadoon, P.W. Chandana Prasad, A. Elchouemi, Incorporating skin color for improved face detection and tracking system, in *2016 IEEE Southwest Symposium on Image Analysis and Interpretation (SSIAI)* (IEEE, 2016), pp. 173–176
10. V. Menezes, V. Patchava, M. Surya Deekshith Gupta, Surveillance and monitoring system using Raspberry Pi and SimpleCV, in *2015 International Conference on Green Computing and Internet of Things (ICGCIoT)* (IEEE, 2015), pp. 1276–1278
11. S. Anand, A. Kumar, M. Tripathi, M.S. Gaur, Human face detection enabled smart stick for visually impaired people, in *Data Science and Big Data Analytics* (Springer, Singapore, 2019), pp. 277–289
12. Jayakumar, A.J. Kumar, S. Muthulakshmi, Raspberry Pi-based surveillance system with IoT, in *Intelligent Embedded Systems* (Springer, Singapore, 2018), pp. 173–185
13. R.E. Abraham, M. Robert Kennedy, Intelligent digital signage system based on gender identification, in *Intelligent Embedded Systems* (Springer, Singapore, 2018), pp. 251–262
14. Y.G. Taneff, Hand washing enforcement system. U.S. Patent 9,030,325 issued 12 May 2015
15. A. Agarwal, R. Singh, A. Gehlot, G. Gupta, M. Choudhary. IOT enabled home automation through nodered and MQTT
16. S.Singh, R. Singh, S.R. Chaudhuri, Design of wireless sensor network node on zigbee for motion detection, in *IMS* (2010), p. 49

Industrial Hazard Prevention Using Raspberry Pi



Prabin Kumar Das, Praveen Kumar Malik, Rajesh Singh, Anita Gehlot, Km Vaishnavi Gupta and Arun Singh

1 Introduction

Being humans of twenty-first century, we always tend to make our life easier, fulfill our needs, and very economical. To achieve this goal, we are very closely interlinked with the various industries (e.g., Nuclear power plants, medicinal factories, gas production lines, etc.). But unfortunately after taking utmost care and surveillance also, we face some failures in the industries (e.g., Gas leak, nuclear radiation emission, etc.), and as a result into hazards and affects the society nearby it in a very harsh unreplenishable way. It lasts up creating a lot of damage to health and wealth. Toxic and flammable components are widely used in industry which when gets uncontrolled causes hazards. A study of past industrial hazards was found that the hazards recorded with highest damage to health and wealth has one thing in common, i.e., the fault lines are detected and proper information about the disaster was provided to the authorities and public with a lot of delay.

P. K. Das (✉) · P. K. Malik · R. Singh · A. Gehlot · K. V. Gupta · A. Singh
Lovely Professional University, Phagwara, Punjab, India
e-mail: prabindas786@gmail.com

P. K. Malik
e-mail: pkmalikmeerut@gmail.com

R. Singh
e-mail: srajssssece@gmail.com

A. Gehlot
e-mail: eranita@gmail.com

K. V. Gupta
e-mail: Vaishnavigupta170299@gmail.com

A. Singh
e-mail: arunmandiarun2001@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

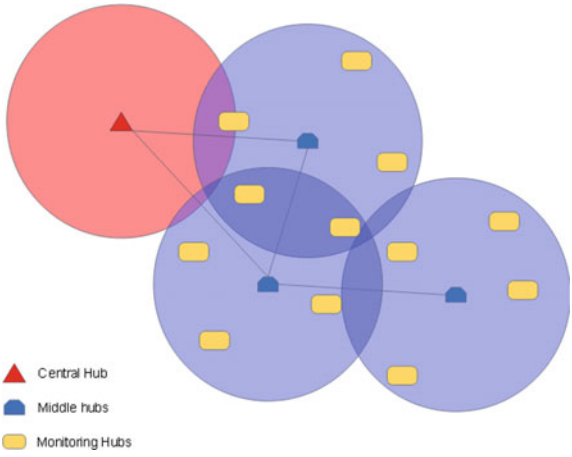
G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_146

Disaster planning and response require complex and more scientific analysis. The phases of the detection and rescue process need a highly efficient mechanism, which can predict the disaster, prevent, and trigger immediate assistance. This proposal focuses on designing nodes powered by solar energy rechargeable cells with pre-installed various parameters analyzing sensors that is to be deployed inside an industry at specified positions following a zonal mapping algorithm and in its nearby areas surrounding it.

2 Review of Literature

Disaster prevention or prediction cannot be achieved through basic mechanism implementation; it needs a very highly specific scientific and analytical approach for its proper and precise implementation of prevention and emergency trigger algorithms [1]. Every field of study related to disasters is nowadays only focused toward analysis of the disaster rather than focusing toward its prevention and prediction mechanisms. We know that analysis of previous disasters can reduce the number of disasters to some extent but will never be able to prevent it in anyway. This paper focuses on all-round aspects of a disaster comprising of disaster planning, training, response as well as relief methodologies [2]. This paper basically focused on the most commonly used as well as very much inflammable gas LPG. This uses some sensors to monitor the level of LPG and also during any event of leakage, it sends an alert signal to the user through a GSM module instantly [3]. This system proposal focuses on detection and immediate emergency triggering for any event of gas leakage like (LPG, isobutane, cigarette smoke, etc.); in case of any leakage event, its output goes low and it is considered as an event of leakage. Instantly, an alert message through GSM module is sent to the user and exhaust are turned on after a delay of few milliseconds [4]. This system proposal includes mechanism for radiation detection. In case of any event of radiation leakage, the device mechanism uses any proximate mobile personal communication for sending an alert message to the authorities and it predicts the possible source of radiation nearby. This basically aims to create a device included inside mobile communication for radiation detection without any human interference [5]. This proposed mechanism hopes to present a device using GPS and other radiation sensors to detect the radiation leakage and provide alert information about the real-time value of the radiation level along with its location to the user through a wireless ZigBee communication protocol. The scheme supports both the continuous and the intermittent measurement modes for making it more reliable and effective [6]. In this proposal aiming toward construction industrial hazards and accidents, a fuzzy risk assessment method has been countered. Various methods have been implemented to analyze various types of hazards and their root cause. Failure mode and effects analysis (FMEA) was performed to assess the risk value of hazard caused based on the fuzzy inference approach [7]. In this paper, various contributing factors like fall size, company size, and individual factors like age, gender, and experience have been studied and graphed for analyzing the major causes of accidents.

Fig. 1 Positioning



Primary measures like fixed barriers, strong handrails, and secondary-type measures like travel restraint systems (safety belt) and fall arrest systems were proposed to prevent or reduce the consequences of these accidents [8]. In this paper, the hazards in chemical industries of Korea due to fire, explosion, or release of harmful gases has been aimed particularly as a critical issue. A PCM system was introduced and was implemented on every industry under which the business owner of such places shall submit a process safety report to government to prevent any accidents, which could inflict an immediate damage on workers or on areas in the vicinity of the workplaces [9]. The aim of the current work is to explicitly link the inherent safety principles of minimization, substitution, moderation, and simplification with strategies for dust explosion prevention and mitigation. Various discussions were performed for implementing various ways to minimize, substitute, moderate, and simplify dust hazards. Particular attention is paid to the relationship between each inherent safety principle, various dust explosibility parameters, alternate methods of processing, selection of process equipment, and development and implementation of safe work procedures [10]. For air quality monitoring, the Air Quality Index (AQI) is important. It is an index prepared by government institutions to categories of risk of air pollution. Lower AQI value represents pleasant air quality. Below is the AQI index value adopted by India. The various algorithms and hardware descriptions are used to read the various gas sensors for monitor and control applications [11–13].

3 Block Diagrams

3.1 Positioning

Figure 1 shows the positioning of the nodes and their interconnections for data sharing. In this, various monitoring hubs are deployed at various specific positions inside the targeted area and connected to a middle hub. All these middle hubs will then be connected to a central hub for final analysis of the data received and logging into dedicated server.

3.2 End Device Node

Figure 2 shows the connection in an end node. These sensors are decided based upon the type of industry and possible elements used in that industry which can cause a hazard. All the sensors sense the environmental parameters in a real-time manner and send them into the microcontroller. The microcontroller is connected to a central hub through ZigBee module for data sharing and further analysis of data.

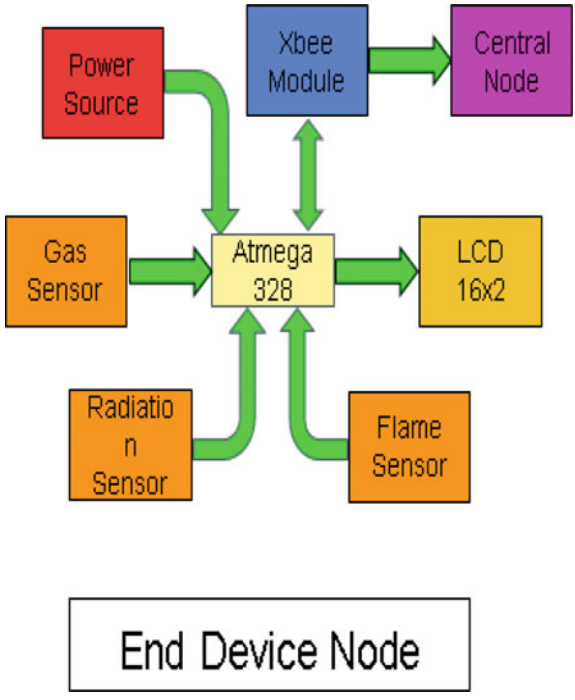


Fig. 2 End node

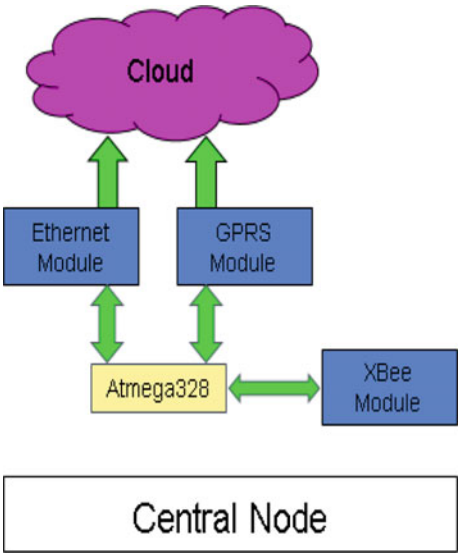


Fig. 3 Central node

3.3 Central Node

Figure 3 shows the central node that receives the data from several end nodes through a XBee module connection and then analyzes it. After successful analysis, it decides the result and logs the data into the dedicated server through Ethernet or GPRS module.

3.3.1 Raspberry Pi

The Raspberry Pi is a tiny and affordable computer that can be used to learn programming. It was originally intended to help spur interest in computing among small school-aged children. This runs totally on open-source software and is very much compatible to be interfaced with any kind of device (Figs. 4 and 5).

3.3.2 Arduino Uno R3

The Arduino Uno is a removable Dual In-line Package (DIP) ATmega328 AVR based easily programmable microcontroller having 20 digital I/O pins (i.e., 6 can be used as PWM outputs and 6 can be used as analog inputs). The R3 is the third and latest version of the basic Arduino Uno. It is a very reliable and easy to use microcontroller for various IoT applications in different fields. Its wide compatibility makes it more reliable.

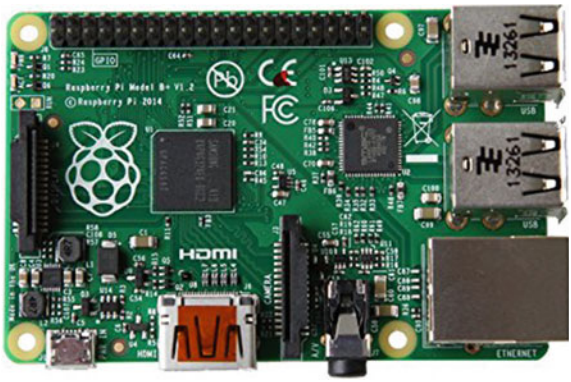


Fig. 4 Raspberry Pi

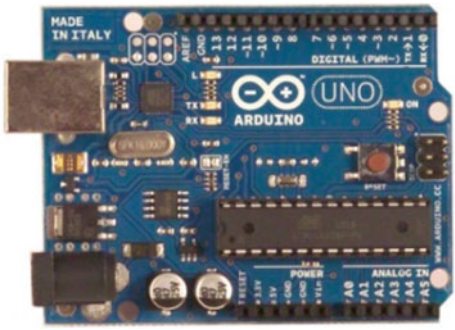


Fig. 5 Arduino Uno

3.3.3 ZigBee Module

ZigBee is an IEEE 802.15.4-based specification for a suite of high-level communication protocols used to create personal area networks with small, low-power digital radios, such as for home automation, medical device data collection, and other low-power low-bandwidth needs, designed for small-scale projects which need wireless connection. Hence, ZigBee is a low-power, low data rate, and close proximity (i.e., personal area) wireless ad hoc network. This is very suitable for implementation in an industrial area (Figs. 6 and 7).

3.3.4 Geiger Counter

A Geiger counter is an instrument used for detecting and measuring ionizing radiation. Also known as a Geiger–Mueller counter (or Geiger–Müller counter), it is widely used in applications such as radiation dosimetry, radiological protection, experimental physics, and the nuclear industry.

Fig. 6 ZigBee module

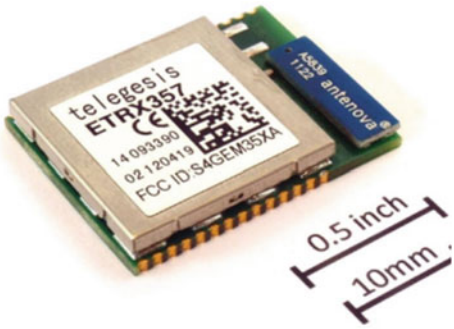


Fig. 7 Geiger counter



3.3.5 MQ Gas Sensor

The MQ series of gas sensors use a small heater inside with an electrochemical sensor. They are sensitive for a range of gases and are used for sensing various gas parameters. This gives an analog output which can be given to a microcontroller for an effective analysis (Figs. 8, 9, and 10).

3.3.6 Flame Sensor

These sensors work with the principle of UV radiation detections. It detects the UV radiation emitted from the event of ignition event and often creates 2–3 s delay to prevent false alarm.

Fig. 8 MQ gas sensor



Fig. 9 Flame sensor

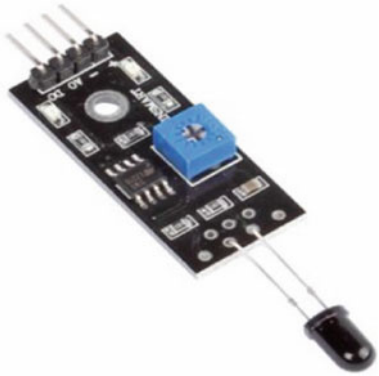
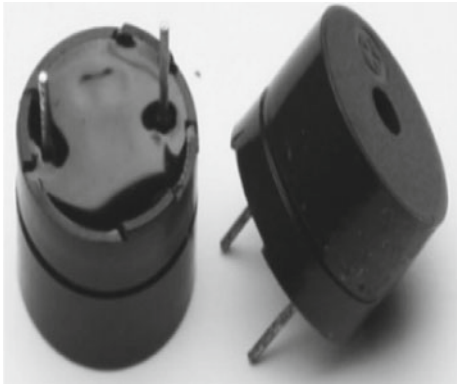


Fig. 10 Buzzer



3.3.7 Buzzer

A buzzer or beeper is a device basically used for alert or alarming applications. These are of different types according to their field of application, i.e., mechanical, electromechanical, or piezoelectric types.

4 Circuit Diagram

Figure 11 shows the circuits and connections of all the equipments used in this device implementation. It consists of circuitry details of main microcontroller Raspberry Pi, Arduino, XBee module, MQ2 gas sensor, Geiger counter, and buzzer.

In Fig. 11 of the proposed device, all the sensors are connected with Raspberry Pi zero W. The microcontroller collects the data sensed by the sensors and logs those into central node through ZigBee module, and then the central node logs the analyzed values into a dedicated server for further analysis and implementation. For interfacing with analog output sensors, we have used Arduino because there is no provision of analog pins input on Raspberry Pi. We have implemented Firmata protocol to create the interaction between Arduino and Raspberry Pi. Basically, the Arduino plays the role as a gate node for the sensor and the Raspberry in this proposed system.

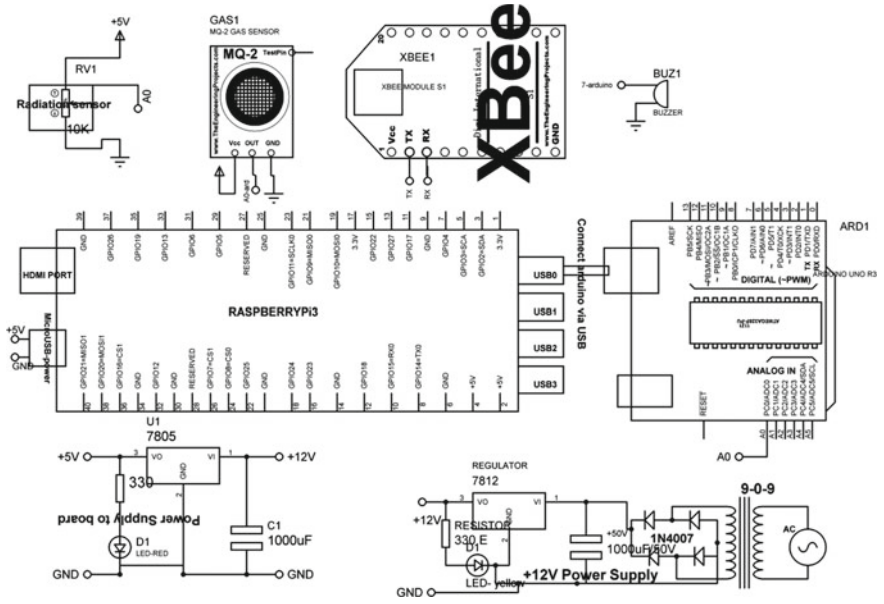
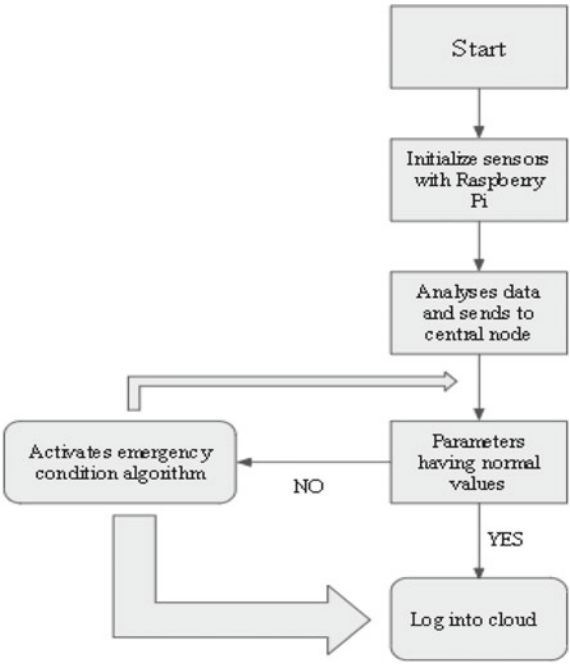


Fig. 11 Circuit details

Fig. 12 Flowchart



5 Software Developments

Once the system is set to ON state, all the deployed sensors start sensing the environmental parameters in a real-time manner and send them on to the microcontroller. The microcontroller receives the data, analyzes the values, and sends them to the central node through ZigBee module. Microcontroller at the central node analyzes the received data and starts logging the real-time values into a dedicated server. This device works on a specific algorithm in which, whenever the values received by the microcontroller show an abrupt rise in a specific parameter, it takes that event as a proximity condition and instantly triggers an alert information comprising of the exact location of abrupt rise and the details of parameters that has risen onto the server as well as to the concerned authorities without any delay. The process is depicted in Fig. 12.

6 Results and Discussion

To get a proper information about its reliability and effectiveness in a practical situation, a prototype of this device has been designed by using sensors like Geiger counter, MQ2, IR flame sensor, and a microcontroller for making end node. This end

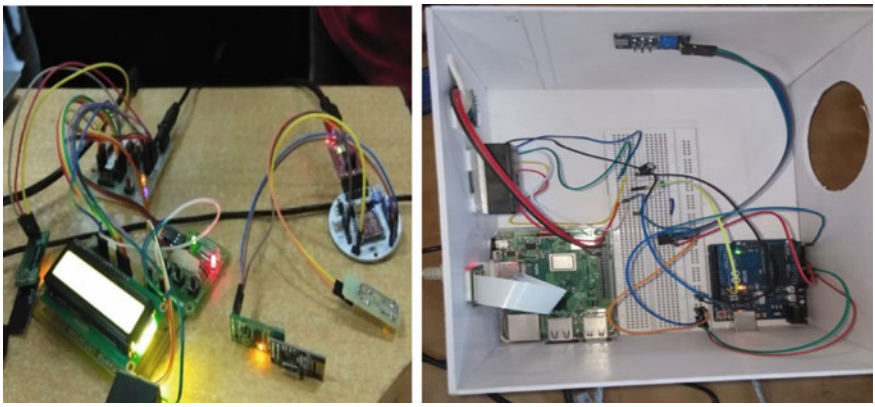


Fig. 13 View of the system

node was then connected to central node through a ZigBee module. Very positive results were obtained from the working of that prototype in every tested situation.

First, all these sensors are deployed at a place resembling condition of an industrial hazard. The system was then turned ON. An Arduino Uno has been used as a gate for taking the analog inputs and feeding them to Raspberry Pi which is used as a microcontroller to take all the sensed values of environmental parameters from these sensors, analyze it, and then send them to the central node. Central node uses a Raspberry Pi to log those values into a dedicated server after analysis through Ethernet module. Based upon the working algorithm whenever the microcontroller analyzes the values received from the sensor and observes an abrupt or abnormal rise in any environmental parameters more than a set max threshold, it takes this event as a proximity condition for hazard and instantly activates its alarm features through the buzzer and also sends an alert warning along with the location of observed abrupt rise and parameter that is rising onto the server instantly as well as to the concerned authorities.

Figures 13 and 14 show the snapshot of the system developed and also values of methane and LPG sensed and recorded during prototype testing into a dedicated server (IOT SARDAR).

7 Conclusion

Various types of scientific approaches are combined together for the implementation of this proposed device, specifically the zonal mapping technique for placement and positioning of the nodes inside a specific industry arena. This creates a lot more impactful implementation of the device by enabling us to detect the zone having the fault line. For the purpose of detecting the analog sensor values, an Arduino interfaced with Raspberry Pi through Firmata protocol has been effectively used.



Fig. 14 Server records

The main reason of using Raspberry Pi in this proposal is because of the usage of constant Internet connectivity to the main central node that can be given easily by an Ethernet cable connected to the Pi. This system also has a scope of improvement for extending its effectiveness, which can be achieved by implementing various aspects of machine learning and artificial intelligence through a specific algorithm in a basic ambition to increase its precise decision-making capability and even emergency help provision capability up to a great extent without any human interference in its mechanism.

Natural disasters are such events that cannot be eliminated by human force but their devastating effect upon the human health and wealth also cannot be ignored. This present proposal of our device can be effectively used for early prediction, instant warning facility, and for a reliable real-time monitoring of industrial hazard-prone areas all over the globe. Due to the implementation of data logging into cloud, it enables the data acquisition and analysis at any point; as a result, it makes the system effective, efficient as well as very reliable. As an add-on, by using the logged data from the server, we can monitor the effects of any industry upon its surrounding and can further advice or order them to improve or change their working procedures and machineries.

References

1. D. Alexander, Disaster and emergency planning for preparedness, response, and recovery, in *Recovery, Risk Management, Response, Preparedness* (2015)
2. M. Masellis, Thermal agent disaster and fire disaster: definition, damage, assessment and relief operations. 978-0-585-33973-3

3. A. Sood, B. Sonkar, A. Rajan, A. Faisal, Micro controller based LPG weighing and detector using GSM module. Int. J. Electr. Electron. Res. **3**(2), 264–269 (2015)
4. T. Ismail, D. Das, J. Saikai, J. Deka, R. Sharma, GSM based gas leakage warning system. IJARCCE4C **3**(4) (2014). ISSN 2278-1021
5. C. Balchunas, D. Rogers, Radiation detection and tracking with GPS-enabled wireless communication system. U.S. Patent Application 10/795,659, filed, 11 May 2006
6. F. Ding, G. Song, K. Yin, J. Li, A. Song, A GPS-enabled wireless sensor network for monitoring radioactive materials. Sens. Actuators A **155**(1), 210–215 (2009)
7. H.-T. Liu, Y.-l. Tsai, A fuzzy risk assessment approach for occupational hazards in the construction industry. Saf. Sci. **50**(4), 1067–1078 (2012)
8. C.-F. Chi, T.-C. Chang, H.-I. Ting, Accident patterns and prevention measures for fatal occupational falls in the construction industry. Appl. Ergon. **36**(4), 391–400 (2005)
9. H. Kwon, The effectiveness of process safety management (PSM) regulation for chemical industry in Korea. J. Loss Prev. Process Ind. **19**(1), 13–16 (2006)
10. P.R. Amyotte, M.J. Pegg, F.I. Khan, Application of inherent safety principles to dust explosion prevention and mitigation. Process Saf. Environ. Prot. **87**(1), 35–39 (2009)
11. R. Singh, R. Samkaria, A. Gehlot, N. Kumar, K. Rawat, A. De, A. Rehman, Algorithm to read various sensors to detect the hazardous parameters in industry. Int. J. Eng. Technol. (UAE) **7**(2) (2018) (Scopus Indexed)
12. A. Gehlot, R. Singh, R. Samkaria, S. Choudhury, A. De, Kamlesh, Air quality and water quality monitoring using XBee and internet of things. Int. J. Eng. Technol. (UAE) **7**(2) (2018) (Scopus Indexed)
13. N.K. Singh, A. Singh, R. Singh, A. Gehlot, Design and development of air quality management devices with sensors and web of things. Int. J. Eng. Technol. (UAE) **7**(2) (2018) (Scopus Indexed)

IoT-Based Traffic Management System Including Emergency Vehicle Priority



Anmol Joshi, Naman Jain and Alok Pandey

1 Introduction

Traffic management system plays a very crucial role in controlling the traffic flows and preventing traffic jams in urban areas. However, current systems are not automated and also, they are not dynamic in nature due to which they are not able to adapt themselves according to ever-changing traffic conditions. Most of the current systems work on the principle of fixed time slots which irrespective of the changes in the traffic density follows the same protocol and thus sometimes causes traffic jams or increase in waiting time of the vehicles. There can only be a data-based solution to make a dynamic system. So, in our model, we have used IR sensors to take input and data collection. IR sensors collect the data about the number of vehicles on the road and then this data is sent to the central processing unit. The central processing unit thus computes the traffic density of each lane according to the fetched data from IR sensors of all four lanes and then takes the decision accordingly.

So clearly in this system, the decision made by the central processing unit changes as the input that it gets from the sensors is updated from time to time. Use of automated traffic control systems is also an economic alternative for logistic, travel, and transportation industries as it tries to minimize the fuel consumption of the vehicles.

In related work [1], they used force resistive sensors to detect the vehicles. It transmits the detected information to the intersection controller via ZigBee. A system was installed in emergency vehicles and provides GPS coordinates to the intersection controller to avoid any waiting time for emergency vehicles at intersections.

A. Joshi · N. Jain
CDAC NETCOM ATC Center, Jaipur, India

A. Pandey (✉)
Scholar Poornima University, Jaipur, India
e-mail: eralok_ece@yahoo.co.in

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_147

Albagul, M. Hrairi et al. in 2006 used MATLAB for simulation and LabVIEW integrated system for interfacing the IR sensors. Complete set of traffic light system was synchronized with the sensors interfaced [2]. Shashikant et al. address the problem using Intelligent and Adaptive Traffic Light Controller (IA-TLC) implemented on FPGA using Verilog [3]. The system has no IoT solutions. IoT solutions are provided by the [1, 4, 5], but in [4, 5] algorithms and system model are not defined properly even there are no priority emergency vehicles. In [6, 7], vehicular sensor networks were utilized to provide promising solutions for traffic management by utilizing localization algorithms to determine the locations of vehicles containing wireless sensor network nodes.

To make such a system, we have placed four IR sensors in each lane, i.e., total 16 IR sensors to collect the data from all four lanes. All these sensors are connected to the central processing unit which is Raspberry Pi 3B+ in our case. The sensors will continuously update Raspberry Pi which will take the decisions based on the implemented algorithm to process the data. The first sensor will be placed at a distance of 10 m from the signal, second one at 20 m, third on 30 m, and fourth at 40 m. Also, in time of extremely dense traffic conditions, an auto-generated email will be sent to the nearest police station through an SMTP server. The hardware is controlled by a python script burn on Raspberry Pi. The testing and validation of our model have been done on python software.

We have used the XBee transmitter and a receiver module for detecting the emergency vehicles. Figure 1 shows the arrangement of the proposed system. We have developed this model at C-DAC campus, Jaipur.

2 Working of the Proposed System

Initially, all the lights will be red and then the CPU will take the input from the IR sensors. The input from sensors of all the four lanes indicates the level of density of traffic on each lane. Then the CPU gives the priority starting from the densest lane to the least dense lane in descending order. After that, time slots are allotted to the respective sorted lanes. All the above processes constitute one cycle of the system. The same cycle is repeated after the execution of the previous one. If the sensor output status of multiple lanes is same, then CPU will assign the same time duration for corresponding lanes. The algorithm of the proposed system is given below.

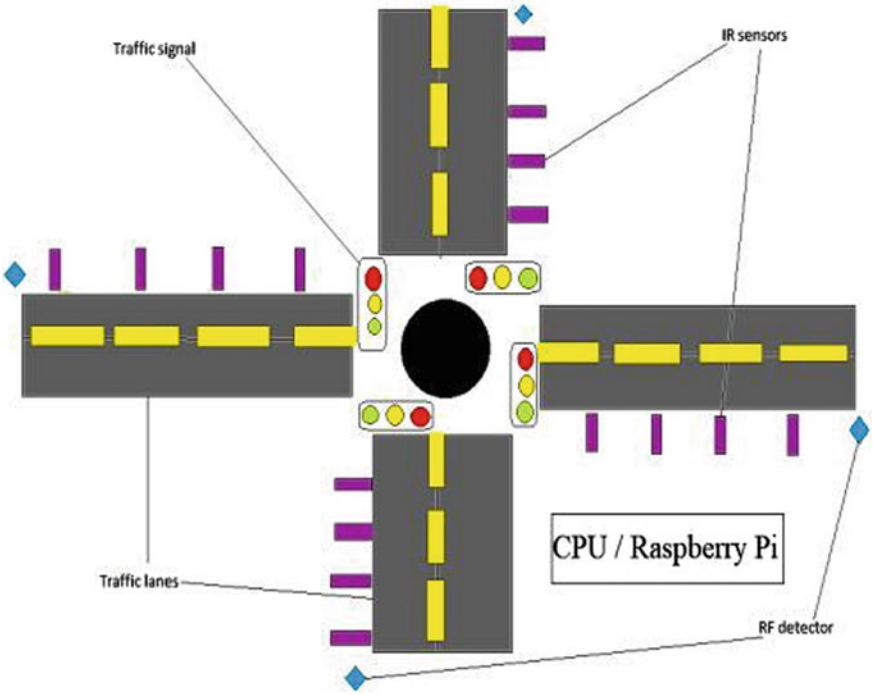


Fig. 1 Proposed traffic management system

Algorithm Intelligent Traffic Management System
Input: IR_SENSOR (IR Sensor Output), **RF_SENSOR** (RF Sensor Output)
Procedure:
Step 1: defining and initializing variables
CURRENT_PATH = I
CURRENT_LIGHT = LIGHT(CURRENT_PATH)
NEXT_PATH(I) = II
NEXT_PATH(II) = III
NEXT_PATH(III) = IV
NEXT_PATH(IV) = I
Step 2: selecting traffic light status on the basis of the path priority
-While (Working_Time) **do**
-if (priority of CURRENT_PATH is NORMAL)
- turn the status of all lights to Red
- turn the status of CURRENT_LIGHT to Green
-if (IR_SENSOR_1 is activated)
- wait for time t1
-End if
-if (IR_SENSOR_2 is activated)
- wait for time t2
-End if
-if (IR_SENSOR_3 is activated)
- wait for time t3
-End if
-if (IR_SENSOR_4 is activated)
- wait for time t4
-End if
- (t1 < t2 < t3 < t4)
- turn LIGHT(CURRENT_PATH) to Yellow
- wait for time t5
- turn LIGHT(CURRENT_PATH) to Red
- CURRENT_PATH = NEXT(CURRENT_PATH)
- CURRENT_LIGHT = LIGHT(CURRENT_PATH)
-END if
-END While

3 Algorithm for Emergency Vehicles

To detect the emergency vehicles (EV), we are using XBee transmitter and receiver module operating at 2.4 GHz frequency band. This module is compatible to connect with Raspberry Pi through Wi-Fi. The transmitting module of XBEE will be placed on moving the emergency vehicle and the receiver antenna will be placed approximately 100 m apart from the position of the last sensor. When an EV will come in contact of receiving antenna at the same time, the receiver antenna will send the signal to Raspberry Pi and Raspberry Pi will turn ON the red lights of remaining three sides.

Only one green light will be ON for the emergency vehicle to pass by. When this vehicle will cross the intersection and will come in the contact of another receiver on the next side, the Raspberry Pi will re-execute the program.

Algorithm for an emergency vehicle is as follows:

Algorithm Intelligent Traffic Management System
Input: IR_SENSOR (IR Sensor Output), **RF_SENSOR** (RF Sensor Output)
Procedure:
Step 1: defining and initializing variables
CURRENT_PATH = I
CURRENT_LIGHT = LIGHT(CURRENT_PATH)
NEXT_PATH(I) = II
NEXT_PATH(II) = III
NEXT_PATH(III) = IV
NEXT_PATH(IV) = I
Step 2: selecting traffic light status on the basis of the path priority
-While (Working_Time) **do**
-if (output of CURRENT_XB_SENSOR is HIGH)
- set priority of CURRENT_PATH is HIGH
- turn the status of all lights to Red
- turn the status of CURRENT_LIGHT to Green
- read all the other XB sensor until anyone except CURRENT_XB_SENSOR goes high
- priority of CURRENT_PATH is NORMAL
-END if
-END While

4 Fault Tolerance

In case of extremely heavy traffic conditions, the system will generate an email and will send it to the nearest traffic police station automatically. The above process is done with the help of the SMTP server.

5 Testing and Validation of the System

We have tested the system in C-DAC Jaipur lab and we have got satisfactory results. We have implemented our programs in Python. We have compared the efficiency of our system with the regular fixed time slot-based systems and found our system to be much more efficient. Let us assume that the fixed time slot-based system assigns the time T_f seconds to each lane. So, the total time taken for one cycle to complete for this system will be $4T_f$ seconds. But on the other hand, in our system, there are variable time slots, which are assigned according to the priority order computed by

the CPU. The values and the buffers for the variable time slots are determined by us experimentally. Let t_1, t_2, t_3, and t_4 denote the variable time slots. So, we can calculate the performance of our system by

$$\left(\left(\sum 4T_f - t_i\right)/4T_f\right) * 100$$

We have performed the experiment eight times and calculated the time duration of ON signal of each lane in each trial. Here, we are assigning 40 s to green signal if traffic density is at level four; similarly, if density is at level three, then the assigned time is 30 s. We have performed the experiment randomly eight times.

Figures 2 and 3 show the comparison between the static system timings and proposed system timings in eight successive trials. For lane 1 total on time was

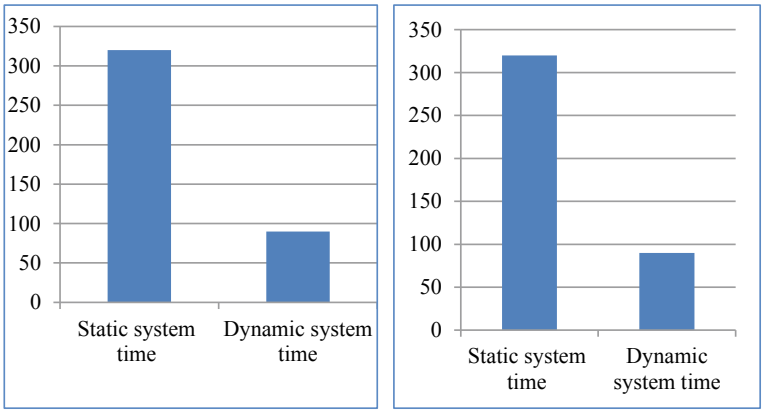


Fig. 2 Static system time versus dynamic system time for lane 1 and lane 2

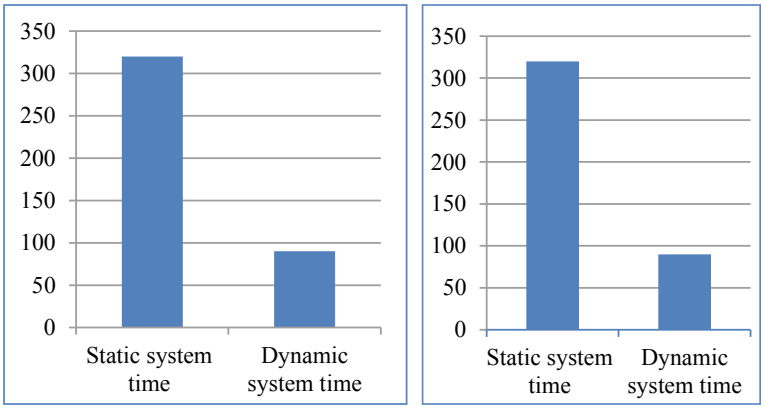


Fig. 3 Static system time versus dynamic system time for lane 3 and lane 4

290 s, for lane 2 time was 240 s, for lane 3 time was 140 s, and for lane 4 time was 90 s. It can be easily observed that in each LANE, system is dynamically saving the halt time of vehicles.

6 Conclusion

In this paper, we have designed an intelligent traffic light management system which is a completely IoT-based solution. This system is controlling the traffic in an intelligent way by gathering the data from multiple sensors and taking decisions accordingly. The system is in working condition and in case of ambiguity, it provides email notification to the nearest police control room.

References

1. A. Khan, F. Ullah, Z. Kaleem, S. Ur Rahman, H. Anwar, Y. Cho, EVP-STC: emergency vehicle priority and self-organizing traffic control at intersections using internet-of-things platform. IEEE Access **6**, 68242–68254 (2018)
2. M. Albagul, W. Hrairi, M.F. Hidayathullah, Design and development of sensor based traffic light system. Am. J. Appl. Sci. **3**(3), 1745–1749 (2006)
3. S.V. Lahade, S.R. Hirekhan, Intelligent and adaptive traffic light controller (IA-TLC) using FPGA, in *2015 International Conference on Industrial Instrumentation and Control (ICIC), Pune* (2015), pp. 618–623
4. L.P.J. Rani, M.K. Kumar, K.S. Naresh, S. Vignesh, Dynamic traffic management system using infrared (IR) and the internet of things (IoT), in *2017 Third International Conference on Science Technology Engineering & Management (ICONSTEM), Chennai* (2017), pp. 353–357
5. Anitha, K.N.R.M. Babu, IOT for ITS: an IOT based dynamic traffic signal control, in *2018 International Conference on Inventive Research in Computing Applications (ICIRCA), Coimbatore* (2018), pp. 532–537
6. R. Sundar, S. Hebbar, V. Golla, Implementing an intelligent traffic control system for congestion control, ambulance clearance, and stolen vehicle detection. IEEE Sens. J. **15**(2), 1109–1113 (2015)
7. E. Sireesha, D. Rakesh, Intelligent traffic light system to prioritized emergency purpose vehicles based on the wireless sensor network. Int. J. Comput. Appl. **40**(12), 24–27 (2012)

Design and Development of IoT-Enabled Portable Healthcare Device for Rural Health Workers



Saurabh Nautiyal and Rekha Devi

1 Introduction

Health care is a very important aspect of any country for ensuring a better quality of life. Many developing countries in the world are going through the public health burden. The economic social and structural disparities have led to immense challenges for healthcare facilities around the world [1]. India, one of the highest growing economies in the world, has a dismal doctor to patient ratio which is less than the WHO prescribed limit 1:1000. The disparity is especially high in rural areas; with about a majority of 132.42 crores, Indians lives in rural areas where the circumstance of medical facilities is not appropriate [2]. Due to lack of basic facilities and bad infrastructure, doctors and pharmacists refrain from doing on assignments in these areas. These areas require special attention from public healthcare authorities. Public health care includes the protection and improvement of community health in terms of the prevention, diagnosis, and treatment of disease, injury, illness, and other physical or mental impairments. In rural areas, lots of healthcare initiatives are being run by the government but the accomplishment and effectiveness of these programs are debatable due to gaps in the implementation [3]. In a general scenario, health centers in rural areas have limited or no patient's medical records' system. This makes periodic public health monitoring challenging, with no medical history of the individual. Also, not all rural inhabitants are serious about their health and they not even visit the health camp in villages. To fulfill the requirement of public health care, the government appoints the health workers as a bridge between rural inhabitants and central medical facilities in rural areas. Sometimes, it becomes inconvenient for health workers to carry and access the medical records of patients when they are working in the field and medical records are maintained at a central medical center.

S. Nautiyal (✉) · R. Devi
ACS Division, Centre for Development of Advanced Computing (C-DAC), (Ministry of Electronics and Information Technology, Government of India), Mohali, India
e-mail: saurabh.nautiyal454@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_148

In this type of healthcare system, there are chances of human errors in maintaining the health records and it makes health system unreliable and cumbrous [4]. Therefore, there is a need of such new technology-based system for health workers who make their routine assignment much easier and effective.

Considering the huge potential of IoT technology, current and undergoing development, in this research paper, a prototype of a healthcare device is developed for health workers who are an important part of the rural healthcare system. The goal of this device is to enable easy and reliable routine health checkups by acquiring the physiological data with different medical sensors interfaced with the device, maintain the medical database of each person at cloud platform, facilitate better health care at door, and enhance the quality of lives residing in rural areas.

2 The IoT-Based Healthcare System in Rural Areas

The Internet of Things (IoT) is a wireless web of electronic devices which are connected with each other to share information and data and yield new information so as to record and analyze it for future utilization [5]. The IoT-based technology provides immense support to the basic healthcare system in many ways like IoT can allow the interaction among the different medical sensors, monitoring devices, and other electronic medical equipment [6]. There are lots of policies and programs being run by the government to improve healthcare services in rural areas. In order to this, the government is bringing IoT technology to its rural healthcare system [7]. The rural healthcare system includes sub-centers, primary health centers, and community health centers [8]. These centers can deliver their healthcare services to rural inhabitants, which are more reliable convenient and less expensive by adopting IoT technology. Figure 1 shows the IoT-based healthcare system.

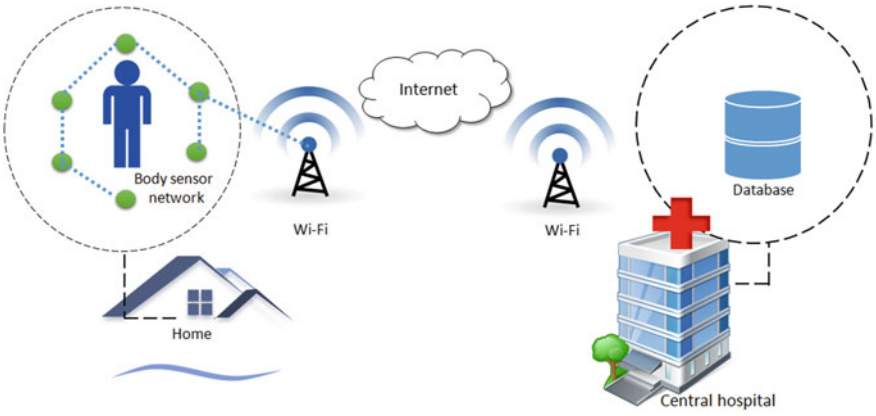


Fig. 1 IoT-based healthcare system

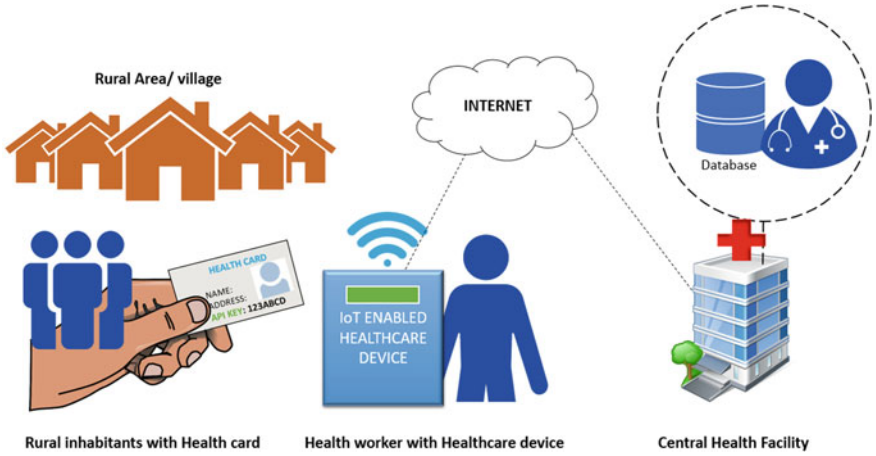


Fig. 2 Proposed electronic healthcare system

3 Proposed System

3.1 Proposed Electronic Healthcare System

In the proposed work, a prototype of a portable healthcare device is developed for health workers, which can be easily used for door-to-door health checkups. In the first visit, health worker makes an individual patient's account in the cloud platform and assign a unique API number on the card to each patient. In subsequent visits, health worker just need to enter the individuals' API key to record the health data in their respective account on the cloud platform. The healthcare device collects and uploads data on the cloud at the same time. The patient's medical history is accumulated at the central database. The decision-making in the proposed system is done by doctors and government medical departments at the centralized medical facility. This system improves the performance of the operational efficiency of the rural healthcare system. The developed device is portable and having simple functioning so it can be easily used by healthcare worker in the field. The proposed healthcare system is as shown in Fig. 2.

3.2 Block Diagram of the Proposed System

The block diagram of the proposed device is as shown in Fig. 3. We have interfaced digital blood pressure sensor, pulse oximetry sensor, and body temperature sensor with a microcontroller. These sensors can collect the individual health condition

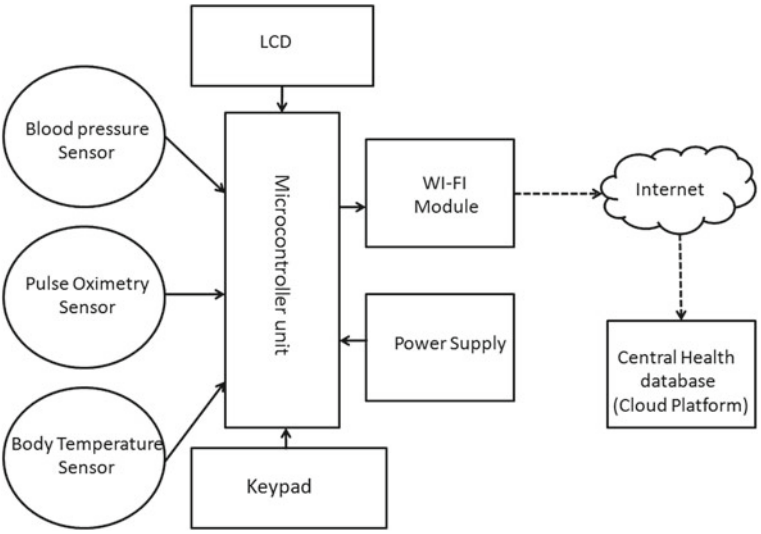


Fig. 3 Block diagram of the proposed system

data and show it on the LCD display. The proposed device is embedded with the Wi-Fi module, which provides an Internet connection to the device. An alphanumeric keypad is used to enter the unique identification number assigned to each person, which helps to send physiological parameters into their respective account in the cloud platform. First of all, the device is switched ON through power bank supply and the device shows a message to enter the API key. The health worker enters the assigned API key of a person using the alphanumeric keypad and then press the enter button. After that, device measures the blood pressure, pulse rate, blood oxygen saturation level, and body temperature simultaneously.

After the measurement of physiological parameters, the send button is pressed by the health worker and all the measured parameters are sent to the cloud platform for further analysis. The flow diagram of the complete process of the proposed healthcare system is shown in Fig. 4.

4 Hardware Implementation

4.1 Developed Prototype

First, we have implemented the hardware using a breadboard and tested the working of the interfaced medical sensor. Different medical sensors are interfaced with a single microcontroller. All the required coding is done by Arduino software. After that, we have interfaced the Wi-Fi module with the microcontroller and tested the

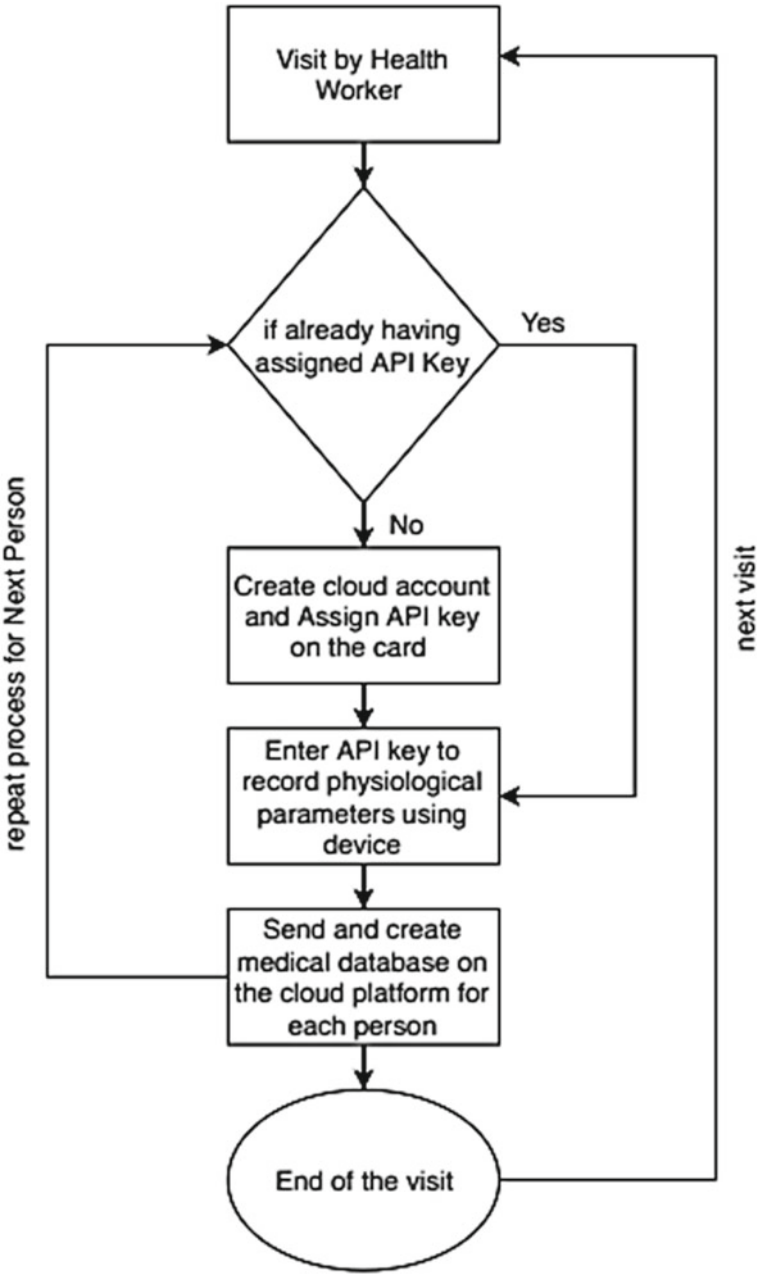


Fig. 4 Flow diagram of the complete system

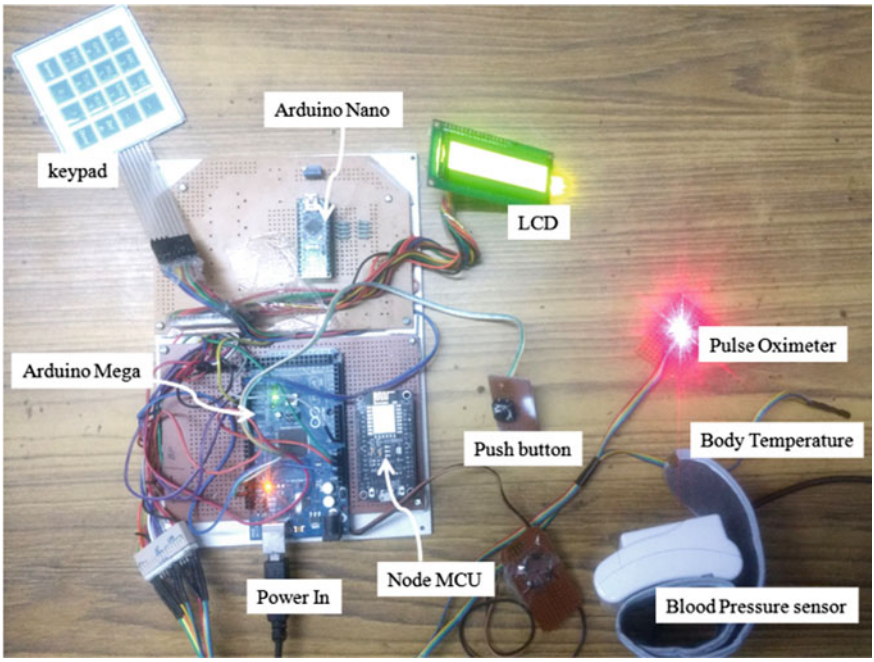


Fig. 5 Internal view of developed prototype

Internet connection for the system. We developed a unique identification-based entry system consisting of a keypad. In this system, a unique API number can be entered to record the medical data. Each person has different API numbers, assigned on the card while making their account in the cloud database. The final hardware is implemented on the printed circuit board. The medical sensors are interfaced with Arduino Mega. The keypad, Wi-Fi module, and LCD display are interfaced with Arduino Nano microcontroller board. Arduino Nano to Arduino Mega microcontroller and Arduino Mega to Wi-Fi module communication are done through serial communication. A push button is added to send the measured parameters to the cloud platform. The whole system is powered by 11000 mAh power bank. The final hardware design of IoT-enabled portable healthcare device is as shown in Figs. 5 and 6.

4.2 Hardware Description

Pulse oximetry sensor: The pulse oximetry sensor uses a noninvasive technique for measuring the oxygen saturation in the blood. The working principle of the pulse oximeter is based on photoplethysmography (PPG). It is an optical technique used to detect the change of blood volume in the microvascular bed of tissue. This sensor displays the oxygen saturation in percentage; it also measures the pulse rate in beats



Fig. 6 Final developed prototype of proposed system

per minute (BPM). In this device, the MAX30100 IC-based pulse oximeter sensor is used [9].

Digital Blood Pressure Sensor: The developed device is interfaced with a digital sphygmomanometer to measure the blood pressure noninvasively using oscillometric methods. The blood pressure sensor gives the range of systolic pressures which is 112–200 mmHg and the range of diastolic pressures is 56–116 mmHg (mean: 81.1 ± 13.4 mmHg). It also displays the pulse rate with blood pressure readings [10].

Body Temperature Sensor: In this research work, LM35 is used for body temperature measurement. It has a wide range of working from −55 to +150 °C. It shows the body temperature value in degree °Celsius (centigrade) [11].

Keypad: A 4 × 4 matrix keypad is used to enter a unique API key. It consists of arrangement of four rows and four columns. The microcontroller detects which button is pressed by the row–column arrangement of membrane switches in keypad [12].

Microcontroller Unit: In this proposed system, two microcontroller boards are used, Arduino Mega board (ATmega2560 microcontroller), and Arduino Nano board (ATmega328P microcontroller). The Arduino Nano is based on ATmega328 or ATmega168 having 14 digital I/O pins and 8 analog input pins. The Arduino Mega

2560 is based on ATmega2560 microcontroller, having 54 digital I/O pins and 16 analog input pins. Both microcontroller boards consist of a USB connection and a power jack [13].

5 Results

We have used the ThingSpeak IoT analytics service that collects and stores sensor data in the cloud and develops IoT applications [14]. First, the unique API was entered using keypad interfaced with the microcontroller and then the input signals from different medical sensors in terms of vital parameters like blood pressure, pulse rate, body temperature, and hemoglobin level were received at the microcontroller for further processing. The microcontroller then sent the received data of a patient to a remote end via the Wi-Fi module. The output is received on the ThingSpeak website.

5.1 Getting Started with Healthcare Device

Step 1: Make an account on "ThingSpeak" cloud platform. Also, create an individual channel for a person and assign the unique API key to a person on the health card as shown in Fig. 7.
Step 2: Use health card to enter assigned API key using the alphanumeric keypad as shown in Fig. 8.
Step 3: Different physiological parameters are measured by using interfaced medical sensors:

Blood Pressure and Pulse Rate Measurement:
After entering the API key, we can wrap the strap of the BP sensor around the cuff. It shows the measured value on the LCD. The sensor displays the blood pressure value in mmHg and also shows the pulse rate value in beats per minute (BPM) as shown in Fig. 9.

Pulse Oxygen Saturation Level Measurement:
After displaying the blood pressure result on the LCD, pulse oximeter starts calculating pulse oxygen level in percentage. To get SpO2 value, place the finger on the LED of the sensor, for one minute. The pulse oximeter shows pulse oxygen saturation level in percentage on the LCD. The pulse oximeter sensors also show the pulse rate on the LCD as shown in Fig. 10.

Body Temperature Measurement:
Measurement of body temperature is done by making physical contact of a body part with temperature sensor. The temperature sensor calculates the temperature value in degree Celsius and displays it on the LCD screen as shown in Fig. 10.

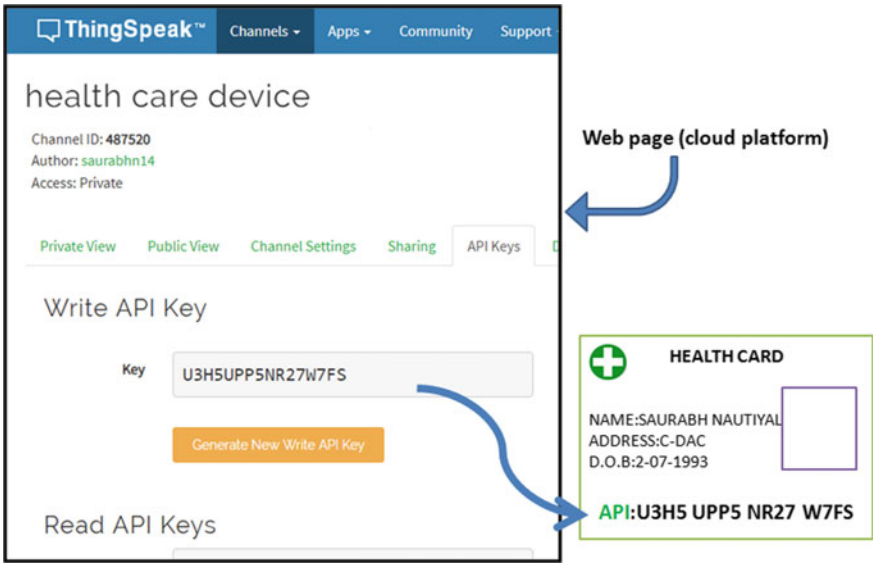


Fig. 7 Generated unique API key for individual person

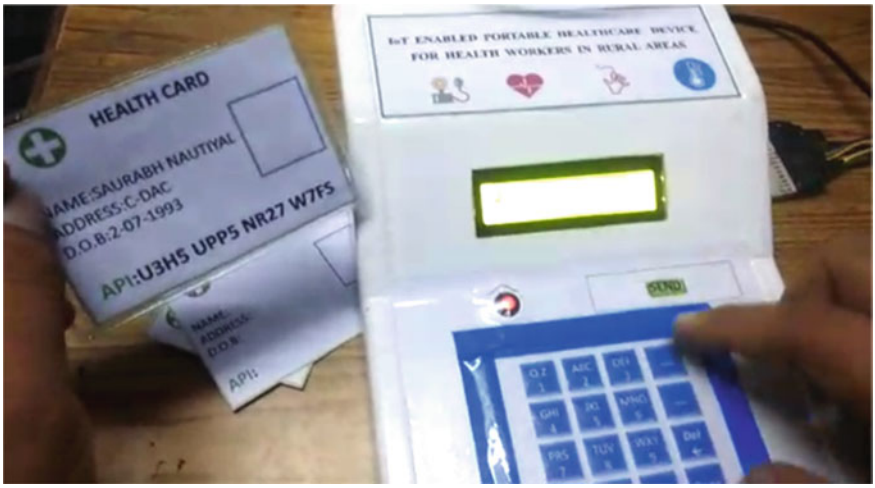


Fig. 8 The individual key entry system

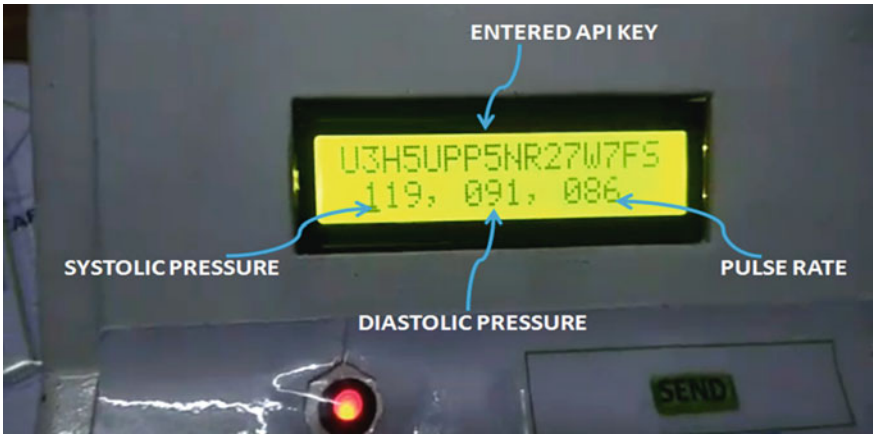


Fig. 9 Blood pressure and pulse rate reading on the device display

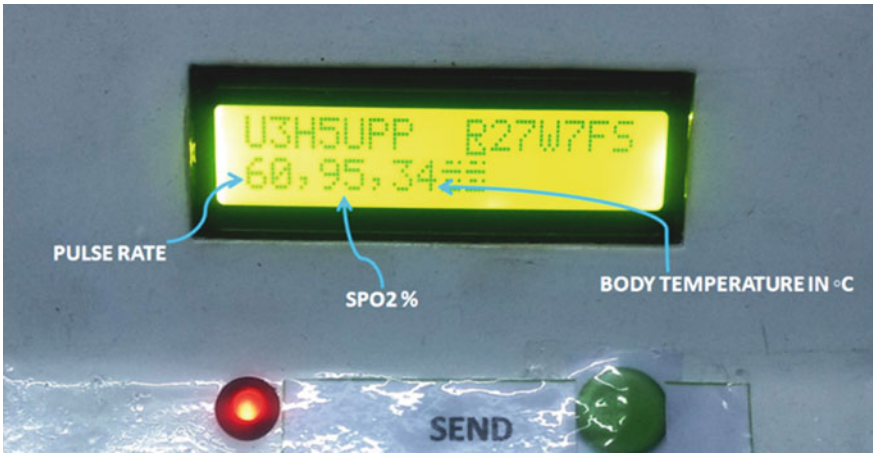


Fig. 10 Pulse oximeter and body temperature reading on the device display

Step 4: After measuring all the parameters, the send button is pressed to post and store the physiological data on the cloud. This data can be updated each time whenever we take readings with the portable device. The complete collected health data on the cloud platform for an individual is as shown in Figs. 11 and 12.

Figures 11 and 12 show the results for the systolic blood pressure and diastolic blood pressure readings received from healthcare device with the specific date and time at the cloud platform. Its measuring unit is mmHg.

Figures 13, 14, and 15 show the data received from the healthcare device in terms of pulse rate in BPM (Beats per minute), pulse oxygen saturation (SpO2)

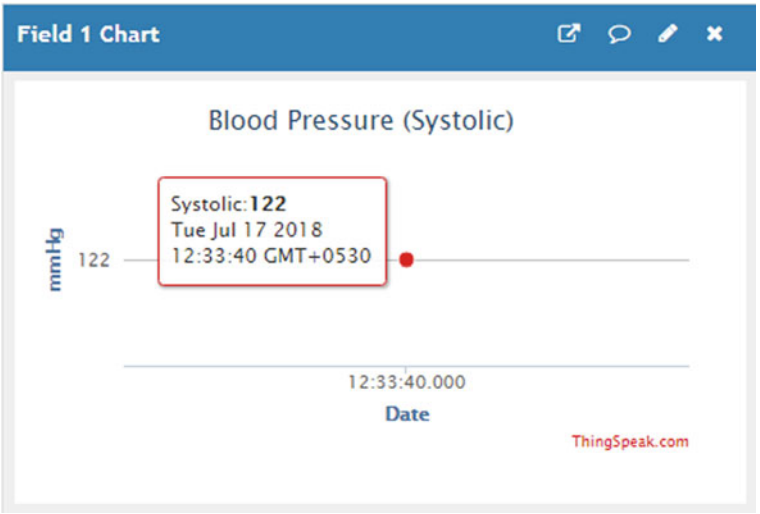


Fig. 11 Systolic blood pressure result

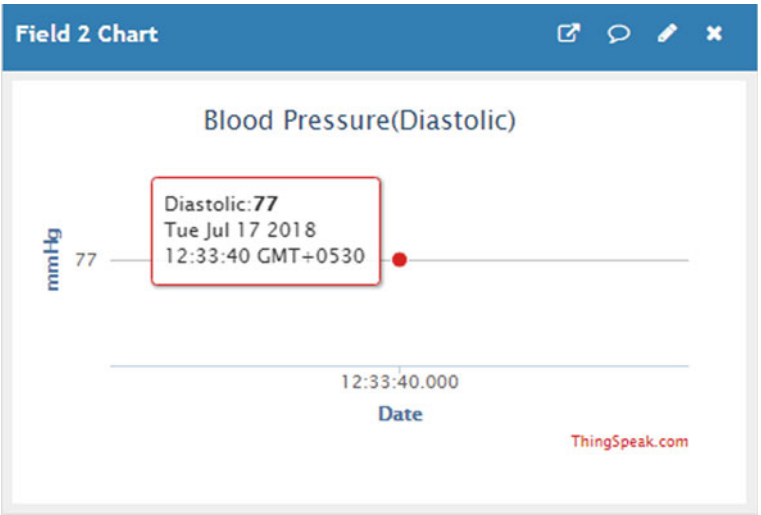


Fig. 12 Diastolic blood pressure result

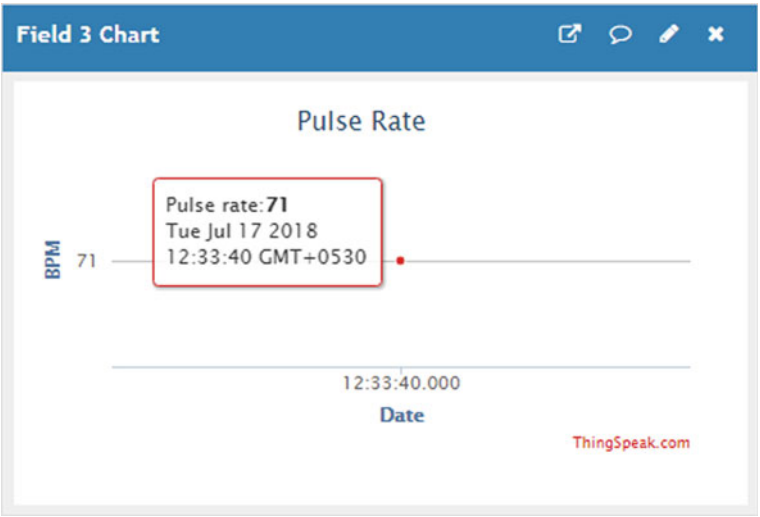


Fig. 13 Pulse rate result

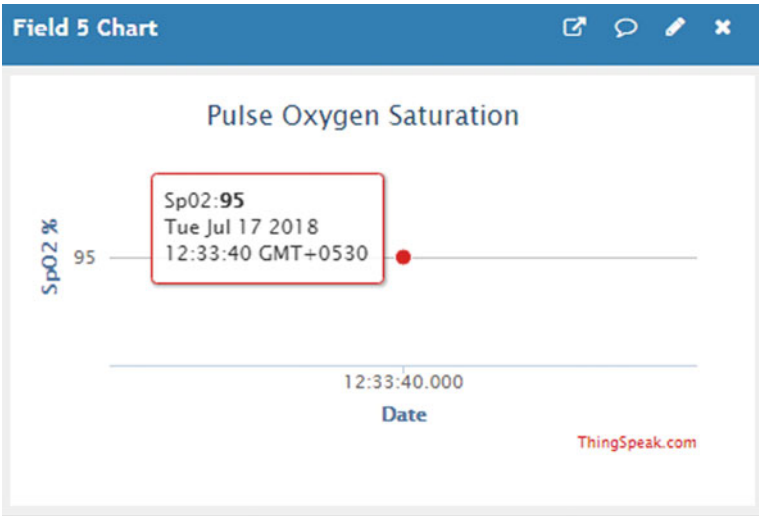


Fig. 14 Pulse oximeter result

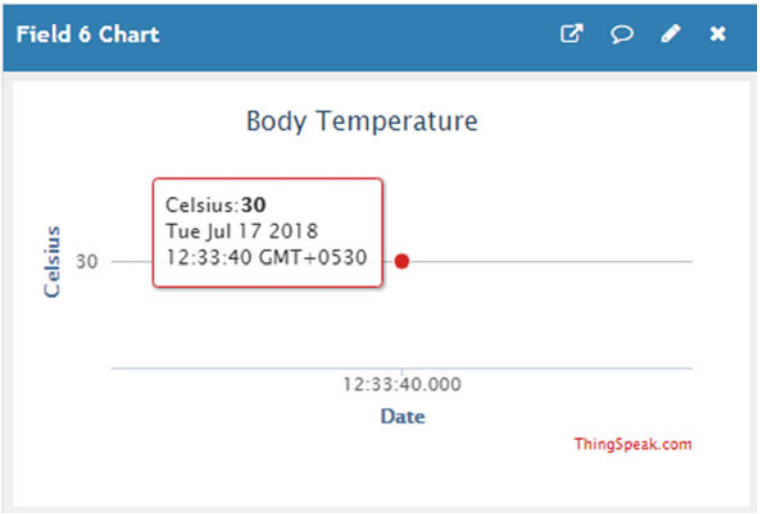


Fig. 15 Body temperature result

level in percentage, and body temperature in degree Celsius, at the cloud platform, respectively.

5.2 System Performance

The system is validated by comparing the measured values using the developed device with the available standard devices. The developed device is tested on three persons, so we have collected the three samples from each standard device as shown in Table 1.

6 Conclusion

This paper majorly focused on the development of an electronic, economic, less complex and portable, healthcare device which is capable of measuring the basic human health parameters and creates a medical database for each person living in the rural areas. Individual's medical data can be shared periodically with the hospital, Government Health Department, and medical personnel anywhere in the world. This device helps in public health monitoring by assisting healthcare workers and also able to create employment opportunities for the people. This system can increase the operational efficiency of the rural health system and deliver high-quality healthcare services to rural inhabitants. The algorithms are implemented such that the

Table 1 Comparison between developed device and standard device

Sample	Health parameter	Developed device reading	Standard device reading	Error %
1	Sys (B.P. in mmHg)	123	119	3.36
	Dia (B.P. in mmHg)	86	88	2.27
	Pulse rate (BPM)	75	73	2.73
	SpO2%	98	97	1.09
	Body temperature (°C)	35	36.6	4.37
2	Sys (B.P. in mmHg)	132	129	2.32
	Dia (B.P. in mmHg)	95	97	2.02
	Pulse rate (BPM)	69	71	2.81
	SpO2%	93	96	3.12
	Body temperature (°C)	38	38.5	1.29
3	Sys (B.P. in mmHg)	122	125	2.4
	Dia (B.P. in mmHg)	77	80	3.75
	Pulse rate (BPM)	71	68	4.41
	SpO2%	95	98	3.06
	Body temperature (°C)	37	36.9	2.73

data acquired will be processed, recorded, and transmitted to the cloud platform. The readings measured by this device are almost coordinating with commercially available equipment in the hospitals. The designed portable device has less complex functioning and ergonomic design, which makes it convenient to use by health workers.

Acknowledgements The authors would like to thank ACS Division C-DAC Mohali for providing the support and facilities. All procedures performed in studies involving human participants were in accordance with the ethical standards. The informed consent was obtained from all individual participants included in the study.

References

1. B. Buehler, R. Ruggiero, K. Mehta, Empowering health workers with technology solutions. IEEE Technol. Soc. Maga. 44–52 (2013)
2. G. Vaani, Rural healthcare: towards a healthy rural India. Ministry of Health and Family Welfare, Government of India, pp. 1–4, July 2015
3. N. Akshay, G. Vamsi Krishna, Design and implementation of real-time bio-signal acquisition system for quality health care services for the population of rural India, in *IEEE International Conference On Recent Trends In Electronics Information Communication Technology*, India, pp. 1315–1319, 20–21 May 2016
4. A.Z. Stanley, L. Zalzala, A. Karimi, Intelligent technologies for self-sustaining RFID-based rural E-health system. IEEE Technol. Soc. Mag. **32**, 36–43 (2013)

5. P. Gupta, D. Agrawal, J. Chhabra, P.K. Dhir, IoT based smart healthcare kit, in *2016 International Conference on Computational Techniques in Information and Communication Technologies (ICCTICT)*, March 2016, pp. 237–242
6. P. Gope, T. Hwang, BSN-care: a secure IoT-based modern healthcare system using body sensor network. IEEE Sens. J. **16** (2016)
7. India pledges to bring high-speed affordable internet access to 150,000 villages, https://mashable.com/2017/02/01/india-150-000-villages-low-tariff-internet/. 01 Feb 2017. Last accessed 6 May 2018
8. Rural Health Survey 2017, Government of India Ministry of Health and Family Welfare Statistics Division, Rural Health Statistics, March 2017
9. U. Timm, E. Lewis, D. McGrath, J. Kraitl, H. Ewald, LED based sensor system for non-invasive measurement of the hemoglobin concentration in human blood, in *13th International Conference on Biomedical Engineering. IFMBE Proceedings*, vol. 23 (Springer, Berlin 2009), pp. 825–828
10. Blood pressure sensor serial output, https://www.sunrom.com/p/blood-pressure-sensor-serial-output. Last accessed 7 June 2018
11. LM35, http://www.ti.com/product/LM35. Last accessed 9 June 2018
12. Interface 4 * 4 matrix keypad with microcontroller, https://embedjournal.com/interface-4x4-matrix-keypad-with-microcontroller/. Last accessed 9 June 2018
13. Arduino home page, https://www.arduino.cc/. Last accessed 21 June 2018
14. IoT Analytics-ThingSpeak Internet of Things, https://www.thingspeak.com/. Last accessed 16 July 2018

Head-Gesture-Based Human–Computer Interface for Disabled People



Sumit Malik, Nitin Kumar and Anuj Jain

1 Introduction

The use of computers has risen very fast in the last few decades. Since a long time, computer developers have not given enough consideration to computer ease of use. Many companies and research foundations are currently working on ease of use in computers [1, 2]. The system purposed in this paper will allow people with disabilities that lead to complete loss of control over voluntary muscles to control the mouse in computer. So, they can do common tasks on a computer.

Krolak and Strumillo present the detection of voluntary eye flicker in their vision-based framework. The proposed method takes into consideration eyeblink detection, estimation of the eyeblink span and understanding of a series of blinks progressively to control a HCI. The recognized eyeblinks are assigned short blinks (shorter than 200 ms) or long blinks (longer than 200 ms). Different short eyeblinks are thought to be unconstrained and were excluded from planned eyeblink code [3]. Divjak and Bischof [4] presented a system to identify fatigue. In this technique, electrodes are used to detect drowsiness. Danisman et al. [5] presented the thought behind this procedure to use electroencephalogram (EEG) signals which show the brain electrical movement and can be recorded as changes in electrical voltages on the scalp. In EEG signal electrooculogram (EOG) which is created because of eye blinking/eye function is detected. To filter through EOG from EEG signals, H∞ adaptive filter is used. Pan et al. [6] presented a constant liveness testing approach against photograph

S. Malik (✉) · N. Kumar · A. Jain
Lovely Professional University, Phagwara 144411, Punjab, India
e-mail: sumit1513151@gmail.com

N. Kumar
e-mail: nitin.14652@lpu.co.in

A. Jain
e-mail: anuj.22631@lpu.co.in

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_149

spoofing in a nonintrusive way for face recognition, which does not require any extra equipment aside from a common web camera.

Kumar et al. [7] presented a simple method for pointing and choice using a blend of eye gaze and keyboard triggers. The framework EyePoint uses a two-step process smoothly merged together in a look-press and look-release action, which makes it conceivable to make up for the precision confinements of the current state of the eye gaze trackers. The framework is also powerful and sufficiently straightforward for normal users to use for their regular computing activities. Singh and Prasad [8] showed multi-modalities for substance passage through the virtual console which is more viable and less strenuous than single methodology framework and furthermore tackles Midas touch issue, which innate in eye-tracker based HCI framework where just stay time is required for choosing character.

This work executed a strong outright virtual head mouse dependent on interpretation of head movements and face gestures caught by web camera [9]. Vertegaal [10] presents Fitts' Law assessment for various eye-tracking and user input gadgets in choice for substantial visual success. Kraichan and Pumrin [11] show a novel user interface that eye and face recognition through Haar-like component with bilateral total variation based super-resolution to resize the low-resolution eye picture by a web camera. The center point for two eyes recognition is mapped to mouse coordinate on a computer display. Left-click and right-click are performed with blinking left eye and right eye, separately. Mackenzie and Ashtiani [12] exhibited another text entry framework, BlinkWrite, which gives a communication gateway to cognitively capable motor-impaired people who cannot use the eye-tracking framework. As opposed to most hand-free frameworks, BlinkWrite enables content to input and rectified using single input method: blink.

Arai and Mardiyanto [13] purposed eye-based HCI by full detail of mouse and keyboard with human eyes of disabled people and for the gadget with portable processing. This uses eye information for gaze estimation detection. As compared with conventional eye-mouse, gaze-mouse doesn't provide robustness to various different kind's of people. Grauman et al. [14] presented two video-based HCI instruments which enact paired switch and run selected command. First, identify user's eye blinks and precisely measure value and then second device distinguishes when user raises his/her eyebrows and after that perform mouse action. The two frameworks introduce their track eye at frame rate and recuperate for the case of error. In face analysis, an important role has been played by automatic facial point detection. For constrained and unconstrained conditions, a few methods have been proposed, but the accuracy of fiducial points is the major problem. It becomes difficult to compare the performance of different frameworks due to the lack of evaluation protocol. Sagonas et al. presented a standardized benchmark for facial landmark localization using the same evaluation protocol to compare the performance of different frameworks on a new-collected dataset [15, 16].

The rest of the paper is organized as follows: In Sect. 2, research methodology has been described. Section 3 contains the equipment, material, and experimental setup of the system. Section 4 contains results and discussion of the user tests. In Sect. 5, we draw conclusions and point out future work.

2 Research Methodology

In this framework, a simple and easy approach, a web camera is used with proper lighting in the sensing and digitizing tasks for viewing the images. Further with a computer processor, Viola–Jones algorithm is applied to these images for bringing out the facial landmarks. On the basis of the features, the final application is made working. Figure 1 shows the basic block diagram of the framework.

The algorithm followed in this experiment is as follows:

(a) The facial landmarks of a given face have been predicted. Facial landmarks are used to localize and represent salient regions of the face, i.e., nose, eyes, eyebrows, mouth, and jawline.

(b) Using these predicted landmarks of the face, features are detected. The two kinds of features extracted are eye aspect ratio (EAR) and mouth aspect ratio (MAR), using EAR to detect eyeblink/wink and MAR to detect yawn (opening mouth).

c) The *eye aspect ratio* (EAR) was introduced by Soukupová and Čech [17]. Six (x, y)-coordinates represent each eye, starting from left corner of eye and then in working clockwise around the rest of the region (Fig. 2).
A relation is present between height and the width of these coordinates. The paper by Soukupová and Čech derives a equation which reflects relation between

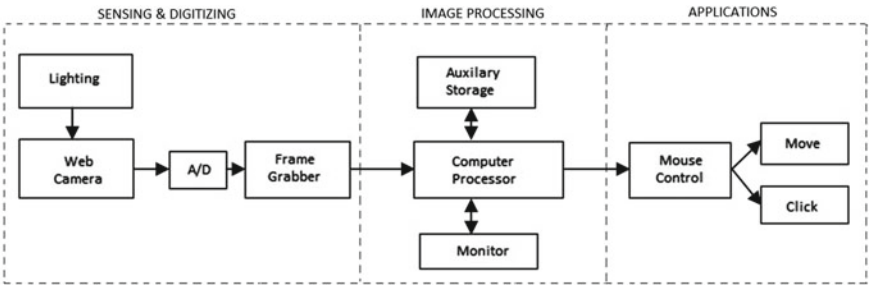


Fig. 1 The block diagram of framework

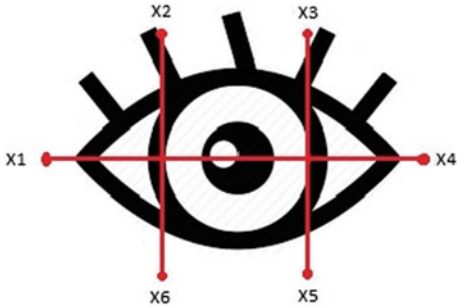


Fig. 2 The six facial landmarks associated with the eye

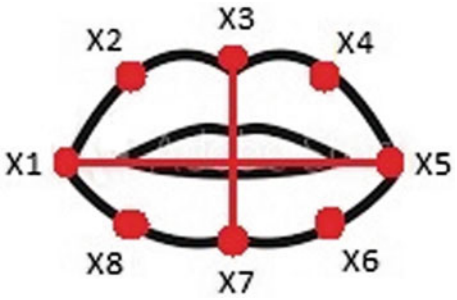


Fig. 3 The eight facial landmarks associated with the mouth

different coordinates known as eye aspect ratio (EAR):

$$\mathbf{EAR} = \frac{||x2 - x6|| + ||x3 - x5||}{2||x1 - x4||} \tag{1}$$

where $x1$, $x2$, …, $x6$ are 2D facial landmark locations.
Similar to EAR, MAR value goes up when the mouth opens. Similar intuitions hold true for this metric as well (Fig. 3).

$$\mathbf{MAR} = \frac{||x2 - x8|| + ||x3 - x7|| + ||x4 - x6||}{2||x1 - x5||} \tag{2}$$

where $x1$, $x2$, …, $x8$ are 2D facial landmark locations.

Further, the mouse cursor is controlled with the help of PyAutoGUI library. The objective of PyAutoGUI is to provide a multi-platform Python module for GUI automation for users. This can simulate actions such as move the cursor, click action, drag operation, keys pressing, press and hold action, and press a keyboard hotkey combination.

3 Equipment, Material, and Experimental Setup

Figure 4 demonstrates the format of the experimental setup of the proposed framework. The framework comprises the following significant parts:

I. A simple webcam,
II. A laptop, and
III. A light source (CFL).

The light sources were put in front around 2 m from the individual. The light is projected at an angle 45° to ground plane. Table 1 shows the detailed particulars of laptop and camera used. The individual sits before the camera at a distance of about 0.6 meters. The face is consistently captured by the camera and process in

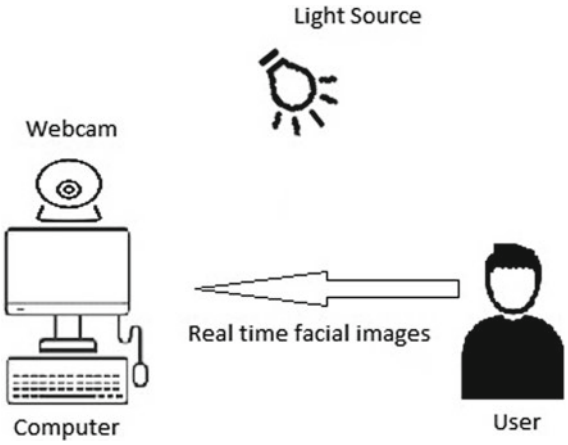


Fig. 4 Experimental setup for real-time recording

Table 1 Different gestures for mouse functions

Function	Gesture
Left/right-click	Left/right eyeblink/wink
Move up/down	Head up/down
Move right/left	Head right/left
Gesture system ON/OFF	Open mouth

Python. The procedure at that point detects and groups the eyeblinks. The framework additionally shows the live facial pictures which indicate recognition of eye pair and mouth. Figure 5 demonstrates a frame capture of the real-time video of the test performed under controlled lighting conditions.

Quantum QHM495LM PC Camera is used for input from the user as it has a frame rate of 30 frames per second that meets well with necessities of our examination work. Additionally, it has a pivoting focal point that will assist the user with adjusting it as indicated by his ease and get the best outcomes. The code is written in Python programming. The below-mentioned modules of Python programming were imported for successful running of the code:

1. Numpy 1.15.1,
2. Imutils 0.5.1,
3. OpenCV-3.4.2,
4. Dlib 19.7.0, and
5. PyautoGUI 0.9.38.

To work the equipment effectively and reliably, it must be set up in the following way:

(Please take note of that the following directions are noted down for Windows 10 and Python 3.6).

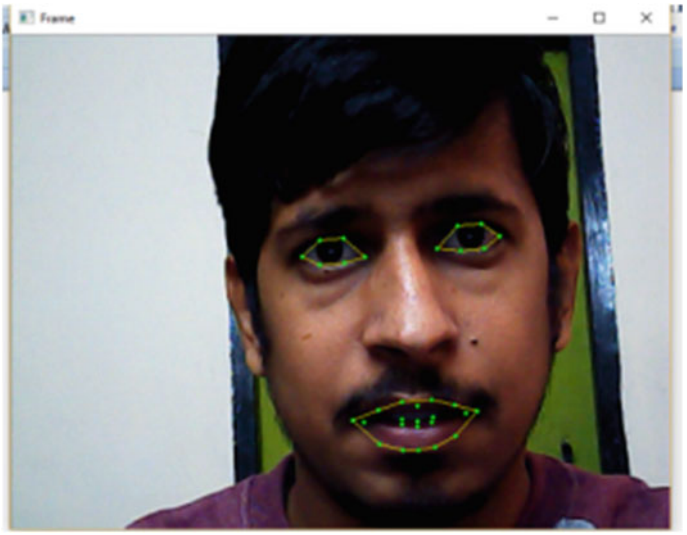


Fig. 5 Screenshot for recorded framework

I. Install the required dependencies by means of pip install command (NumPy, SciPy, OpenCv, Matplotlib).
II. Have the program source code and the scripts in separate directory on the user's PC.
III. Connect the web camera to the USB port of the subject's PC.
IV. Wait before launching the eye-tracking application as camera setup requires some time. Camera status can be checked by "Device Manager".
V. Launch the Python application (Table 2).

Table 2 Specifications of the hardware and software

Particular	Specification
Computer	Intel® Core i5-4210U CPU @ 1.70–2.40 GHz, RAM 4.0 GB 64-bit operating system Windows 10
Camera	Quantum QHM495LM PC Camera Image resolution 25MP, 640 × 480
Python	Python 3.6.4, 64-bit Computer vision library
Lighting condition	Natural lighting condition Controlled lighting condition
Light source	8 W, 400 lm CFL lamp

4 Results and Discussion

Basically, an experiment was performed to know the distance of user from camera that provides maximum gestures—accuracy. The distance was set as 20 cm to 100 cm (1 m). The variation of gesture accuracy w. r. t. distance is shown in Fig. 6. The value for which system gives maximum gesture—accuracy is optimum value which is 0.6 m.

The performance of the proposed system was checked by performing experiment in which five users, who can perform all of gesture, participated voluntarily. The members were advised about the undertaking to be performed and were approached to perform 20 gestures at random. The trial was performed for two lighting conditions:

1. Controlled lighting condition (Fig. 6).
2. Natural lighting condition (Fig. 7).

A proper light source is needed to capture object by the camera. In regular condition, the light fall upon the object depends absolutely on ecological lighting condition. In controlled conditions, a CFL is utilized for extend light on the user's face to make the lighting condition uniform and consistent. The screen gives a specific measure of light on the user's face in any condition. The accompanying parameters were considered to assessment of the presented methodology to gesture detection when recorded continuously:

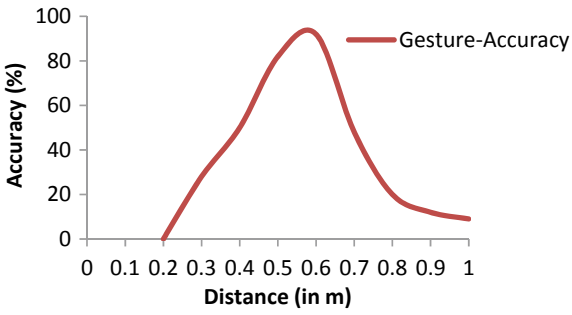


Fig. 6 Gesture accuracy w.r.t distance from the camera under controlled lighting

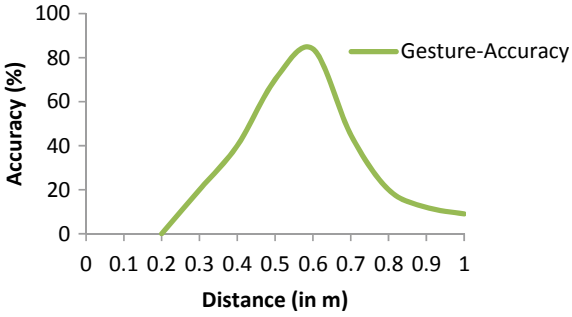


Fig. 7 Gesture accuracy w.r.t distance from camera under natural lighting

(a) Gesture accuracy (GA),
(b) False alarm rate (FAR), and
(c) Overall success rate (OSR).

These parameters are characterized as below:
Gesture accuracy (GA)

$$\mathrm{GA} = \frac{\mathrm{PT}}{\mathrm{PT} + \mathrm{NF}} \tag{3}$$

False alarm rate (FAR)

$$\mathrm{FAR} = \frac{\mathrm{PF}}{\mathrm{PT} + \mathrm{PF}} \tag{4}$$

Overall success rate (OSR)

$$\mathrm{OSR} = \frac{\mathrm{GA}}{\mathrm{GA} + \mathrm{FAR}} \tag{5}$$

where PT is defined as positive true that correctly detects gestures, PF is the positive false that gesture reported as correct gesture but not, NF is the negative false that gesture not detects by the algorithm, and NT is the negative true gestures that are correctly reported as any gesture.

Under controlled light condition and natural light condition, the experiment was conducted. For controlled light condition, gesture accuracy got 92% for performing mouth gestures. Though, in natural light condition, gesture accuracy got 84% for performing, the overall success rate acquired is 88% for performing mouth gesture. The false alert rate is seen to be 0.

The presented framework provides good execution under controlled light condition when contrasted with natural light condition in performing gesture. Unfortunately, natural light condition results are to a great degree poor and then the framework loses the efficiency of gesture accuracy.

5 Summary and Conclusions

A basic and economic real-time gesture detection approach has been presented. This work is a result of gesture-based HCI framework, in which gestures are utilized to perform different mouse tasks, i.e., move, left-click, and right-click. Thus, system detects three types of gestures: mouth gesture, head move left/right/up/down, and left and right wink. Under controlled light condition and natural light condition, the experiment was conducted.

In controlled light condition, the recognition precision of the presented strategy was obtained 92% for performing mouth gesture. The proposed strategy provides an

overall success rate of 88% for performing eye mouth gesture. Though, in natural light condition, the gesture accuracy got 84% for performing mouth gesture, the overall execution of presented technique is comparatively good under controlled light when contrasted with natural light condition.

The gestures are ended up being extremely helpful in performing different mouse activities under controlled light condition and when user sits at a separation about 0.6 m. The precision of action of moving and enacting mouse control task found high. The real worry of this framework is its poor execution while identifying eye winks which should be tended to in future.

Ethical Standards All procedures performed in studies involving human participants (Sumit Malik, Nitin Kumar, and Dr. Anuj Jain) were in accordance with the ethical standards of international conference on intelligent computing and smart communication, ICSC-2009.

References

1. W. Wetzlinger, A. Auinger, M. Dörflinger, Comparing effectiveness, efficiency, ease of use, usability and user experience when using tablets and laptops, *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence Notes Bioinformatics)*, vol. 8517, PART 1 (2014), pp. 402–412
2. M. Igbaria, J. Iivari, H. Maragahh, Why do individuals use computer technology? A finnish case study. Inf. Manag. **29**(5), 227–238 (1995)
3. A. Krolak, P. Strumillo, Eye-blink detection system for human–computer interaction. Univers Access Inf. Soc. **11**, 409–419 (2012)
4. M. Divjak, H. Bischof, Eye blink based fatigue detection for prevention of computer vision syndrome, in *IAPR Conference on Machine Vision Applications* (2009), pp. 350–353
5. T. Danisman, I.M. Bilasco, C. Djeraba, N. Ihaddadene, Drowsy driver detection system using eye blink patterns, in *International Conference on Machine and Web Intelligence* (2010), pp. 230–233
6. G. Pan, L. Sun, Z. Wu, Eyeblink-based anti-spoofing in face recognition from a generic web-camera, in *11th IEEE International Conference on Computer Vision (ICCV'07)* (2007)
7. M. Kumar, A. Paepcke, T. Winograd, EyePoint?: practical pointing and selection using gaze and keyboard, in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'07)* (2007), pp. 421–430
8. J.V. Singh, G. Prasad, Enhancing an eye-tracker based human–computer interface with multi-modal accessibility applied for text entry. Int. J. Comput. Appl. **130**, 16–22 (2015)
9. T.P.A. Guillamet, M.T.M. Teixidó, A.F. Viso, C.R.J. Palacín, Implementation of a robust absolute virtual head mouse combining face detection, template matching and optical flow algorithms. Telecommun. Syst. **52**, 1479–1489 (2013)
10. R. Vertegaal, A Fitts' law comparison of eye tracking and manual input in the selection of visual targets, in *The 10th International Conference on Multimodal Interfaces (ICMI'08)* (2008), pp. 241–248
11. C. Kraichan, S. Pumrin, Face and eye tracking for controlling computer functions, in *2014 11th International Conference on Electrical Engineering/Electronics, Computers, Telecommunication and Information Technology* (2014)
12. I.S. Mackenzie, B. Ashtiani, BlinkWrite?: efficient text entry using eye blinks. Univers Access Inf. Soc. **10**, 69–80 (2011)
13. K. Arai, R. Mardiyanto, Eye-based HCI with full specification of mouse and keyboard using pupil knowledge in the gaze estimation, in *Proceedings—2011 8th International Conference on Information Technology: New Generations (ITNG 2011)* (2010), pp. 423–428

14. K. Grauman, M. Betke, J. Lombardi et al., Communication via eye blinks and eyebrow raises?: video-based human–computer interfaces. Univers Access Inf. Soc. **2**, 359–373 (2003)
15. P. Viola, M. Jones, Rapid object detection using a boosted cascade of simple features, in *Proceedings of IEEE International Conference on Computer Vision & Pattern Recognition (CVPR)* (IEEE, 2001), pp. 511–518
16. C. Sagonas, G. Tzimiropoulos, S. Zafeiriou, M. Pantic, 300 faces in-the-wild challenge: the first facial landmark localization challenge, in *Proceedings of the 2013 IEEE International Conference on Computer Vision Workshops*, 2–8 Dec 2013, pp. 397–403
17. T. Soukupová, J. Cech, *Real-Time Eye Blink Detection Using Facial Landmarks* (2016)

Intelligent and Hybrid Control Techniques for Robotic Manipulator



Mandeep Singh and M. K. Shukla

1 Introduction

For decades, a number of research papers on robotic manipulator control techniques have been written. Many of these control techniques depend on a dynamic robot model and are referred to as a model control system. The control algorithms available include adaptive control [1], PID control [2], sliding mode control [3], and robust control mode [4].

However, a human operator can do the same task without knowing the mathematical model. Therefore, there is a need for other methods of control that mimic the behavior of a human and do not require large conventional comparison control calculations. This resulted in the study of intelligent systems that use the latest improvements in the field of artificial intelligence [5].

The advent of soft computing technologies or computer intelligence based on biological and human intelligence is one of the most exciting and important areas of engineering. This technology is generally expected to play a significant role in the growth of intelligent robotic systems, machine systems, and mechatronic systems [6]. Intelligent control techniques of robotic manipulator involve computations related to artificial neural network (ANN) [7], fuzzy logic technique (FL) [8], and genetic algorithm (GA) [9].

In addition, some control problems cannot be solved by using an intelligent (ANN, FL, and GA) control techniques. This limitation is the main reason for the creation of hybrid control techniques. In a hybrid control technique, two or more intelligent control techniques are blended to overcome the limitations of each individual technique

M. Singh · M. K. Shukla (✉)
School of Electronics and Electrical Engineering, Lovely Professional University, Phagwara 144411, Punjab, India
e-mail: manoj.22223@lpu.co.in

M. Singh
e-mail: msdhanjalece26@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_150

[10]. The hybrid control techniques (neuro-fuzzy, neuro-genetic, and neuro-fuzzy-genetic) have been discussed in this paper

The paper is arranged as follows: Sect. 2 gives an introduction about different intelligent control techniques of robotic manipulator. Section 3 presents the hybrid control techniques, and finally the paper is concluded in Sect. 3.

2 Intelligent Control Techniques

Intelligent control refers to a control system with behavior and intelligence similar to that of human beings, such as parallel processing of information, associative memory, environmental adaptation, learning, generalization, and self-organization. It represents an integration of modern control theory, intelligent system engineering, and information system engineering [11].

The basic problem in robot control is the desired trajectory. In robotics literature, this control problem is known as motion control problem or trajectory problem. Traditional robotic control methods rely heavily on precise mathematical modeling, analysis, and synthesis. These techniques are useful for robot control which works in a structured environment.

However, in order to work in unstructured environments, robots need to do more complex work without adequate analytical model. The most challenging problem is the uncertainties in the environment. These uncertainties are mainly due to uncertain prediction about sensor defects and environmental characteristics and its mobility [12].

As a result of this, the development of intelligent and hybrid control techniques for robotic manipulator technologies is based on biological processes, such as evolutionary techniques based on the tools of FL, NN, and GA, artificial intelligence, and field of expert systems design.

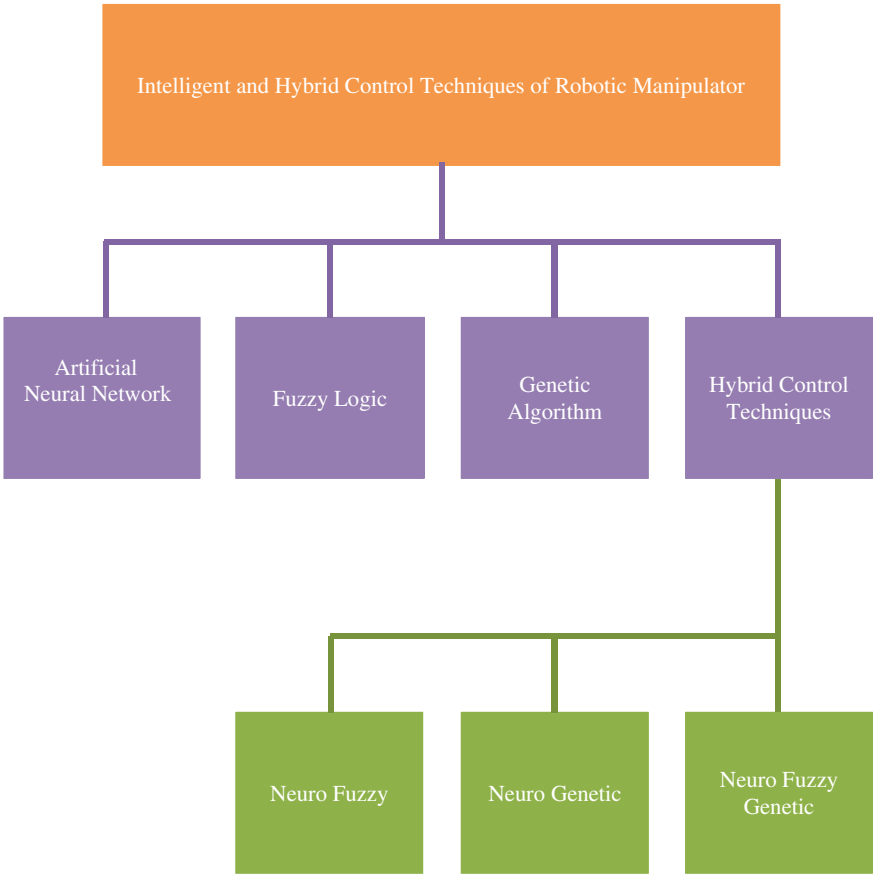


Fig. 1 Types of intelligent control techniques of robotic manipulator

A classification of different techniques is depicted in Fig. 1. This section focuses on the intelligent control methods in the field of robotic manipulator control, for example, artificial neural network method, fuzzy logic method, and genetic control method.

2.1 Artificial Neural Network (ANN)

The ANN is a learning algorithm that stimulates the process of learning in the human beings. In ANN, there are basically three layers: input, hidden, and the output layers. The goal of the neural network application is to provide a set of parameters that reflect correspondence between user input and input sent to the manipulator. There are two types of ANN training methods for manipulation control, namely, online

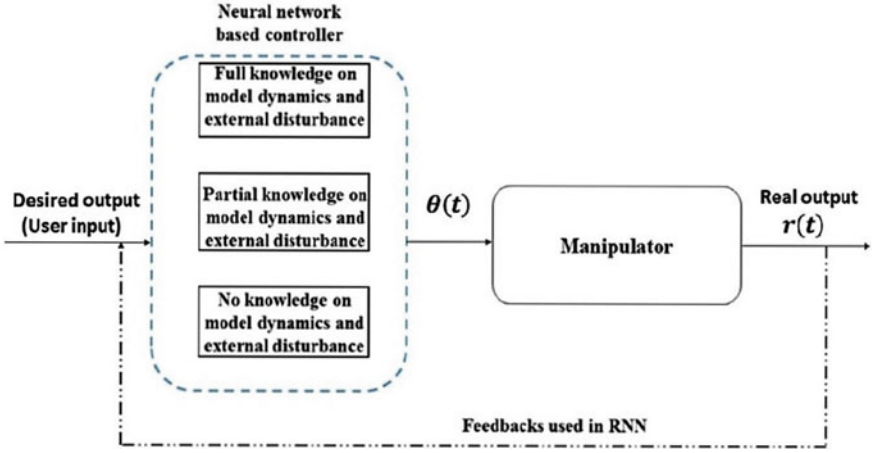


Fig. 2 ANN-based controller [13]

training and offline training problems. These two methods can be applied gradually to a particular task, depending on their performance [13]. A neural-network-based natural workflow for controlling a robotic manipulator is given in Fig. 2.

The n-link robot manipulator tracking control with flexible joints is discussed in [14]. In order to address system uncertainties and improve system robustness, a full-state feedback NN controller is proposed. The control of redundant neural-network-based manipulators is discussed. There are two limitations identified in existing recurrent neural network solutions, and amended models to overcome these limitations are established [15].

2.2 Fuzzy Logic (FL)

In comparison to humans, robotic machines are not capable of solving problems with imprecise information in which restriction of the plant model and the control (for example, linearity) is required. FL and fuzzy sets were introduced in order to emulate this ability. These rules reflect a person's knowledge. Fuzzy controllers are adaptive and nonlinear in nature, ensuring reliable operation depending on changes in parameters and the effect of load disturbance [16]. A fuzzy logic-based closed-loop control architecture for robotic manipulator is shown in Fig. 3.

A two-degree-of-freedom robot was controlled by FL and particle swarm optimization (PSO) by [18]. For a given path, the parameters of the control of FL are optimized with PSO with three different cost functions. In order to compare FL control with other controllers, the PID controller is adjusted with PSO optimization.

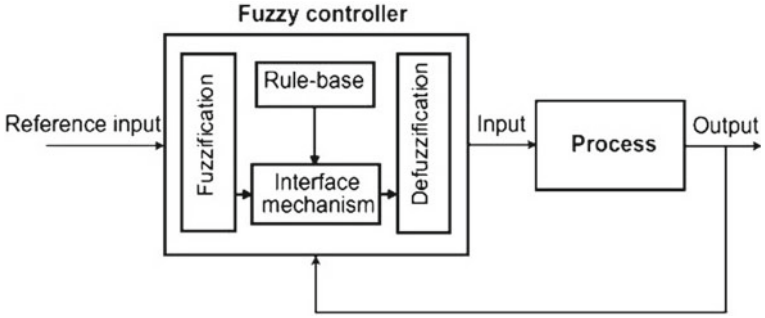


Fig. 3 Fuzzy logic controller [17]

The literature on controlling a robot manipulator with two DOFs using FLC is discussed in [19]. An overview shows that traditional control is inefficient and reliable for controlling a robot arm with different uncertainties, such as model uncertainty, variable payload, and parameter variation. The literature has shown that the use of fuzzy logic control with classical control methods increases the reliability and efficiency of the system.

The design and control of the flexible-joint single-link robot manipulator are provided by [20]. To achieve rapid trajectory tracking performance and to remove link vibrations, a cascade fuzzy logic controller has been used. The logic structure includes three different controllers in the cascade.

2.3 Genetic Algorithm (GA)

GA is the research method that is based on mechanisms of natural selection and natural genetics. They form a class of heuristic methods and computer models for adaptation and evolution. GA mimics and develops the genetic dynamics underlying natural evolution to find the best possible solutions to common optimization problems. The genetic algorithm can be considered as a stochastic search algorithm for optimizing nonlinear functions using natural genetics and mechanism of natural selection. It combines the fittest artificial survival with nature-abstracted genetic operators to provide a surprisingly robust search mechanism for an extensive range of search problems. GA is global and population-based research techniques. In general, the performance of analyzes based on genetic algorithms is superior to that of conventional techniques [17].

A genetic algorithm has been proposed by [21] to optimize point-to-point trajectory for a three-link robotic arm. The aim of the proposed genetic algorithm is to reduce travel time and travel space without exceeding a predefined maximum torque without an obstacle in the working area of the robot.

A new robot manipulator trajectory planning method is proposed by [22] whose workplace contains many obstacles. A GA is established for the generation of robot path in order to find valid and best solutions in the workspace. A new way to plan the trajectory of a mobile robot is proposed by [23] on the basis of a new GA. The main characteristic of the method is its chromosome variable length. The search for a robot environment involves obstacles and goals. The location of the robot, its destination, and the obstacle characterize the workplace. The proposed system can produce collision-free routes in static and dynamic environments.

3 Hybrid Control Techniques

The hybrid control technique has proven its worth in designing the control system for robot manipulator. ANN, FL, and GA are the complementary methodologies for the design of intelligent robotic manipulator control systems. In last few years, researchers have suggested integrations of these methodologies to take benefit of the merits and eliminate their drawbacks. These techniques include the combination of ANN and FL and the integration of these two techniques with genetic algorithm. The combination of ANN, FL, and GA can be realized in three different ways and these are as follows.

3.1 Neuro-fuzzy

A fuzzy-neural system is a system that combines the features of ANN and FL system. The combination of ANN control and FL control can solve many complex nonlinear problems. The hybrid combination of ANN and FL method was used for 3-DOF robotic manipulator in [24].

The neuro-fuzzy controller is introduced by [25] to control the robotic arm position. A five-layer neural network adjusts fuzzy logic input–output parameters. To train this network, a hybrid learning method is used. A new neuro-fuzzy method consisting of an FLC and a recurring neural network is suggested by [26]. The system was implemented to calculate reverse cinematic transformations in a robot manipulator with two links. It was shown to combine the benefit of both disciplines with the integration of fuzzy controllers and neural networks.

3.2 Neuro-genetic

A neuro-genetic system is that combines the features of AAN and GA. In the control of robotic manipulator, the solution to the problem of inverse kinematics is fundamental. In recent years, due to ANN, an inverse kinematics problem in robotic

manipulator has received a lot of attention. However, the accuracy of the results requires improvements for some sensitive tasks. In [27–29], ANN and GA are used together to solve a robotic manipulator's inverse kinematics problem.

3.3 Neuro-fuzzy-genetic

A neuro-fuzzy-genetic system is a system that combines the features of ANN, FL, and GA. A neuro-fuzzy control optimized by GAs is presented for a robot manipulator by [30]. For this purpose, the robot was modeled. A regulator which combines FL and ANN in the form of a multilayer network is used to automatically extract rules where human expertise is required. A straight-line fuzzy subset, such as interpolation functions, is used to avoid the problem of positions and fuzzy subsets forms. The most appropriate form of the fuzzy subsets was found to optimize these functions and minimize a performance criterion. A new form of flexible-joint rigid-link robotic controller design based on ANN, FL, and GA is suggested by [31].

The path planning technique of mobile robot using an ANN, FL, and GA is introduced by [32]. The ANN is trained to select a path from n-paths. FL is used to avoid collision when all n-paths are blocked by obstructions. GA are used as optimizers for selecting optimal locations in obstacle-free directions and places by selecting fuzzy rules from fuzzy rule base of the FL system.

4 Comparison

There are three different control techniques to compare traditional, intelligent, and hybrid control. Each of these techniques has its advantages and disadvantages. Therefore, it is suitable to hybrid these three techniques to remove one's weakness (Table 1).

5 Conclusion

Intelligent and hybrid control techniques are preferred over conventional methods for controlling robot manipulator. Traditional control techniques depend on robotic dynamic model. Therefore, there are other methods that are necessary for control that emulates the human operator's behavior and even does not require many traditional control calculations. In this paper, the introduction of intelligent and hybrid control techniques has been presented, and a brief review of the literature has been provided. The essence is that intelligent and hybrid control techniques are very useful for controlling robot manipulation and it will have the ability to cope with robotic manipulator complex mathematical equations and dynamic models.

Table 1 Comparison of traditional, intelligent, and hybrid control techniques

Control techniques	Controller	Advantages	Disadvantages
Intelligent	ANN	1. It is a fast method for the creation of a controller without the knowledge of system dynamics 2. A powerful tool for ill-defined systems 3. A neural network learns itself and there is no need for the programming 4. Able to manage a large number of data and input variables	1. Needs supervised and offline training of the weights 2. Learning can be slow 3. Require high processing time if a neural network is large
	FL	1. It is similar to human logic 2. The exact mathematical model of the physical system is not required 3. Easy to design and apply	1. Difficult to choose control parameters due to lack of knowledge about the system 2. Not able to receive feedback for implementation of learning strategy 3. Unable to determine the numbers of fuzzy rules and the membership function parameters in advance
	GA	1. It can work without user input 2. Same encoding, change the fitness function 3. Generally, it is very good for optimizing problems 4. Determine unknown solutions	1. Genetic algorithms are complicated 2. Computational time is high 3. Slower than some other methods 4. Networking will increase complexity

(continued)

Table 1 (continued)

Control techniques	Controller	Advantages	Disadvantages
Hybrid control	Neuro-fuzzy Neuro-fuzzy systems incorporate knowledge representation of fuzzy logic with ANN's learning capabilities	1. The power of ANN comes from the distributed processing capacity of a large number of simple elements in computation 2. It is a higher level logic	1. FL systems do not possess the capabilities of learning, adaptation, or distributed computing 2. FL system does not have the ability to learn itself 3. ANN lacks the ability to represent knowledge like human beings
	Neuro-genetic	1. It encodes NN's parameters as a string of network properties, which are chromosomes 2. It has the ability to find neighborhood solutions as compared to other traditional search strategies	1. For a given network, it needs a lot of memory to handle and manipulate chromosomes
	Neuro-fuzzy-genetic A neuro-fuzzy-genetic system is a system that combines the features of ANN, FL, and GA	1. Best suitable for path planning technique of mobile robot 2. ANN is trained to select a path from n-paths 3. FL is used to avoid a collision when all n-paths are blocked by obstructions 4. GA is used as optimizers for selecting optimal locations in obstacle-free directions and places by selecting fuzzy rules from fuzzy rule base of the FL system	1. Time-consuming 2. Difficult to implement 3. Complex in nature

References

1. H. Chaudhary, V. Panwar, R.P.N. Sukavanam, Adaptive neuro fuzzy based hybrid force/position control for an industrial robot manipulator (2014)
2. S.G. Ahmad, A.S. Elbanna, Dynamic modelling with a modified PID controller of a three link rigid dynamic modelling with a modified PID controller of a three link rigid manipulator (2018)
3. R. García-Rodríguez, V. Parra-Vega, Task-space neuro-sliding mode control of robot manipulators under jacobian uncertainties. Int. J. Control Autom. Syst. **9**(5), 895–904 (2011)
4. F. Piltan, A. Nabaee, M. Ebrahimi, M. Bazregar, Design robust fuzzy sliding mode control technique for robot manipulator systems with modeling uncertainties. Int. J. Inf. Technol. Comput. Sci. **5**(8), 123–135 (2013)
5. N.K. Sinha, C. Ls, 4l7 introduction, vol. 7 (1998), pp. 356–363
6. S.S. Roy, D.K. Pratihar, A genetic-fuzzy approach for optimal path-planning of a robotic manipulator among static obstacles. J. Inst. Eng. (India), Part CP Comput. Eng. Div. **84**(May), 15–22 (2003)
7. L. Wang, T. Chai, C. Yang, Neural-network-based contouring control for robotic manipulators in operational space. IEEE Trans. Control Syst. Technol. **20**(4), 1073–1080 (2012)
8. K.A. Khoiy, F. Davatgarzadeh, M. Taheri, D. Branch, A review on fuzzy-logic method to control robotic manipulator systems. **1**, 40–47 (2013)
9. G. Marcos, J.A.T. Machado, T. Azevedo-perdicoúlis, Trajectory planning of redundant manipulators using genetic algorithms. Commun. Nonlinear Sci. Numer. Simul. **14**(7), 2858–2869 (2009)
10. J.A. Saleh, F. Karray, Soft computing techniques in intelligent wall following control for a car-like mobile robot (2009, 2014)
11. Y. Li, Intelligent control with application to robotics (2000, 2014)
12. S. Alavandar, M.J. Nigam, Tracking control of 3-DOF robot manipulator using genetic algorithm m tuned fuzzy PID controller (2005), pp. 15–24
13. L. Jin, S. Li, J. Yu, J. He, Robot manipulator control using neural networks : a survey R. Neurocomputing **0**, 1–12 (2018)
14. W. He, Z. Yan, Y. Sun, Y. Ou, C. Sun, Neural-learning-based control for a constrained (2018), pp. 1–11
15. S. Li, Y. Zhang, L. Jin, Kinematic control of redundant manipulators using neural networks (2016), pp. 1–12
16. M.I. Al-saedi, H. Wu, H. Handroos, ANFIS and fuzzy tuning of PID controller for trajectory tracking of a flexible hydraulically driven parallel robot machine. J. Autom. Control Eng. **1**(2), 1–8 (2013)
17. H.N. Rahimi, M. Nazemizadeh, Dynamic analysis and intelligent control techniques for flexible manipulators: a review. Adv. Robot. **28**(2), 63–76 (2014)
18. Z. Bingül, O. Karahan, A fuzzy logic Controller tuned with PSO for 2 DOF robot trajectory control. Expert Syst. Appl. **38**(1), 1017–1031 (2011)
19. K. Lochan, B.K. Roy, F. Detection, in *Proceedings of Fourth International Conference on Soft Computing for Problem Solving*, vol. 335, Dec 2015
20. I.H. Akyüz, Z. Bingül, S. Kizir, Cascade fuzzy logic control of a single-link flexible-joint manipulator. Turkish J. Electr. Eng. Comput. Sci. **20**(5), 713–726 (2012)
21. B. Kazem, A. Mahdi, A. Oudah, Motion planning for a robot arm by using genetic algorithm. JJMIE **2**(3), 131–136 (2008)
22. L. Tian, C. Collins, An effective robot trajectory planning method using a genetic algorithm. Mechatronics **14**(5), 455–470 (2004)
23. J. Tu, S.X. Yang, Genetic algorithm based path planning for a mobile robot, in *2003 IEEE International Conference on Robotics and Automation (Cat. No.03CH37422)*, vol. 1, no. 0 (2003), pp. 1221–1226
24. S. Manjaree, V. Agarwal, B.C. Nakra, Inverse kinematics using neuro-fuzzy intelligent technique for robotic manipulator. **4**, 160–165 (2013)

25. I. Branch, A new method for position control of a 2-DOF robot arm using neuro-fuzzy controller. **5**(3), 2253–2257 (2012)
26. F. Terms, Neuro-fuzzy system for robotics applications. **2063** (2016)
27. R. Köker, A genetic algorithm approach to a neural-network-based inverse kinematics solution of robotic manipulators based on error minimization. Inf. Sci. (Ny) **222**, 528–543 (2013)
28. R. Köker, Full Length research paper a neuro-genetic approach to the inverse kinematics solution of robotic manipulators. **6**(13), 2784–2794 (2011)
29. P. Kalra, N.R. Prakash, A neuro-genetic algorithm approach for solving the inverse kinematics of robotic manipulators, in *Proceedings of IEEE International Conference on Systems, Man and Cybernetics*, vol. 2 (2003), pp. 1979–1984
30. S. Refou, K. Benmahammed, Control of a manipulator robot by neuro-fuzzy subsets form approach control optimized by the genetic algorithms, pp. 1–13 (2018)
31. G. Mester, Neuro-fuzzy-genetic controller design for robot manipulators. Integr. Vlsi J. 87–92 (1995)
32. T. Khelchandra, J. Huang, S. Debnath, Path planning of mobile robot with neuro-genetic-fuzzy technique in static environment, in *International Conference on Intelligent Systems Design and Applications*, April 2012, pp. 149–154

Health Monitoring Gadgets



Jasjit Singh, Ankur Kohli, Bhupendra Singh, Aishwarya Prasad Bhatoye and Sauhardh Prasad Bhatoye

1 Introduction

Health defines the vitality in every human being which architect his/her living in day-to-day routine, but previous years have seen rapid increase in health problem and mortality rate of people above age 50. In fact, doctors in countryside are working at their best potential but people sometime could not detect disease at initial stage which led to the rise of health number (Blood sugar, Blood pressure, Heart rate).

With advancement in technology, Government of India is proving Internet facility throughout the country. With this, the horizon of IoT embedded devices has broaden and we can easily gather the health monitor data on the IoT ThingsSpeak server database, analyse the result and automatically send health parameters in the form of message to the person's cell phone; if these parameters exceed the threshold value, then health message will be sent to the doctor with patient's id [1].

With the help of this, doctor will be able to analyse the record of person's health parameters on the ThingsSpeak channel and provide him/her with suitable medication.

We also provide user an opportunity to register this device with seven health care centres so that immediate action can be taken during critical health condition of the patient.

Figure 1 shows how the data from the sensors is received from NuttyFi IoT board which transmits it to the ThingsSpeak server channels; a trigger will be set by server

J. Singh (✉) · A. Kohli
School of Engineering Studies, University of Petroleum and Energy Studies, Dehradun, India
e-mail: jasjitsingh4999@gmail.com

B. Singh
SME, Dehradun, India

A. P. Bhatoye · S. P. Bhatoye
Shree Dev Bhoomi Institute of Science & Technology, Dehradun, India

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_151

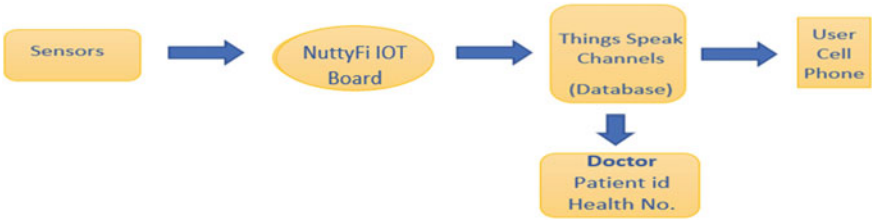


Fig. 1 Block diagram of flow of data from sensors to user cell phone

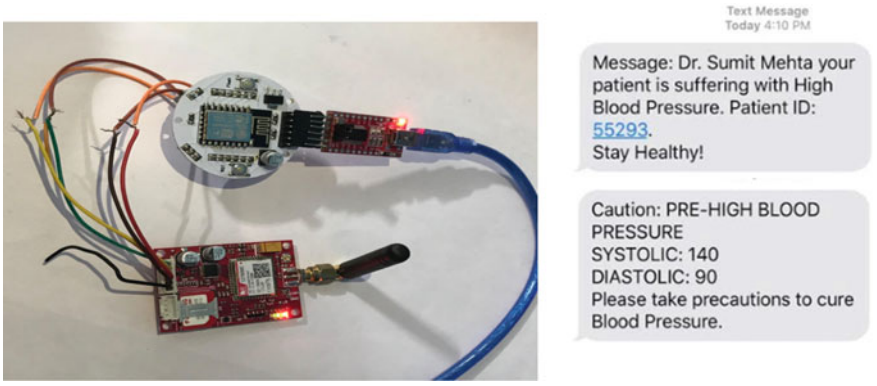


Fig. 2 View of GSM SIM 800 and B.P. (blood pressure) rate send to user and doctor (during critical conditions)

which will send alert notification to the user cell phone through GSM SIM 800A and the doctor in case the health number reaches the critical levels (Fig. 2).

2 Hardware Development

The implementation of different hardware components by first drawing different block diagrams on the plain sheet with reference to specific application of each component, then testing on the breadboard to validate the working of different sensors, and finally designing the PCB layout and preparation of final board will be installed in the healthcare gadget (Figs. 3 and 4) [2].

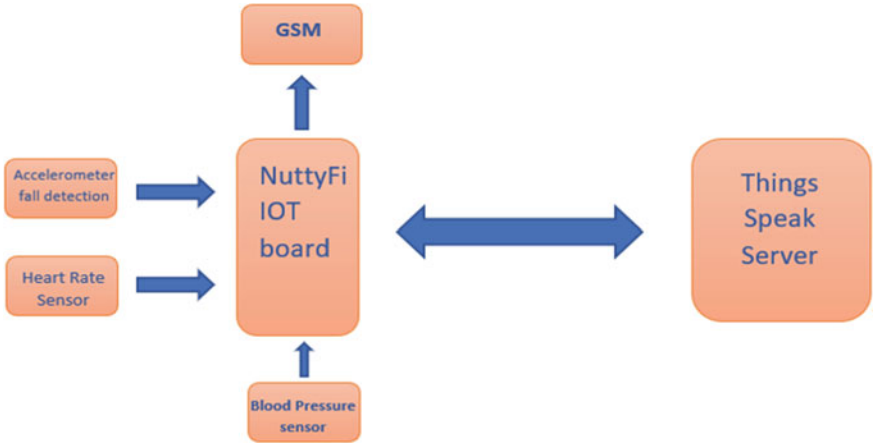


Fig. 3 Block diagram of sensors interfacing with ThingsSpeak server

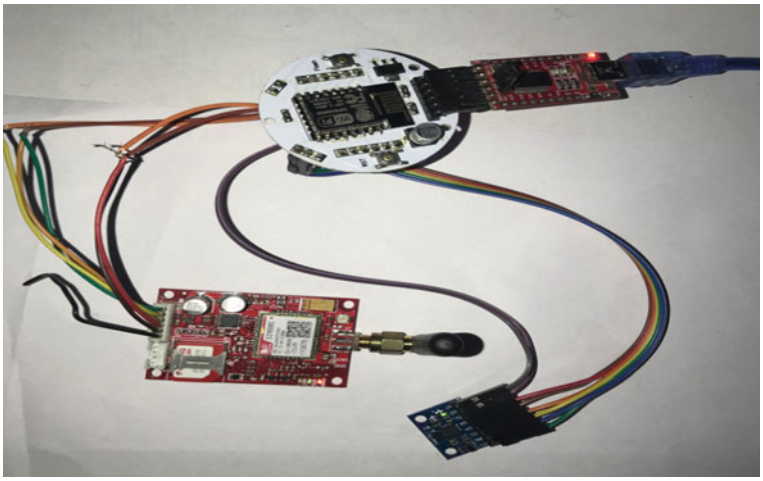


Fig. 4 Hardware view of system

2.1 ECG (Electrocardiogram) AD8232

This sensor converts heart pulse rate (systolic and diastolic) into electric signal which is charted as electrocardiogram (ECG); this electrocardiogram is obtained on the ThingsSpeak server channel [3]. The ECG sensor is divided into two intervals—PR interval, an electric impulse travelling from right atrium to the left atrium, generates the wave which forces it to contract and drain blood into the right ventricle, and QT interval, a signature 'beep' is generated in the cardiac monitor because during this interval both the ventricles begin to pump (Fig. 5) [4].



Fig. 5 ECG sensor graph

2.2 Accelerometer

It is used to detect sudden fall detection, i.e. when person faints under critical health condition, in such case ThingsSpeak server analyses all the health number and sets trigger by sending alert message to the registered health care. We have monitored different data by placing device at different positions on the body (Fig. 6) [5, 6].

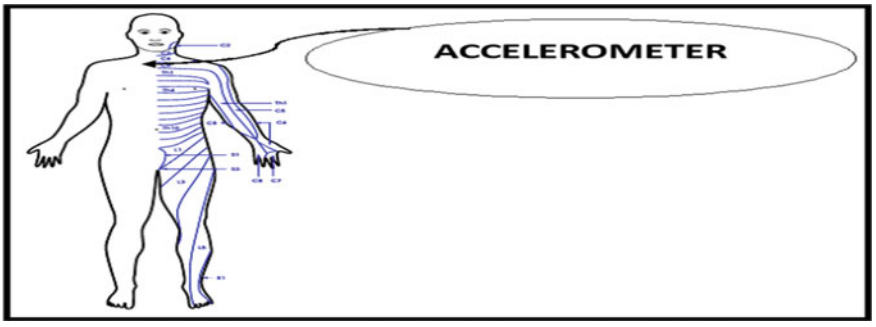


Fig. 6 Accelerometer paced on the top right part of the body

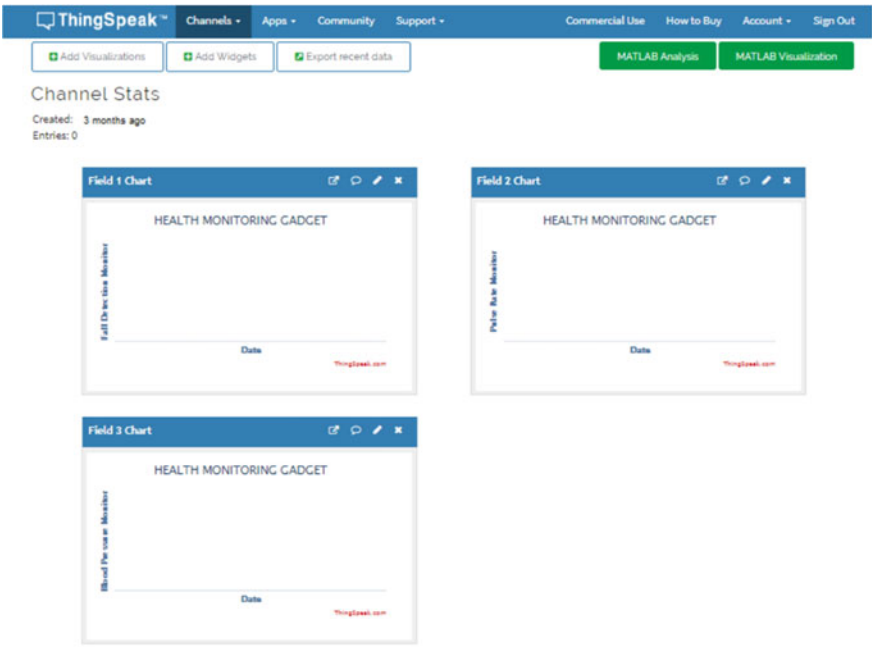


Fig. 7 ThingsSpeak server health monitoring channels

3 ThingsSpeak Server

IT is an IoT platform with MATLAB analytics which allow embedded device to communicate data on cloud through ThingsSpeak channels [7, 8]. It allows feasibility in analysing the results stored in its database, manipulates it and writes on the other channel. It provides processing of data in MATLAB analytics and triggers the response as in the case of critical health number; the server triggers an alert message to the user's cell phone (Fig. 7) [9, 10].

4 Conclusion and Future Scope

Health gadget IoT device provides feasibility to users for monitoring their health parameters periodically and save all the results in ThingsSpeak database. The device is capable of triggering alert signal in the form of text message, to create awareness about health numbers to user. It too provides doctors' complete analysis of patient health parameters during the time of patient's treatment so as to provide him with suitable medication [11].

In future, implementation of health gadgets would reduce mortality rate for the people of age above 50 due to sudden attack which as it has automatic triggering capability to report registered health care [12].

References

1. G. Anita, S. Rajesh, M. Mamta, S. Bhupender, S. Ravindra, G. Vikasgarg, Hands on approach on application of internet of things on various domain of engineering. Int. Res. Publ. House (2018)
2. G. Kortuem, F. Kawsar, D. Fitton, V. Sundramoorthy, Smart objects as building blocks for the internet of things. IEEE Int. Comput. 46–50 (2010)
3. L. Tan, N. Wang, Future internet: the internet of things, in *Proceedings of the 3rd International Conferences on Advanced Computer Theory and Engineering (ICACTE)*, vol. 5 (Aug 2010), pp. V5-376–V5-380
4. A.J. Jara, A.F. Alcolea, M.A. Zamora, A.F.J. Skarmeta, M. Alsaedy, Drugs interaction checker based on IOT, in *Proceedings of the Internet of Things (IOT)* (Nov/Dec 2010), pp. 1–8
5. Z. Shelby, C. Bormann, *6LoWPAN: The Wireless Embedded Internet* (Wiley, London, UK, 2009
6. G. Sebestyen, A. Hangan, S. Oniga, Z. Gal, eHealth solutions in the context of internet of things, in *Proceedings of the IEEE International Conference on Automation, Quality and Testing, Robotics* (May 2014), pp. 1–6
7. G. Mantas, D. Lymberropoulos, N. Komninos, A new framework for ubiquitous context aware healthcare applications, in *Proceedings of the 10th IEEE International Conference on Information Technology Applications in Biomedicine (ITAB)* (Nov 2010), pp. 1–4
8. Z.J. Guan, Somatic data blood glucose collection transmission devices for internet of things (Mar 2013)
9. M.L. Liu, L. Tao, Z. Yan, Internet of things-based electrocardiogram monitoring system (Nov 2012)
10. L.M.R. Tarouco et al., Internet of things in healthcare: interoperatibility and security issues, in *Proceedings of the IEEE International Conference on Communications (ICC)* (Jun 2012), pp. 6121–6125
11. D.Y. Lin, Integrated internet of things application system for prison (Jan 2013)
12. E. Gayat et al., Performance evaluation of a noninvasive hemoglobin monitoring device. Ann. Emergency Med. **57**(5), 331–336 (2011)

Evaluating Input Stream for IoT Commodity Eyes



Shubham Asthana, Rajiv Pandey and Archana Sahai

1 Introduction

With the modern advancements in technology and convenient access to it, almost every place is covered by a CCTV (**C**lose **C**ircuit **T**ele**V**ision) camera. Some of these also get a real-time monitoring, but the numbers of such cameras are very low. Generally, the CCTV cameras installed are just for the passive purposes such as reporting the crime after it has happened and/or just for the recording purposes. The cameras also are working independently; every household, every institution, and/or society has its own CCTV system. Had these cameras been interconnected and sending data to a central hub, a Police dispatch center, etc.? for instance, they could be monitored and also a suitable machine learning algorithm may be devised to reduce the human involvements even further. The perspective of this paper is to integrate the security commodities or security peripherals to make the conventional policing a more Hi-tech and futuristic through IoT and commodity eyes (Commodity Cameras), in other words, an architectural implementation of upgrading the conventional security system to make it at par with the world if not better. Since security is a major concern, India should empower smart securities and policies.

Conventionally or what is happening till date is that the police patrol the city with their cars and other law enforcement vehicles consuming lots of fuel (Petrol or Diesel) and employing a huge manpower for this purpose. On the other hand in the societies of private houses, public have devised close circuit cameras to protect

S. Asthana · R. Pandey (✉) · A. Sahai
Amity University, Lucknow 226028, Uttar Pradesh, India
e-mail: rpandey@lko.amity.edu

S. Asthana
e-mail: shubhamasthana123@yahoo.co.in

A. Sahai
e-mail: asahai@amity.edu

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_152

their places. But the camera proves out to be just a tool of evidence after the crime or burglary has been committed as there is no continuous monitoring system. In some cases, people connect the cameras to Internet to keep a watch on the place/premises under the camera scope even when far away, by tools like remote login or private broadcasts. Even in this case a human intervention for monitoring and reporting is required. The monitoring in this case is not being done by any legal authority, so in cases of any mishap the local authorities like police or fire brigade have to be alerted by the person monitoring, which is time-consuming.

The modern advancement in technology has given mobile and ubiquitous computing, and the rapid growth of internetworking application and personal gadgets like fitness bands, smartwatches to smart houses or offices has escalated the Internet service called Internet of Things (IoT) originally devised in 1999. IoT is massively used technology nowadays because of its productivity in mobile computing, nanotechnology, and embedded system communications. It is expected that the number of IoT devices will reach more than 50 billion devices by 2020 [1].

According to the definition issued by British Standards Institute (BSI) smart city is:

> The effective integration of physical, digital and human system in the built environment to deliver sustainable, prosperous and inclusive future for its citizens [2].

2 Camera Features

CCTV footage is the first thing any agency reaches for while beginning any investigation. It also is a crucial source of evidence for the judiciary systems as well. In earlier days, during the period of analog cameras, there used to be many ambiguity and non-uniformities due to the diverse camera feed and data formats. In the present time, on the contrary, the analog cameras are also connected to Digital Video Recorders (DVR) or Network Video Recorders (NVR) which in turn converts the data from any analog camera to a digital format video for storage and transmission over LAN (if any).

IP cameras directly capture the data in digital format and transmit over the LAN without any requirement of external hardware. The differences on the basis of data format and streaming are also described in subsequent section.

3 Activation Function

Identity	Binary Step Function
Logistic or Sigmoid	Tanh
ArcTan	Rectified Linear Unit (ReLU)
Leaky ReLU	Softmax

Activation functions are used to contain a wide range of input to a predefined range of output, generally from zero to one. These functions are implied for the accuracy of the model being developed also used to impart a nonlinearity.

3.1 Identity Function

It is the simplest function as its output is same as the input provided. Mathematically, if the input provided is X (let) then the output produced will also be X. Following is the formula for it (Fig. 1):

$$f(x) = x$$

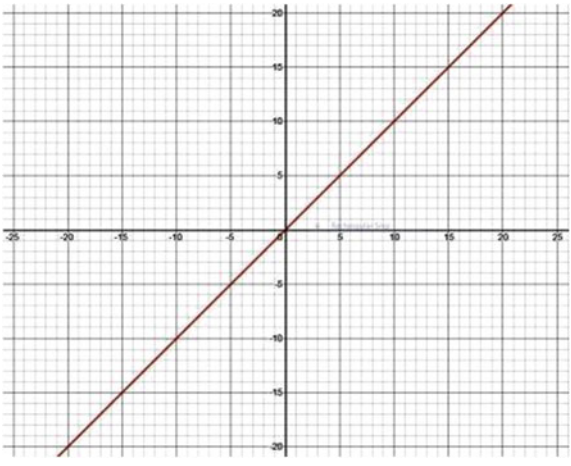


Fig. 1 Identity function

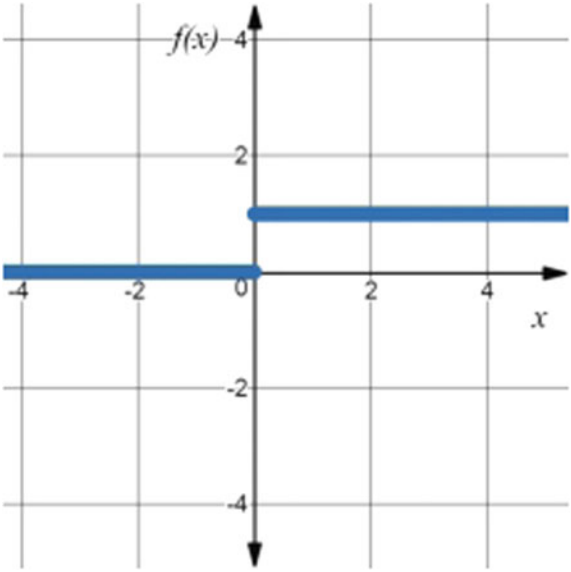


Fig. 2 Binary step activation function

3.2 Binary Step Activation Function

In binary step activation function, the output is mapped to only two values, zero and one, that is binary values. The computation of this function depends on the magnitude of the input only, regardless of the magnitude. All the input values pertaining to the negative side would be mapped to 0 (zero) and all the nonnegatives are mapped to 1 (one). This function is very useful when a binary classification is required. Following are the graph and function of binary step activation function (Fig. 2).

$$f(x) = \begin{cases} 0, & x < 0 \\ 1, & x \geq 0 \end{cases}$$

3.3 Logistical or Sigmoid Function

The logistical or sigmoid function may look similar to the binary step activation function in a manner that this function also maps any given input between 0 and 1. But it is different from the binary step function because the binary function only has two output values (0 and 1) and so is a discontinuous graph but sigmoid function, on the other hand, has a continuous output range between 0 and 1 and is a continuous graph (Fig. 3).

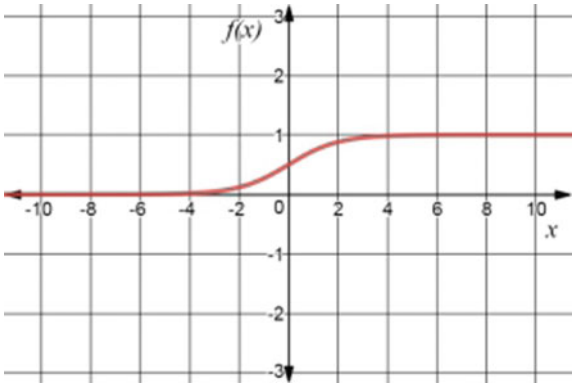


Fig. 3 The logistical or sigmoid function

$$f(x) = \frac{1}{1 + e^{-x}}$$

3.4 Tanh Activation Function

The Tanh function may look similar to the logistical or sigmoid activation function visually but logistical or sigmoid maps any given input between 0 and 1; however, Tanh function on the other hand has an output range between -1 and 1 and is also a continuous graph. The accuracy of both the functions may differ within projects; hence, it is always advised to try out both the functions and select the more suitable one accordingly. Following is the formula for it (Fig. 4):

$$f(x) = \frac{2}{1 + e^{-2x}} - 1$$

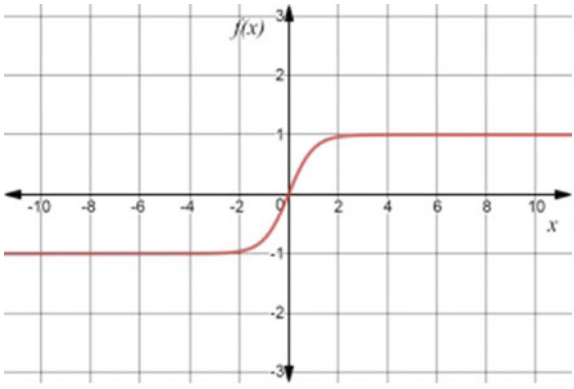


Fig. 4 The Tanh function

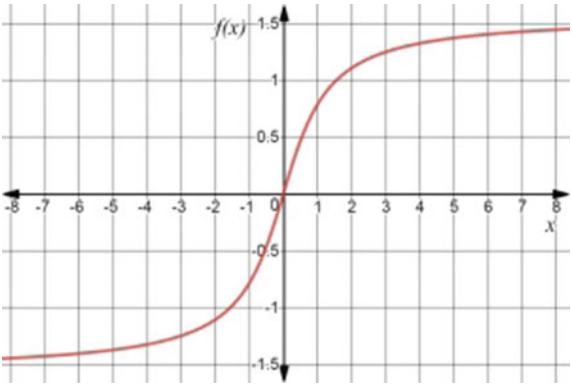


Fig. 5 The ArcTan function

3.5 ArcTan Function

The ArcTan Function is just the tan inverse function applied to the input variable x. The range of trigonometric function $\tan^{-1}(x)$ is somewhere between $-\pi/2$ and $\pi/2$. The simplicity of this function makes it easier to upgrade or modify the function as and when required. The ArcTan function can be seen as an alternative or replacement to the Tanh or the Sigmoid functions. Following is the formula for ArcTan function (Fig. 5):

$$f(x) = \text{ArcTan}(x)$$
$$\text{OR}$$
$$f(x) = \tan^{-1}(x)$$

3.6 ReLU

ReLU is the most popular and widely used function among all the other activation functions in the field of computer vision, speech recognition, neural network, and deep learning. As seen in the below ReLU curve, it only operates on nonnegative values. For all the negative inputs, the output is zero and for all the positive values the output is same as the input value. This activation function is most useful when we want to operate only on positive inputs and ignore the negative input values (Fig. 6).

$$f(x) = \begin{cases} 0, & x < 0 \\ x, & x \geq 0 \end{cases}$$

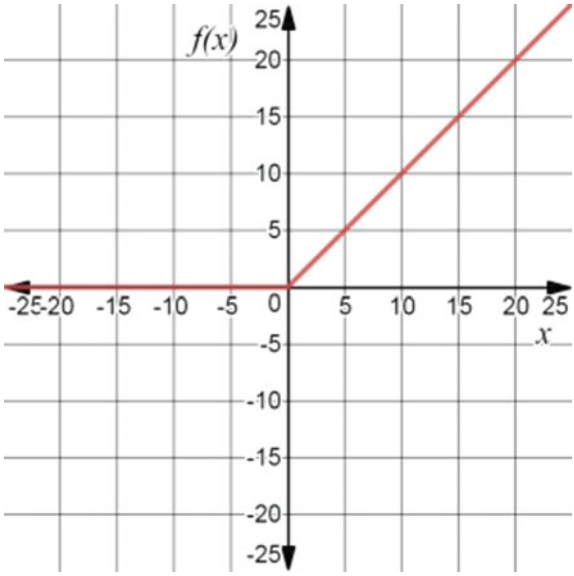


Fig. 6 ReLU

3.7 Leaky ReLU

The Leaky ReLU function a job very similar to the ReLU function, all the positive input values are mapped to the output having same magnitude as the input. In addition to this, it does not make the negative values completely zero; rather it just reduces their (input's) magnitude by a factor of 1 percent (1%) (Fig. 7).

$$f(x) = \begin{cases} 0.01x, & x < 0 \\ x, & x \geq 0 \end{cases}$$

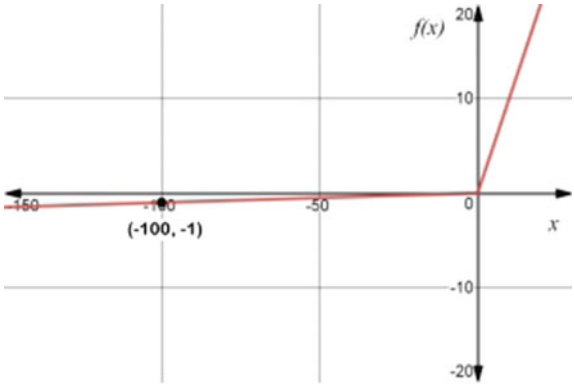


Fig. 7 The Leaky ReLU function

3.8 SoftMax Function

Last but not the least, the SoftMax function is the most complex classifier discussed so far. The distinction of this function is that it imparts a probability to the input in sample; besides the output of this function is now also dependent on the sample size and the individual values of sample element due to the denominator of the equation which is the linear sum of all the input values, this restricts its use within small input sizes. For these reasons, it is applied on the last layer of the network, so as to refine the outputs from the previous layers. When the output size (obtained on applying any of the above-described functions on the whole input) is contained in single digit (that is, say around 6–7 outputs are obtained to select from), in such cases, SoftMax is applied to obtain probability distribution among these outputs and the most probable occurrence (classification where the probability of a class is maximum) is selected (Table 1).

$$f(x)_j = \frac{e^{x_j}}{\sum_{i=1}^{K} e^{x_i}} \quad \text{for } j = 1, \ldots, K$$

Above described are some of the activation functions that are useful and quite famous for deciding the accuracy of neural network model similar to the one needed for this project.

Table 1 Summary of functions

Name	Function	Output range
Identity	$f(x) = x$	R
Binary step activation	$f(x) = \begin{cases} 0, & x < 0 \\ 1, & x \geq 0 \end{cases}$	{0, 1}
Logistic or Sigmoid	$f(x) = \frac{1}{1+e^{-x}}$	(0, 1)
Tanh	$f(x) = \frac{2}{1+e^{-2x}} - 1$	(−1, 1)
ArcTan	$\tan^{-1} x$	$\left(\frac{-\pi}{2}, \frac{\pi}{2}\right)$
ReLU	$f(x) = \begin{cases} 0, & x < 0 \\ x & x \geq 0 \end{cases}$	$R^+ + \{0\}$
Leaky ReLU	$f(x) = \begin{cases} 0.01x, & x < 0 \\ x, & x \geq 0 \end{cases}$	$(-\infty, \infty)$
SoftMax	$f(x)_j = \frac{e^{x_j}}{\sum_{i=1}^{K} e^{x_i}}; \forall \text{j} \in [1,\text{K}]$	(0, 1)

4 File Format and CODEC

Digital video recorders began to replace older analog CCTV recorders about 15 years ago, and they promised a lot: better quality, ease of use, and reliability. One important feature was lost in the changeover: a standard data format. With Digital Video Recorders (DVR) file system, the storage mechanism and encoding techniques for the video data were to be resolved by the manufacturers. Owing the absence of a universal regulation or globally accepted industry standards, all the manufacturer started developing their own file formats and codec the need for media players were also obliged individually. The result was thousands of proprietary file format and ever-growing codec.

The letters or numbers after the final '.' in the video file name are called the file extension; these extensions are the key to identifying the file format, further to anticipate the possible codec used. If the file ends in 'AVI', the next step is to determine what codec the file requires. If it is 'mpeg', 'mpg', 'flv', or one of the other open formats, the file will likely play in a standard movie player such as VLC or Windows Media Player. If it is none of these, the file format is most likely closed and will only play in the manufacturers own software.

AVI is the most common file format in CCTV, but within AVI, most of the manufacturers use their own, nonstandard, codecs. The AVI files have their codec written into the file header. It is possible to look at the codec by opening the file in HxD, that is, a HEX viewer/editor (look at byte position 188); to make it more simple, there are free tools available online. There are more than 850 different codecs available, which are only for AVI file system. Manufacturers provide with the codec they are using for the products. So the user can download a specific codec and install it to any existing media player to run the video coded using this codec.

5 Architectural Model of IoT Eye

The detailed analysis of the input streams as discussed in the above sections shall be the input for our proposed system figure. The system shall capture the input assigned appropriate feature and thus perform an efficient analysis and generate alarms as found suitable (Fig. 8).

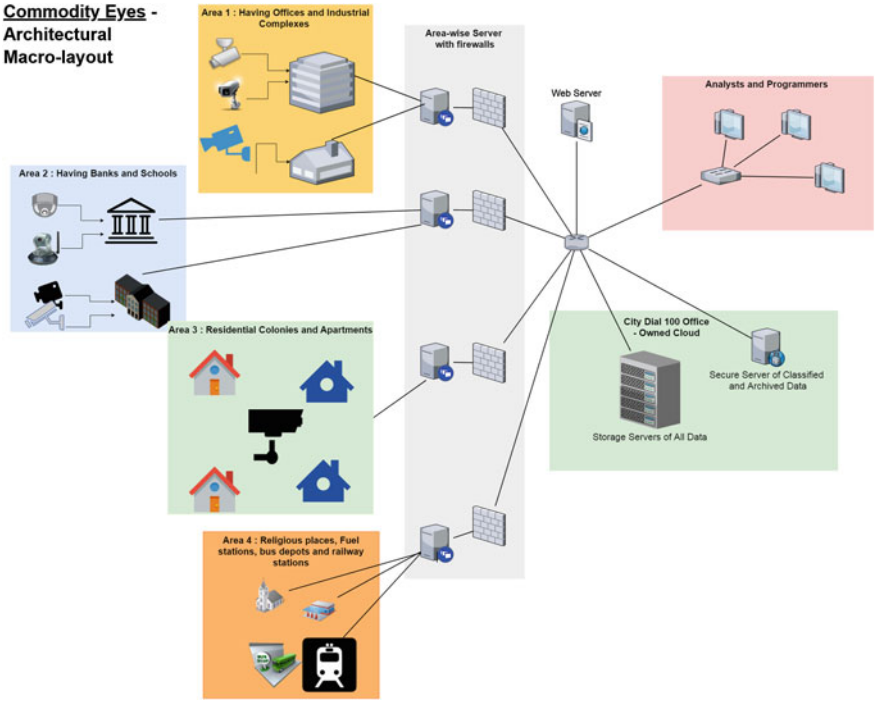


Fig. 8 Proposed architectural model for smart policing

The macro-architectural layout presented here shall be seen as the vision toward designing a city-spread smart surveillance system. The intent for writing this paper is to explain the first step, that is, data collection, and the challenges associated with it.

The wholesome idea of the project is to explore possibilities of big data features and techniques, cloud, and IoT to devise a better and smarter surveillance and security establishment. When orchestrated and integrated in a smart and intelligent manner, this implementation idea can make the smarter cities a lot safer too.

However, the detailed architectural implementation and the response generation algorithm to be implemented are beyond the scope of this paper and shall be discussed in subsequent publications.

The broader idea of implementation in this project is that the camera (or commodity eyes, here) will be combined with an RF transmitter (if not already equipped) to connect with respective area-wise networks. These networks will have a machine learning algorithm set up on a central node to analyze the live feeds from various sources.

In case of any untoward or objectionable incident getting detected, this central node would generate alarms. These alarms shall have three significant levels based on the severity of the event, for example, any mild happening could result in a

text-based alert message being generated for the respective management authorities like owners or society secretary. Further, the next level of alarm shall be for more severe incidents which require a moderate degree of response, for instance, alerting the local authorities like public safety departments (Police and/or fire). The most critical events would require utmost concern and so the real-time feeds shall be transmitted to the higher authorities for monitoring and swift decision-making. The cameras and their features are well described in the previous sections, and hence this section shall concentrate on the concentrators and their connectivity. The cameras shall have a unique ID or UID; these UID(s) would act as a primary key identifier to have a unique identification of record in the central database of cameras. These records will provide the system to obtain the information about the locality, also the information regarding the person or groups to be contacted for level-1 notification generation. While designing and devising an integrated architecture of this system, it was discovered that some commodity cameras lack the ability to effectively transmit the video feeds over the web. One solution to this problem is to connect the whole arrangement with a network like LoRaWAN with upgrade modules mote to make them LoRa compatible; however, the actual connectivity implementation is beyond the scope of this paper and shall be discussed in subsequent papers.

6 Actuating Devices/Concentrators Nodes

Concentrator nodes act as gateways in the proposed architecture of this paper. The location of installation of these actuators is at stage 2, named as area-wise server with firewall; these actuators would be in two layers. The first layer shall be equipped with server-side scripting and function as a central hub (area-wise) to collect video feeds transmitted by the commodity cameras from their respective area and domains, for example, an actuator acting as a concentrator of school camera feeds in a particular shall only collect the feeds coming from the school transmitters; for the residential commodity camera feeds from the same area, a different actuator shall act as a concentrator. These concentrators will then analyze the feeds and send the details to the next concentrator, in case any alarm is generated. The higher level actuator shall be connected to the application server, and the level of action to be taken will be decided at this actuator and the IRC messages or other actions shall be initialized at this stage by the utilization of the application server in connection with this actuator node. These gateways/actuators are serving as a bridge in this architecture between the commodity eye (cameras) and the core network, which includes the database servers, data analysts, and the actual artificially intelligent processors.

7 Conclusion

This paper discussed processes and various challenges that are involved in the phases of data collection (here, CCTV surveillance videos) due to the vast heterogeneity among manufacturers, the data for integration phase to unify the data from multiple sources to provide as an input to the alarm generation algorithm (to be described in subsequent papers). Various activation functions that are to be used in the actual algorithm were also discussed, which shall be used for data cleaning phase of the proposed system. An architectural overview was also presented in the previous sections of this paper. The actual algorithmic implementation for video analysis and alarm generation is beyond the scope of this paper and shall be discussed in subsequent papers for future work and research on security and computational intelligence using IoT.

References

1. A. Franklin, The future of CCTV in road monitoring
2. J.K. Park, K.S. Han, Intensity classification background model based on the tracing scheme for deep learning based CCTV pedestrian detection, in *2018 IEEE 9th International Conference on Mechanical and Intelligent Manufacturing Technologies*, 19 April 2018
3. The Internet of Things—Architecture: Final architectural reference model for the IoT v3.0
4. BRK1552 by, Kevin Miller, Principal Program Manager, Azure IoT
5. Internet of Things factsheet Architecture
6. G.Y. Park, *Large Scale Intelligent Surveillance System* (2013)
7. Internet of Things: How Much are We Exposed to Cyber Threats?—InfoSec Resources, http://resources.infosecinstitute.com/internet-thingsmuch-exposed-cyber-threats/
8. W. Cho, D.H. Kim, J.K. Paik, Gait recognition using multiple feature detection. Eng. Korea—SP **44**(6), 84–92 (2007)
9. X. Bellekens, A. Seeam, K. Nieradzinska, C. Tachtatzis, A. Cleary, R. Atkinson, I. Andonovic, Cyber-physical-security model for safety-critical IoT infrastructures
10. S. Institute, SANS InstituteInfoSec Reading Room tu-evolution, the history and Evolution of Intrusion Detection (2001)
11. C. Han, Y. Lv, D. Yang, Y. Hao, An intrusion detection system based on neural network, in *2011 International Conference on Mechatronic Science, Electric Engineering and Computer (MEC)* (2011)
12. ISO/IEC 27005:2011—Information technology—Security techniques—Information security risk management
13. ISO/IEC 27000:2014—Information technology—Security techniques—Overview and vocabulary
14. A new approach to investigate IoT threats based on a four layer model
15. O. Vermesan, et al., Internet of things strategic research and innovation Agenda

Application of Sensors Using IoT for Waste Management System



Akhil Chitreddy, Kailash Gogineni, Vattikuti Anirudh, P. V. Akhilesh, Konda Krishna Vamsi and P. Swarna Latha

1 Introduction

Internet of things (IoT) is the application in which the role of the Internet is a must. It is one of the popular domains of advanced technology which stores all your data and information in the cloud with real-time quick access to data. As the information is stored in the cloud and people are connected to the Internet, it will provide a great access to the people who are using the same application from any part in and around the world. Such kind of advancement conjointly equips sensors and routers for gathering and causation of information across the net. IoT conjoints certain technologies like smart grids, intelligent transportation, smart homes, and smart cities deals with sensors. It is discovered that usually the garbage gets collected, thanks to the irregular removal of garbage gift within the dustbin. The garbage bins are going to be monitored by the system which informs in real time regarding the amount of garbage collected within the garbage bins through an Internet page. For this function, the system uses the inaudible sensors present over the bins to distinguish the level of garbage and examine it with the depth of garbage bins. GSM module is employed to communicate the data, which is during this case the amount of garbage to the mobile whenever most of the program is dead and passes onto the Arduino to perform the various materials hooked up thereto via the breadboard through the connecting wires. Conjointly for the data to be accessible from anyplace, this project's integration with ThingSpeak is finished whenever a plot is created for time against the amount of garbage.

A. Chitreddy (✉) · K. Gogineni · V. Anirudh · P. V. Akhilesh · K. K. Vamsi · P. Swarna Latha (✉)
Vellore Institute of Technology, Vellore, Tamil Nadu, India
e-mail: akhilchitreddy@yahoo.com

P. Swarna Latha
e-mail: hgswarna@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_153

- Scale back human resources and efforts besides the improvement of any wise town vision is the main aim of this project.
- This project principally helps the general public to cut back many health hazards by keeping the town clean.

2 Comparative Study

Unhygienic environment has been there in most of the cities owing to the overflowed bins. This can more result in the rise of various kinds of unknown diseases. This can degrade the quality of living. To beat these things, an efficient smart garbage management technique should be developed. Because the scope of IoT is developing day by day, effective ways are discovered simply. Numerous styles were projected and have advantages as well as disadvantages. Taking this drawback into account, we have enforced an answer to a similar scenario; however, in our college whenever the dustbins once stuffed get overflowed thereby spreading everywhere around the place making unsanitary smell and diseases to some extent. This implementation of the thought of "IoT Garbage observance system" involves a small controller (Arduino Uno) connected to a radical sonic detector that has waves decisive of the peak of garbage within the instrumentality and GSM module to send the information to the involved person's mobile. An LDR detector is additionally enclosed as a trigger to ascertain if the least bit of the rubbish is gift, reducing the energy consumption of ultrasonic detector to find the amount even within the absence of garbage.

This idea of sensible waste management system may well be varied and altered in many ways which involves the utilization of various communication protocols like RFID, GSM module, and additionally the utilization of various detectors like IR sensor and gas detector that determines in what means the waste disposal affects the close surroundings. A number of connected work area units are given below. In [1], it includes a waste bin containing solids' observance scheme garbage bin set that is within the public places, and then camera is set for the dustbin location. The camera shoots image of the garbage bin. GPS, radio-frequency identification (RFID), and GIS send the image to the workstation. The camera and the RFID area unit are mounted within the truck; once the truck comes back closer to the bin, RFID reader communicates with RFID tag and sends the info. A [2], number of the connected devices work in the area unit as follows. The ultrasonic detector can show us the varied levels of garbage within the dustbins and sends its result ahead once its intensity exceeds. These details of the area unit are given to the Arduino, and also the controller offers the main points to the transmitter module (Wi-Fi module). At the location of the receiver, a mobile telephone set must be connected to the Wi-Fi router. In [3], the global system for mobile communication (GSM), ZigBee, and ARM7 are employed to make the integrated system for observing the dustbins remotely. The sensors area unit present within the common dustbins is present in the general public areas. Once the rubbish is reached to a level of the detector indication, the details are given to

the ARM seven controller. The controller can offer indication to the driving force of garbage pickup truck on that garbage bin which is totally stuffed and needs immediate attention. ARM seven can offer indication by causation SMS and victimization of GSM technology. In [4], the projected system indicates that the amount of garbage within the dustbins has been detected with the assistance of detector systems, and the communication is done to the licensed room via GSM system. For interfacing the detector system with GSM system, microcontroller is employed. A GUI has been additionally incorporated to watch the required info related with the rubbish for various elected locations. In [5], description of the application of our model of "Smart Bin" in the management of the waste assortment system of a complete town is dispensed. The detector networks that enabled sensible bins connected through the cellular network can generate an oversized quantity of knowledge that is more analyzed and unreal at the important time to realize sight regarding the standing of waste around the town. They additionally analyzed the particular performance of the system. In [6], this method can stop the overflow of ashbin and create the setting neat and clean. It will scale back the wastage of time and energy of the human. Victimization 8051 microcontroller, inaudible detector, GSM modem, and resistor are used. Plus it'll additionally stop the prevalence of any malady. The truck drivers simply get the knowledge regarding the clearing method and do their work at once. In [7], the sensible waste management victimization IoT is enforced. This projected system assures us the gathering of garbage shortly once the rubbish level has reached its most level victimization—cloud, ESP8266, UV sensor, Raspberry pi, load cell, wetness detector, and garbage aggregation vans. The system can also give correct reports, increasing the potency of the system. In [8], the aim is to enhance a route designing scheme for a metropolitan space with accessible live GPS data collected from all vehicles within the network. We have a tendency to evaluate the standard of the obtained journeys; if we have a tendency to ignore the feed and also the quantity, we have a tendency to increase the quality by incorporation of the feed in the route designing scheme. Raptorial bird at the side of illustrious and novel extensions has been tested by the U.S. The key insight of the work which we have done is that we will improve the standard of the computed journeys, particularly for brief varied queries. In [9], the algorithmic program, that is, the graph primarily based on the resolution matter of route improvement has been projected. The paper cares principally with best route director in the Federal Republic of Nigeria touristy.

3 Methodology

In this paper, we tend to send the data regarding the amount through the GSM module connected to the Arduino, besides communication interfaces like serial port and USB for connecting to different devices. The ultrasonic sensing element is employed to notice the peak of the garbage stuffed at completely different intervals of your time in line with the time delay. The LDR detector is employed as a trigger to notice the presence of garbage within the garbage can. This is often a significant step during

this project to cut back the energy consumption of the ultrasonic sensor, i.e., the detector solely determines the amount of garbage once the trigger or the LDR detector glows. The Arduino Uno board has been employed as microcontroller. Interfacing is finished among GSM module and Arduino board, therefore sending the message to the involved authority if the amount of garbage exceeds 75%.

Another feature introduced inside this project is the ThingSpeak which is the online cloud DB which could be used to get data online.

This system has a microcontroller and communication module with sensor for data acquisition. The components used in our system are as follows:

3.1 Hardware Components:

- Arduino Uno (Microcontroller),
- Ultrasonic sensor,
- SIM 900A GSM module,
- Resistors,
- Garbage can,
- GPS system (extra feature), and
- Breadboard and connecting wires.
- Software Components:
- Arduino programmable software (in PC),
- Cloud (ThingSpeak) as an online DB, and
- Google API, Google Maps (Fig. 1).

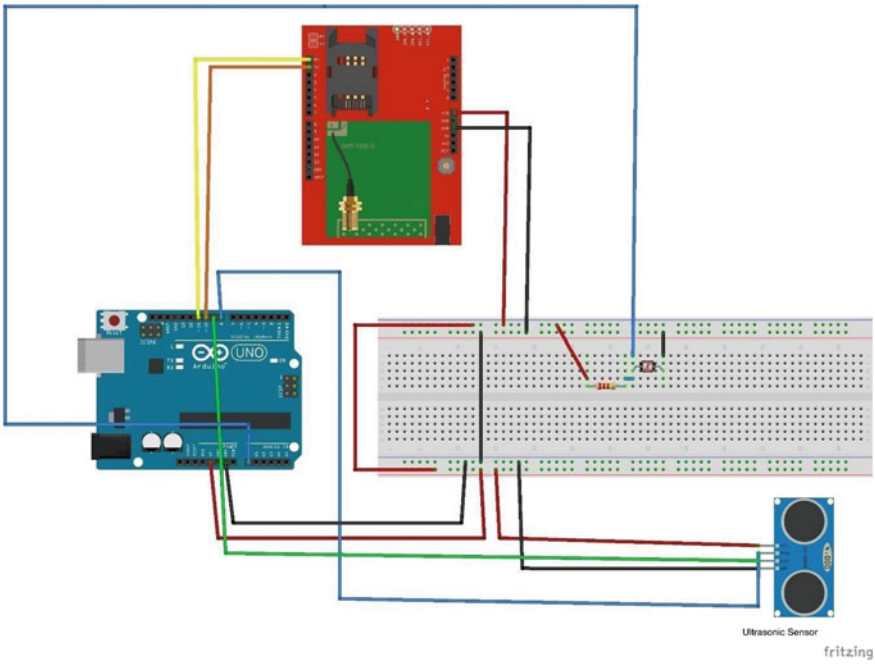


Fig. 1 Connection model

The GSM module is connected with the Arduino by connecting TX pin of the board to RX pin of the modem and vice versa. Trigger and echo pins of sensor are connected with digital pins 8 and 9 of Arduino board. The LDR, ultrasonic, and GSM module are connected to the 5 V supply provided by the Arduino (Fig. 2).

4 Procedure

First, we connect the Arduino and the outer components via the designated pins represented in the code. Then, we open the Arduino software where we select the type of board, the baud rate, and the libraries (in this case the software serial library) which can be downloaded from the Internet. Then after verification and compilation, we upload the code through the RX pin into the Arduino and then open the serial monitor in the software. We connect the monitor to the GSM module which is displayed.

Another feature added, i.e., its integration with the cloud online DB (ThingSpeak), really brings out the smartness in implementing the project. The percentage of the garbage level is monitored and plotted on a graph with the API key being specifically embedded into the code.

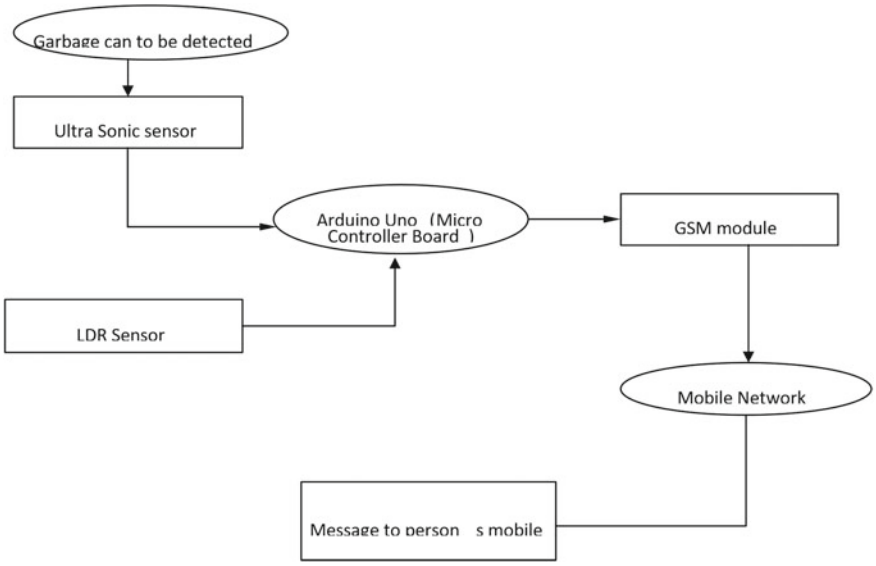


Fig. 2 Architectural design

4.1 Innovative Idea

We have implemented a GPS tracker idea without actually using a GPS module. This code lets us know the location of the garbage and also helps us to determine the shortest distance between the garbage collector and the position of the garbage. This idea inhibits us from automatically knowing the location of the garbage but has a procedure as follows. We have included six buildings with each building having an ID. Whenever we enter the source ID of the garbage, it quickly gives us the shortest path from the location to any building, helping the collector in moving faster to the location.

5 Test Case Generation and Analysis

We have to perform testing to identify and correct any errors that will be occurred in the project. It will increase the quality of a product. Testing will be of two steps. They are unit testing and system testing.

5.1 Unit Testing

Test ID	Scenario of the test	Test steps	Test data	Expected results	Actual results	Pass/fail
T01	Test compatibility of Arduino hardware with software	(1) Connect Arduino to laptop (2) Execute and copy the code to Arduino	Connecting Arduino to laptop	Code should be executed	Code is executed properly	Pass
T02	Check the ultrasonic sensor with the obstacle	If there is any obstacle, there should be an indication	Put an obstacle in front of the ultrasonic sensor	Ultrasonic sensor indicating that there is an obstacle	Ultrasonic sensor indicating that there is an obstacle	Pass
T03	Check the ultrasonic sensor without any obstacle	If there is no obstacle, there should be no indication	Having no obstacle in front of the ultrasonic sensor	Ultrasonic sensor indicating that there is no obstacle	Ultrasonic sensor indicating that there is no obstacle	Pass
T04	Compile Arduino code	Compile and run all the divisions of the Arduino code	Arduino code	Compilation and execution successfully	Compilation and execution successfully	Pass
T05	Check GSM module	(1) Condition should reach threshold value (2) GSM module should alert the registered mobile	GSM module with given condition more than the threshold value	GSM module should alert the registered mobile	GSM module sent a message to the mobile	Pass
T06	Test the breadboard	Connect wires with a circuit	Connect the wires with a basic output	It should give correct results	Breadboard is working fine	Pass

5.2 System Testing

Test ID	Scenario of the test	Test steps	Test data	Expected results	Actual results	Pass/fail
T01	Dustbin is empty	Dustbin is empty	Ultrasonic sensor is directed to the dustbin	Update the level of dustbin	The level of the dustbin is updated to 0 level	Pass
T02	Dustbin is at 75% level	Check when dustbin level is at 75%	Ultrasonic sensor is directed to the dustbin	Update the level of dustbin and send no alerts to the mobile	The level of the dustbin is updated to 75%, and there are no alerts sent to the mobile	Pass
T03	Dustbin is at 80% level	Dustbin is at 80% level (threshold level)	Dustbin is at threshold level	Send alert to the registered mobile about threshold level of the dustbin	GSM sends the alert that the dustbin is at threshold level	Pass
T04	Dustbin is at 90% level	Dustbin is at more than threshold level	Dustbin is at more than threshold level	Ultrasonic sensor will give an alert buzzer sound	Ultrasonic sensor will sound to alert 90% level	Pass
T05	Dustbin is at 0 level after getting filled	(1) Dustbin is emptied after getting filled (2) Check the dustbin level	Dustbin is at 0 level after removing the dust from the dustbin	Update the level of the dustbin at 0 level	The level of the dustbin is updated to 0 level	Pass

6 Results

See Figs. 3, 4, 5, and 6.

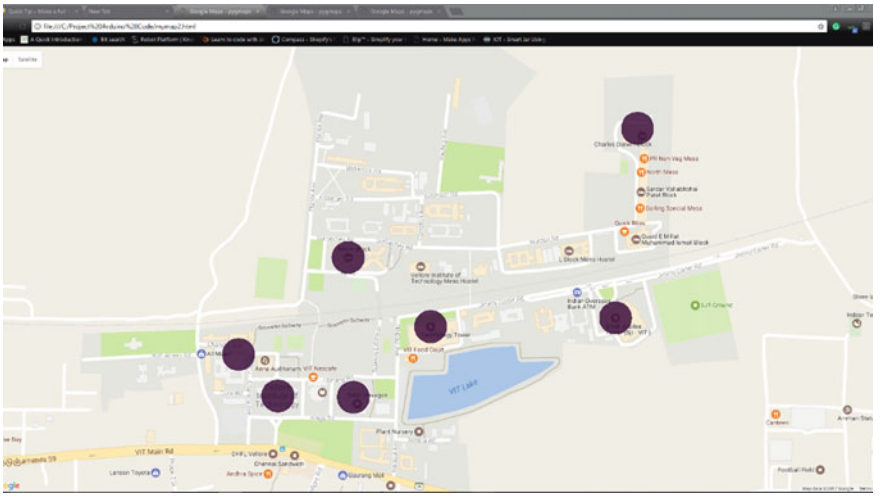


Fig. 3 Google Maps (Nodes for path to garbage)

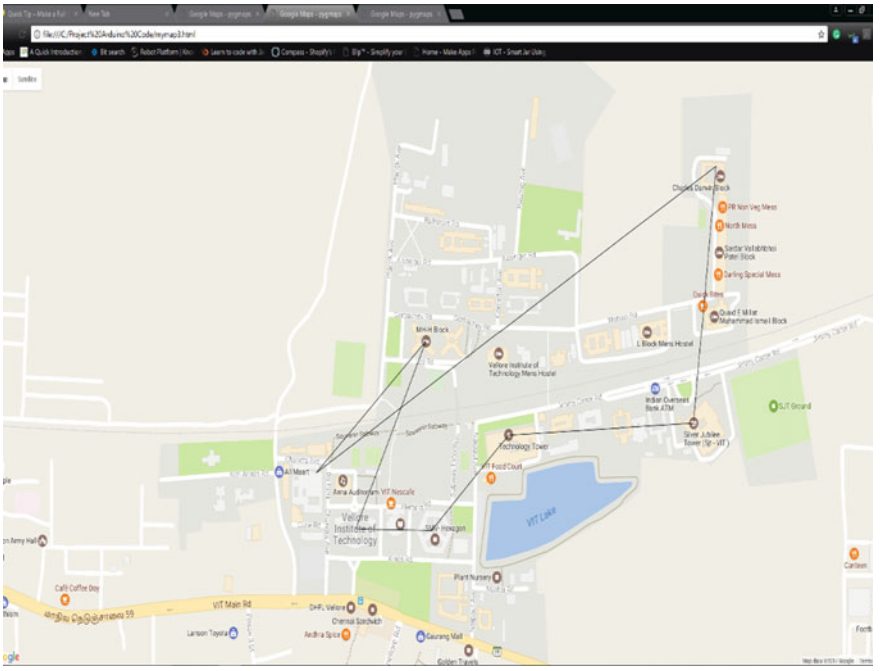


Fig. 4 Google Maps (Paths for travel to garbage)

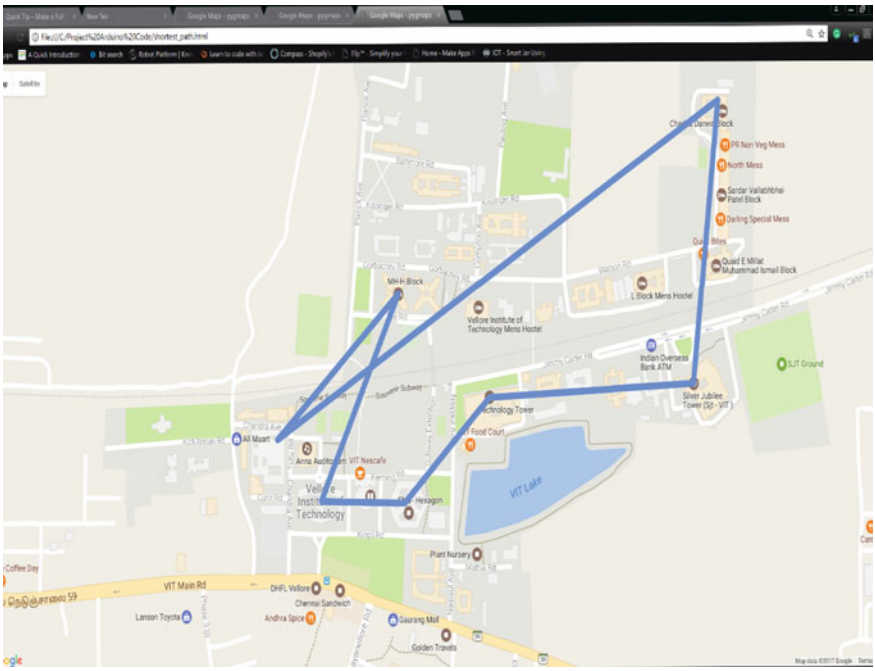


Fig. 5 Google Maps (Shortest path from source to destination)

7 Conclusion

The garbage level will be monitored by the system and would be sent via the mobile network communication module to the garbage collector, and preventive action could be taken about the overflowing of garbage in any place. Also, its integration with ThingSpeak could be an extra feature in monitoring the level from time to time by observing the graph, and statistical prediction could be applied for further smartness in the project. In future, there would be an email transfer facility adopted inside the project where the message alert would be received from any place. Our work incorporates many nodes inside the Google Maps so that the shortest path to be taken by the concerned person would be even more flexible.

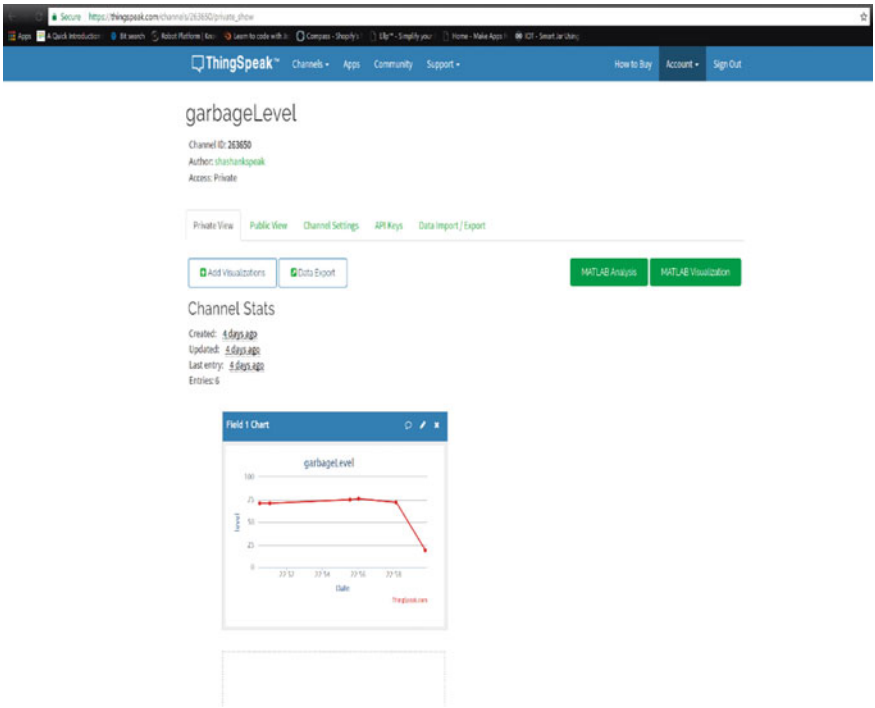


Fig. 6 Level of garbage on the cloud (ThingSpeak)

References

1. S.M. Chaware, S. Dighe, A. Joshi, N. Bajare, R. Korke, Smart garbage monitoring system using Internet of Things (IOT). Int. J. Innov. Res. Electr. Electron. Instrum. Control Eng. ISO 3297:2007 Certified **5**(1) (2017)
2. I. Anoop, A. Jain, S. Pathak, G. Yadav, IOT based Smart Waste Management. Int. J. Adv. Res. Comput. Commun. Eng. ISO 3297:2007 Certified **6**(1) (2017)
3. A. Dev, M. Jasrotia, M. Nadaf, R. Shah, IoT based smart garbage detection system. Int. Res. J. Eng. Technol. **03**(12) (2016)
4. P.V. Parkash, IoT based waste management for smart city. Int. J. Innov. Res. Comput. Commun. Eng. **4**(2) (2016)
5. S.S. Navghane, M.S. Killedar, V.M. Rohokale, IoT based smart garbage and waste collection bin. IJARECE **5**(5) (2016)
6. I. Hong, S. Park, B. Lee, J. Lee, D. Jeong, S. Park, IoT-based smart garbage system for efficient food waste management. Sci. World J. (2014)
7. S.A. Mahajan, A. Kokane, A. Shewale, M. Shinde, S. Ingale, Smart waste management system using IoT. Int. J. Adv. Eng. Res. Sci. (IJAERS) **4**(4) (2017)
8. D. Delling, G.F. Italiano, T. Pajor, Better transit routing by exploiting vehicle GPS data
9. E.E. Ogheneovo, E. Seetam, A heuristic graph-based shortest path algorithm for optimizing routing problems. Am. J. Eng. Res. **5**(11). e-ISSN: 2320-0847 p-ISSN: 2320-0936

Nodal Price Determination for Radial Distribution System Using Load Flow Approach



Ishita Jain and Ashwani Kumar

1 Introduction

Deregulation of the energy industry is the restructuring of the rules and economic incentives that the governments have decided to manage and promote in the energy industry. With the introduction of deregulation, prices of electricity are becoming more important in distribution system. The node price is a method of determining the prices charged at certain clearing prices at multiple locations in the transmission network [1]. It is seen that this technique can also be applied to the distribution system [2]. Nodal pricing is an economically efficient pricing mechanism to determine marginal energy costs and loss costs. The different technical and economical advantages of the integration of distributed generation into distributed system have been studied in the past [3]. In [4], DGs are examined to allot the losses which use direct loss coefficient to analyze the nodal prices. The effect of renewable energy penetration on the nodal price is studied in [5]. In [6, 7], the effect of solar power inclusion on the nodal price is examined. In [8], the author discussed the dynamic tariff concept to determine congestion-related issues which are obtained by regulating the prices at various nodes; this concept is further studied in [9].

The behavior of nodal prices with ZIP load is studied in [10] with the formulation of load flow. In this paper, extension of [10] is done, where nodal price analysis is described considering different load patterns such as ZIP and RIC including the seasonal load fluctuations. This paper is structured into four different sections. Second section narrates the methodology of nodal pricing technique with the use of marginal loss coefficients (MLCs). Third section gives the results and discusses the comparison

I. Jain (✉) · A. Kumar
National Institute of Technology Kurukshetra, Kurukshetra 136119, Haryana, India
e-mail: Ijain467@gmail.com

A. Kumar
e-mail: ashwaks@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_154

of nodal prices considering different load conditions, where the fourth section gives the conclusion of the paper. IEEE 69 bus RDS system is taken as test system in this paper, and corresponding program was instigated and run in MATLAB 7.0.4.

2 Nodal Pricing Methodology

The distribution system's nodal prices can be obtained in the same way as the transmission system. MLCs are calculated first to get the nodal prices on each distribution system bus. MLCs are those coefficients that show incremental or marginal deviation of net active power loss as a result of variations in active and reactive power injections in certain network nodes [11].

$$\rho_{Pi} = \frac{\partial L}{\partial P_i} \tag{1}$$

$$\rho_{Qi} = \frac{\partial L}{\partial Q_i} \tag{2}$$

where L is the net power loss, ρ_{Pi} is active power MLC, and ρ_{Qi} is the reactive power MLC at node i.

The point of connection between the transmission and the distribution systems is called as power supply point. λ is defined as electricity price (USD/MWh) of the active power at power supply point, so the nodal prices of active and reactive power at different nodes are

$$C_P = \lambda + \lambda \cdot \rho_{Pi} = \lambda(1 + \rho_{Pi}) \tag{3}$$

$$C_{Qi} = \lambda \cdot \rho_{Qi} \tag{4}$$

where C_{Pi} and C_{Qi} are the nodal prices for the active and reactive power, respectively, at node i. The reactive power nodal price is considered zero at power supply point.

2.1 Determination of Marginal Loss Coefficients

Jacobian matrix method is used to determine the MLCs [12]. The method was introduced by the author in [4]. It is considered that there is no direct link between the power injections and the power losses. So, the intermediary variable as the magnitude of bus voltages and bus angles which are attained from AC load flow are used to calculate the MLCs.

The following equations describe the abovementioned method:

$$\begin{bmatrix} \rho_{Pi} \\ \rho_{Qi} \end{bmatrix} = \begin{bmatrix} \frac{\partial P_1}{\partial \theta_1} & \frac{\partial P_2}{\partial \theta_1} & \cdots & \frac{\partial P_n}{\partial \theta_1} \\ \vdots & & \ddots & \vdots \\ \frac{\partial P_1}{\partial \theta_n} & \frac{\partial P_2}{\partial \theta_n} & \cdots & \frac{\partial P_n}{\partial \theta_n} \\ \hline \frac{\partial P_1}{\partial V_1} & \frac{\partial P_2}{\partial V_1} & \cdots & \frac{\partial P_n}{\partial V_1} \\ \vdots & & \ddots & \vdots \\ \frac{\partial P_1}{\partial V_n} & \frac{\partial P_2}{\partial V_n} & \cdots & \frac{\partial P_n}{\partial V_n} \end{bmatrix} \begin{bmatrix} \frac{\partial Q_1}{\partial \theta_1} & \frac{\partial Q_2}{\partial \theta_1} & \cdots & \frac{\partial Q_n}{\partial \theta_1} \\ \vdots & & \ddots & \vdots \\ \frac{\partial Q_1}{\partial \theta_n} & \frac{\partial Q_2}{\partial \theta_n} & \cdots & \frac{\partial Q_n}{\partial \theta_n} \\ \hline \frac{\partial Q_1}{\partial V_1} & \frac{\partial Q_2}{\partial V_1} & \cdots & \frac{\partial Q_n}{\partial V_1} \\ \vdots & & \ddots & \vdots \\ \frac{\partial Q_1}{\partial V_n} & \frac{\partial Q_2}{\partial V_n} & \cdots & \frac{\partial Q_n}{\partial V_n} \end{bmatrix}^{-1} \begin{bmatrix} \frac{\partial L}{\partial \theta_1} \\ \vdots \\ \frac{\partial L}{\partial \theta_n} \\ \hline \frac{\partial L}{\partial V_1} \\ \vdots \\ \frac{\partial L}{\partial V_n} \end{bmatrix} \tag{5}$$

$$\frac{\partial P_i}{\partial \theta_j} = V_i V_j \left[G_{ij} \sin \theta_{ij} - B_{ij} \cos \theta_{ij} \right] \tag{6}$$

$$\frac{\partial P_i}{\partial \theta_i} = -B_{ii} V_i^2 - \sum_{j=1}^{n} V_i V_j \left[G_{ij} \sin \theta_{ij} - B_{ij} \cos \theta_{ij} \right] \tag{7}$$

$$\frac{\partial P_i}{\partial V_j} = V_i \left[G_{ij} \cos \theta_{ij} + B_{ij} \sin \theta_{ij} \right] \tag{8}$$

$$\frac{\partial P_i}{\partial V_i} = G_{ii} V_i + \sum_{j=1}^{n} V_j \left[G_{ij} \cos \theta_{ij} + B_{ij} \sin \theta_{ij} \right] \tag{9}$$

$$\frac{\partial Q_i}{\partial \theta_j} = -V_i V_j \left[G_{ij} \cos \theta_{ij} + B_{ij} \sin \theta_{ij} \right] \tag{10}$$

$$\frac{\partial Q_i}{\partial \theta_i} = -G_{ii} V_i^2 + \sum_{j=1}^{n} V_i V_j \left[G_{ij} \cos \theta_{ij} + B_{ij} \sin \theta_{ij} \right] \tag{11}$$

$$\frac{\partial Q_i}{\partial V_j} = V_i \left[G_{ij} \sin \theta_{ij} - B_{ij} \cos \theta_{ij} \right] \tag{12}$$

$$\frac{\partial Q_i}{\partial V_j} = -B_{ii} V_i + \sum_{j=1}^{n} V_j \left[G_{ij} \sin \theta_{ij} - B_{ij} \cos \theta_{ij} \right] \tag{13}$$

The following equation describes the distribution system's total loss:

$$L = \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} G_{ij} \left[V_i^2 + V_j^2 - 2 V_i V_j \cos \theta_{ij} \right] \tag{14}$$

From the equation above, the loss derivative can be found with respect to voltage and angle as follows:

$$\frac{\partial L}{\partial \theta_i} = 2 \sum_{j=1}^{n} V_i V_j G_{ij} \sin \theta_{ij} \tag{15}$$

$$\frac{\partial L}{\partial V_i} = 2\sum_{j=1}^{n} G_{ij}\left[V_i - V_j \cos\theta_{ij}\right] \tag{16}$$

2.2 Reconciliated Marginal Loss Coefficients

Using the above analysis, the total approximate losses can be obtained as follows:

$$L_{approx} = \sum_{i=1}^{n}\left[\rho_{Pi} P_i + \rho_{Qi} Q_i\right] \tag{17}$$

where P_i and Q_i are the active and reactive power injections at node i, respectively. It is observed in [4] that the values of losses were almost double using MLCs than actual losses. Hence, MLCs should be adjusted in an order to evaluate exact price of the losses, which is obtained using reconciliation factor (RF).

$$Rf = \frac{L}{L_{approx}} \tag{18}$$

The new nodal prices for active and reactive power are acquired using reconciliation factor as shown below:

$$C_{Pi} = \lambda(1 + Rf \cdot \rho_{Pi}) \tag{19}$$

$$C_{Qi} = \lambda \cdot Rf \cdot \rho_{Qi} \tag{20}$$

2.3 Load Modeling

The load model which is realistic time-varying and time-dependent on voltage is used in this paper. It is assumed that the load is a combination of industrial, commercial, and residential loads.

$$P_{Di}(t) = P_{Di0}(t)\left[\alpha(t)\left(\frac{V_i(t)}{V_0}\right)^{npr} + \beta(t)\left(\frac{V_i(t)}{V_0}\right)^{npc} + \gamma(t)\left(\frac{V_i(t)}{V_0}\right)^{npi}\right] \tag{21}$$

$$Q_{Di}(t) = Q_{Di0}(t)\left[\alpha(t)\left(\frac{V_i(t)}{V_0}\right)^{nqr} + \beta(t)\left(\frac{V_i(t)}{V_0}\right)^{nqc} + \gamma(t)\left(\frac{V_i(t)}{V_0}\right)^{nqi}\right] \tag{22}$$

where the parameters $\alpha(t)$, $\beta(t)$, and $\gamma(t)$ are the fractions of residential, industrial, and commercial loads at time t, at node *i*, and *npr*, *npc*, and *npi* are varied to obtain various load types as constant current, constant impedance, and constant power. This can be presented by a single entity known as ZIP load [13]. In addition to these load models, seasonal load variation [14] in IEEE 69 bus radial distribution system is also included.

3 Results and Discussions

The IEEE 69 bus RDS is simulated in MATLAB 7.0.4. The real and reactive prices for all the load models described in Sect. 2.3 are calculated. The nodal prices for real and reactive power are obtained for ZIP and RIC load considering seasonal load variations. The coefficients of given load models are given in [13].

The comparison of nodal prices is done by noticing the results for any randomly chosen hour. Figure 1a, b shows the active power and reactive power nodal price comparison for 24th hour of the study with normal ZIP and RIC load. The maximum active power nodal price is observed at bus 65, and the value is 52.1023 USD/MWh. This value is reduced for realistic ZIP and RIC load which is 45.1164 USD/MWh and 44.8395 USD/MWh, respectively. It is observed that the prices are higher in the case of ZIP load compared to RIC load. A similar observation is shown in Fig. 1b for reactive power node prices. Now the nodal prices of active and reactive powers

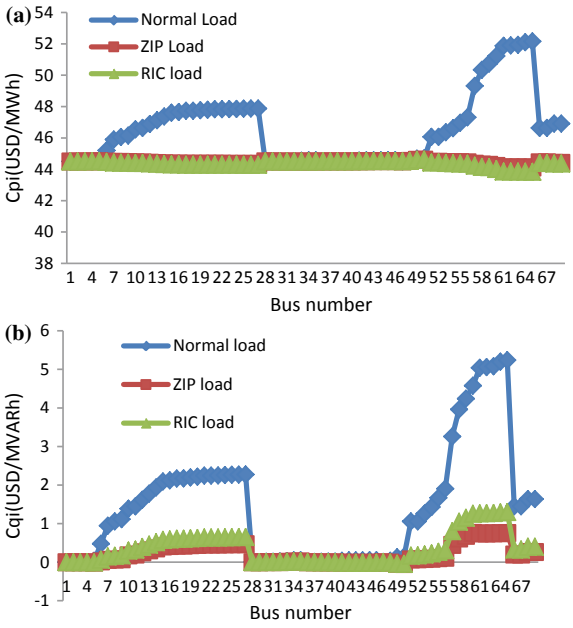


Fig. 1 **a** Active and **b** reactive power nodal price variation at 24th hour

for various seasonal loads for 24th hour are compared in Figs. 2, 3, and 4. In Fig. 2a, comparison of normal load with seasonal load variation is shown, from which it is concluded that active power nodal prices are lowest from bus 1 to bus 5, moderate from bus 27 to 49, and highest from bus 59 to 64. Nodal costs are higher in winter season followed by summer and spring seasons. The same observation is seen in

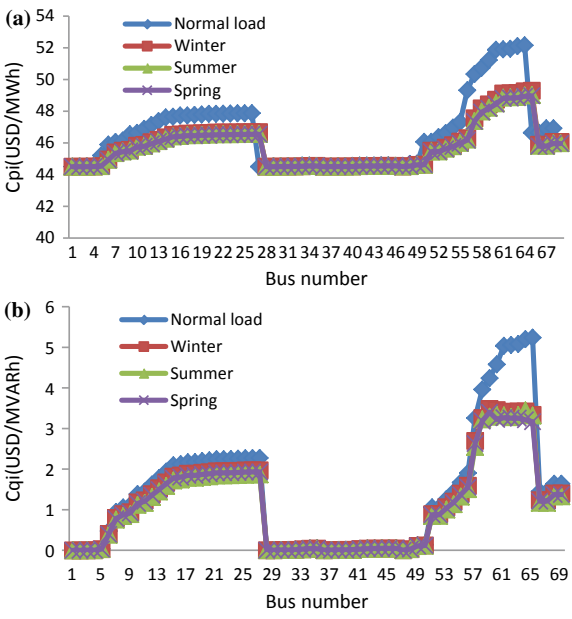


Fig. 2 **a** Active and **b** reactive power nodal prices for normal load considering seasonal variations at 24th hour

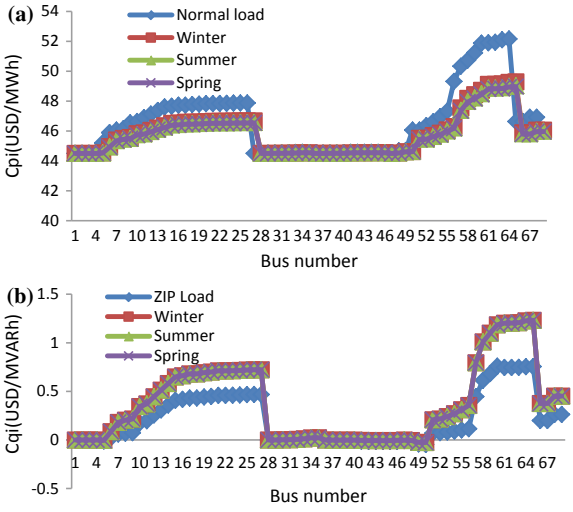


Fig. 3 **a** Active and **b** reactive power nodal prices for zip load considering seasonal variations at 24th hour

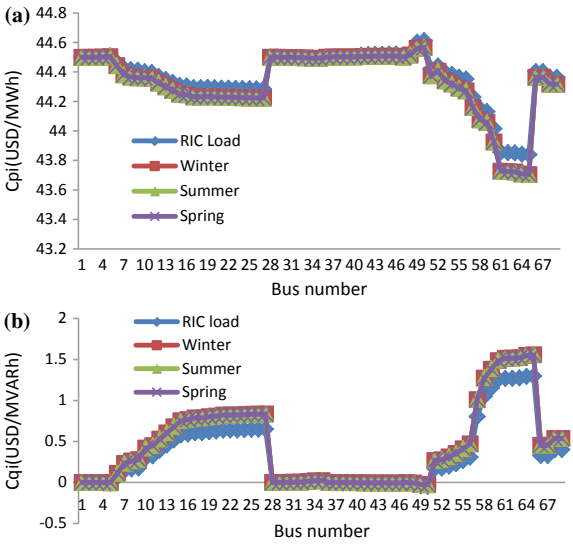


Fig. 4 **a** Active and **b** reactive power nodal prices for RIC load considering seasonal variations at 24th hour

case of reactive power nodal prices for normal load. Figures 3 and 4 show the comparison of nodal prices for realistic ZIP and RIC load, respectively. In these cases, negative values of reactive power nodal prices are also observed in some buses, which represents the supply of reactive power at that particular bus. In Figs. 5 and 6, active and reactive nodal prices are shown for complete 24 h scenario considering seasonal variations.

4 Conclusion

This paper depicts the method for obtaining nodal prices in radial distribution systems, taking into account the marginal loss coefficients of the model. Nodal prices are attained with realistic ZIP and RIC loads. The variation of seasonal load is also considered. It is noticed that nodal prices vary according to load model along with seasonal load changes. It is observed that at peak hours nodal prices are higher for 65th bus and for ZIP load the nodal prices are higher compared to RIC load. For seasonal variations such as summer, spring, and winter seasons, nodal price variations are studied.

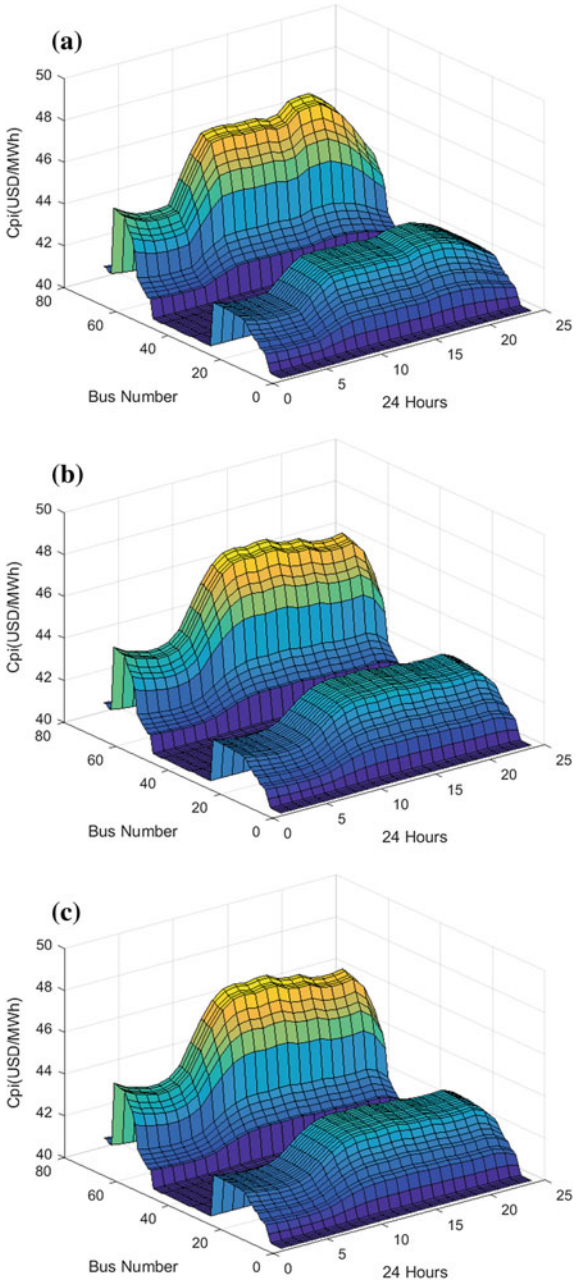


Fig. 5 Active power nodal prices for **a** winter, **b** summer, and **c** spring seasons

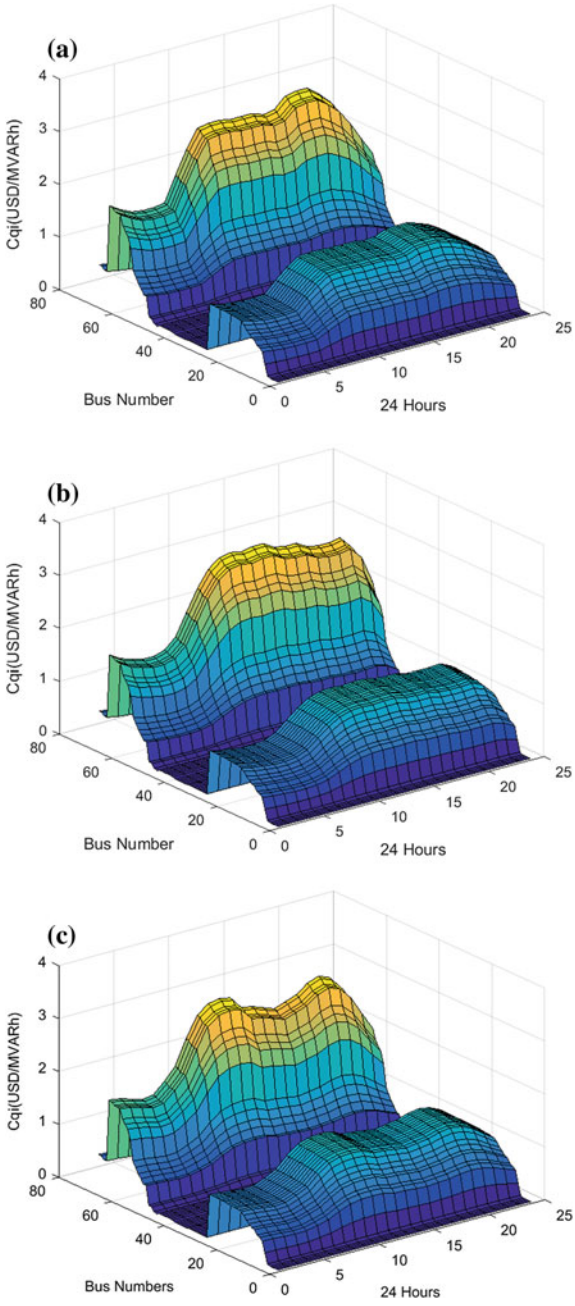


Fig. 6 Reactive power nodal prices for **a** winter, **b** summer, and **c** spring seasons

References

1. F.C. Schweppe, M.C. Caramanis, R.D. Tabors, R.E. Bohn, *Spot pricing of electricity* (Springer, Berlin, Germany, 1988)
2. M. Ghayeni, R. Ghazi, Transmission network cost allocation with nodal pricing approach based on Ramsay pricing concept. IET Gener. Transm. Distrib. **5**, 384–392. F. Author, S. Author, T. Author, Book title, 2nd edn. (Publisher, Location, 1999)
3. L.F. Ochoa, A. Padilha-Feltrin, G.P. Harrison, Evaluating distributed generation impacts with a multiobjective index. IEEE Trans. Power Deliv. **21**(3), 1452–1458 (2006)
4. J. Mutale, G. Strbac, S. Curcic, N. Jenkins, Allocation of losses in distribution systems with embedded generation, in *IEE Proceedings: Generation, Transmission and Distribution*, vol. 147 (2000), pp. 7–14
5. Q. Zhao, P. Wang, L. Goel, Y. Ding, Impacts of renewable energy penetration on nodal price and nodal reliability in deregulated power system, in *IEEE Power and Energy Society General Meeting* (2011), pp. 1–6
6. Q. Zhao, P. Wang, Y. Ding, L.K. Goel, Impacts of solar power penetration on nodal prices and nodal reliability, in *IPEC Conference Proceeding* (2010), pp. 1134–1139
7. P.M. Sotkiewicz, J.M. Vignolo, The value of intermittent wind DG under nodal prices and amp–mile tariffs, in *6th IEEE/PES Transmission and Distribution: Latin America Conference and Exposition (T&D—LA)* (2012), pp. 1–7
8. N. O'Connell, Q. Wu, J. Ostergaard, A.H. Nielsen, S.T. Cha, Y. Ding, Day ahead tariff for the alleviation of distribution grid congestion from electric vehicles. Electr. Power Syst. Res. **92**, 106–114 (2012)
9. R. Li, Q. Wu, S.S. Oren, Distributional locational marginal pricing for optimal electric vehicle charging management. IEEE Trans. Power Syst. **29**, 203–211 (2014)
10. K. Sooraj Narayan, A. Kumar, Distribution system nodal pricing analysis with realistic ZIP load and variable wind power source, in *IEEE India International Conference (INDICON—2015)* (2015)
11. P.M. Sotkiewicz, J.M. Vignolo, Towards a cost causation-based tariff for distribution network with DG. IEEE Trans. Power Syst. **22**, 1051–1060 (2007)
12. P.M. De Oliveira-De Jesus, M.T. Ponce de Leao, Cost loss allocation in distribution networks with high penetration of distribution renewable generation—a comparative study, in *Proceedings of International Conference* on *Renewable Energy and Power Quality*, Zaragoza, Mar 2005
13. V.V.S.N. Murthy, A. Kumar, Mesh distribution system analysis in presence of distributed generation with time varying load model. Int. J. Electr. Power Energy Syst. **62**, 836–854 (2014)
14. https://www.ee.washington.edu/research/pstca/rts/rts79/ieeerts79.txt

Smart Quadripod Walking Stick for the Aid and Security of Visually Challenged and Elderly People



Amitejash Rout, Kshaunish Roy, Swapnil Chhatre, Rhythm Sahu and H. Parveen Sultana

1 Introduction

Humans have five senses: tongue to taste, nose to smell, ears to hear, skin to feel, and finally eyes to see. Our eyesight is one of the most important senses. This is because about 80% of what we sense comes from sense of sight. Life is difficult for those who have difficulty in seeing things, i.e., they are partially or completely visually impaired. Life for such people is very troublesome as they face a lot of difficulties in doing their day-to-day jobs. They usually depend on their walker/blind stick. This assists them to some extent; however, it is quite annoying and tiring to hover the stick on the path ahead and identifying the presence of obstacles. If the person is in some area where the ambient lightings are low or it is dark, the problem for these people further increases. This is because other people nearby may not notice him and collide with him. Also, if there is any kind of emergency situation, he cannot always look out for his mobile phone and send someone his location in panic situation. Often people belonging to this category are visually challenged people and elderly people.

In this paper, smart quadripod walking stick has been proposed and designed for visually impaired people. The stick can identify obstacles in front of it (up to a

A. Rout · K. Roy · S. Chhatre (✉) · R. Sahu · H. Parveen Sultana
Vellore Institute of Technology, Vellore 632014, Tamil Nadu, India
e-mail: chhatre.swapnil@gmail.com

A. Rout
e-mail: amitejashrout@gmail.com

K. Roy
e-mail: kshaunish.roy@gmail.com

R. Sahu
e-mail: rhythmasahu@gmail.com

H. Parveen Sultana
e-mail: hparveensultana@vit.ac.in

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_155

range of 4 m). It is programmed to provide a haptic feedback to the person holding the stick, if the obstacle is at a distance of less than 50 cm. The haptic feedback is sent using a vibrating motor. When the amount of light in the surrounding area is low, the stick automatically starts its LED light as an indication of caution to people around the person using the stick. Also, if the user faces any emergency situation, he can simply press the Save Our Souls (SOS) button located near the handle of the stick. This action will send the location of the stick as an SMS to a specific mobile number. It will be a Google Maps link, generated based on the GPS coordinates. The emergency contact can easily navigate to the location and help the person.

2 Literature Survey

A lot of work has been done in this area; every day new methods are proposed and better system is built. One of the designs for contributing to the above purpose was designed by authors of [1]. The proposed model helps the blind person by working in both indoor and outdoor environment. It is made up of various components like ultrasonic sensor, buzzer, GPS and GSM module, vibratory motor, and Raspberry Pi as a microcontroller. It achieves its objective of helping the blind people by detecting obstacles using the ultrasonic sensor, GPS for navigation, and GSM for sending messages in case of emergency to help the blind person in getting desired help.

The authors of [2] tried to implement a walking stick for the blind by using the ultrasonic sensor, moisture sensor as sensors, and buzzer as the actuator which is implemented on a microprocessor-based model. The purpose of the ultrasonic sensor is both to detect the distance so that it can warn the user of the danger and to inform the user in case the user forgets where he has placed the cane. The only shortcoming of this model is the unused moisture sensor which has so much scope in detecting the surrounding environment and providing more detailed feedback to the user.

In [3], authors have proposed a model which contains an array of ultrasonic sensor. Further sensors include moisture sensor, RF tag, and vibration motor and buzzer as actuators. The collection of ultrasonic sensors are used for detecting potholes, obstacles above the knee, below the knee, and the distance between the user and the obstacle. The moisture sensor is used for identifying water presence. The RF and the buzzer combine together to help the user to find his missing stick in case he misplaces. The only drawback of this model is that it might become much heavy for the user's convenience.

Another model designed by the authors in [4] tried to help the blind persons by using ultrasonic sensors for detecting obstacles of any height in any direction. It also calculates the distance. The person is made alert of the obstacle by the system using haptic feedback in the form of vibrations or in the form of audio signals using earphone. Based upon the intensity of the vibration, type of audio is selected and transmitted to the earphone. The earphone and the walking stick communicate using a Bluetooth module. The whole system is implemented using the Arduino Uno. The

drawbacks of this model are the usage of very few sensors. More sensors could have been used for providing a more realistic view of the surroundings of the user.

The model proposed by the authors in [3] consisted of camera and ultrasonic sensor and headphone for providing the feedback based on the data obtained from the sensors. The entire setup was implemented using Raspberry Pi in which camera captures the surrounding environment of the user and on detection of obstacles warns the user by sending feedback using the headphone output. Distance is measured by using the ultrasonic sensor. The drawback of the model is the addition of headphones which adds to the complexity of the model. The drawback could have been avoided by using a buzzer or haptic feedback.

3 Processors and Sensors Used

3.1 Arduino Uno R3

It is one of the most popular, open-source microcontrollers used for the development of several projects. The fact that it is open source is the main reason for the huge and strong community it has. It costs around Rs. 600 (Fig. 1).

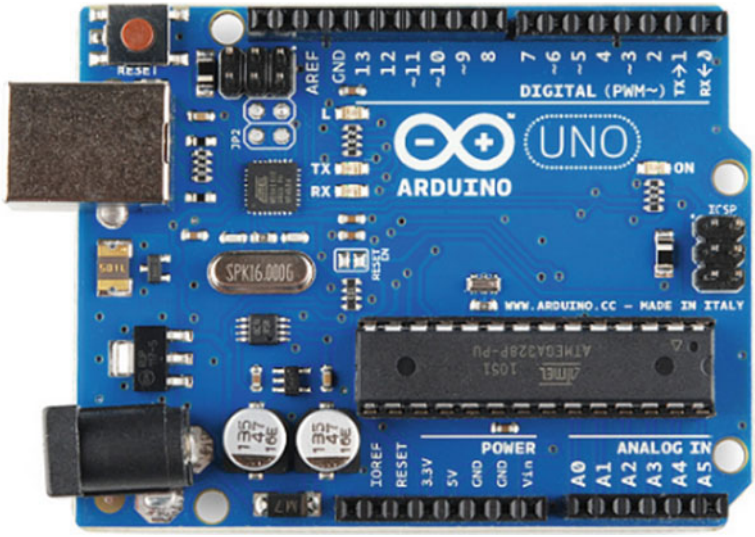


Fig. 1 Arduino Uno R3



Fig. 2 Ultrasonic sensor

3.2 Ultrasonic Sensor

HC-SRO4 ultrasonic sensor is used in this stick. It has a range of 2–400 cm. It requires 5 V in VCC pin (Fig. 2).

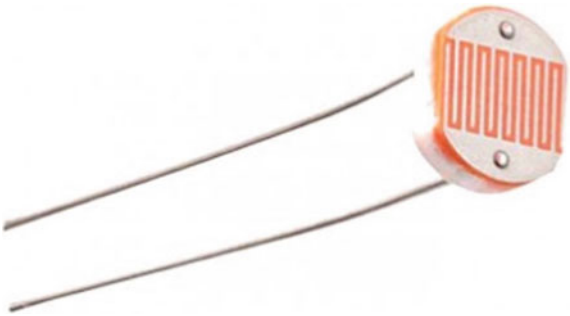
3.3 Vibrating Motor

It is an eccentric rotating mass vibration motor (ERM). The motor uses a small unbalanced weight on a DC motor. When the weight rotates, a force is created that translates to vibrations. They include miniaturized DC coreless motors with eccentric masses, both in coin form and in cylindrical form. It includes a linear resonant actuator, also a few eccentric rotating mass vibration motors. These vibration motors are specifically designed for haptic feedback applications (Fig. 3).



Fig. 3 Vibrating motor

Fig. 4 Photoresistor



3.4 Photoresistor (LDR Sensor)

The photoresistor, as the name suggests, can vary its resistance value depending on the light intensity which is incident on it. When it is dark, resistance is in megaohms, while in bright light, it reduces to a few kilo-ohms (Fig. 4).

3.5 GPS Module

QUESTAR G702-001UB is used in this stick for GPS module. Its working frequency is 1575.42 MHz. It has an accuracy of up to 5 m (Fig. 5).

Fig. 5 GPS module



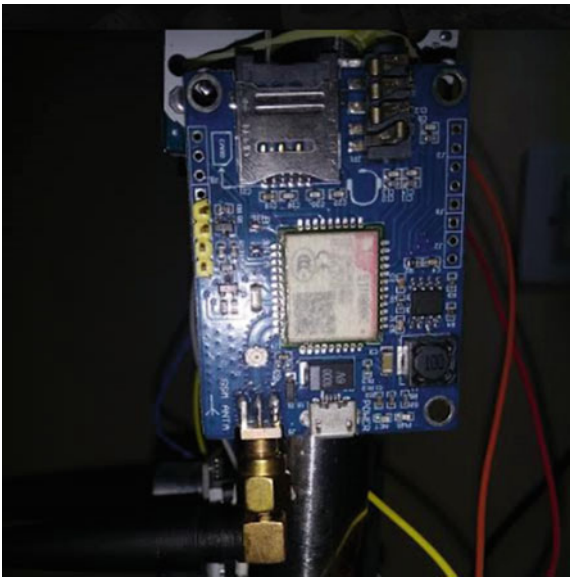


Fig. 6 GSM module

3.6 GSM Module

SIM800 GSM module is used for sending and receiving the messages. It can be powered directly using a micro USB port (Fig. 6).

3.7 Push Button

It is a simple push button, which will be the SOS or the panic button on the stick (Fig. 7).

3.8 Other Electronic Components

- LED,
- Diode,
- NPN transistor,
- Capacitor, and
- Resistors.

Fig. 7 Push button



4 Architecture of Smart Walking Stick

The Smart Quadripod Walking Stick has been broken down the task into three modules.

1. The first is interfacing Arduino, ultrasonic, and vibrating motor. The actuator is the vibrating motor that depends on the value fetched by the ultrasonic sensor.
2. The second module is interfacing LDR and LED sensors. Here, the actuator is LED and it depends on the value provided by the LDR sensor.
3. The last module is interfacing GSM, GPS, and push button (SOS button) module with Arduino.

First, ultrasonic sensor is interfaced with Arduino. The ultrasonic sensor will act as input for the next process, which is interfacing vibrating motor. Depending upon the distance fetched from the ultrasonic sensor, if the distance is less than 50 cm, the vibrating motor will vibrate and thus provide a haptic feedback on the handle of the walking stick. The next component, which is LDR sensor, is interfaced and it acts as an input for deciding if the LED is to be turned ON or remain OFF. LED is the actuator. If the value fetched from the LDR sensor is less than 400 lx, it will consider that the ambience is dark and hence turn ON the LED light. If it is greater than 400 lx, it means that it is bright and hence LED will remain OFF. Finally, the last components, which is the GPS and GSM modules, are integrated. The GPS sensor holds the GPS coordinates. When the SOS button is pressed, the GSM module is activated and it sends an SMS containing the GPS coordinates of the stick to a specific number. Figure 8 shows the connections and overall circuit of the smart quadripod walking stick.

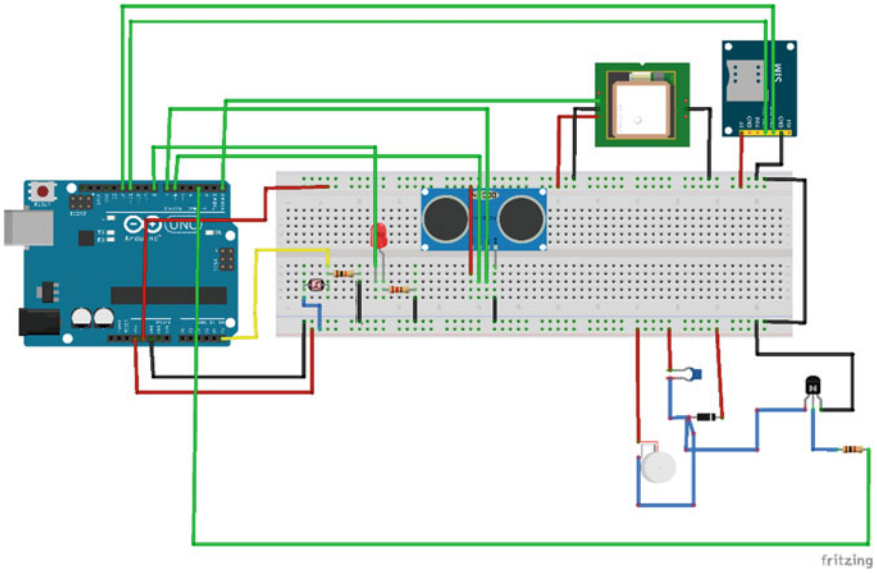


Fig. 8 Architecture of smart quadripod walking stick

5 Methodology

5.1 Interfacing Ultrasonic Sensor and Arduino

- Two male to male jumper wires W1 and W2 are connected to the VCC and GND of the ultrasonic sensor, respectively.
- W1 and W2 are connected to the top part of the breadboard, from which two wires W1′ and W2′ are taken and connected to the 5 V and GND of the Arduino, respectively.
- After that, two male to male jumper wires A1 and A2 are taken and are connected to the TRIG and ECHO of the sensor, respectively
- A1 and A2 are then directly connected to the digital pins 6 and 7, respectively (or any other digital pin is ok).

5.2 Adding Vibrating Motor to Arduino

- Connect diode, capacitor, and vibrating motor in parallel.
- Connect one end of the diode, capacitor, and vibrating motor to 3.3 V power output pin of the Arduino board.

- Connect one end of the diode, capacitor, and vibrating motor to end to N-end of NPN transistor.
- Connect P-end of the NPN transistor to pin 3 of the Arduino board with a resistor in between.
- Connect the second N-end of the NPN transistor to GND pin of the Arduino board.

5.3 Adding LDR and LED to the Circuit

- Connect 3.3 V power and GND connections to the Arduino.
- Place the LDR (Light-Dependant Resistor) sensor on the breadboard.
- Place a 10 KΩ resistance, such that it shares a common leg with the LDR sensor.
- Connect A5 (or any other analog pin) of Arduino at the junction with 10 K resistor.
- Connect the other end of the 10 K resistor to the negative rail of the breadboard.
- Place LED on breadboard.
- Connect any of the digital pin of Arduino (say pin8) to the positive (Long leg) of the LED.
- Place the 220 Ω resistor such that it shares a common junction with the negative leg of LED.
- Connect negative rail with the other end of 220 Ω resistor (Black wire), i.e., Ground.

5.4 Interfacing GSM and GPS Module with the Arduino

- The VDD of the GPS module is connected to 5 V supply.
- The GND and the PPS pins of the GPS module are connected to ground.
- The Tx pin of GPS module is connected to pin1(Rx) of the Arduino.
- Remove the Tx line while uploading the code.
- This wire sends the data received by the GPS to the Arduino.
- The Rx of GSM is connected to digital pin7(Tx) of Arduino.
- The Tx of GSM is connected to digital pin8(Rx) of Arduino.
- The GND of GSM is connected to GND of Arduino.

5.5 Interfacing Push Button

- Place on leg the 10 KΩ resistor and the wire pointed with green arrow in parallel with one of the sides of the push button.
- Connect the wire shown with green pointer to any of the digital pin of the Arduino.
- Place the wire shown with blue pointer to another leg of the resistor and connect it to the ground pin or negative rail.
- Place the wire shown with the orange pointer to another side of a push button and connect it to the 5 V. Positive rail or the 5 V pin of the Arduino.
- When the button is not pressed and the circuit is open, the current flows from 5 V to GND.
- When the button is pressed and the circuit is closed, the current flows from 5 V to any digital pin of Arduino.

5.6 Assembled Hardware of the Smart Walking Stick

See Fig. 9.

6 Comparison with the Existing Models

Table 1 gives a comparative analysis of similar models.

7 Results and Discussions

7.1 Ultrasonic Sensor and Vibrating Motor

The ultrasonic sensor determines the distance between the stick and the obstacle and returns it to the module handling the vibrating motor. If the distance becomes less than 50 cm, haptic feedback is triggered which can be felt on the handle of the stick (Fig. 10).

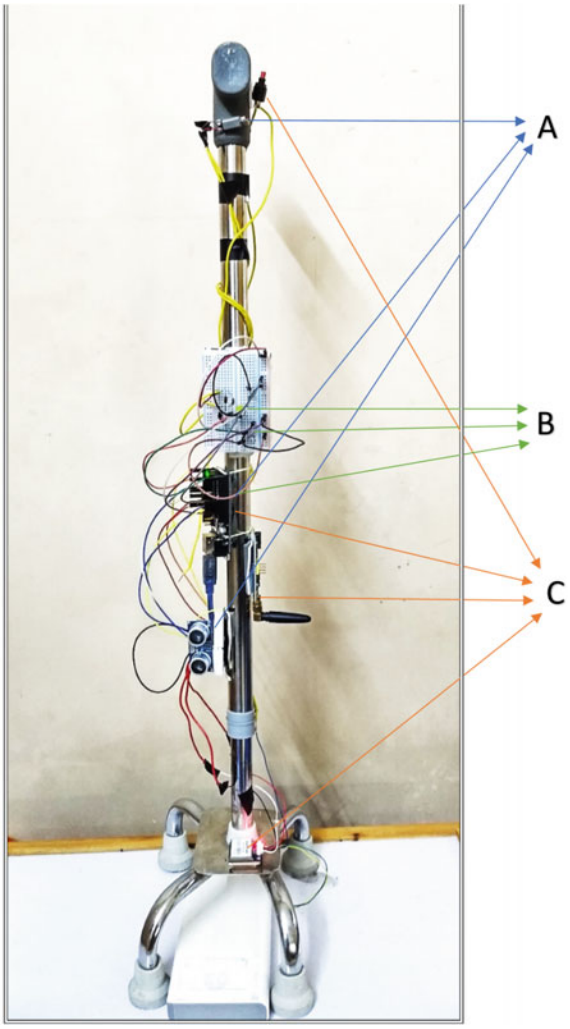


Fig. 9 Module A contains a vibrating motor and ultrasonic sensor. Module B contains an LED and an LDR. Module C contains GPS module, GSM module, and a push button. All the modules are interfaced with the Arduino board

7.2 Photoresistor and LED

When the intensity of light falling on the stick is less than 400 lx, it causes the LED to glow and when it falls below 400 lx, the LED turns OFF. The result is shown in Fig. 11.

Table 1 Comparison with the existing model

Authors	Approach	Processor/Microcontroller used	Sensors used	Actuators used	Drawbacks
Dabir et al. [5]	The ultrasonic sensor warns the user about obstacles; the three-axis accelerometer is used for fall detection and if the pulse rate is abnormal, then it sends the GPS location of the user to the family members using GSM–GPRS module	Raspberry Pi 3b	Ultrasonic sensors, three-axis accelerometer, pulse sensor, GSM–GPRS module, GPS module	Push button, buzzer	The fall detection system is prone to a number of false positives, for example, if the user accidentally drops the stick
Sen et al. [6]	A body-mounted artificial navigation system which detects obstacles using ultrasonic sensors and GPS module for indoor and outdoor navigations	Not mentioned	Ultrasonic sensors, GPS module	Buzzers	Not suitable for novice users. Sensing accuracy can be affected due to change in temperature
Dey et al. [7]	The stick has ultrasonic sensors' obstacle detection	PIC microcontroller 16F877A	Ultrasonic sensors	Buzzer	The range for detecting obstacles is too small (5–35 cm)

(continued)

Table 1 (continued)

Authors	Approach	Processor/Microcontroller used	Sensors used	Actuators used	Drawbacks
Mala et al. [8]	The system uses a navigation gadget and a Bluetooth headset to guide the user	ATMEGA microcontroller	Ultrasonic sensors, GPS module	Bluetooth headset	Not suitable for people with hearing impairment
Mohapatra et al. [9]	The ultrasonic sensor detects obstacles and the microcontroller sends an alert to the user. On pressing the distress button, the stick streams the video call on an Android phone	Raspberry Pi 3b	Ultrasonic sensors, Raspberry Pi Camera	Push button, Android App	Requires internet connection all the time

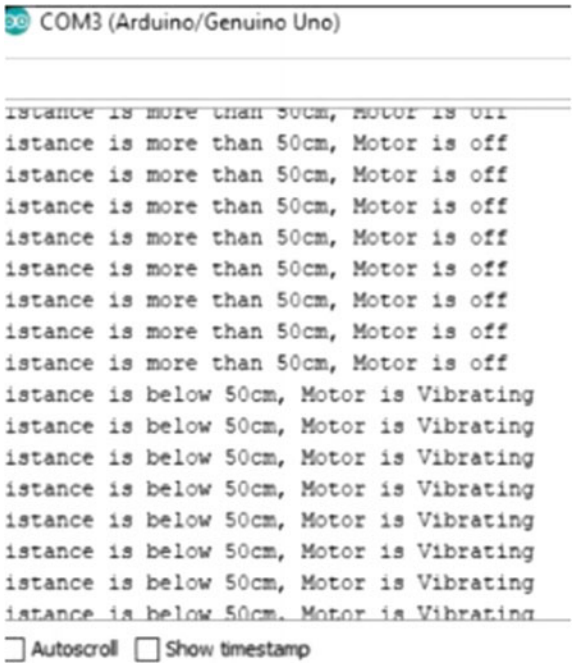


Fig. 10 Results from ultrasonic sensor and vibrating motor

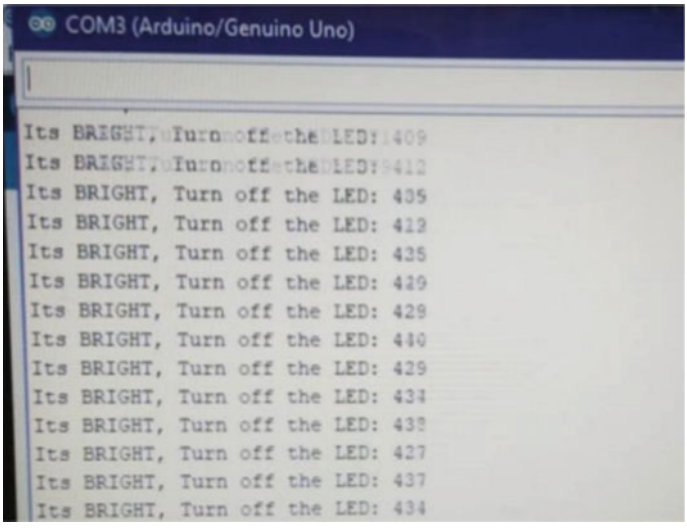


Fig. 11 Result from LDR and LED

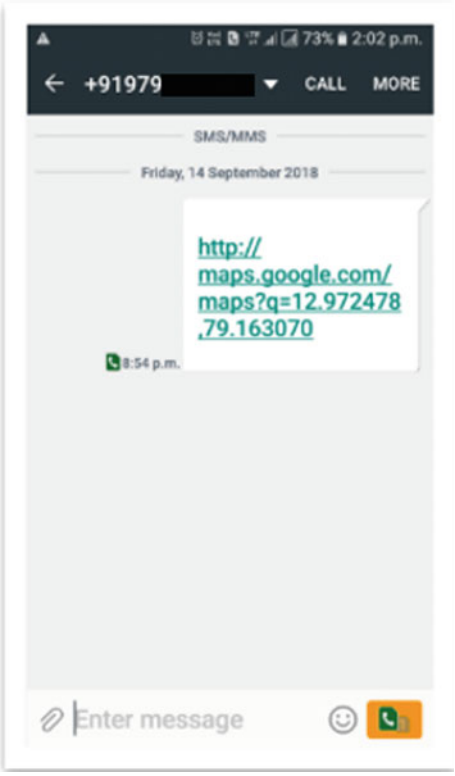


Fig. 12 SMS received from GSM module

7.3 GSM Module and GPS Module

Once the panic button is pressed, the SOS message containing the GPS coordinates of the location of the stick would be sent to the registered number. Figure 11 shows the SMS received from the stick. It will be a link containing the GPS coordinates. On clicking the link, it would prompt to open it on Google Maps. Figure 12 shows the coordinates opened on Google Maps (Fig. 13).

7.4 Setting Up ThingSpeak

The data is collected from the smart quadripod walking stick and is analyzed using Thingspeak (Fig. 14).

The name of the channel is "Smart_Quadripod_Walking_Stick".

The field values entered are stored in the cloud. The four field values entered were obstacle_distance, vibration_status, LDR_status, and LED_status (Fig. 15).

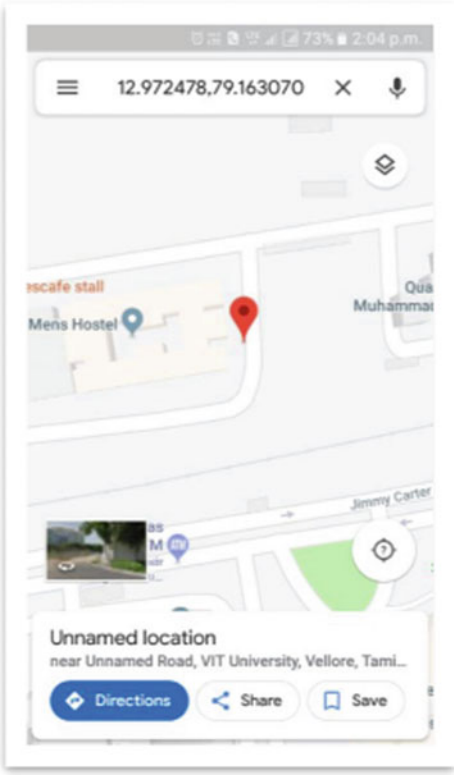


Fig. 13 GPS coordinates opened on Google Maps

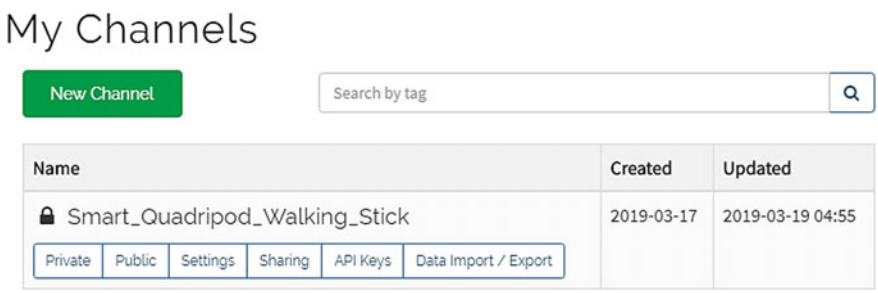


Fig. 14 Channel for analysis

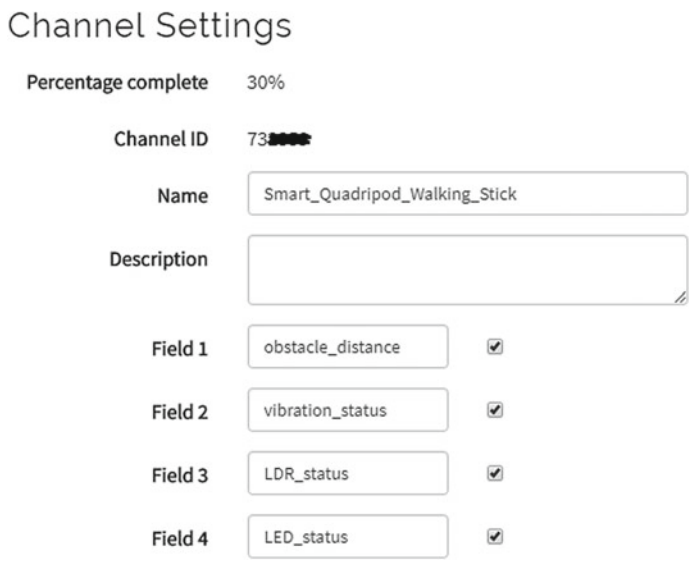


Fig. 15 Channel settings

```
response = urlopen('https://api.thingspeak.com/update?
api_key=9[redacted]VKE&field1='+str(distance)+'&field2='+str
(Vibration_status)+'&field3='+str(LDR_status)+'&field4='+str
(LED_status))

html = response.read()
```

Fig. 16 Code snippet to upload data from stick to ThingSpeak

The API was added to python code (Fig. 16).

Upon execution of the code, the data is uploaded to the cloud for analysis purposes. Values in the cloud are updated, and graph is plotted displaying the working of the components together.

7.5 Analysis

Figures 17 and 18 show how the ultrasonic sensor and the vibrating motor work together to warn the visually impaired user about an obstacle ahead. The threshold distance to trigger the vibrating motor is 50 cm, below which the vibrating motor will vibrate rapidly. Upon careful inspection of Figs. 17 and 18, a few seconds past 19:22, the ultrasonic sensor senses an object which is closer than 50 cm, leading to

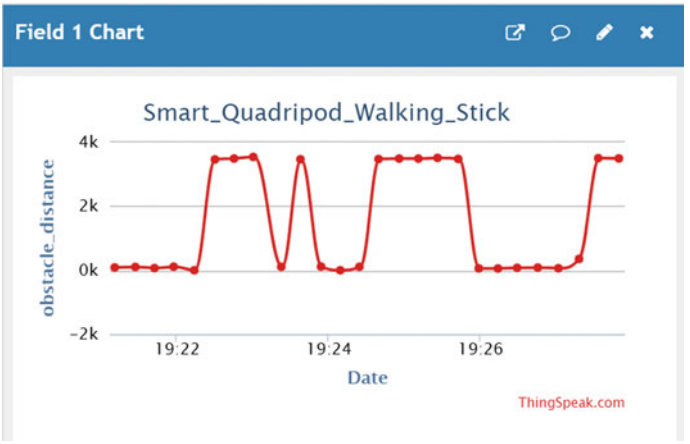


Fig. 17 Distance from obstacle graph

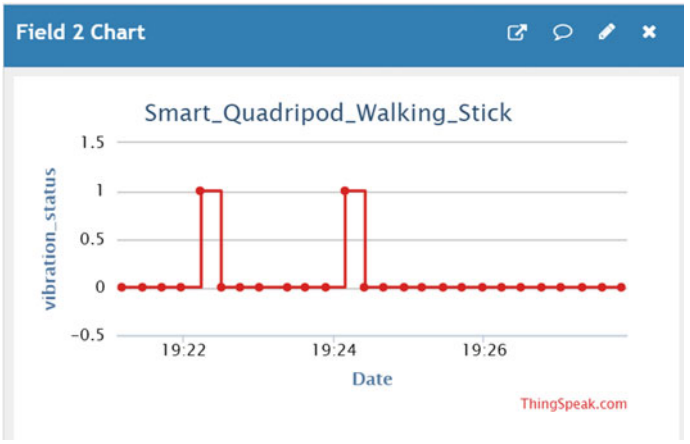


Fig. 18 Status of vibrating motor versus time

the triggering of vibrating motor. But after that, the distance from the object increases to more than 50 cm, causing the vibrating motor to turn OFF. Similarly, few seconds past 19:24, the ultrasonic sensor again senses an object which is closer than 50 cm, leading to the triggering of the vibrating motor. But after that, the distance from the object increases to more than 50 cm, causing the vibrating motor to turn OFF.

Figures 19 and 20 show how the LDR and the LED work together to warn others around the visually impaired user about the user's presence in a low light environment. The threshold luminance to turn ON the LED is 400 cd per unit square meter (cd/m^2), below which the LED will be ON. Upon careful inspection of Figs. 19 and 20, a few seconds before 19:22, the luminance is above 400 cd/m^2, so the LED turns OFF.

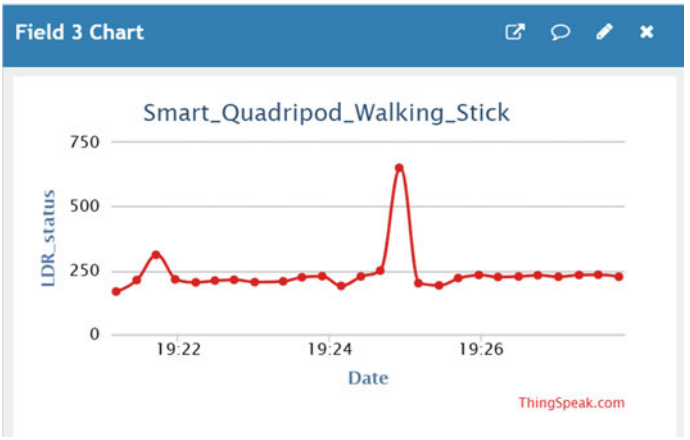


Fig. 19 Luminance versus time graph

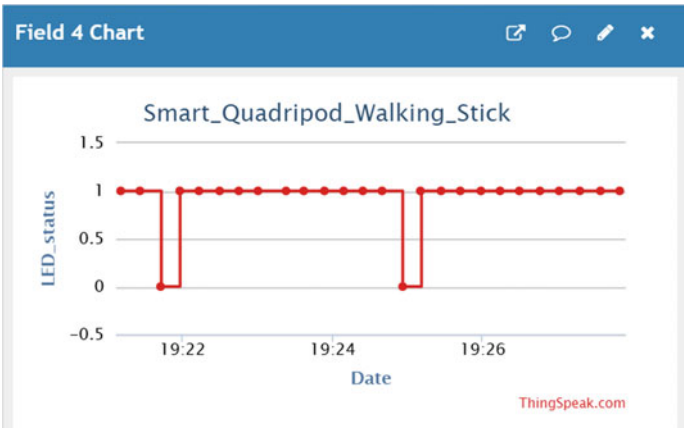


Fig. 20 LED status versus time

But after that, the luminance drops below 400 cd/m^2, so the LED turns ON again. Similarly, just before a few seconds before 19:25, the luminance is above 400 cd/m^2, so the LED turns OFF. But after that, the luminance drops below 400 cd/m^2, so the LED turns ON again.

8 Conclusion

The proposed smart quadripod walking stick is designed and implemented successfully. The ultrasonic sensor can detect objects coming in front of it and provides a

haptic feedback to the user when an object gets closer than 50 cm to the stick. The ultrasonic sensor has a range of 1 cm to 400 cm. The haptic feedback is provided by the vibrating motor, which is mounted near the handle of the walking stick. As long as the object is at a distance less than 50 cm, it will continue to vibrate. The LDR sensor detects the ambient lighting conditions and determines the action of LED. 400 lx has been taken as the threshold value. When the light intensity is less than 400 lx, it is assumed that ambience is getting dark and hence turns the LED light ON. When it is light intensity is greater than 400 lx, it turns the LED light OFF. The panic button is mounted on the handle of the stick. The GPS sensor is always powered and has locked the GPS coordinates. When it is pressed, the locked GPS coordinates are sent to the registered number using GSM module.

9 Recommendations

People may stumble into the user not just from the front but also may come sideways. This often occurs in a busy street or a crossroad. Movement of people from sideways can pose a problem to the blinds.

Attaching an extra ultrasonic sensor can resolve this issue. It would be mounted in such a way that it will be rotating and covering an angle of 45° in clock and anticlockwise direction which can be used to identify the obstacles coming toward the user from oblique directions, thus providing a better assessment of objects which are not only directly in front of him but also which could probably come toward him sideways. Sometimes, the user might forget where exactly he has kept his walking stick. Addition of a buzzer can resolve it. In this case, he may send an SMS to number, which is in GSM module of the stick, which will cause the buzzer on the stick to buzz and the LED to glow.

References

1. A. Gupta, A. Singh, D. Upadhyay, Rohtash, Smart blind walker stick. Int. J. Adv. Res. Innov. Ideas Educ. **4**(2), 2364–2368 (2018)
2. V.V. Nahar, J.L. Nikam, P.K. Deore, Smart blind walking stick. Int. J. Modern Trends Eng. Res. **2349**, 645–652 (2016)
3. H. Sharma, M. Tripathi, A. Kumar, M.S. Gaur, Embedded assistive stick for visually impaired persons, in *2018 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT)* (IEEE, 2018), pp. 1–6
4. S. Sharma, M. Gupta, A. Kumar, M. Tripathi, M.S. Gaur, Multiple distance sensors based smart stick for visually impaired people, in *2017 IEEE 7th Annual Computing and Communication Workshop and Conference (CCWC)* (IEEE, 2017), pp. 1–5
5. A. Dabir, R. Solkar, M. Kumbhar, G. Narayanan, GPS and IOT equipped smart walking stick, in 2018 *International Conference on Communication and Signal Processing (ICCSP)* (IEEE, 2018), pp. 0322–0326

6. A. Sen, K. Sen, J. Das, Ultrasonic blind stick for completely blind people to avoid any kind of obstacles, in *2018 IEEE Sensors* (IEEE, 2018), pp. 1–4
7. N. Dey, A. Paul, P. Ghosh, C. Mukherjee, R. De, S. Dey, Ultrasonic sensor based smart blind stick, in *2018 International Conference on Current Trends towards Converging Technologies (ICCTCT)* (IEEE, 2018), pp. 1–4
8. N.S. Mala, S.S. Thushara, S. Subbiah, Navigation gadget for visually impaired based on IoT, in *2017 2nd International Conference on Computing and Communications Technologies (ICCCT)* (IEEE, 2017), pp. 334–338
9. S. Mohapatra, S. Ham Rout, V. Tripathi, T. Saxena, Y. Karuna, Smart walking stick for blind integrated with SOS navigation system, in *2018 2nd International Conference on Trends in Electronics and Informatics (ICOEI)* (IEEE, 2018), pp. 441–447

Real-Time Tracking and Lane Line Detection Technique for an Autonomous Ground Vehicle System



Tamal Datta, S. K. Mishra and S. K. Swain

1 Introduction

The vision-based ground vehicle navigation and guidance system is a major area of interest among researchers from various fields as well as among different industries. Application of such system ranges from auto-parking system, industrial robot control to self-driving car, rocket carrier, etc. Major parts of such systems are tracking, lane detection, path generation, road sign recognition, controller design, etc. The sophistications of technology behind these applications are not same and vary from each other, for example, all the subsystems of self-driving car have to be more robust and precise [1] than the subsystems employed for traffic monitoring.

The work presented here is having two major parts of self-driving car design project; therefore, we have chosen the simplest and fastest approaches to avoid all kinds of synchronisation problems. We have proposed techniques for real-time tracking of vehicle and lane line detection. Mainly Lidar and ultrasonic sensor are used in traditional intelligent driving systems. However, these traditional techniques are less reliable [2]. The more convenient way is computer vision and image processing-based techniques [3]. Advantages of computer vision-based vehicle detection techniques are presented in [4]. Here, robustness and less processing time of image-based techniques have been emphasised. Hidden Markov model-based approach to distinguish foreground and shadow, followed by particle filter-based tracking, has been

T. Datta (✉) · S. K. Mishra · S. K. Swain
Birla Institute of Technology, Mesra, Ranchi, India
e-mail: ttamalddatta@gmail.com

S. K. Mishra
e-mail: sudhansu.nit@gmail.com

S. K. Swain
e-mail: swain.subrat01@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_156

presented in [5]. In [6], three-step Kalman filter-based approach for position prediction is introduced. Here, prediction is done on recorded video and its applicability on real-time video is not guaranteed. In [7], the authors introduced modified background subtraction with 'low-rank + sparse' decomposition method. Real-time algorithm for pedestrian tracking is presented in [8]; here, blob tracking has been done using extended Kalman filter and hardware implementation has been performed using static roadside camera. In [9], the particle filter is applied for robust detection and tracking of object by considering the colour information. Gaussian distribution-based pixel modelling is used in [10]; this approach works fast for large image, and also hence suitable for real-time applications.

Lane line detection on urban street is presented in [11]; here, detection is done using top-view images. RANSAC algorithm has been used. In [12], colour information-based lane boundary segmentation is presented. Here, a threshold value is applied to distinguish lane lines after getting region of interest. This approach requires less computation, and hence is easy for formulation and fast in operation. In [13], lane detection using Hough transform and smoothing of the output of line tracker by Kalman filter is presented. Hough transform-based lane boundary detection with a parabolic model is presented in [14]. Here, a multi-iteration strategy is used to improve conventional Hough transform. In [15], the authors presented detailed process of Hough transform to detect different shapes like circle, line, parabola, etc. The authors provide a review about efficient implementation of Hough transform in [16]. In paper [16], the relation between Hough transform and other Hough transform-based transform methods is compared.

To address few modifications in existing approaches, we have proposed the following sequels of steps. To detect the car in real-time frame, background subtraction and for suppressing noise, median filtering have been used. Here, we have used fixed threshold value because most of the thresholding techniques behave unpredictably in case of vacant image frame (when no object is there in image frame). Furthermore, fixed threshold value reduces the execution time. We have found 15 as suitable threshold value for our purpose. We have employed Kalman filter for the estimation of car's spatio-temporal coordinates. We have enhanced processing speed and reduced computational complexity by tracking the centroid of the detected vehicle instead of tracking entire vehicle. Here, the velocity of car is measured in terms of pixels per image frame. Velocity of vehicle is calculated from the difference between current position of centroid and previous position of centroid. For lane detection, we have used an onboard camera to capture the front road. To get the region of interest, we have used a triangular-shaped mask function. Conventional Hough transform fails in case of infinite slope and therefore we have employed Hough transform in polar form.

Implementation of proposed approach is carried out on a robotic car of 2 feet by 1 feet dimension in laboratory environment. All the image acquisition is done using 5 MP webcam. Proposed approach results in reliable tracking and proper detection of lane lines. Our approaches have plenty of applications like vehicle navigation, industrial robot control, etc. This paper is organised as follows: Sect. 2 presents the overview of our project. In Sect. 3, all the subparts of vehicle tracking are presented.

It consists of details of object detection and Kalman filter-based tracking. In Sect. 4, details of lane detection method are elaborated. In Sect. 5, algorithms of our proposed approach are presented. Hardware details and analysis of results are presented in the next following sections.

2 Overview of Proposed Approach

With the vision to navigate a car towards its destination, a local positioning system has been considered. A roadside static camera has been used to detect and track the position of our car, in sequence of image frames, and hence, the local position of car is found. After getting current position of car, we aim to guide it to the predefined destination point. In each image frame, pixel position of destination point and pixel position of current location are known; using this information, an optimised path between these two points can be drawn. If the car be aligned along that generated path, it will reach its destination. In case of real-world scenario, car should not go off-road or out of lane lines; therefore, we considered another onboard camera to take care of lane lines. If in any case car moves out from its lane lines, then it will be detected by onboard camera and required decisions can be taken.

In Fig. 1, camera (1) is the static camera and camera (2) is the onboard camera. Corresponding lines represent the range of camera. Local position of b is predefined. Detection and tracking of car are performed using camera (2). Hence, position of 'a' is found out. After pixel-wise path from 'a' to 'b' is found, car needs to be aligned to the pixel positions of that defined path. Importance of camera (2) comes in picture

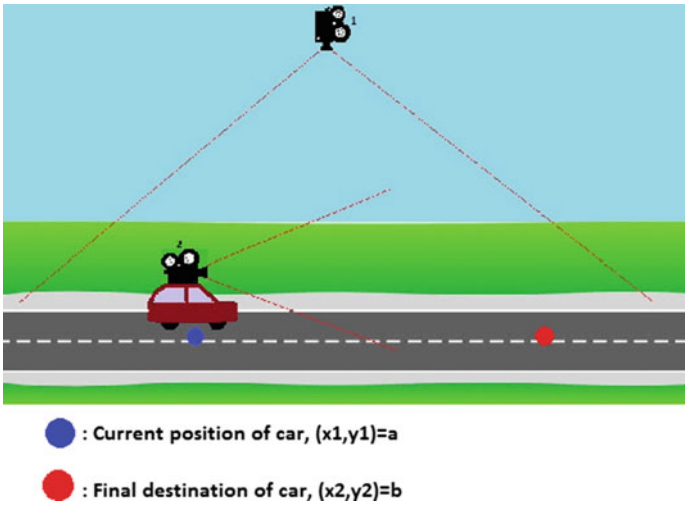


Fig. 1 Overview of approach

if the generated path guides car to go off-road. Camera (2) checks the lane lines in sequence of each image frame and guide car to stay in proper lane.

In this paper, two major parts of proposed approach have been explained. In part one, a method for real-time detection and tracking of car using static camera has been performed. Similarly, in part two, real-time lane line detection has been performed using onboard camera.

3 Tracking of Car Using Roadside Static Camera

3.1 Object Detection

Spatial methods are the easiest methods for identifying region of interest (ROI) in image frame [17]. Few well-known techniques of this category are optical flow field, grey comparison, frame difference, background subtraction, edge detection, etc. Optical flow method holds good when there is a relative motion between camera and object of interest. This method gives a velocity vector between camera and ROI. In this project, static roadside camera has been considered; furthermore, mentioned technique consists of very complex calculation, and hence, it is not suitable for real-time application. Therefore, this technique is not used in our project. More convenient technique for our purpose is frame difference. It is just the subtraction of two successive image frames followed by thresholding. This approach results in too much noise unless pre- and post-processing are done properly. Edge detection is purely gradient-based method, i.e. the gradient between two adjacent pixels is used to extract the demarcation lines between object and background. Robert edge detection, Sobel edge detection and Canny edge detection are few popular edge detection algorithms. We found simple background subtraction seemed to be the easiest technique to serve our purpose. Unlike frame difference, background subtraction is not sensitive to speed of vehicle. Furthermore, any static object can be detected using this approach (Fig. 2).

To make the background subtraction method adaptive, methods like histogram, Gaussian average, average, etc. can be applied. In histogram approach, most frequent value of a pixel in an image is considered as the background value of that particular pixel. Average method performs in more convenient way than other methods. Furthermore, it is faster; thereby average method has been adapted in this project [4, 18].

$$B_{i+1} = (1 - \beta)B_i + \beta F_i \tag{1}$$

The above equation defines the average method where B is the background, F is the foreground and β is the adaptation rate. Speed of adaptation increases with the increase of β. Finally, the ROI is found by

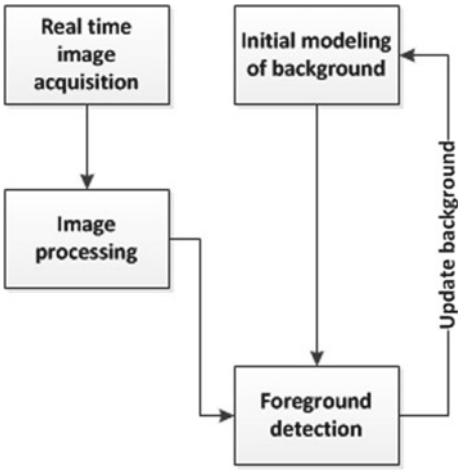


Fig. 2 Block diagram of background subtraction

$$F_i(x, y) = \begin{cases} 0, \ d_i(x, y) < Threshold \\ 1, \ d_i(x, y) > Threshold \end{cases} \tag{2}$$

where

$$d_i(x, y) = |I_i(x, y) - B(x, y)|, \ I_i \text{ is } i\text{th image frame.} \tag{3}$$

After the detection of foreground, noises in form of blobs were generated, and median filter has been used to take care of this.

3.2 Tracking Using Kalman Filter [19–21]

Object tracking is the method of estimating the states of a system at any instance in the video. The initial instance of the object is predefined. Kalman filter (KF) and Kanade–Lucas–Tomasi (KLT) are few widely used algorithms for mentioned purpose. In this project, KF has been used for the estimation of car position. Kalman filter is a set of linear unbiased minimum error covariance sequential state estimation algorithms. It is used to estimate the state of linear system, where state of the system is distributed by a Gaussian. The main concern of this project is to estimate the position of the car; therefore, the state matrix (iv) consists of position and velocity along x- and y-axes.

$$X(K) = \left[x(k) \ y(k) \ \dot{x}(k) \ \dot{y}(k) \right] \tag{4}$$

Detected car consists of many pixels and tracking of each pixel is not possible because it may lead to degradation of performance. Therefore, only centroid pixel has

been tracked instead of all pixels. Car moves on a 2D surface; therefore, its motion has been explained by equation of 2D motion (v).

$$X(k + \Delta T) = X(k) + V * \Delta T + a * \Delta T^2 \tag{5}$$

This project considers the car to be non-accelerating; therefore, a is zero in this case. Final system matrix of the vehicular motion is (vii). State transition is modelled by

$$X(k) = AX(k-1) + BU(k) + W(k) \tag{6}$$

$$A = \begin{bmatrix} 1 & 0 & \Delta T & 0 \\ 0 & 1 & 0 & \Delta T \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{7}$$

$$B = \begin{bmatrix} \Delta T^2 & 0 \\ 0 & \Delta T^2 \\ \Delta T & 0 \\ 0 & \Delta T \end{bmatrix} \tag{8}$$

$$U = \begin{bmatrix} a_x \\ a_y \end{bmatrix} \tag{9}$$

Here, U is considered zero as the acceleration of car is taken zero. Here, estimation is done at each frame, and therefore, ΔT is taken as 1. Initial covariance matrix (Q) is given by

$$Q = B \begin{bmatrix} \Delta x^2 & 0 \\ 0 & \Delta y^2 \end{bmatrix} B^T \tag{10}$$

The overall algorithm of Kalman filter is given by (Fig. 3):
where U is the control variable matrix, W is the state noise matrix, Q is the process noise covariance matrix, K is the Kalman gain, R is the sensor noise covariance matrix (Measurement error), X is the state matrix and P is the process covariance matrix. $X0$ and $P0$ represent the state matrix and process covariance matrix, at initial condition (Fig. 4).

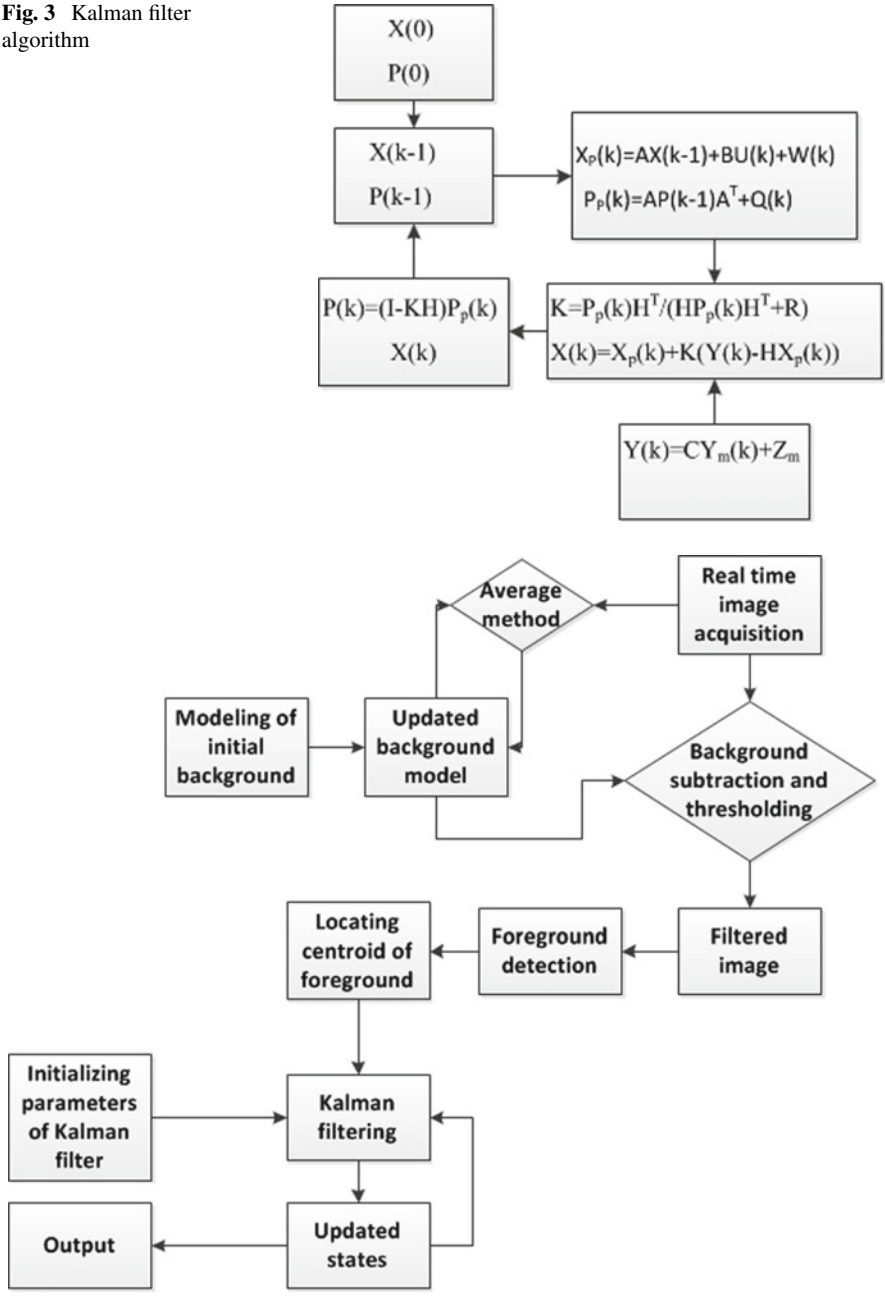


Fig. 3 Kalman filter algorithm

Fig. 4 Flowchart of our approach for vehicle tracking

4 Lane Line Detection

Lane line detection is done using onboard camera. Camera captures image of the road ahead; this image is then processed to distinguish lane lines from the entire image. Pixel position of these lane lines guides processor to generate relevant signals that keep the car on the lane.

4.1 Required Pre-processing

Captured multichannel image requires more processing power. Therefore, conversion of image to single-channel greyscale image is done followed by canny edge detection algorithm.

4.2 Mask Generation and ROI

Mask is a triangular-shaped white patch on black background. Road ahead of vehicle usually takes shape similar to triangle for which a triangular mask has been considered. The size of mask image is same as the image captured. In mask, all pixels of white patch have value 255, and all the pixels of black portion have value 0. Thereby, bitwise logical AND operation between this mask and canny image results in the region of interest (Fig. 5).

Region of interest consists of broken parts of lane lines, that is, all pixels of lane line are not identified.

0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	255	0	0	0	0	0	0
0	0	0	0	0	0	255	255	255	0	0	0	0	0
0	0	0	0	0	255	255	255	255	255	0	0	0	0
0	0	0	0	255	255	255	255	255	255	255	0	0	0
0	0	0	255	255	255	255	255	255	255	255	255	0	0
0	0	255	255	255	255	255	255	255	255	255	255	255	0
0	255	255	255	255	255	255	255	255	255	255	255	255	255



Fig. 5 Pixel values of triangular mask

4.3 Hough Transform [15, 16]

Any straight line ($y = mx + c$) can be represented as $p = x\cos\theta + y\sin\theta$ in polar form. In Cartesian coordinate, any two points are considered (Fig. 6); there are many lines that pass through each point. All lines passing through each point takes a sinusoidal shape in respective polar plane. If two such sinusoids in polar plane intersect at a point (point C in Fig. 7), then the straight line with that p and θ value intersects both points (points a and b in Fig. 6) in Cartesian plane. Similar is applicable to multiple points also, that is, using Hough transform we can get parameters of a straight line that intersects multiple points, for example, if we know the coordinates of points a

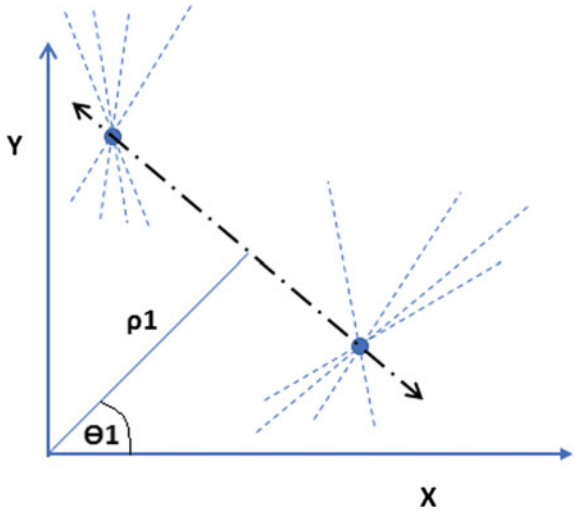


Fig. 6 Cartesian plane

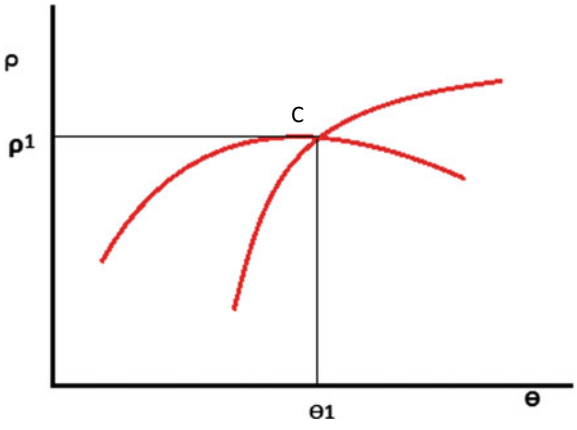


Fig. 7 Polar plane

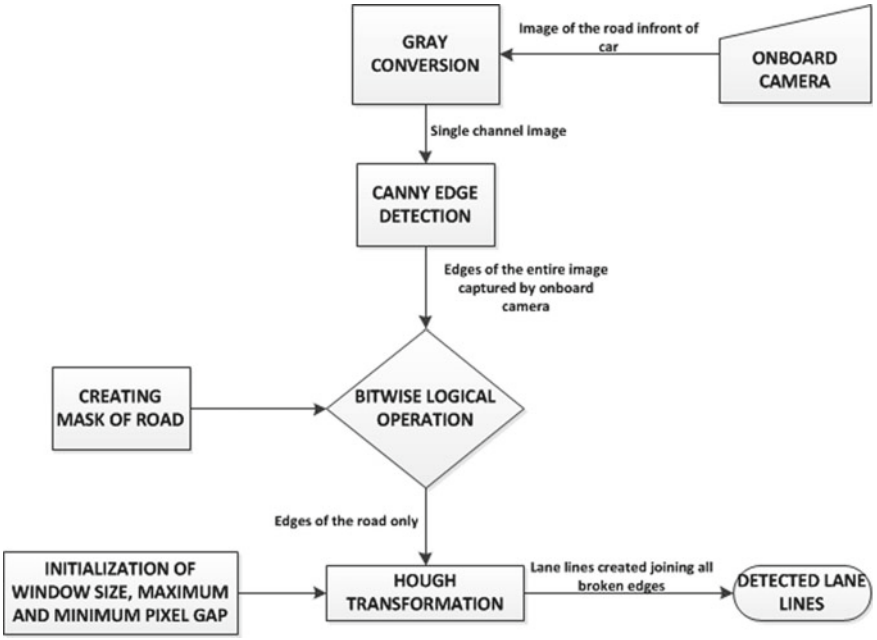


Fig. 8 Flowchart of our approach for lane-line detection

and b in Fig. 6, then using Hough transform we get the parameters of straight line (bold dotted line in Fig. 6) that passes through both a and b.

In Fig. 6, two points are considered. Blue lines represent the set of infinite straight lines that can pass through each point. Green line represents the line that interests both the points. Figure 7 shows the polar representation of Fig. 6. Two sinusoids intersect at $p1$, $\theta 1$, which are parameters of green line in Cartesian plane. In ROI, we found broken lane lines. Here, Hough transform is used for joining all these broken points to form proper lane lines (Fig. 8).

5 Algorithm

5.1 Vehicle Tracking

The steps of vehicle tracking algorithm is as follows:

Step 1. Initialize raspberry pi.
Step 2. Initialize camera connected to raspberry pi as an object.
Step 3. Initialize background modelling.
Step 4. Initialize system matrix and initial position of car.

Step 5. For each successive image frame in sequence:
{
Step 6. Acquisition of real-time image and image processing.
Step 7. Foreground = Subtraction of modelled background from current frame of image and filtering.
Step 8. Update background model.
Step 9. Locate centroid of foreground.
Step 10. Velocity of the car(Pixels/frame) = Difference between position of centroid in current frame and position of centroid in previous image frame.
Step 11. Apply Kalman filter to estimate the current position (Comment: It is the estimated position).
}
Step 12. Plot the variation between real-time captured data and estimated data.

5.2 Lane Line Detection

The steps of lane line detection algorithm are as follows:

Step 1: Initialize raspberry pi.
Step 2: Import computer vision library.
Step 3: Import scientific computing library.
Step 4: Import road image.
Step 5: RGB to grey conversion of imported image.
Step 6: Define higher and lower threshold.
Step 7: Canny edge detection.
Step 8: Generate a 2D null array of the same size of imported image.
Step 9: Define triangular mask.
Step 10: Bitwise AND operation between mask and Canny image.
Step 11: Define window size, min line length, max line gap.
Step 12: Hough Transform.
Step 13: If detected pair of point is not null:

for i in nl Comment: nl = no. of detected lines in Hough transform
  for $x1$, $y1$, $x2$, $y2$ in i
    Draw line using ($x1$, $y1$), ($x2$, $y2$) coordinate

Step 14: Weighted sum of drawn line and imported image.
Step 15: Output.

5.3 Hardware Setup

In laboratory environment, a four-wheel drive VEEROBOT robotic car has been used for tracking of car. Camera is the only sensing device used here. Real-time images were taken using a 'Logitech 5 MP' webcam, and the relevant processing is carried out using 'Raspberry pi 3'.

6 Experimental Results

6.1 Vehicle Tracking

Real-time captured image has been processed, and background subtraction with threshold value 15 is performed. Figure 9 shows the output of background subtraction.

Very essential part of this background subtraction is background modelling. Figure 10 shows how the speed of background adaptation increases with the increase of adaption rate.

Figure 11 shows variation between estimated and measured position. In Fig. 11, car is moving towards a direction where indexes along x- and y-axes are increasing but pixel-wise displacement along y-axis is less than displacement along x-axis.

Here, red line indicated measured values of centroid, and green line represents estimated values of centroid. Figure 11 shows that measured curve and estimated curve follow almost the same pattern. Initially, error is more but it reduces with time. At the very last of the curve (frame > 80), it shows large variation because after that



Fig. 9 Detected car and its centroid

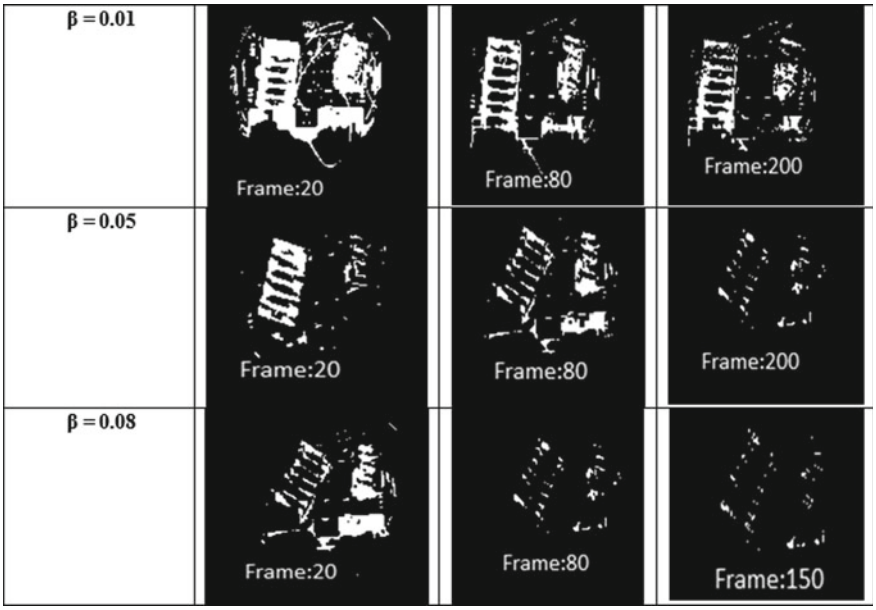


Fig. 10 Background modelling output

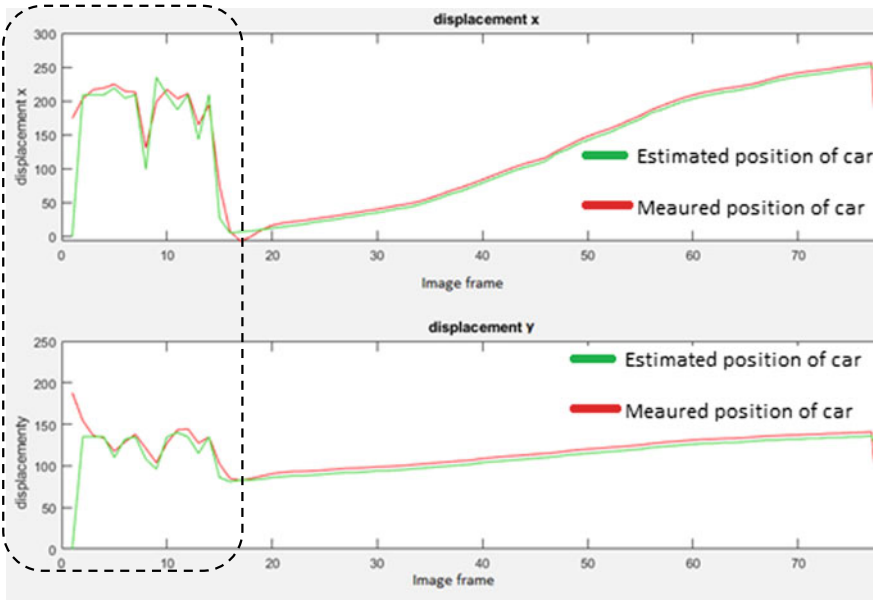


Fig. 11 Kalman filter estimation result

instant car has moved out of the coverage range of camera. At beginning (dotted region in Fig. 11), car was outside of camera range which caused the variation in plot up to around 20th image frame.

6.2 Lane Line Detection

Lane detection algorithm has been implemented using multiple road images available online. Entire processing has been performed using Raspberry pi-3. Figure 12 shows a sample image on which lane line detection algorithm has been applied.

Figure 13 shows the edges of Fig. 12; multichannel to single-channel conversion followed by Canny edge detection results is shown in Fig. 13. Figure 14 shows extracted region of interest. It has been found by performing bitwise AND between



Fig. 12 Road image; *Source* https://static1.squarespace.com/static/54326055e4b0fe7e4d69deea/t/5bde960540ec9a449db463c8/1541314065142/PO01RoadMain.jpeg

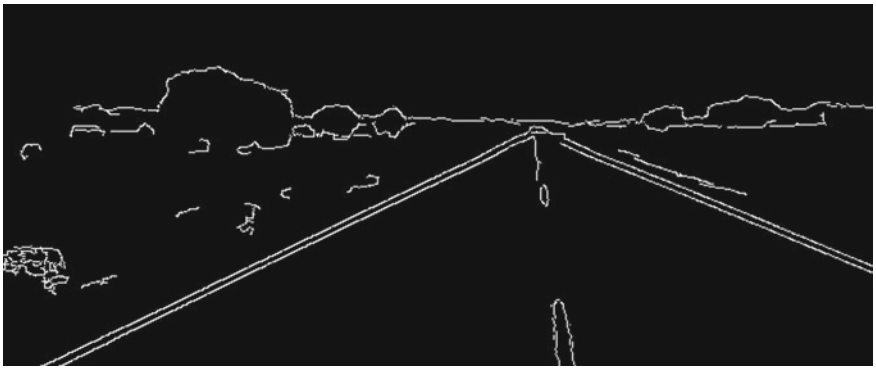


Fig. 13 Edges of the input image

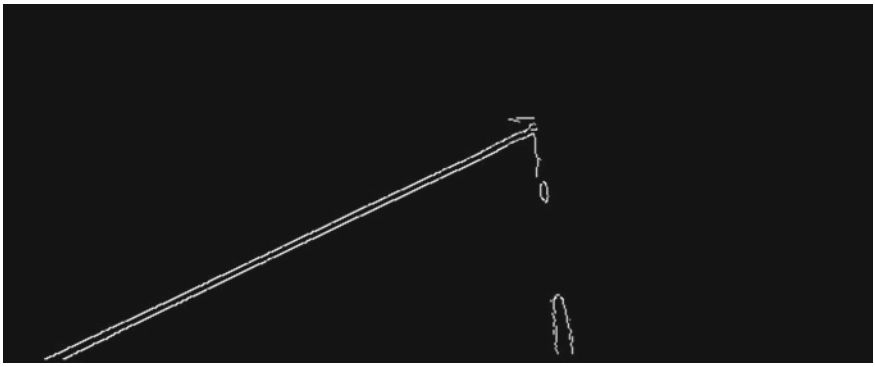


Fig. 14 Region of interest



Fig. 15 Detected lane lines

Fig. 13 and generated mask. Application of Hough transform on ROI followed by superimposing with original image results in Fig. 15. Here, lane lines are clearly marked.

7 Conclusion

Two major fields for making a ground vehicle autonomous have been elaborately discussed in this paper. Real-time images captured by the camera mounted above the vehicle may be distorted due to the effects of occlusion, lighting condition and vibration. Although these images are not reliable for vehicular navigation, we successfully applied an approach by estimating the position of vehicle in image frame. This estimated position is durable enough to navigate the vehicle. The very basic condition of self-driving car is maintaining proper lanes. We proposed very easy

and fast image processing-based algorithm for lane line detection. Here, we have implemented the algorithm on a open-source image but similar can be easily performed using real-time captured image. Both approaches are very convenient for real-time operation. The total approach has been tested on a model car in laboratory environment and found to be working smoothly. This is a very simple approach to achieve our goal, and in future more intelligent techniques may be investigated for implementing it in high-speed advanced processing unit.

In future, we expect to estimate the position of multiple cars simultaneously under single camera range. In case of lane line detection, we have used triangular mask, which is a compulsion to this approach because size and shape of road vary from region to region. We look forward to propose more robust and convenient technique for lane line detection.

References

1. S. Thrun et al., Stanley: the robot that won the DARPA grand challenge. J. Field Robot. **23**(9), 661–692 (2006)
2. H.M. Atiq et al., Vehicle detection and shape recognition using optical sensors: a review, in *2010 Second International Conference Machine Learning and Computing (ICMLC)* (IEEE, Bangalore, India, 2010)
3. Y. Liu et al., A survey of vision-based vehicle detection and tracking techniques in ITS, in *2013 IEEE International Conference Vehicular Electronics and Safety (ICVES)* (IEEE, Dongguan, China, 2013)
4. K. Wu et al., Overview of video-based vehicle detection technologies, in *2011 6th International Conference Computer Science & Education (ICCSE)* (IEEE, Singapore, 2011)
5. J. Rittscher et al., A probabilistic background model for tracking, in *European Conference on Computer Vision* (Springer, Berlin, Heidelberg, 2000)
6. L. Xie et al., Real-time vehicles tracking based on Kalman filter in a video-based ITS, in *2005 International Conference Communications, Circuits and Systems, 2005. Proceedings* (IEEE, Hong Kong, China, 2005)
7. Honghong Yang, Qu Shiru, Real-time vehicle detection and counting in complex traffic scenes using background subtraction model with low-rank decomposition. IET Intell. Transp. Syst. **12**(1), 75–85 (2017)
8. O. Masoud, N.P. Papanikolopoulos, A novel method for tracking and counting pedestrians in real-time using a single camera. IEEE Trans. Veh. Technol. **50**(5), 1267–1278 (2001)
9. A. Kumar, S.K. Mishra, P.P. Dash, Robust detection and tracking of object by particle filter using color information, in *Fourth IEEE International Conference Computing, Communications and Networking Technologies (ICCCNT)*, India (2013), pp. 1–6
10. C. Stauffer, W.E.L. Grimson, Adaptive background mixture models for real-time tracking, in *Annual conference on Computer Vision and Pattern Recognition (CVPR)* (IEEE, USA, 1999)
11. M. Aly, Real time detection of lane markers in urban streets, in *IEEE Intelligent Vehicles*, Eindhoven, Netherlands (2008), pp. 7–12
12. K.-Y. Chiu, S.-F. Lin, Lane detection using color-based segmentation, in *Intelligent Vehicles Symposium*, Las Vegas, NV, USA (2005), pp. 706–711
13. A. Borkar, M. Hayes, M.T. Smith, Robust lane detection and tracking with RANSAC and Kalman filter, in *16th International Conference on Image Processing (ICIP)* (IEEE Cairo, Egypt, 2009), pp. 3261–3264

14. B. Yu, A.K. Jain, Lane boundary detection using a multiresolution Hough transform, in *Proceedings of International Conference on Image Processing*, Santa Barbara, CA, USA (1997), pp. 748–751
15. D.H. Ballard, Generalizing the Hough transform to detect arbitrary shapes. Pattern Recognit. **13**(2), 111–122 (1981)
16. J. Illingworth, J. Kittler, A survey of the Hough transform. Comput. Vis. Graph. Image Process. **44**(1), 87–116 (1988)
17. Q. Hu et al., Fast detection of multiple objects in traffic scenes with a common detection framework. IEEE Trans. Intell. Transp. Syst. **17**(4), 1002–1014 (2016)
18. S.S. Sengar, M. Susanta, Foreground detection via background subtraction and improved three-frame differencing. Arab. J. Sci. Eng. **42**(8), 3621–3633 (2017)
19. Youngrock Yoon, Akio Kosaka, Avinash C. Kak, A new Kalman-filter-based framework for fast and accurate visual tracking of rigid objects. IEEE Trans. Robot. **24**(5), 1238–1251 (2008)
20. A.J. Lipton, H. Fujiyoshi, R.S. Patil, Moving target classification and tracking from real-time video, in *Fourth IEEE Workshop on Applications of Computer Vision, 1998. WACV'98. Proceedings* (IEEE, Princeton, NJ, USA, 1998)
21. Á. Odry et al., Kalman filter for mobile-robot attitude estimation: novel optimized and adaptive solutions. Mech. Syst. Signal Process. **110**, 569–589 (2018)

Wireless Controlled Lake Cleaning System

Sudhanshu Kumar, Saket Kumar, Rajkumar Viral and H. P. Singh

1 Introduction

Water is one of the most essential elements needed to sustain life. 70.78% of earth is covered with water. 97% of this is saltwater and is not fit for human use. Only 3% of this is freshwater which is suitable for daily use, but almost 98.8% of this is trapped as glaciers or underground water. The remaining 1.2% is distributed in the forms of lakes, ponds, rivers, streams, and other surface water sources. Out of this 1.2% of freshwater, 87% is present in lakes [3]. These data bring us to a conclusion that very less freshwater is present on the surface of earth which can be easily extracted for everyday use. With the everyday rising population and industries, it is becoming difficult to manage our freshwater resources. According to the Cambridge dictionary, "relating to, living in, or consisting of water that does not contain salt" [2] is called freshwater sources. Human use freshwater for many purposes and discharge it when it gets polluted. Polluted water consists of many types of salts and bacteria, which makes it unfit for consumption. Few uses of freshwater which is discharged after use are as follows:

- Domestic
  - Drinking,

S. Kumar · S. Kumar · R. Viral (✉) · H. P. Singh
Amity University Uttar Pradesh, Noida, UP, India
e-mail: rviral@amity.edu

S. Kumar
e-mail: shudhanshu67.sk@gmail.com

S. Kumar
e-mail: skumar39@amity.edu

H. P. Singh
e-mail: hpsingh2@amity.edu

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_157

  - Bathing, and
  - Washing.

- Agriculture
  - Farming,
  - Gardening, and
  - Fisheries.

- Industrial
  - Manufacturing,
  - Dissolving, and
  - Cooling.

The abovementioned use requires freshwater, which is extracted from lakes, rivers, and ponds. As soon as the water gets polluted after use, it is declared unfit for consumption.

Solid waste is often dumped in water, which pollutes the water body. Solid waste is seen floating on the surface of water bodies. In India, a large part of water is polluted due to the untreated sewage that flows directly from the households to the nearest water body. According to the central pollution control board of India, "This is mainly due to discharge of domestic wastewater mostly in untreated form from the urban centers of the country. The municipal corporations at large are not able to treat increasing the load of municipal sewage flowing into water bodies without treatment" [7]. Industries and agricultural runoffs are a major concern. Farmers are usually observed to use unknown quantity of pesticides and urea. Excess pesticides and urea are carried away to the nearby water body during rainfall. Small- and medium-scale industries and workshops discharge large quantities of untreated water which pollutes the nearby water bodies [9]. The pollutants are chemical waste and solid waste. The quantity of pollutants is large to such an extent that the water bodies are not capable of dilution. This results in decrease in dissolved oxygen and causes bacterial contamination and water-borne diseases. According to the online blog published by NDTV on October 6, 2018, "thousands of snails and fishes have died and were found floating on the banks of the historical Madivala lake in south Bengaluru" [1].

Aquatic animals are harmed in an unpredictable way. Some of the causes of the decline of aquatic animals are as follows:

- Rapidly changing climate,
- Toxic algae,
- Nano-materials,
- Microplastics, and
- Lowering of calcium concentrations.

Survival of aquatic animals is dependent on pH value of the water body. Freshwater aquatic animals cannot survive in saltwater, with few exceptions. Chemicals and heavy metals alter the pH of the water body. Area of water exposed to atmosphere

Table 1 Water quality standards

Water use		Standards			
		pH	BOD	SS	DO
1	Water supply class-1, conservation of natural environment	6.5–8.5	1	25	7.5
2	Water supply class-2, Fishery class-2, bathing	6.5–8.5	2	25	7.5
3	Water supply class-3, fishery class-2	6.5–8.5	3	25	5
4	Fishery class-3, industrial water class-1	6.5–8.5	5	50	2
5	Industrial water class-2, agricultural water	6.5–8.5	8	100	2
6	Industrial water class-3	6.0–8.5	10	*	2

*No floating particles

governs the pH and the dissolved oxygen in the water body. Floating solid waste reduces this contact area, thus harming the aquatic animals. Given below is a chart showing normal water quality standards for rivers and lakes on a basis of daily average [11]. These parameters are calculated for different uses of water in different industries. In some industries, the water required should be free of all ions and minerals, while in other industries impure water can be used. In agriculture, the water used must be free of harmful chemicals and ions which can clog the surface over a period of time. The biological oxygen demand, pH value, suspended solids, and dissolved oxygen have to be appropriate proportion for survival of animals or its use in an industry. Even in an industry, the water quality required varies. Water supply is differentiated on basis of certain parameters and designated into various classes as discussed in Table 1.

Removal of solid waste from surface of water bodies helps to maintain the biodiversity and also increases tourism. Traditional approach for removal of floating solid waste is done by manually entering the polluted water body in small boats and the collecting the trash by fish traps. This increases the risk of water-borne diseases and drowning.

2 System Description

2.1 ATmega328 Microcontroller

Microcontroller is the brain of the entire system. The microcontroller used in this system is an “ATmega328p”, developed by Microchip technologies. This microcontroller is used by various other embedded system designers to develop their own single-board computers. Arduino is one such company which develops open-source software and hardware. Arduino uses the ATmega328p chip to develop their single-board computer named “Arduino”. The Arduino board used here is the “Arduino Uno” [5]. Arduino Uno was used because of its low cost and also the fact that it can

Table 2 Features of Arduino

S. No.	Feature	Range
1	Microcontroller	ATmega328
2	Operating voltage	5 V
3	Input voltage	7–24 V
4	Analog input pins	8
5	Flash Memory	32 KB
6	EEPROM	1 KB
7	Digital I/O pins	14
8	SRAM	2 KB
9	Clock speed	16 MHz
10	DC current per I/O pin	40 mA

be easily programmed by using C language. For a manual control system, Arduino Uno provides enough computing power and speed. Arduino released an integrated development environment for their single-board computers. Scripts can be written easily and can be burned in the ROM of the microcontroller. Various features of Arduino Uno are shown in Table 2.

2.2 Relays

Relays are electromechanical switches used for a variety of purposes. In this system, relays are used to switch on or off the motors and to reverse the direction of the motors installed in the system.

2.3 DC Geared Motors

Motors working on direct current provides a high initial torque and has the ability to reverse direction of rotation. This reversal of rotation helps to guide the direction of motion of the system. The motors are waterproofed by the help of sealants. Two motors are used to propel the system in either direction. One motor is used to continuously rotate the conveyor belt. One motor is used to rotate the conveyor system which helps in saving power and also to trap heavy waste.

2.4 Conveyor System

A conveyor system helps to trap the trash and carry it up to the trash bin. Grooved PVC belt rotates freely on the rotating axle fixed with a DC motor. The conveyor system is fixed rigidly in the front of the system which picks up the solid waste.

2.5 Catamaran Hull Structure

Catamaran hulls of equal sizes are used in the system. These are also called double-hull structures. Double-hull design provides several advantages over mono-hull designs. Catamaran hull reduces the effect of buoyancy principle and provides stability to the system under dynamic weight. Catamarans are lightweight structures and can easily traverse on the surface of water by minimal thrust. This structure had been used to achieve very high speeds in racing boats and ferries due to its low weight and less drag [6]. Hollow PVC tubes have been used to design the catamaran structure. Silicone-based sealants are used to waterproof the hull.

2.6 Trans-receiver

The system is wirelessly controlled and avoids the risk of water-borne diseases and drowning. A remote control system was designed which transmits and receives signals. The trans-receiver used for the system was HC-12 trans-receiver. Different features of HC-12 are shown in Table 3 [4].

Table 3 Features of trans-receiver

S. No.	Feature	Range
1	Max. transmission distance	1000 m
2	Number of channels	100
3	Frequency range	433.4–473.0 MHz
4	Max. transmitting power	100 mW
5	Supply voltage	3.2–5.5 V
6	Receiving sensitivity	−117 to −100 dBm
7	Idle current	16 mA

Table 4 Operating battery specifications

S. No.	Feature	Range
1	Minimum voltage	9.7 V
2	Maximum voltage	14.4 V
3	Capacity	2200 mAh
4	Connection type	Series
5	Weight	800 g

2.7 Lithium–Ion Polymer Battery

The system is electrically powered, by using rechargeable lithium–ion polymer batteries. These batteries are based on lithium–ion technology and use a form of polymer electrolyte [8]. This polymer electrolyte prevents leakage. Lithium polymer batteries have high specific energy. Li-Po batteries are used in applications where weight is a critical issue and the system requires large amount of electrical power. Weight of components is an important factor in this system. Li-Po batteries provide highest specific energy in today's world, which means it stores very high amount of power per kilogram by weight. Every cell provides a maximum voltage of 4.2 V at full charge. These batteries can go through deep discharge cycles and can be recharged 1000 times before the storage capacity falls below 80%. However, Li-Po batteries must not be deeply discharged frequently as it adversely affects its backup. Higher voltages can be obtained by using multiple cells in series. Every cell has its own battery management system. The battery management system continuously checks the battery voltage and prevents from overcharging and discharging. Features of the Li-Po used in the project are mentioned in Table 4.

3 Experimental Setup Design

The design parameters of the proposed system are given in Table 5. The loading capacity is the key parameter of the proposed design. Based on the load capacity, the other design parameters are fixed using some basic necessary analytical expressions. For instance, the proposed design can carry 5 kg of load capacity, and other parameters like drive motors rpm and torque, and catamaran hull size are calculated.

3.1 Block Diagram of Proposed System

The floating system consists of four motors where two motors are used for propulsion and the other two motors are used for conveyor arm movement. An HC-12 module is used as a wireless trans-receiver. Water sensor is used to test the presence of water leakage inside the tubes and hull. Ultrasonic sensor acts as a draft and helps

Table 5 Design parameter of proposed system

S. No.	Design parameter	Value
1	Overall dimensions	1000 mm × 559 mm × 294 mm
2	Net weight without payload	8 kg
3	Maximum weight with payload	13 kg
4	Area coverage per hour	100 m^2
5	Average battery backup while cleaning	1.5 h
6	Control system	Wireless remote control
7	Drive motors	300 rpm 1.2 N-m torque
8	Wireless module	HC-12 wireless trans-receiver
9	Conveyor belt RPM	150 RPM

to determine the distance between the waterline and the bottom of the hull. A basic block diagram of the design is shown in Fig. 1.

The remote control unit is the main command center of the whole system. As the system is manually controlled, the user alertness plays an important role in increasing

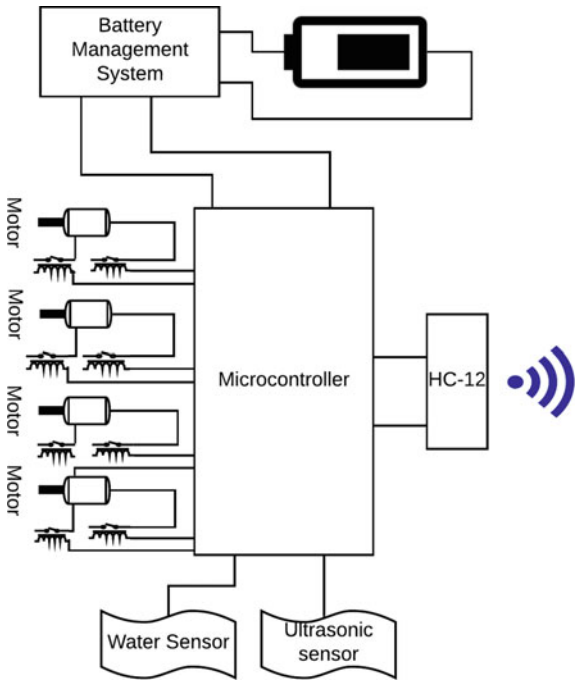


Fig. 1 Block diagram for floating system

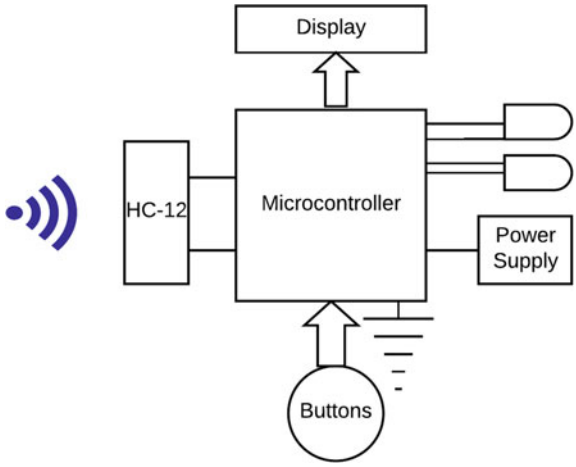


Fig. 2 Block diagram for remote control system

the efficiency of the system. The remote control takes user input from the buttons, encodes them, and then transmits a packet of data via the HC-12 module. An OLED display is provided to facilitate the user interface. LEDs are provided to indicate healthy and dangerous conditions. A basic block diagram is shown in Fig. 2.

3.2 Flowchart of Proposed System

An efficient algorithm is the backbone of the system. This helps to increase the overall efficiency and realize the dangers in time. The algorithm serves as a feedback loop, reading data from every sensor and the data received by the wireless module. These data help the system to generate certain possibilities of the system health. Health condition of the system decides whether the system will clean the trash or notify the dangers to the user so the system can be brought back to safety. If the battery falls below 15%, the system switches to power saving mode and notify the user about the condition. A basic flowchart of the running algorithm is shown in Fig. 3.

The remote control is the only link between the system and the user. Probability of the data packets reaching the floating system decreases as the system travels far from the user. To reduce the probability of packets lost while transmission, a data echo algorithm was designed. To increase the safety of the data transmitted, encoding was employed before any data is sent to the floating system. The flowchart for the remote control system is shown in Fig. 4.

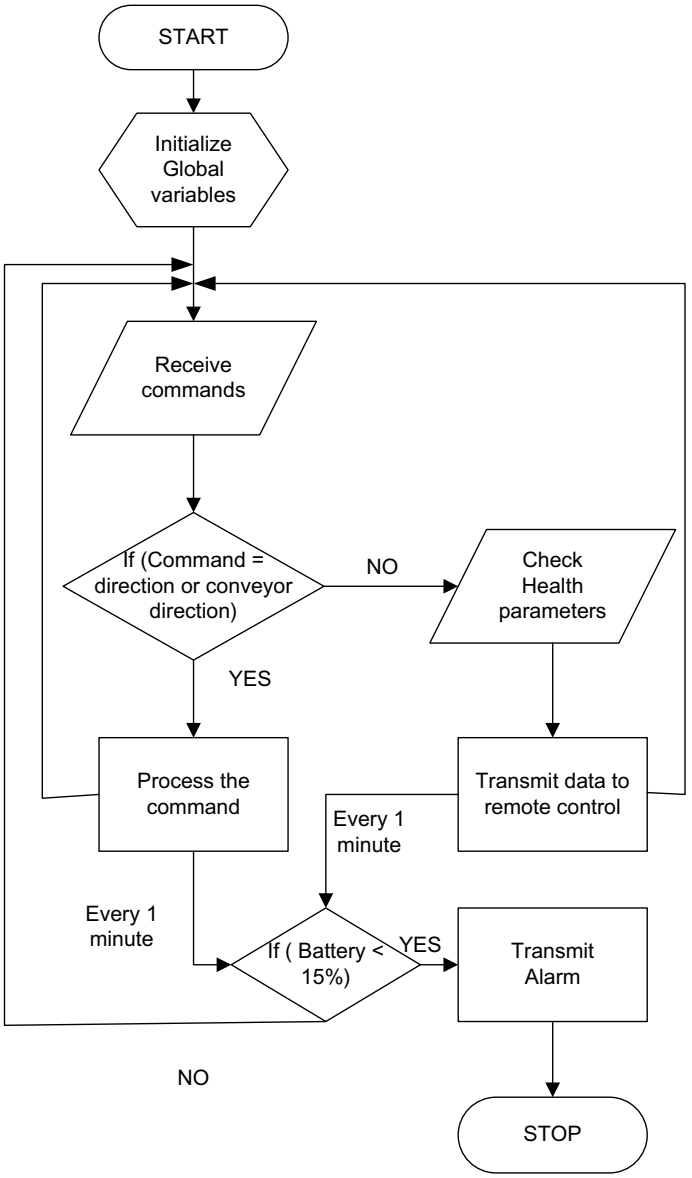


Fig. 3 Flowchart for the floating system

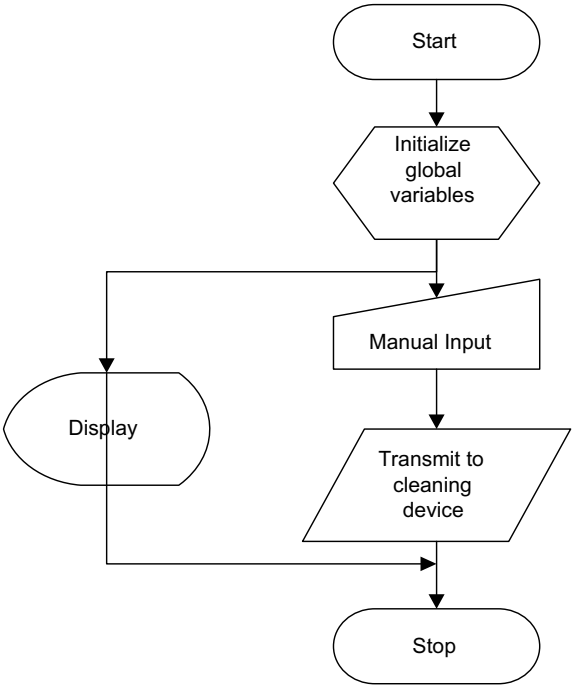


Fig. 4 Flowchart for the remote control system

3.3 Chassis

The chassis of the system is made of hollow PVC tubes [12]. The length of tube is 2 feet and the diameter is 2.5 inches. Hollow tube creates an air chamber which supports the tube in floating. The two tubes are connected with the help of lightweight aluminum bars to make a "H"-shaped structure. The chassis has a double-hull catamaran design which helps in stability. Waterproofing was done by the help of silicone-based sealants.

3.4 Conveyor Assembly

Floating trash collection is the main aim of the system. To collect the trash from the water surface, a conveyor assembly was designed [14]. The tail end section of the conveyor assembly dips inside the water and collect trash. Grooved PVC belt rotates freely which carries load (trash) from one end to the other [15]. The head section of the conveyor consists of a DC motor connected to a roller. This roller in turn rotates the belt. The transfer point of the belt is open and it lets the load fall freely. The load is

dropped into an open plastic box, which stores the load until the system returns back to the land surface. The conveyor belt is able to carry loads up to 500 g at any given time. The tail end of the conveyor belt is 10 inches wide. conveyor track produces drag, and thus more energy is required to push the system forward. To reduce this drag, a provision, to raise the conveyor track, has been made.

3.5 Propulsion

The system travels on the surface of the water body [10]. To push the system in any direction, thrust is required. This thrust is generated by side paddlewheels connected to the system. The paddle wheel is connected to a waterproof DC motor which helps the paddles to rotate continually, thus producing enough thrust to move the system forward. PVC-based plastic polymer was used to design the paddle. The diameter of the paddle is 1 foot. The motors' shaft is connected to the center bore. Whole motor and paddle assembly is covered by PVC sheets to avoid spillage of water. On top of the cover, IP65-rated PVC boxes are provided, to house the electronics and the battery. The whole system was designed to keep the weight as low as possible.

3.6 Wireless Control

The system is designed to avoid contact between humans and the polluted water [13]. To facilitate this, the system is wirelessly controlled by a remote controller. HC-12 trans-receiver is used to communicate between the system and the remote control box. Features of HC-12 trans-receiver module are shown in Table 2. Special commands for the direction of propulsion and functioning of conveyor belt were developed. These commands are transmitted wirelessly to the floating system. The system receives these commands and acts accordingly. After completing the received commands, it sends back a message and then gets ready to receive a new command. A display panel on the handheld remote facilitates the user to know the current command transmitted and the task assigned to the system. This handheld remote is powered by 500 mAh, 7.4 V Li-Po batteries.

4 Experimental Results

The system has been tested on various sites to validate its working. The floating ability of the system had been tested first, in a sewage treatment plant, sedimentation tank inside Amity University, Noida. Second test was performed to test the systems, full capability of free movement, and then perform waste collection. Hooks and strings were attached to the system for safety, and the size of the test area chosen was kept small. Tests were performed only once on both sites. Proper safety procedures

were followed to prevent any casualties or damage. Details of the test sites and the results are given below. Testing of proposed system is depicted in Table 6.

Testing site 1:
Amity STP, Noida campus

Test Objective: To check the floating ability of the system and conveyor movement.
STP pond area: 50 m^2.
Type of waste available: Algae, foam, suspended solids.
Selected area: 5 m^2 to be cleaned

Testing site 2:
Okhla bird sanctuary, sector 128, Noida

Test Objective: To test the system movement in water and perform waste collection.
Lake catchment area: 50 m^2.
Type of waste available: Leaves, petals, polythene bags, algae, weeds, suspended solids.
Selected area: 20 m^2 to be cleaned.

The actual experimental test setup is developed and tested in the laboratory as shown in Fig. 5. The computational algorithm described in Sect. 2.3 is burn on the Arduino controller units. Several tests are performed in order to get the given site in order to check the performance of the developed system.

The proposed design is a prototype which was tested with a maximum weight load of 3.8 kg within 15 min. An industrial designed system with a better battery backup and longer range of wireless control will help to clean larger areas. The system designed is not suitable for cleaning very large lakes and ponds filled with large quantities of trash. A weed harvester device used to clean large lakes and ponds is already commercially available. Such kind of weed harvesters are complex in design and have huge cost of operation for daily or weekly use. Therefore, these are not feasible for small lakes and ponds in particular reference to India. However, the developed system is capable of cleaning small areas of few meters and can be used again after recharging.

Table 6 Testing procedure

S. No.	Descriptors	Testing site 1	Testing site 2
1	Total catchment area	50 m^2	3.5 km^2
2	Selected area to clean	5 m^2	20 m^2
3	Type of waste present	Foam, suspended solids, algae	Leaves, petals, polythene bags, algae, weeds, suspended solids
4	Time taken to clean the selected area	Not applicable	15 min
5	Total waste collected	Not applicable	3.8 kg



Fig. 5 Experimental setup

5 Conclusion

Regular cleaning of floating trash of river/lakes/ponds is necessary for the survival of aquatic animals and also to keep the water body clean. India is facing a large problem of sewage treatment and water treatment. Most of the public entertainment places have small water bodies for entertainment and recreation which can be easily cleaned by the developed system on a daily or weekly basis. The commercially available cleaning system is costly and not suitable place like India. Consequently, this paper reported a low cost, simple in design, and operation to clean the small lakes/ponds/rivers. Try to keep the cost of operation of the developed system as lowest as possible, and hence it can be used on a daily or a weekly basis on small water bodies. The developed system can be used in places such as ponds, swimming pools, public entertainment pools, zoological parks, riverside after idol dispersion, and other necessary places. The system aims to reduce the dangers that workers face while manually cleaning the water bodies. Employment is not affected by the system and requires a trained worker which helps in improving the technical knowledge of workers. This system can be implemented in larger water bodies after necessary design extensions.

References

1. https://www.ndtv.com/bangalore-news/thousands-of-snails-dying-in-bengalurus-madivala-lake-baffles-experts-1927732. Accessed Mar 2019
2. https://dictionary.cambridge.org/dictionary/english/freshwater. Accessed Mar 2019
3. https://www.e-education.psu.edu/earth103/node/701. Accessed Mar 2019

4. https://probots.co.in/hc12-long-range-1km-wireless-rf-transceiver-module-433mhz-si4463.html. Accessed Mar 2019
5. https://www.arduino.cc/. Accessed Mar 2019
6. https://www.yachtsinternational.com/owners-lounge/sail-debate-monohull-vs-catamaran. Accessed Mar 2019
7. http://cpcb.nic.in/nwmp-monitoring-network/. Accessed Mar 2019
8. https://www.genstattu.com/blog/the-lipo-battery-characteristics-and-applications/. Accessed Mar 2019
9. A. Dwivedi, Researches in water pollution: a review (2017). https://doi.org/10.13140/rg.2.2.12094.08002
10. D. Harte, N. Bose, R. Clifford, T. Roberts, G. Davidson, An application of paddlewheel propulsion to a high-speed craft (2019)
11. Y. Magara, Classification of water quality standards, in *Water Quality and Standards—Classification of Water Quality Standards*, vol. I (2002)
12. S.Md.S.Md. Rafique, A. Langde, Design and fabrication of river cleaning machine. IJSART **3**(11), 8–18 (2017)
13. M. Mohamed Idhris, M. Elamparthi, C. Manoj Kumar, N. Nithyavathy, K. Suganeswaran, S. Arunkumar, Design and fabrication of remote controlled sewage cleaning machine. IJETT **45**(2) (2017)
14. A.M. Ballade, V.S. Garde, A.S. Lahane, P.V. Boob, Design and fabrication of river cleaning system. IJMTER **04**(2), 2349–9745 (2017)
15. P.M. Sirsat, I.A. Khan, P.V. Jadhav, P.T. Date, *Design and fabrication of River Waste Cleaning Machine* (2017)

Multipurpose Voice Control-Based Command Instruction System

Deepak Ranjan, Rajkumar Viral, Saket Kumar, Gaurav Yadav and Prateek Kumar

1 Introduction

Technology has made life simple and easy lives with capability. Most of things in our homes to workplace are controlled by technologies either it is refrigerator, washing machines, ACs, or laptops [1]. Everyday, a technology is replaced by a new technology which is more advanced, upgraded, and helpful. The present time world is dependent on technology and we cannot deny our dependence on technology. For instance, it can make handicapped people lead to a more comfortable life. Many diseases have been cured with the help of available technology. Technology has let us know what is going on around the world in a fraction of second. It has made humans more interactive and social. Therefore, technology played a key role in everything around us [2–4].

Similar progressive development has been noticed in voice-based technology, broadly, in telecommunication, aerospace, transportation, navigation, internet, Internet of things (IOT), biomedical, automation, broadcasting, education, industrial, and various domestic and commercial sectors. A voice-based technology for an automated industrial operation is shown in Fig. 1 [5].

D. Ranjan · R. Viral (✉) · S. Kumar · G. Yadav · P. Kumar
Amity University Uttar Pradesh, Noida, UP, India
e-mail: rviral@amity.edu

D. Ranjan
e-mail: deepakranjan1113@gmail.com

S. Kumar
e-mail: skumar39@amity.edu

G. Yadav
e-mail: gyadav@amity.edu

P. Kumar
e-mail: prateek.pln@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_158



Fig. 1 Voice recognition application in automated industrial operations [5]

Among above-based application areas, some major and more advanced applications of voice signal can be found in telecommunication/wireless communication as a voice recognition technology [5]. Today, most of smartphone, notebook, laptop, and tablet use this feature efficiently in order to interact, just for information sharing and set of instruction or command execution [6]. The latest voice recognition technology application can be seen in all Android and Windows-based smartphone such as "Ok Google". In this system, user can directly give commands/instruction by speak and inbuilt Android application responds accordingly within second of time [6].

Similarly, interface can be noticed in PCs and desktop by using an external or inbuilt microphone [7]. Contrary to this, the system can also perform by giving certain voice instructions or commands and can change it into text information in many applications. Based on this, a number of other manufacturer and software companies developed different voice recognition logistics for various purposes, making our lives more convenient [8].

Principally, voice recognition is the process of converting an acoustic signal, captured by a microphone or a telephone, to a set of words as shown in Fig. 2 [9]. In other words, it is an ability of computer or machine that can program in a such a way to receive and interpret transcription or to perceive and execute spoken commands [10]. Voice recognition has gained now much importance with advancement of artificial intelligence (AI) such as Amazon's Alexa, Apple's Siri, and Microsoft's Cortana [9]. Thus, a voice recognition system makes consumer/users to interact with modern technology commonly by speaking to it, making hands-free request or commands, reminders, and other usual tasks [11].

Human speech production mechanism has gradually developed in many years which result in a vocal system that transfer information with a minimum effort. Moreover, today there are several good examples that can be found where one might encounter voice recognition systems, such as automated phone system, Google Voice,

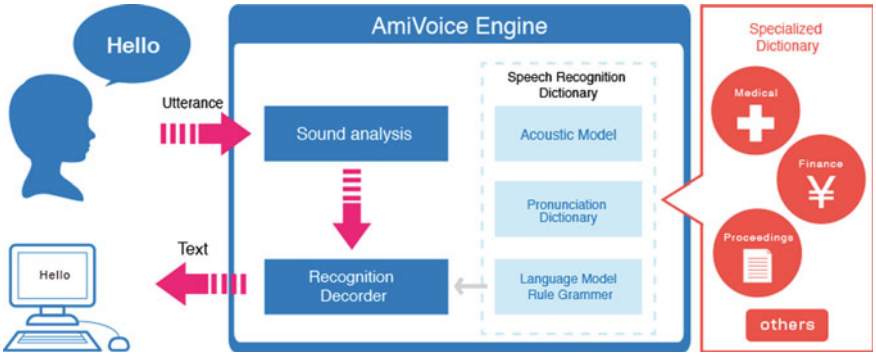


Fig. 2 Voice recognition process [9]

digital assistants (Amazon Echo, Apple's Siri, Google Assistant, etc.), car Bluetooth, etc. [11]. Based on input of commands and application, a voice recognition system can be of the following types [2–5].

a. Speaker-dependent systems: It needs training before to be used, which further needs to execute a series of words and phrases [12].
b. Speaker-independent systems: This system identifies most of the users' voice/sounds without any training [4].
c. Discrete speech recognition: This system takes pause in between each word so that each word separately identifies [6].
d. Continuous speech recognition: This system recognizes the words in normal rate of speaking [7].
e. Natural language: This system can easily understand the voice and revert back the answers of commands or queries [6].

Number of research article and studies have been published that majorly discussed the application of AI tools, optimization techniques, advanced software and hardware, various processing algorithms, and embedded system. Some of the interesting works are elaborated here to identify certain inferences.

Cui and Xue [6] used a technology for voice recognition that is based on the SPCE061A single chip for door control circuit. An animal voice identification system based on animal voice patterns recognition algorithm has been developed using zero-cross-rate Mel frequency cepstral coefficient and dynamic time wrapping algorithm by Yeo et al. [7]. Kanawade et al. [8], developed a gesture and voice recognition system that can be used for storytelling. Voice recognition is used to change the background sound/music dynamically during story's narration. Johnson et al. [9] presented a state-of-the-art review on voice-to-text applications. Abdalrahman et al. [10] proposed a cascade voice biometric system which utilize text-dependent and text-independent speaker and voice recognition process. They proposed the 91.2% approximate efficiency of the system.

Mohammadi and Kain [11] presented an overview of different voice communication systems and its real-world applications. Consequently, it has been noticed from

above studies that most of the works focus on application-based voice recognition system in different fields of utilization, though some studies also used AI-based and computer optimization techniques and tools in successful design of such system in modern and futuristic gadgets and logistics.

Therefore, this paper aims to develop and implement the voice recognition system for another very useful application, that is, voice-to-text commands/instructions. The proposed system can efficiently recognize the voice of user(s) and then can display it on liquid crystal display (LCD) or by other means. This system can be used in the restaurants for customer's order or call to waiter, on food delivery counters, big hotels for room and other services, reception counters, education and academics, traffic signboard, street road circle notice display, public information display, hospital railways and bus stops, airports, etc.

The proposed system uses two devices, namely, input device that can receive any voice command and the output device that is a display unit. An advanced microcontroller is used to recognize the voice and that can process it to text-based commands/instructions. The proposed system is easy to use and install and low in cost.

2 System Description

The quality of a speech recognition systems is assessed according to two factors: its accuracy (error rate in converting spoken words to digital data) and speed (how well the software can keep up with a human speaker). Commonly, such software is used for automatic translations, dictation, hands-free computing, medical transcription, robotics, automated customer service, and much more. If one has ever paid a bill over the phone using an automated system, then they have probably get benefited from speech recognition system [3].

Speech recognition technology has made huge strides within the last decade. However, speech recognition has its weaknesses and nagging problems. Current technology is a long way away from recognizing conversational speech. Despite its shortcomings, speech recognition is quickly growing in popularity. Within the next few years, experts say that speech recognition will be the norm in phone networks over the world. Its spread will be aided by the fact that voice is the only option for controlling automated services in places where touch-tone phones are uncommon.

Voice recognition works by analyzing the features of speech that differ between individuals. Everyone has a unique pattern of speech stemming from their anatomy (the size and shape of the mouth and throat) and behavioral patterns (their voice's pitch, their speaking style, accent, etc. [3]).

Most commonly, voice recognition technology is used to verify a speaker's identity or determine what speakers want from pantry that is identified. Speaker verification is the process of using a person's voice to verify that what they are speaking and what they want.

To do so, the proposed system majorly comprises two important parts: the input system (combination of voice recognition module and one wireless transmitter that is

NRF24L01 and Arduino controller) or transmitter section and output device (combination of ATmega328 controller and LCD unit). If any command/instruction request uses microphone placed after hearing the sound of voice assistant, then the system will translate it to the machine language and compare the voice instruction which is already saved in the system and if these instructions will be saved, then these instructions can transmit the command to the display at the receiver end. The receiver units with display create an alarm of feedback command and according to that comment, the user will get the response. The brief details of each system component/hardware with their relative technical specification and cumulative working are given in the next subsection.

2.1 A Subsection Sample

Transmitter section consists of microphone, voice recognition system, Arduino controller, buttons, speaker, LED, NRF24L01, wireless transmitter, etc. The system included microphone which receives voice instructions from user. The feedback of voice instruction is then processed through voice recognition system. The instructions are converted into digital format, and it is compared with the already saved instructions. The voice recognition system sends digital signal to the Arduino controller saved in library file. With the help of this library file, Arduino identifies the digital signal of voice recognition system and further transmits the instructions through NRF24L01 wireless transmitter. Arduino controls the functioning of speaker, and entire system gets activated when the button will be in active mode. When button is pressed, Arduino controller sends an instruction to the speaker and plays the presaved instructions and then enables the user to give commands.

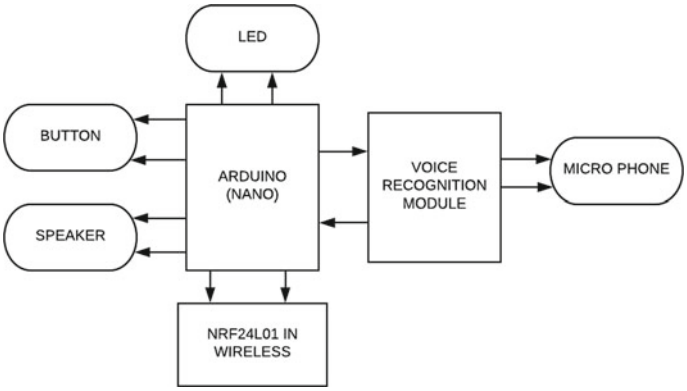


Fig. 3 Transmitter unit block diagram

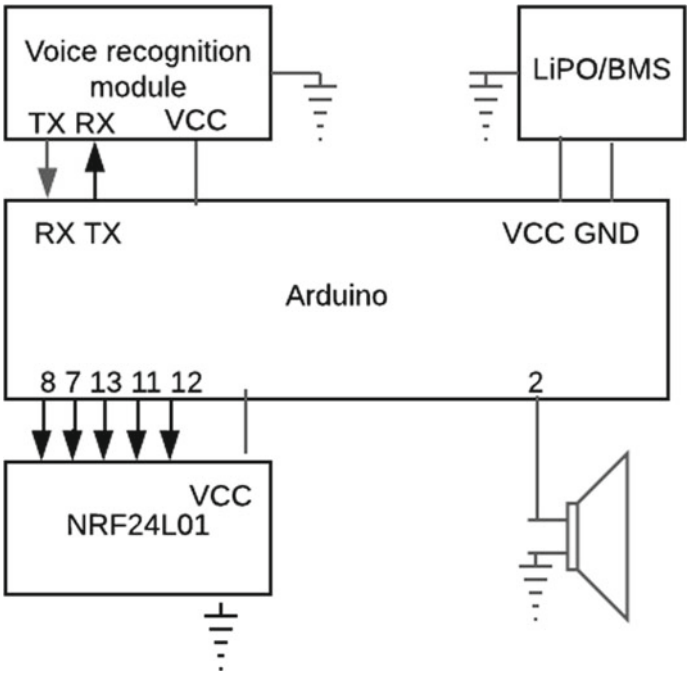


Fig. 4 Transmitter unit circuit connection

After completing the whole process, Arduino will send the instruction through NRF24L01 to the receiver unit. The entire transmitter section is shown in Fig. 3 (Transmitter section), and circuit diagram of transmitter is outlined in Fig. 4.

2.2 Receiver Section

The signal received by NRF24L01 is further processed by Arduino controller. According to the instructions received through receiver, Arduino controller displays all the instruction on a suitable display unit. Also, it activates an alarm which represents that the signal is received. Over receiving the signal, Arduino controller glows LED which represents that receiver unit is receiving the signal. The buttons used in receiver section help the small display unit to display messages or instructions. The receiver is shown in Fig. 5, and the circuit diagram receiver unit is delineated in Fig. 6.

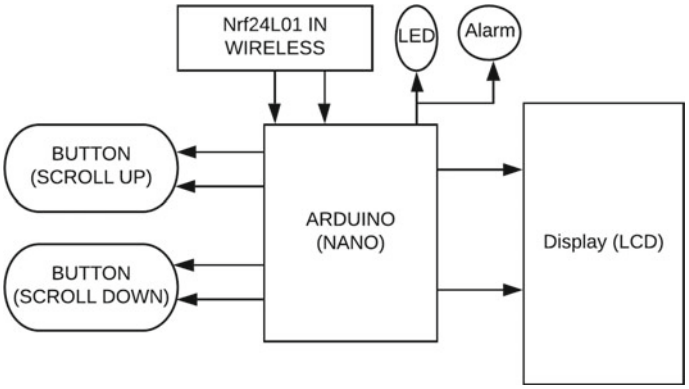


Fig. 5 Receiver unit block diagram

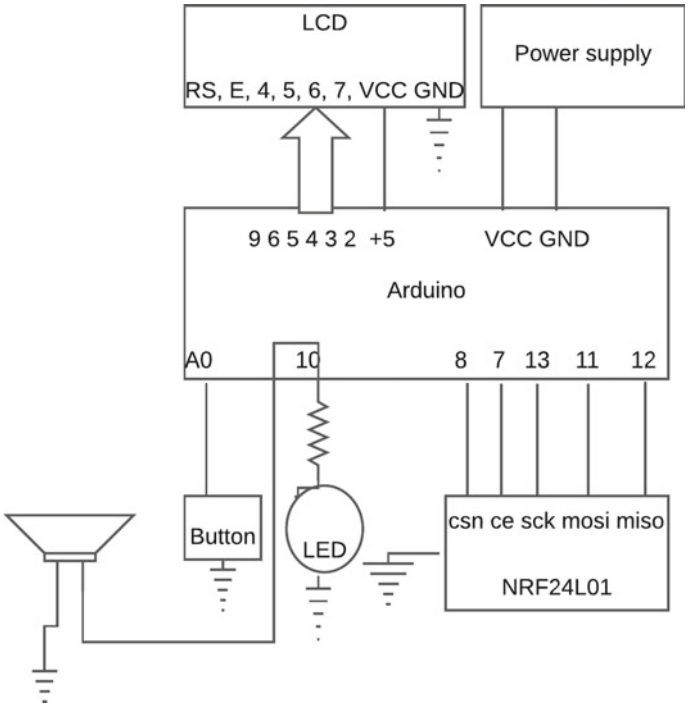


Fig. 6 Receiver unit block diagram

2.3 Computational Algorithm

The algorithm used for the proposed system is explained below:

Step 1. Initialize all the parameters.
Step 2. If (Signal receiver == 1)
Microphone = enable
Speaker = enable (play predefined instruction)
After beep user can give voice command.
Else (step 1)
Step 3. Voice recognition system = enable
Compare signal (Predefined == new instruction)
Step 4. If (Signal received == 1 instruction)
NRF24L01 == enable
Signal transmitted to receiver unit else (step 3)
Step 5. If (Signal received from NRF24L01 == 1)
LED == enable
Else LED == Blinking mode
Step 6. Arduino enable LCD 16*2
Initialize all parameters of LCD.
Display instruction on LCD.
(Alarm beep == enable)
Step 7. If (more instruction == received)
(Scroll mode == enable)
Button activated
UP button ==+ (Up)
Down button == –(Down)
Else (Display single instruction)
Step8. Ready for text instruction.

The flowchart of proposed system is shown in Fig. 7. The flowchart has the transmitter and receiver section' computational algorithm. The transmitter algorithm receives commands from user and compares it with saved instructions coming from microphone and return logical values to the Arduino controller. Arduino controller transmits signal by the help of NRF24L01 transceiver.

In parallel of transmitter section, the receiver's flowchart is also depicted in Fig. 7. Receiver unit process the received instructions and according to the instruction Arduino controller the instruction and activate alarm system is providing beep mode.

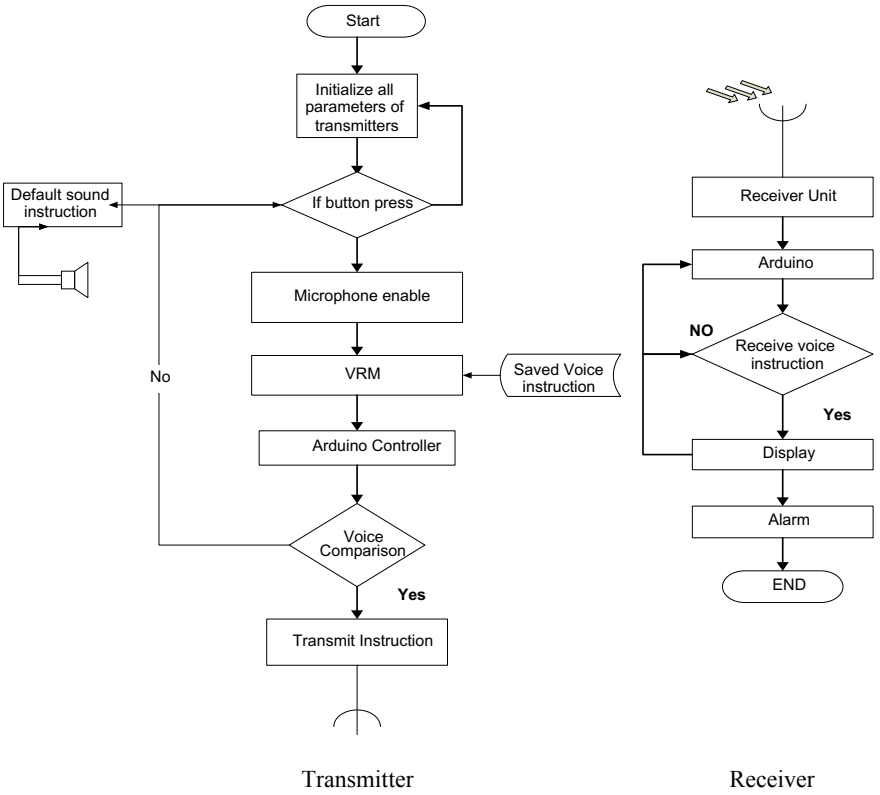


Fig. 7 Flowchart of proposed voice recognition system

3 Hardware Modules

Different components and their working are demonstrated in this section with their relative specifications.

3.1 Component Specifications

Power supply—All components and hardware are used in proposed system under the rating of 5 V DC. To operate all the components, it requires a step-down DC voltage with 1 A current. Few power supply systems are available in the market with 12 V voltage, and current rating of 1 A to reduce this voltage regulator 7805 has been used with current in the range of 12 V–5 V DC. After applying this voltage driver circuit, all the components can be operated voice recognition module.

Voice recognition module—Speak (Voice) recognition module V3-Arduino compatible is a compact and easy-control speaking recognition board. Speak (Voice) recognition module V3 product is a speaker-dependent voice recognition module. It supports up to 80 voice commands and maximum seven voice commands could work at the same time. Any type of voice command can be saved in the module for further verification, when user repeats this command and the module. This board can be controlled in two ways: serial part (Full function) [15, 16].

Arduino controller—Famous microcontroller manufacturers are Micro Chip, Atmel, Intel, Analog devices, and more. A microcontroller board contains onboard power supply, USB port to communicate with PC, and an Atmel microcontroller chip. It simplifies the process of creating any control system by providing the standard board that can be programmed and connected to the system without any need of sophisticated PCB design and implementation. It is an open-source hardware and anyone accordingly can be used. The microcontroller ATMEGA328 pin configuration live A/D I/D pins, clock speed, flash memory, and power supply [15, 17].

Communication system NRF24L01—NRF24L01 is a wireless trans-receiver which transmits data from one place to other at a distance of 1 km. This module has inbuilt antenna and its gain peak is 2 Dbi [18].

LCD 16*2 Display—The display unit is implemented to display all the parameters of the system. The instruction which is transmitted by receiver and controlled by Arduino is displayed on LCD (16 * 2). 16 * 2 represents 16 bit and to represent lines on the display. The maximum line that can be scrolled this display is 6. So, it can show only six instructions at a time. The scrolling function is provided by interfacing of buttons through Arduino [15].

Secondary Supply System—This system consists of an Li–Po battery and rechargeable circuitry. This battery provides the power to the transmitter section only. The receiver section is supplied by power supply.

Miscellaneous Components—In this category, alarm system, LED indicator, and microphone are connected to the proposed system. Passive buzzer is used for the alarm system. These buzzers are of different types that can be piezoelectric buzzers and electromechanical buzzers. These buzzers consist of electrical oscillating voltage to audio signals. The electrets microphone is used for the proposed system. It is a condenser-type microphone in which a diaphragm and metal conductor plate are fixed. The lower the distance of these two plates, the higher the capacitance gain. This mechanism is responsible for the audio signal recording. Light-emitting diode (LED) is used as indicator at various places and systems. However, the proposed system also uses voice instructions as an indicator.

Table 1 Various hardware specifications

Component (s)	Specification
Arduino Nano ATMEGA328	8 analog input and output pins, 22 digital input and output pins, 8-bit microcontroller 32 K of flash memory, 1 K of EEPROM, and 2 K of internal SRAM
Power Supply	12 V, 1 A
Voice recognition module	It supports 80 voice commands in all max 7 voice commands work at the same time. Digital interference 5 V TTL level for UART interface and GPIO, Analog Interface: 3.5 mm mono-channel microphone connector + microphone pin interface. Arduino library is supplied
Speaker	Loudspeaker: 8 Ω, 0.5 W
Voltage regulator IC7805	Input voltage range 7–35 V, Current rating, I = 1A, V_{min} = 4.8 V, Vmax = 5 V
Wireless transceiver NRF24L01	Operating range is about 1 km and antenna gain peak is 2 Dbi
Display	LCD16 × 2 Parallel LCD display, it has six lines
Battery	BMS battery, rechargeable battery, voltage up to 3.7 V
Alarm	Operating voltage 3.3–5 V, it is 5 V DC part active buzzer. Positive external 3.3–5 V voltage, Negative external GND out external microcontroller IO port
LED	5 mm white, Maximum Current 20 mA
Microphone	Electret microphone, operating voltage and frequency 2–10 V and 20–6,000 Hz, Recommended operating voltage 2 V, Current consumption 0.5

3.2 Component Specifications

Table 1 illustrates the various hardware components used in the proposed system and their specifications as demonstrated in Table 1 [13–18].

4 Experimental Results

The actual experimental test setup is developed and tested in the laboratory as shown in Fig. 8 (input and output devices). The computational algorithm described in Sect. 2.3 is burnt on the Arduino controller units (in transmitter and receiver sections). Several tests are performed in order to get the voice recognition of the respective person/user, and desired output is obtained on the LCD display unit.

Using microphone (input unit), one can send the commands by speaking after assistant sound to help. The entire conversation between the user and the display unit is away from the user or the service staff. The actual test results of this conversation are shown in Fig. 9a–d.

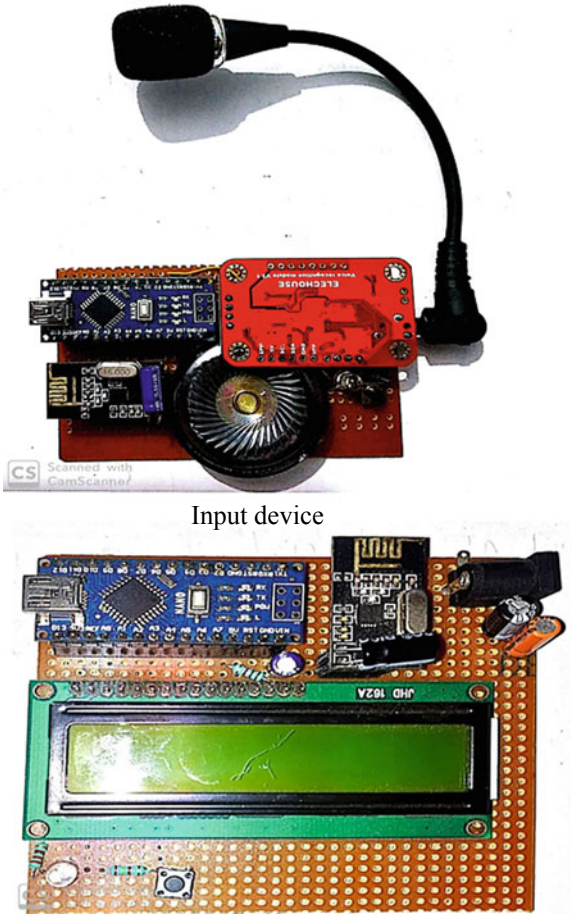


Fig. 8 Experimental setup (input and output device)

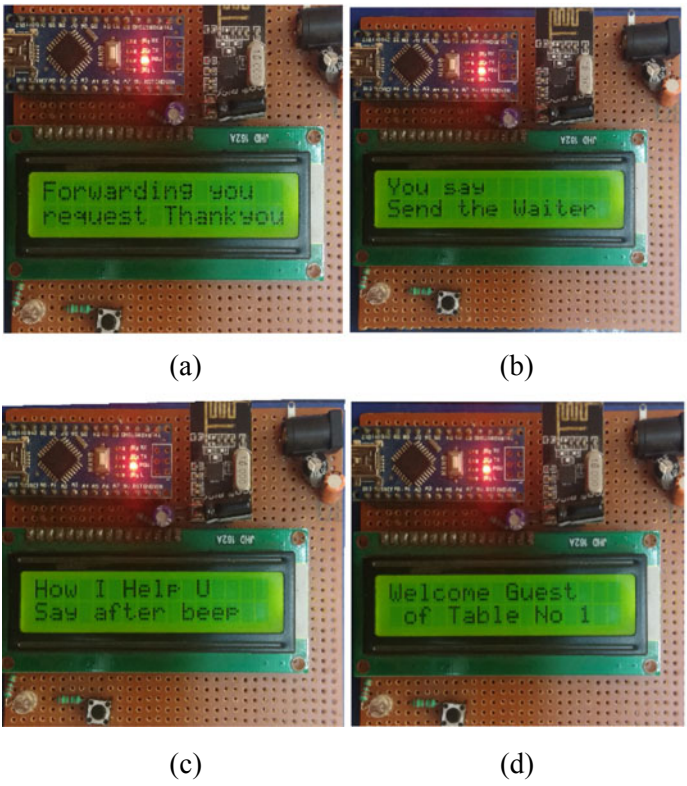
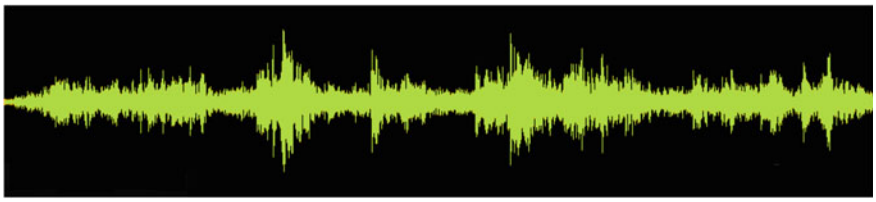


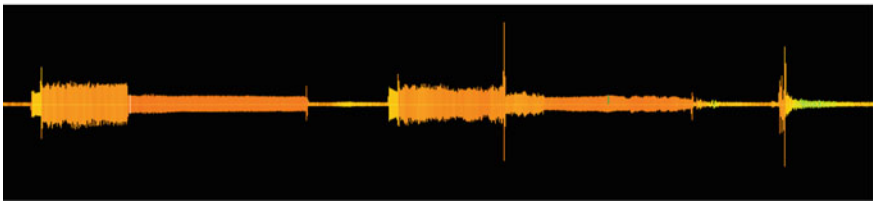
Fig. 9 **a–d** Test results of voice conversation between user and service staff



(a) How may I help you?



(b) Please Speak One



(c) Forwarding your message

Fig. 10 **a–c** Test results of voice conversation between user and service staff

The voice recognition patterns of this conversation are also plotted to show the process of speech recognition to the transmitter section as shown in Fig. 10a–c. These plots also verify the successful transmission of voice signals that are received appropriately at receiver end or not.

5 Conclusion

This paper proposes smart voice recognition system for the various applications at restaurants, reception desks, hotels/motels, customers' outlets, food courts, and many others. Using proposed system, customer(s) can order by voice to waiter or service staff from the seating place or room. This system consists of voice transmitter and output receiver devices. This voice order or command is processed by Arduino nanocontroller and displayed it on the appropriate display unit that can read or noted by waiter or service staff and one can get haste-free service or delivery in a sort of time. The experimental test setup has been successfully tested, and different results are obtained which validate its supremacy.

References

1. A. Bala, A. Kumar, N. Birla, Voice command recognition system based on MFCC and DTW. Int. J. Eng. Sci. Technol. **2**(2), 7335–7342 (2010)
2. R.L Clayton, D.L.S Winter, Speech data entry, result of test of voice recognition for survey data collection. J. Off. Stat. **8**(3) 377–388 (1992)
3. K. Kannan, Dr. J. Selvakumar, Arduino based voice controlled robot Int. Res. J. Eng. Technol. (IRJET) **2**(1) (2015). ISSN 2395-0072
4. M.I. Malik, T. Bashir, Mr. O. Farooq Khan, Voice controlled wheelchair system. Int. J. Comput. Sci. Mob. Comput. **6** (411–419) (2017)
5. V.H. Arul, M. Ramalatha, A study on speech recognition technology. J. Comput. Technol, **3**(7), 2278—3814 (2014)
6. B. Cui, T. Xue, Design and realization of an intelligent access control system based on voice recognition, in *ISECS International Colloquium on Computing, Communication, Control, and Management*, Sanya, China, vol. 1, pp. 229–232
7. C.Y. Yeo, S.A.R. Al-Haddad, C.K. Ng, Animal voice recognition for identification (ID) detection system, Penang, Malaysia, pp. 198–201, (2011)
8. A. Kanawade, S. Varvadekar, D R Kalbande, P. Desai, Gesture and voice recognition in story telling application, in *International Conference on Smart City and Emerging Technology (ICSCET)*, Mumbai, India, pp. 1–5 (2018)
9. F. Johnson, S. Garza, K. Gutierrez, Research on the use of voice to text applications for professional writing, in *IEEE International Professional Communication Conference (IPCC)*, Austin, TX, USA, pp. 1–5 (2016)
10. R.S.A. Abdalrahman, B. Bolat, N. Kahraman, A cascaded voice biometric system. Procedia Comput. Sci. **131**, 1223–1228 (2018)
11. S.H. Mohammadi, A Kain, An overview of voice conversion systems. Speech Commun. **88**, 65–82, (2017)
12. http://www.vensi.com/voice-recognition-applications-to-automate-industrial-operations/. Accessed March 2019
13. https://www.advanced-media.co.jp/english/aboutus/amivoice. Accessed March 2019
14. https://robokits.co.in/. Accessed March 2019
15. https://robu.in/. Accessed March 2019
16. https://github.com/elechouse/VoiceRecognitionV3.gi. Accessed March 2019
17. https://www.arduino.cc/. Accessed March 2019
18. https://www.robimek.com/arduino-ile-nrf24l01-rf-modul-kullanimi/. Accessed March 2019

IoT-Based Cross-Functional Agribot



Suraj Sudhakar and Sharmila Chidaravalli

1 Introduction

India is a land of agriculture. Agriculture constitutes a major portion of Indian economy, i.e. about 15–20% of the gross domestic product. In fact, India is the largest producer of rice, wheat, pulses and spices in the world.

In recent times, due to urbanization, deforestation and soil erosion, many farmers migrate to urban areas along with their families, find new occupation and, therefore, lead to lack of farmers for agricultural practices. It is also difficult for small-scale farmers to carry out all agricultural works. Therefore, automating agricultural tasks such as ploughing, seeding, levelling and irrigating land is necessary. Many researchers have worked on automating farming by harvesting crops using image processing and predicting the exact date to harvest crops based on certain parameters which fall under the domain of artificial intelligence and robotics.

Agribot is a robot which is used to substitute complex tasks which involve manpower. The robot is designed to aid farmers by simplifying their tasks. The project is developed using technologies such as IoT and robotics [1]. It is a wireless controlled robot which is used to perform the tasks of ploughing, seeding, levelling and watering the land. Most of the existing systems use fuel to perform tasks, for example, tractors, seeders and water pumps, while the agribot harnesses solar energy for its working [2].

S. Sudhakar · S. Chidaravalli (✉)
Global Academy of Technology, Bengaluru 560091, India
e-mail: c.sharmila11@gmail.com

S. Sudhakar
e-mail: surajsudhakar3@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_159

2 Literature Review

Over the past decade, various measures have been taken to automate the agricultural industry. The main motive of this is to reduce human intervention and ensure efficient usage of resources. An agricultural robot can be an efficient replacement in performing regular agricultural tasks. The following survey of previously designed agribots gives an insight into how to optimize and improve the existing system of robotic culture.

According to the work carried out by Akhila et al. [3], a robot can be designed and developed to plough the field, sow seeds and cover the seeds with soil. The system makes use of a stepper motor, a DC motor, PSoC and relay as its controller. In line with Dasari et al. [4], a robot can be used in the agricultural field for surveillance purposes. The robot will carry a camera and send the data being captured to a computer through a Wi-Fi network. It also performs ploughing, distribution of seeds and irrigation with the help of servomotors. Additionally, it consists of an ultrasonic sensor to avoid obstacles and Arduino Uno is used as the central controller.

Palepu et al. [5] have developed a vision-based agricultural robot, wherein the robot navigates with the help of optical sensors. The location of the robot can be identified with a map generated using Global Positioning System (GPS). Durga et al. [6] proposed a robot that performs ploughing, seeding and irrigation and can be controlled via Bluetooth. Shahik et al. [7] proposed a system that is controlled by Bluetooth and magnetometer. It performs ploughing, seeding and levelling, and indicates the starting of irrigation.

Based on the idea of eliminating the usage of batteries, Rahul et al. [8] made a proposal to harness solar energy when irrigation was not performed. The proposed robot was to make use of wireless sensor networks and GPRS technology and with the help of several sensors, plant health was monitored. It also had a smart weather-based irrigation system that sensed the moisture content in the soil and irrigated accordingly.

Amritanshu et al. [9] developed a system using Dual-Tone Multi-Frequency (DTMF). Using this technology, the agricultural robot could be controlled via a cell phone and could communicate over long distances. It included automatic detection of obstacles and moisture sensors for sensing the moistness of the soil.

Observing the research that has been conducted, it is evident that a wireless agribot that can be used without batteries is the best system to be used as an agribot. Therefore, in this paper, we propose a system that makes use of a solar panel to power the robot that can perform ploughing, seeding, irrigation and levelling functions. The robot will be remotely controlled through a mobile phone connected using Bluetooth.

3 Proposed System

The proposed system aims to provide a compact eco-friendly agricultural robot which is used to perform labour-intensive agricultural tasks such as ploughing, seeding, levelling and watering (Fig. 1).

3.1 Hardware Components

A solar panel is used to harness solar energy where direct sunlight is converted into useful electrical energy. The solar panel is mounted on top of the system where it can imbibe sunlight for optimum use. The ends of the solar panel are connected to a 12 V DC power supply which acts as the powerhouse for the entire system. A crystal oscillator is used to produce a periodic electronic analog signal. ARM7 microcontroller LPC2148 is the kernel of the system. It is a 32-bit processor with least power consumption when compared to other boards. It has a RAM onboard ranging from 8 to 40 and 512 kb flash storage. The board requires an external power supply. A 32-bit LCD display LC1621 is used to show commands under execution. Two separate motor drivers L293D are used to control four DC motors. Motor 1 is connected to a cylindrical tube which comprises a hole used for seeding. Motor 2 is attached to a connecting rod which is equipped with blades used for ploughing land. Motors 3 and 4 are used to provide locomotion to the system, to move in a direction either forward, backward, left or right. The motor drivers provide different sets of

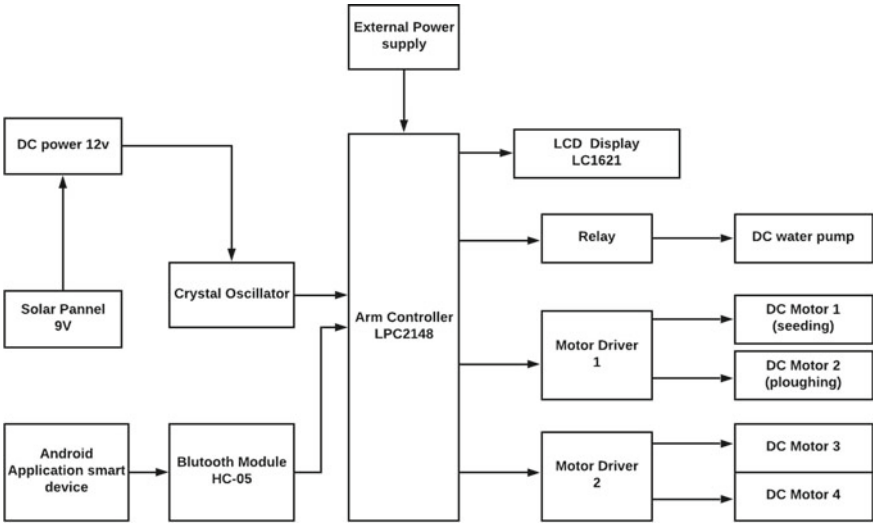


Fig. 1 Architecture of the proposed system

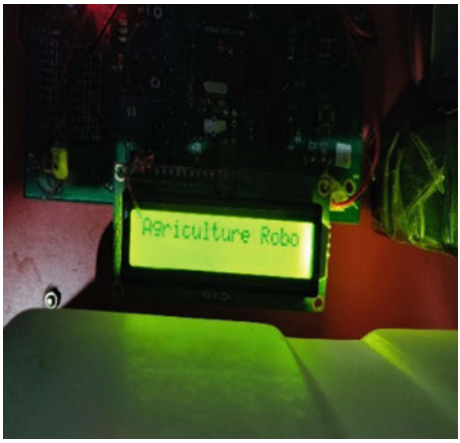


Fig. 2 LCD display unit on power ON

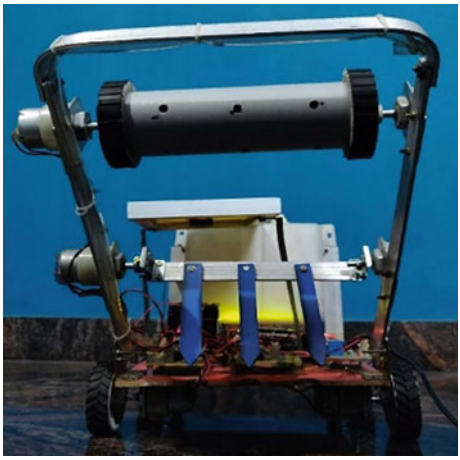


Fig. 3 Anterior view of agribot

signals for each possible outcome to the respective motors. A relay is used to convert AC which flows through the MCU as a DC input to the water pump which pumps water from the tank and thus irrigates the land (Figs. 2, 3 and 4).

3.2 Software Components

An Android application is used to provide inputs to the system. A single terminal is modulated continuously to perform a task. Commands are sent through the application via a wireless Bluetooth connection to the microcontroller. A Bluetooth module is used which establishes an intranet connection with a 2.4 GHz radio



Fig. 4 Side views of agribot

Table 1 Comparison of existing agricultural mechanisms and agribot

	Parameters	Farmer (Human)	Tractor	Seeder	Agribot
1.	Manpower	High	Limited	No	No
2.	Time	High	Limited	Low	Low
3.	Seeding Technique	High	Manual	Automatic	Automatic
4.	Power supply	Manually	Very high	Less	Low
5.	Expenditure	More	More	Initial investment	Lesser
6.	Rate of pollution	More	More	No	Nil

transceiver and baseband. The module supports bidirectional communication (Table 1 and Figs. 5, 6, 7).

4 Methodology

The ARM microcontroller LPC2148 consists of two 10-bit analog-to-digital converters. When the user runs the mobile application, an intranet connection is established via the Bluetooth terminal based on Serial Port Protocol (SPP). The following commands and actions can be executed on the agribot (Fig. 8).

4.1 Locomotion of the Bot

The commands to move forward or backward is *F and *B, respectively. When the user inputs the command, it is transferred via Bluetooth to the MCU as an analog signal. The analog signal is then converted into a digital input with the help of analog-to-digital converter which is processed in the MCU. The output is fed into motor driver 2 which controls the DC motors 1 and 2. On receiving the signal, the two motors rotate either clockwise or anticlockwise by applying the same amount

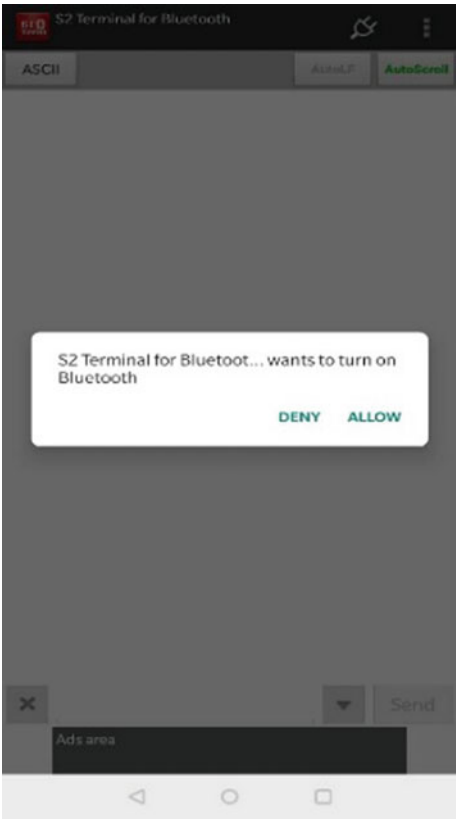


Fig. 5 Android application to control the agribot

of force in the same direction. Thus, it creates a linear motion. The body continues to remain in its state unless a stopping criterion is met.

*L and *R are the commands used for left and right motions, respectively. When a user inputs these commands, it is transferred via Bluetooth to the MCU similar to the forward/backward commands. The MCU converts the analog signal into digital input which is processed within the MCU. The output is fed to L293D motor driver 2. If the body has to turn left, then only the DC motor 3 (right) will be activated. Similarly, if the body has to turn right, then DC motor 4 will be activated. However, in such cases, both the motors are independent. The body continues to remain in its state until the next criterion is met.

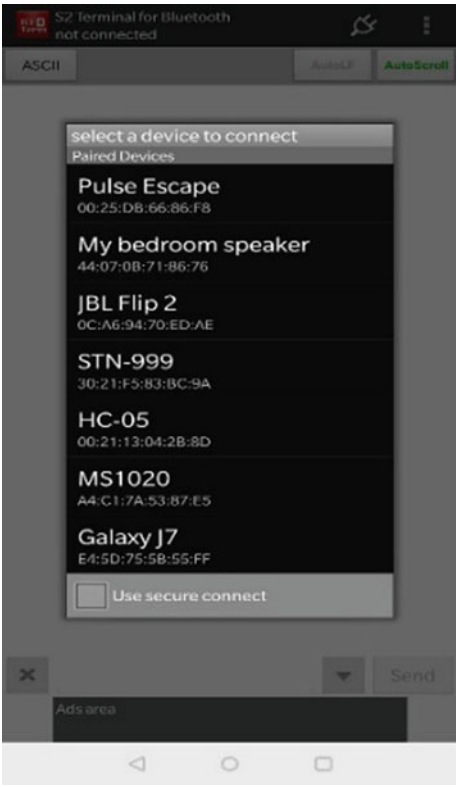


Fig. 6 Connecting to Bluetooth HC-05

4.2 Ploughing

A single DC motor is used. A connecting rod which is well equipped with ploughing blades is then welded perpendicularly to the rod [10]. The two ends of the motor are connected to L293D motor driver 1. The motor and the connecting rod rotate simultaneously, and the blades acquire a circular motion which is perpendicular to that of the motor. Hence, the blades move in either clockwise or anticlockwise direction. The commands used are *1 to plough clockwise and *2 for the opposite. However, both the operations can be stopped by pressing *3 which is the halt command for this task (Figs. 9, 10 and Table 2).



Fig. 7 Command prompt for agribot

4.3 Seeding

A hollow cylinder with one end that has a lid to add the seeds is used for seeding [10]. The other end of the cylinder is connected to DC motor 1. Evenly spaced holes are drilled on the cylinder, the holes spaced at optimum distance between seeds for plant growth. The command to start seeding operation is *4 which results in the cylinder rotating in clockwise direction and *5 for anticlockwise direction. As the motor rotates, the cylinder starts to drop seeds through the holes in equal intervals. *6 is the halt command for this task (Figs. 11, 12 and Table 3).

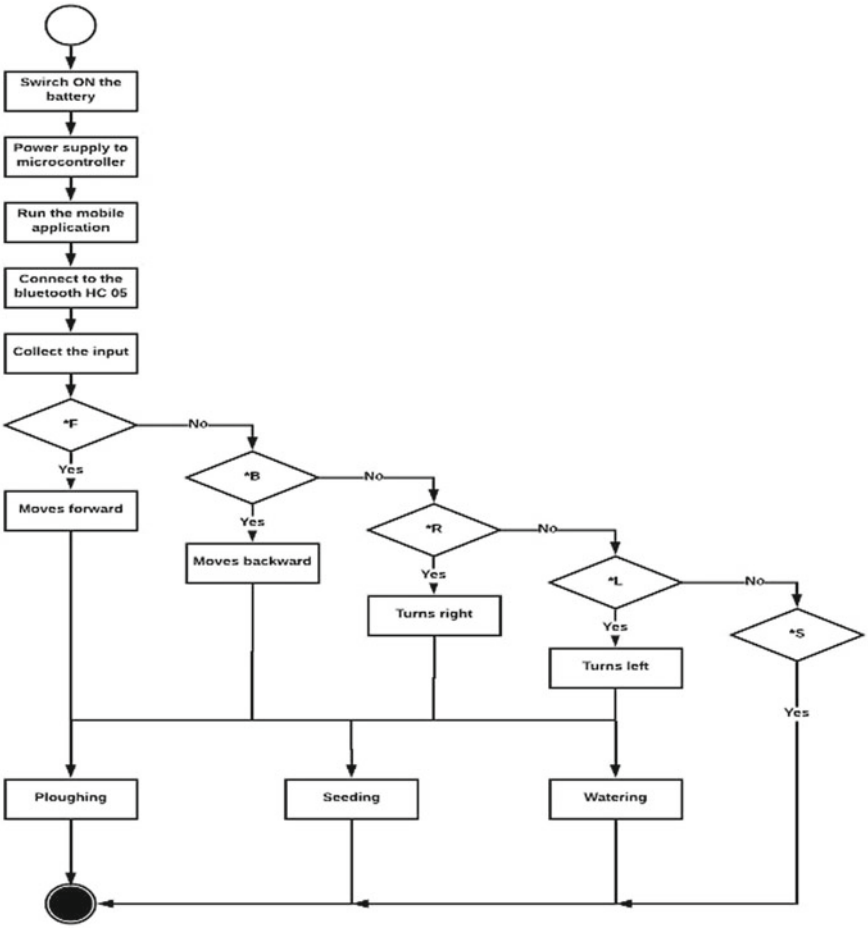


Fig. 8 Flowchart representing the procedures of the agribot

4.4 Watering

The system contains a compact 5-l water can that needs to be filled when required. On inputting *7, an electronic DC pump is activated to push water from the can through thin watering pipes which are attached frontally. The pipes have a set of distributed holes through which the agribot waters land efficiently. Once the task is completed, the user must input *8 to turn off the DC pump from pumping water (Figs. 13, 14 and Table 4).

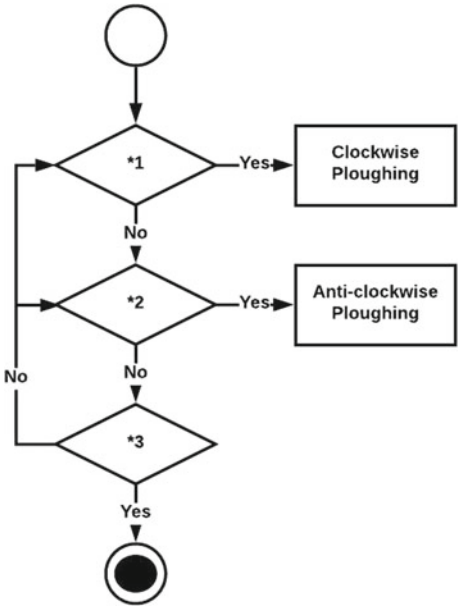


Fig. 9 Ploughing commands



Fig. 10 Agribot performing ploughing procedure

Table 2 Ploughing function analysis

Sl. no	Parameters	Value obtained
1.	Torque of Motor 2	4 kg cm
2.	Speed of Motor 2	10 RPM
3.	Distance covered by agribot	10.8 m
4.	Time taken to plough	53 s

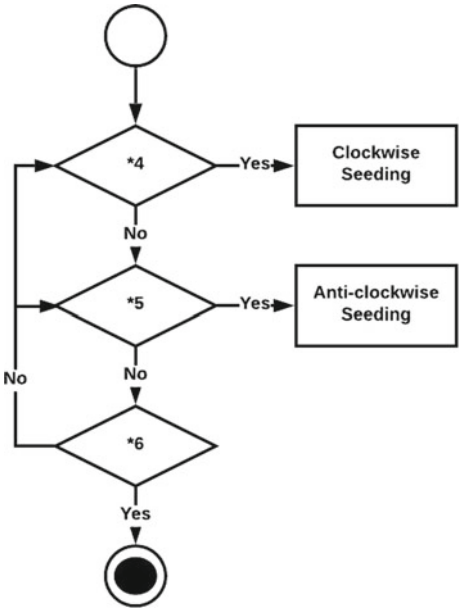


Fig. 11 Seeding commands



Fig. 12 Agribot sowing seeds

Table 3 Seeding function analysis

Sl. no	Parameters	Value obtained
1.	Torque of Motor 1	4 kg cm
2.	Speed of Motor 1	10 RPM
3.	Area covered by the agribot	5 m × 5 m
4.	Number of seeds sowed per minute	45
5.	Time taken to sow seeds	1 min

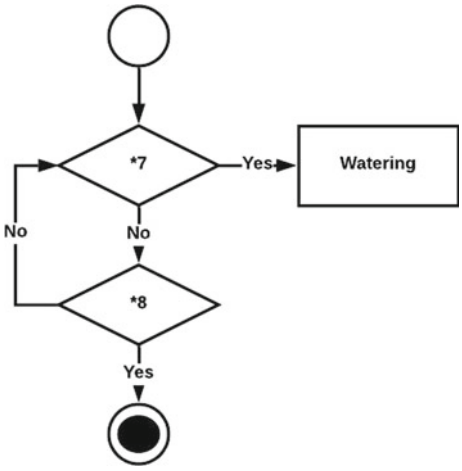


Fig. 13 Commands for watering

4.5 Levelling

The system is equipped with a small T-shaped wooden plank which is attached towards the posterior end of the agribot. As the Bot moves forward, the soil gets levelled by the plank, resulting in an even distribution of soil (Fig. 15 and Tables 5, 6).

5 Future Enhancement

Agribot can be made autonomous by specifying a few crucial parameters as input by implementing deep neural networks. This can further be extended to swarm intelligence where multiple agribots work on the same land with a swarming behaviour using algorithms such as ant colony optimization or particle swarm optimization. By integrating image processing outcomes and machine learning algorithms such as decision trees and Naïve Bayes, the system could be successful in discrimination of ripe fruits from the unripe ones. The same procedure can also be used to predict the exact time and date to harvest crops for best possible yield.



Fig. 14 Agribot irrigating the soil

Table 4 Watering function analysis

Sl. no	Parameters	Value obtained
1.	Flow rate of electric pump	80 L/h
2.	Maximum lift	40–110 mm
3.	Area covered by the agribot	5 m × 5 m
4	Time taken for watering	1 min

6 Conclusion

A technological solution is provided for the issue of lower agricultural manpower in this paper. Agribot is a multipurpose robot that is designed for an agricultural purpose which can carry out elementary operations such as ploughing, seeding, levelling and watering the land, resulting in a reduction of human intervention.

The proposed system has been successfully implemented and tested under rigorous conditions. Agribot has been developed in C. This system will have a huge impact in today's agriculture precision as it will increase the efficiency and accuracy of farming. Agribot is eco-friendly; it does not depend on any nonrenewable energy



Fig. 15 Agribot levelling the soil

Table 5 Ground levelling function analysis

Sl. no	Parameters	Value obtained
1.	Distance covered by the agribot	10.8 m
2.	Time taken for levelling	1 min
3.	Maximum weight for levelling	300 g

Table 6 Solar cell charging function

Sl. no	Parameters	Value obtained
1.	Working voltage	9 V
2.	Expected working voltage	8.8 V
3.	Operating current	220 mA

for carrying out its tasks. Agribot is a compact and economical system which does not require huge capital investment. If it is used for agricultural practices on a daily basis, the agribot can be considered a farmer's best friend.

References

1. A. Lalwani, M. Bhide, S.K. Shah: A review: autonomous agribot for smart farming, in *46th IRF International Conference* (2015)
2. D.S. Rahul, S.K. Sudarshan, K. Meghana, K.N. Nandan, R. Kirthana, P. Sure, IoT based solar powered agribot for irrigation and farm monitoring, in *Proceedings of the Second International Conference on Inventive Systems and Control, ICISC* (2018)
3. A. Gollakota, M.B. Srinivas, Agribot—a multipurpose agricultural robot, in *2011 Annual IEEE India Conference*

4. D.N. Vinod, T. Singh, Autonomous farming and surveillance agribot in adjacent boundary, in *2018 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, Bangalore (2018), pp. 1–7
5. P.V. Santhi, N. Kapileswar, V.K.R. Chenchela, C.H.V.S. Prasad, Sensor and vision based autonomous agribot for sowing seeds, in *2017 International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS)*, Chennai (2017), pp. 242–245
6. K.D. Sowjanya, R. Sindhu, M. Parijatham, K. Srikanth, P. Bhargav, Multipurpose autonomous agricultural robot, in *2017 International Conference of Electronics, Communication and Aerospace Technology (ICECA)*, Coimbatore (2017), pp. 696–699
7. K. Shaik, E. Prajwal, B. Sujeshkumar, M. Bonu, V.R. Balapanuri, GPS based autonomous agricultural robot, in *2018 International Conference on Design Innovations for 3Cs Compute Communicate Control (ICDI3C)*, Bangalore (2018), pp. 100–105
8. D.S. Rahul, S.K. Sudarshan, K. Meghana, K.N. Nandan, R. Kirthana, P. Sure, IoT based solar powered Agribot for irrigation and farm monitoring: agribot for irrigation and farm monitoring, in *2018 2nd International Conference on Inventive Systems and Control (ICISC)*, Coimbatore (2018), pp. 826–831
9. A. Srivastava, S. Vijay Vargiy, A. Negi, P. Shrivastava, A. Singh, DTMF based intelligent farming robotic vehicle, in *International Conference on Embedded Systems—(ICES 2014)*, Amrita Vishwa Vidhyapeetham, Coimbatore, Tamil Nadu, India
10. B.S. Shivaprasad, M.N. Ravishankara, B.N. Shoba, Design and implementation of seeding and fertilizing agriculture robot. Int. J. Appl. Innov. Eng. Manag. (IJAIEM) **3**(6) (2014)

A Novel Interfacing Scheme for Analog Sensors with AMLCD Using Raspberry Pi



Peeyush Garg, Ajay Shankar, Mahipal Bhukya and Vinay Gupta

1 Introduction

The analog sensor means that the device can provide continuous output signal of analogous type, and this analog-output signal of sensor is proportional to the measured quantity. Nowadays, variety of analog sensors are available for physical signal measurement, for example, voltage sensor, accelerometer, pressure sensor, light sensor, temperature sensor, humidity sensor, etc.

In the designed system, the measured values by the sensor will be fetched to the Raspberry Pi, which will be programmed using code written in Python and then the data will be visualized on an AMLCD panel compatible with it. Both digital and analog sensors can be interfaced with Raspberry Pi using the code written in Python or by wrapping a C code. GPIO is the library used for the interfacing of sensors with the GPIO pins of the hardware.

The algorithm associated with the block diagram of interfacing system as shown in Fig. 1 is given below:

Step 1—Booting the KERNEL.
Step 2—Connecting the sensor.
Step 3—Observing the sensor value from terminal.
Step 4—Observing the sensor value from Python program.

P. Garg (✉) · A. Shankar · M. Bhukya · V. Gupta
Department of Electrical Engineering, Manipal University Jaipur, Jaipur, India
e-mail: Peeyush01garg@gmail.com

A. Shankar
e-mail: asjaiswal@hotmail.com

M. Bhukya
e-mail: mahipalbhukya@gmail.com

V. Gupta
e-mail: vinra_20@rediff.com

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8_160

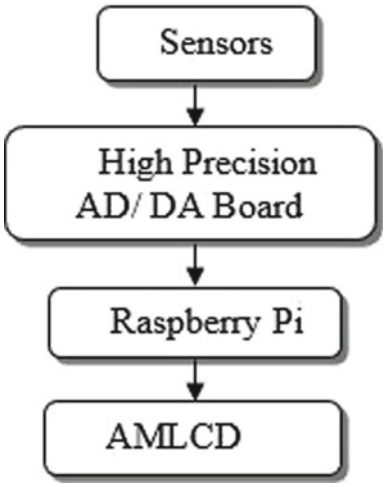


Fig. 1 Interfacing block diagram of system

Step 5—Displaying results on AMLCD panel.

Prerequisites for getting started with Raspberry Pi includes the Raspbian OS, PuTTY, and VNC server. The Raspbian Jessie pixel has been used as the operating system on the Raspberry Pi. Win32DiskImager was used to burn the SD card with the OS image. PuTTY is an open-source terminal emulator, serial console, and network file exchange software. Virtual network computing (VNC) is a graphical user interface for sharing data virtually between two systems and makes use of the remote frame buffer protocol (RFB) to control another computer remotely.

2 Literature Survey

Abileah and Green, Beaverton, USA, [1] conducted a study to show design and applications of optical sensors embedded with AMLCD panel. The paper explained about the various optical methods including finger, stylus and LED pen, laser, etc. The cover glass plate can facilitate not only as a protective plate but also as a light guide in night mode. It is also not compromised with display quality. It features are sleek and lightweight, having linear and accurate touch sensitivity over the display panel.

Tsai et al. [2] proposed the design of delta-sigma analog-to-digital (A/D) converter. It was implemented with the thin-film transistors (TFTs) on glass substrate and packed using 3-μm low-temperature polycrystalline silicon process. It was also verified experimentally that the probability in the Boolean output bitstream from the delta-sigma modulator correctly fits in the analog-input voltage ratio. The bitstream can be transformed into 8-bit Boolean code under the operating voltage of 10 V with thin-film transistors (TFTs) on the glass substrate.

Swaroop [3] paper showed the measurement of real-time temperature using Raspberry Pi. The measured temperature can be showed on Raspberry Pi kit using command. The suggested method aimed at uninterrupted monitoring of the real-time temperature in a cost-effective way by setting fixed interval and monitored by Raspberry Pi. The sensor used was DS18B20 temperature sensor. The sensor comes in metal can package and small in size. The measurement of temperature is quite precise comparatively. Jumper wires are used to connect the sensors to the Raspberry Pi. The Raspberry Pi-based system is used to store and display the real-time temperature.

Yu et al. [4] developed a 4-inch (QVGA, 320 × 240, 262,144 colors) amorphous silicon thin-film-transistor liquid-crystal display (TFT-LCD) which has an embedded color image scanner.

The Raspberry Pi as shown in Fig. 2 is a chain of comparatively smaller single-board computers developed by the Raspberry Pi foundation. The Raspberry Pi 3 means third-generation Raspberry Pi. The Raspberry Pi 3 is about the same size as the previous RPi 2 and Pi 1 models and also has full compatibility with previous module.

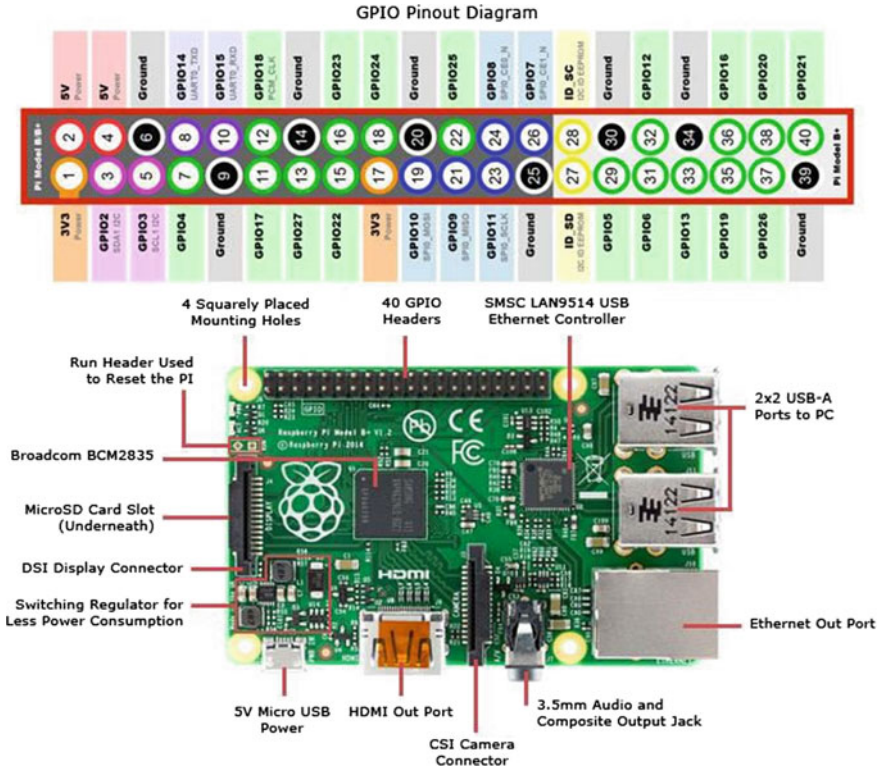


Fig. 2 Raspberry Pi 3 Model B and pin diagram [5]

The Raspberry Pi 3 is with a quad-core Cortex-A53 processor. It featured 10 times better performance when compared with Raspberry Pi 1. It is also observed that the speed of Raspberry Pi 3 is to be around 80% quicker than the Raspberry Pi 2 during parallel processing.

3 Raspberry Pi Communication Protocols

Raspberry Pi has three communication protocols included in its GPIO pins. These are turned off in normal mode. The configuration of RPi protocol needs to be turned on to use these communication protocols. These protocols provide extra functionality to the Pi. The serial peripheral protocol (SPI) is a method to communicate with digital device serially. This protocol was first introduced by Motorola in the late 1980s, and then it became the universal standard. This is especially used to send and receive data serially for short-distance communication especially around the embedded system module. The peripherals such as liquid-crystal display (LCD), keyboard, and joysticks are generally connected to the processor unit using SPI protocol. Generally, master–slave configuration is used to efficiently develop the multi-peripheral system. The GPIO pins 7, 8, 9, 10, and 11 on RPi are used as SPI pins. These pins include serial clock (SCLK), chip select, master out slave in, master in slave out, and master out master in.

The I2C protocol also used as connect peripheral IC to microcontroller or microprocessor board. This protocol was introduced by Philips Semiconductor. The I2C pins comprise a preset 1.8 Kilo-ohms pull-up resistor to 3.3 V supply. This means that they are not suitable for use as general-purpose IO where a pull-up is not required.

UART (Universal Asynchronous Receiver/transmitter): There are two types of UART protocols available on Pi, one associated with the Bluetooth control (AMA0) and one for the GPIO pins Rx and Tx (S0). The UART used here is the GPIO Mini-UART (S0). Settings need to be changed to stop using AMA0 and use S0 with full control and without any loss of frequency. This port is used to interface the LCD which will be discussed later.

The mini-UART is a secondary UART projected to be utilized as a console. The mini-UART has the specification 7- or 8-bit processor, single start and stop bit, no parity bit, eight symbol deep FIFOs for receive and transmit, 16,550 registers.

4 Waveshare High-Precision AD/DA Board

An analog-to-digital converter (ADC) is a system hardware which converts an analog signal into digital signal further used by the RPi. The image captured by camera and sound received by the microphone is of analog type, the analog signal is fed into ADC in the form of voltage signal with specific magnitude. Then, the ADC converts it into equivalent digital information. The Digital to Analog (DAC) has the

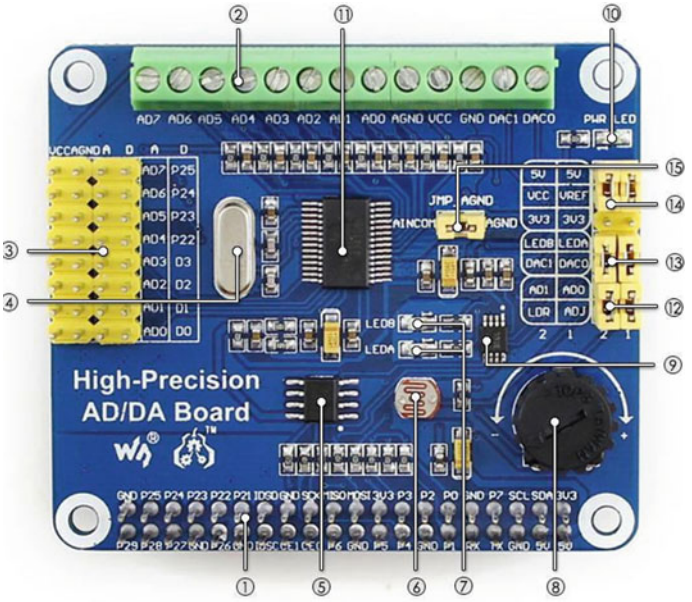


Fig. 3 High-precision AD/DA board

reverse functionality of ADC. The high-precision AD/DC has high bandwidth and high signal-to-noise ratio. The high-precision AD/DA board is shown in Fig. 3. The low-noise ADC (ADS1256), used in this board, has 24-bit analog-to-digital (A/D) converters. It provides whole high-resolution measurement.

5 4D Systems uLCD-43P—AMLCD

The 4D uLCD-43, an intelligent graphics display, is the epitome of displays at delivering a varied array of structures in a single, small, and cost-effective entity. The PICASO processor-based system comprises highly optimized virtual core engine: Extensible Virtual Engine (EVE). Wide-ranging peripherals are comprised of the system, which open the area of working for the user to use it in many projects. The front and rear views of AMLCD are shown in Fig. 4.

The uLCD-43 plays a significant role in designing any product that requires a brilliant color scheme, moving pictures on a 4.3″ widescreen display. The display module is well-designed amalgamation of 4.3″ 480 × 272 pixels, 65 K true color LCD screen, audio amplifier and speaker, micro-SD card connector, all along with a group of general-purpose input/output pins (GPIOs), which includes I2C and serial UART communications. The uLCD-43 has two dedicated hardware asynchronous serial

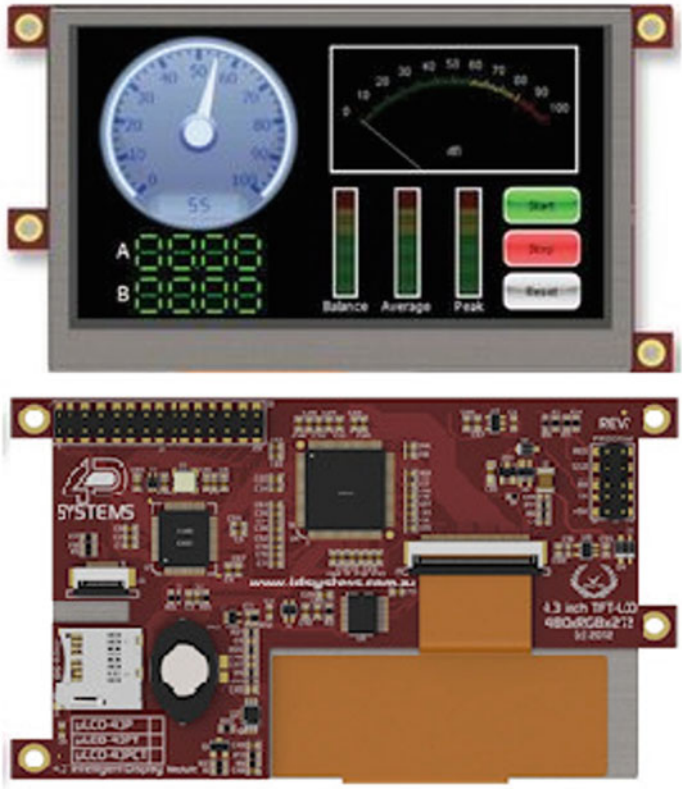


Fig. 4 4D system uLCD-43 (front and rearview)

ports COM0 and COM1 serial port that can transmit/receive signal with external devices in serial mode.

The Workshop4 IDE supports numerous development environments along with high functionality to provide diverse user requirements as well as expertise level. Workshop4 is all-inclusive software IDE for Microsoft Windows. Four things are pooled together by the IDE to develop complete 4DGL application code and these are editor, compiler, linker, and downloader. All user application code is done in the Workshop4 IDE. ViSi Genie is used for this project for the interfacing of RPi with LCD.

6 Interfacing with ADC

The waveshare high-precision AD/DA board is connected with Raspberry Pi board, and then all sensors are connected to it. After completing installation of library, the

WiringPI test will be conducted. The 4D workshop IDE software is used to create project, and then add uLCD-43p as AMLCD. The AMLCD is connected to PC with the 4D programming cable. The GUI is designed with the help of the various graphics like Gauge, Slides, Texts, Switches, Pictures, etc.

After this, the image of the design needs to be copied onto a micro-SD card which goes in the display panel by clicking on the button that says build copy/load. The card should be FAT formatted or else the image will not be copied on the card. The baud rate of the LCD must be set correctly and should match the baud rate used in the code. The SD card is put back into the display and the GUI will be displayed on the LCD as shown in Figs. 5 and 6. The display can be checked using a debugger called the gtx.

Outputs of Sensors on AMLCD

The designed system is used IR sensor, LDR, potentiometer ultrasonic sensors as analog sensors, which is going to be connected and utilized for the purpose of measure the analog signals. The IR sensor is used to check the existing of obstacle, LDR to measure the light intensity, ultrasonic sensor to measure the distance and potentiometer to measure the angular displacement. The results is well presented on AMLCD.

The IR sensor is interfaced with the system, and the output received at terminal is shown in Figs. 7 and 8.

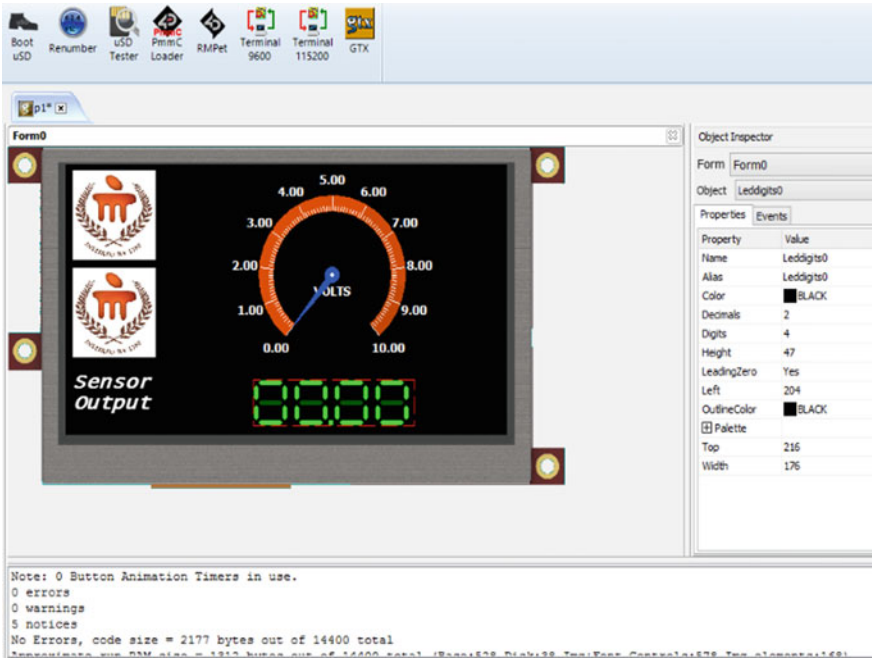


Fig. 5 GUI on LCD panel on 4D workshop IDE

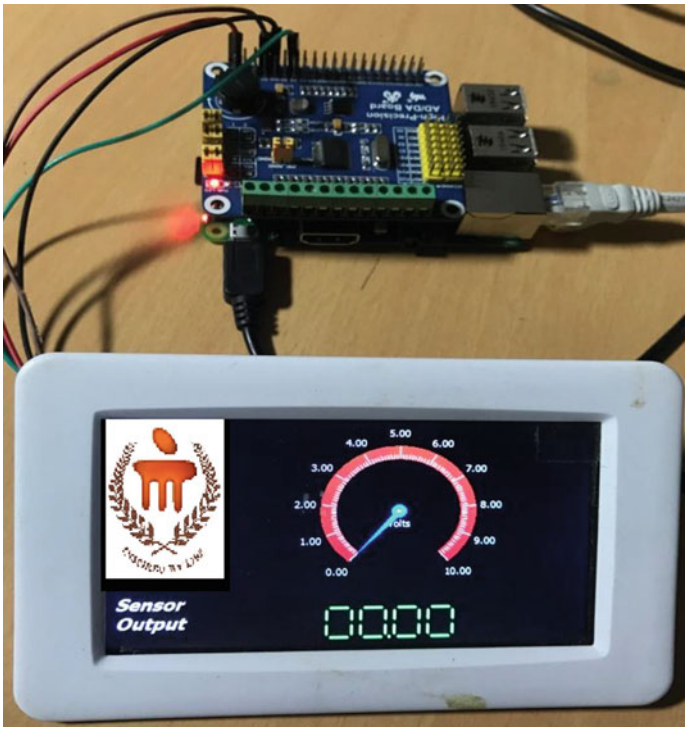


Fig. 6 LCD connection with RPi and ADC

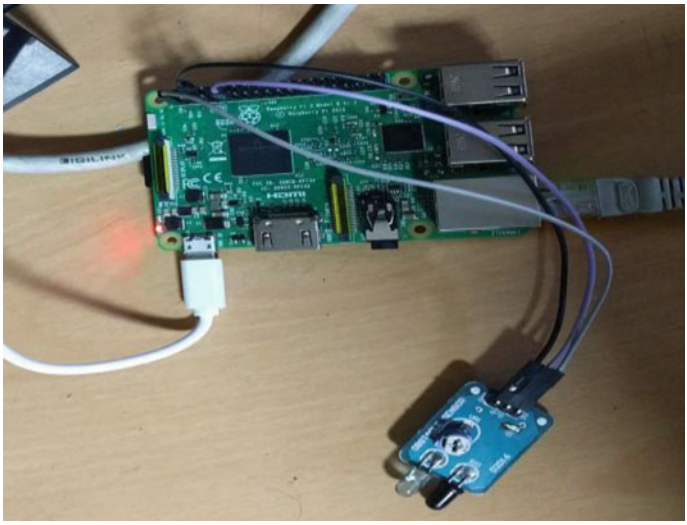


Fig. 7 IR sensor connection diagram

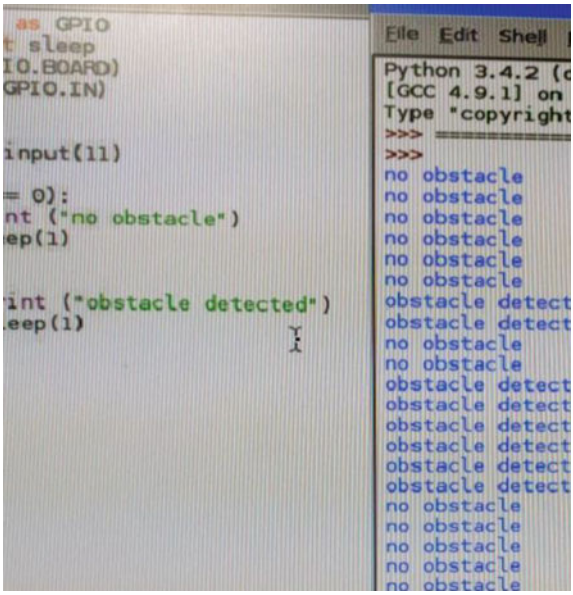


Fig. 8 Output of the code for IR sensor

The potentiometer is connected to channel 0 and its output can be varied. Here, for instance, the output comes out to be 4.75 and displayed on the MLCD as shown in Fig. 9.

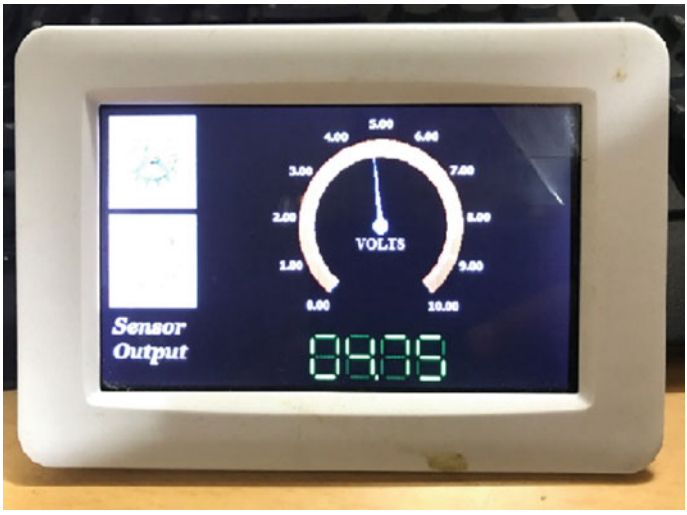


Fig. 9 Output of potentiometer on AMLCD

```
pi@raspberrypi ~ $ sudo python range_sensor.py
Distance Measurement In Progress
Waiting For Sensor To Settle
Distance: 12.52 cm
pi@raspberrypi ~ $
```

Fig. 10 Output of ultrasonic sensor



Fig. 11 Output of LDR on AMLCD

The ultrasonic sensor is also interfaced and output in terms of distance is reflected on terminal as shown in Fig. 10.

The LDR is channel 1 and the output can be varied as the light falling on the sensor changes. The output here is 3.74 which is displayed on the LCD as shown in Fig. 11.

7 Conclusion

The work was cohesion of software and hardware to produce the essential results. Raspberry Pi uses python to interface with any possible hardware expansion compatible with the device. Digital sensors can be directly interfaced with Pi but as there is no analog I/O pins available so therefore an ADC expansion board was used. All the

required libraries are to be installed for working out the integration of the hardware. The 4D display panel was used for displaying output with custom GUI and for easier interpretation of the data.

The sensor programming is free flowing which helps to manipulate the output as per the wish of the user and helps to integrate more sensors or hardware with Pi. Python has a memory management. It has a large and comprehensive standard library, and therefore it proves to be a better language to code on this powerhouse SoC.

References

1. A. Abileah, P. Green, Optical sensors embedded within AMLCD panel: design and applications, Planar Systems, Beaverton, OR, USA, 5
2. C.-C. Tsai, T.-M. Wang, M.-D. Ker, Implementation of delta–sigma analog-to-digital converter in LTPS process, vol. **9** (2009)
3. P. Swaroop, Y. Sheshank Reddy, E. Syed Saif, Sasikala, The real time temperature sensing using Raspberry Pi, Department of Electronics and Communication Engineering Saveetha School of Engineering, vol. 6 (2015)
4. J.-H. Yu, K. Choo, H. Kang, Y. Kim, D. Lee, I. Kang, I. Chung, P-7: 4 inch a-Si TFT-LCD with an embedded color image scanner, in *SID Symposium Digest of Technical Papers* (2007)
5. E. Upton, G. Halfacree, *Raspberry Pi User Guide*. Wiley (2014)

Author Index

A
Acharya, D. S., 937
Adhikari, Manoj Singh, 1389
Agarwal, Alok, 827, 1375
Agarwal, Bhavna, 423
Agarwal, Shobit, 633
Agarwal, Tanvi, 779
Aggarwal, Akshai, 105
Aghwariya, Mahesh Kumar, 779, 815
Agrawal, Bulbul, 1027
Agrawal, Harsh, 479
Agrawal, Mansi, 479
Agrawal, Nikhil, 1197
Ahmed, Suhaib, 1423, 1433
Akhilesh, P. V., 1565
Anirudh, Vattikuti, 1565
Anjana, S., 451
Arora, Shaveta, 995
Ashok, Alaknanda, 559
Asthana, Shubham, 1553
Awate, Vikarm, 795

B
Baba, Majid Irfan, 1423, 1433
Baghel, Amit, 1285
Bagwari, Ashish, 1277
Bahad, Pritika, 235
Bahadur, Promila, 271
Balakrishna, Bhukya, 1327
Bali, Bandana, 463
Banerjee, Mahesh, 423
Banerjee, Mudita, 473
Bansal, Dipali, 1443
Bansal, Praveen, 1197
Bansal, Vakul, 815
Baudh, Rishabh Kumar, 533, 689
Bedi, Harpreet Singh, 1093
Beliya, A. K., 811
Bhardwaj, Kaushal, 1017, 1219
Bhargav, Cherry, 1093
Bhatia, Shipra, 583
Bhatoye, Aishwarya Prasad, 1547
Bhatoye, Sauhardh Prasad, 1547
Bhat, Soha Maqbool, 1423, 1433
Bhattacharyya, D. K., 311
Bhukya, Mahipal, 1673
Boggavarapu, Aditya, 713

C
Chaubey, Nirbhay, 105
Chaudhari, Narendra S., 359
Chaudhary, Pravesh, 545
Chauhan, Ankita, 1139
Chauhan, Arvind, 1057
Chaurasia, Amit, 907, 1107
Chaurasia, Amita, 907
Chhatre, Swapnil, 1587
Chidaravalli, Sharmila, 1657
Chilwal, Bhavna, 1, 19
Chitreddy, Akhil, 1565
Chopra, Seema, 811

D
Dagdee, Nirmal, 183
Dalvi, Omkar, 245
Das, Arundhati, 1017, 1219
Das, B. K., 841
Dash, Nivedita, 575
Das, Prabin Kumar, 1461, 1487

© Springer Nature Singapore Pte Ltd. 2020

G. Singh Tomar et al. (eds.), *International Conference on Intelligent Computing and Smart Communication 2019*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-15-0633-8

Datta, Tamal, 1609
Dave, Mayank, 27
Deepak Nair, M. V., 583
Denish, Konjengbam, 1337
Devi, Rekha, 1509
Dewangan, Jaydeep, 625
Dhar, Rudra Sankar, 1383, 1405
Dixit, Manish, 1027
Dubey, Neha, 805
Dubey, Shiksha, 255
Dubey, Shradha, 1027
Dubey, Vikas, 805
Dwivedi, Ajay Kumar, 697
Dwivedi, Umesh Kumar, 907

G
Gaharwar, Mohit, 625, 713
Ganesh, Aman, 1347
Garg, Peeyush, 1673
Garg, Pranav, 1009
Gehlot, Anita, 1451, 1461, 1475, 1487
Ghosh, S. M., 287
Gogineni, Kailash, 1565
Gola, Kamal Kumar, 167
Govada, Roja Rani, 147
Goveas, Neena, 1009
Govinda, K., 85, 195, 207, 415
Goyal, Aashish, 915
Goyal, Archit, 479
Goyal, Samta Jain, 301, 349
Gupta, Anamika, 605, 657
Gupta, Anshu, 811
Gupta, Ashish, 985
Gupta, Ganesh, 115, 1413
Gupta, Hemant Kumar, 705
Gupta, Neeraj, 1413
Gupta, Raghav, 1093
Gupta, Rashmi, 1413
Gupta, Reetu, 183
Gupta, Rekha, 985
Gupta, Vaishnavi Km., 1461, 1487
Gupta, Vinay, 1673
Gupta, Vrinda, 127

H
Hakim, Najeeb-ud-Din, 1251
Harbola, Ayusha, 949
Hashmi, Farukh Md., 1451
Hijam, Deena, 61
Hinge, Abhiraj, 1009

I
Indu, S., 1211
Irungbam, Amit Kumar, 1317

J
Jadon, Rakesh Singh, 301, 349
Jaglan, Vivek, 115
Jain, Anuj, 407, 1525
Jain, Ishita, 1577
Jain, Naman, 1501
Jain, Palak, 589
Jain, Shubham Kumar, 907
Jain, Vasu, 633
Jana, Subrata, 867
Jaswal, Ram Avtar, 949
Jha, A. N., 889
Jindal, Poonam, 1231
Joshi, Anmol, 1501
Joshi, Ashish, 167
Joshi, Vikalp, 1389

K
Kabra, Ashwin, 383
Kakran, Sandeep, 1327
Kalita, Jugal K., 311
Kalra, Dheeraj, 1065, 1293, 1311
Kalyani, D., 85
Kamal, Rajeev, 495
Kamboj, Robin, 127
Kamboj, Vikram Kumar, 1171
Kansal, Vaishali, 27
Kanti, Jyotshana, 1277
Kanumuri, Deepak, 37
Kanungo, Priyesh, 183
Kashyap, Rani, 451
Katiyar, Hitanshu, 559
Kaul, Siddarth, 407
Kaur, Jagjeet, 805
Kaushik, Sakashi, 1127
Khandare, Nikhil B., 245, 359
Khan, Gulista, 167
Khan, Huma, 287
Khan, Ikhlaq Ahmed, 841
Khanna, Bhavika, 1093
Khanna, Kavita, 995
Khanna, Rajesh, 513
Khare, Bharat Bhushan, 721
Kharkongor, Carynthia, 1259
Khiangte, Lalthanpuii, 1383, 1405
Khosla, Anita, 473
Kohli, Ankur, 1547

Krishna, Vijaya A., 195
Kudtarkar, Darshan, 1057
Kumar, Akshi, 371, 1269
Kumar, Ashwani, 47, 729, 1365, 1577
Kumar, Devendra, 1065, 1293, 1311
Kumar, Divesh, 1065, 1293, 1311
Kumare, Jamvant Singh, 1077
Kumari, Soni, 159
Kumar, Kuleen, 1383
Kumar, Mahendra, 533
Kumar, Nitin, 1525
Kumar, Prashant, 1413
Kumar, Prateek, 1641
Kumar, Praveen, 815
Kumar, Rajesh, 1241, 1347
Kumar, Saket, 1627, 1641
Kumar, Sandeep, 643
Kumar, Sudhanshu, 1627
Kumar, Sunil, 327
Kumar, Tarun, 495, 559
Kumar, Trivesh, 523
Kumar, V., 85
Kumar, Vipin, 1347
Kumar, Vivek, 267, 605
Kushwah, K. K., 795
Kusuma Kumari, E., 677

L
Lata, Suman, 159
Lavanya, K., 7, 451

M
Mahapatra, Sheila, 889
Maheshwari, Ankur, 915
Malik, Nitin, 889
Malik, Praveen Kumar, 1487
Malik, Sumit, 1525
Manivannan, S. S., 207
Manzoor, Insha, 1423, 1433
Marriwala, Nikhil, 1151
Meel, Priyanka, 479
Meetei, Huidrom Hilengamba, 1337
Mehta, Deepak, 337
Meitei, Ingudam Chitrasen, 1317, 1337
Mishra, Manish Kumar, 805
Mishra, P. K., 1, 19
Mishra, Rupesh Kumar, 159
Mishra, Sarat, 433
Mishra, S. K., 937, 1609
Mishra, Sudhansu Kumar, 433
Misra, Rajul, 827, 1375
Misra, Yatharth Shankar, 749
Mohan, Akhilesh, 769
Mohan, Himanshu, 503
Mukherjee, Soumonos, 219

N
Nafees, Naira, 1423, 1433
Nain, Garima, 985
Nair, Latha R., 789
Nandi, Ayani, 1171
Naresh, B., 615
Nath, B., 1259
Nautiyal, Saurabh, 1509
Nikam, Valmik, 245
Nimmagadda, Sailaja, 147

P
Pahuja, G. L., 925
Pal, Shweta, 1231
Pandey, Alok, 1501
Pandey, Rajiv, 1553
Pandey, Rudra Narayan, 433
Pandit, Anala A., 255, 245, 393
Pant, Isha, 73
Panwar, Ravi, 513, 523, 597
Parmar, Monika, 625
Parsediya, D. K., 605
Parveen Sultana, H., 1587
Patel, Manish, 105
Patel, Raju, 1389
Patra, Swarnajyoti, 1017, 1219
Pawar, Shivaji D., 463
Pimpale, Bhakti, 255
Pinto, Verlyn, 1049
Prasad, Ravi Kant, 533
Prashanth, M. C., 973
Pruthi, Jyotika, 995
Puri, Vishal, 1423, 1433

R
Rafi Lone, Mohd., 1251
Rafique, Umair, 633
Raghav, Ashok K., 115
Rahi, O. P., 37
Rajesh, Arunacharam, 95
Rajkumar, R., 219
Rakhi, 925
Ramasubbareddy, Somula, 85, 195, 207, 415
Ranjan, Deepak, 1641
Ranjan, Priya, 327
Rathore, Rahul, 167

Rathour, Navjot, 1475
Ravikumar, M., 961, 973
Ray, Suvendra Kumar, 319
Reetu, 1117
Rout, Amitejash, 1587
Roy, Kshaunish, 1587

S
Sachdeva, Nitin, 371
Sahai, Archana, 1553
Saharia, Sarat, 61
Sahoo, Nigama Prasan, 795
Sahu, Arpit, 513, 523, 597
Sahu, Mayank, 1285
Sahu, O. P., 1151
Sahu, Rhythm, 1587
Sahu, Sonal, 533
Samantaray, S. D., 423
Sampathkumar, S., 973
Sarswat, Narender Kumar, 815, 849
Satapathy, Siddhartha Sankar, 319
Saxena, Amit, 827, 1375
Saxena, Anurag, 489, 721
Saxena, Preeti, 235
Sen, Piyali, 319
Shah, Arati Kumari, 1241
Shankar, Ajay, 1673
Sharma, Abhinav, 495, 559
Sharma, Anand, 697
Sharma, Bhupendra, 1107
Sharma, Dinesh, 565, 677
Sharma, Garima, 879
Sharma, Gaurav, 95
Sharma, Harshita, 1269
Sharma, Kamal Kr., 463
Sharma, Pooja, 311
Sharma, Purnima K., 565
Sharma, Raghavendra, 705
Sharma, Ragini, 779
Sharma, Santosh, 657
Sharma, Sumit, 915
Sharma, Taruna, 857
Sharma, Veena, 37
Sharma, V. K., 615
Shashank, 1211
Shimray, Benjamin A., 867, 1317
Shinghal, Kshitij, 827, 1375
Shivakumar, G., 961
Shivaprasad, B. J., 973
Shrawne, Seema, 383
Shrivastav, Laxmi, 657
Shrivastava, Laxmi, 879
Shukla, M. K., 1535
Shukla, Rahul, 551
Shukla, Shashank, 1049
Singh, Arun, 1461, 1487
Singh, Ashutosh Kumar, 489, 697
Singh, Bhupendra, 1547
Singh, H. P., 1627
Singh, Jasjit, 1547
Singh, Mandeep, 1535
Singh, Mukesh, 915
Singh, Naorem Ajesh, 1337
Singh, Omveer, 857
Singh, Praveen Kumar, 879
Singh, Rajesh, 1451, 1461, 1475, 1487
Singh, Rajni Ranjan, 137
Singh, Rishi Raj, 769
Singh, Saurabh, 689, 759
Singh, Sunil Kumar, 575, 589
Singh, Thounaojam Bebekananda, 1337
Singh, Vinod Kumar, 489, 503, 615, 721
Sridhar, Banothu, 729
Srinivas, Aditya Sai T., 207
Suchu, Tanveer Kaur, 513
Sudhakaran, Shreya, 327
Sudhakar, Suraj, 1657
Sumaiya Thaseen, I., 7
Surapaneni, Ravi Kishan, 147
Swain, Mahendra, 1451
Swain, S. K., 1609
Swarna Latha, P., 1565

T
Talluri, Salman Raju, 643, 665
Thakral, Shaveta, 1443
Thakre, Vandana Vikas, 705
Thakur, Ankur, 643, 665
Thasneen, Sumaiya, 451
Tiwari, Ashutosh, 47, 1365
Tiwari, Garima, 551
Tiwari, Ratnesh, 795, 811
Toksha, Gaurav, 393
Tomar, Deepak Singh, 137
Tomar, Geetam Singh, 1277
Tripathi, Babita, 841
Tripathi, Shubha, 795
Tripathi, Suman Lata, 1395
Tripathy, Malay Ranjan, 327
Tyagi, Shivam, 713

U
Umale, Niraj, 1197

Upadhyay, Anand, 1049, 1057
Upadhyay, Arvind Kumar, 301, 349

V
Vaishali, 1107
Vamsi, Konda Krishna, 1565
Varghese, Prabha Elizabeth, 789
Vashishtha, Anuradha, 1077
Verma, Chaman, 337
Verma, Pankaj, 1139
Verma, P. K., 665
Verma, Ramesh Kumar, 749
Verma, Shekhar, 1395
Verma, S. K., 73
Verma, Sudhanshu, 689, 759
Vijaya Krishna, A., 415
Vijay, Ravika, 545
Vijay Sai, T., 565
Viral, Rajkumar, 1627, 1641
Vohra, Anil, 1151

W
Wani, Tanveer Ahmad, 841
Waris, Abdul, 319

Y
Yadav, Akansha, 759
Yadav, Ashok, 503
Yadav, Ashwani Kumar, 1107
Yadav, Gaurav, 1641
Yadav, Rajesh, 1117
Yadav, Ravi, 523
Yadav, Rekha, 1117, 1127