

Thangaprakash Sengodan  
M. Murugappan  
Sanjay Misra *Editors*

# Advances in Electrical and Computer Technologies

Select Proceedings of ICAECT 2021

# Lecture Notes in Electrical Engineering

## Volume 881

### Series Editors

Leopoldo Angrisani, Department of Electrical and Information Technologies Engineering, University of Napoli Federico II, Naples, Italy

Marco Arteaga, Departament de Control y Robótica, Universidad Nacional Autónoma de México, Coyoacán, Mexico

Bijaya Ketan Panigrahi, Electrical Engineering, Indian Institute of Technology Delhi, New Delhi, Delhi, India  
Samarjit Chakraborty, Fakultät für Elektrotechnik und Informationstechnik, TU München, Munich, Germany

Jiming Chen, Zhejiang University, Hangzhou, Zhejiang, China

Shanben Chen, Materials Science and Engineering, Shanghai Jiao Tong University, Shanghai, China

Tan Kay Chen, Department of Electrical and Computer Engineering, National University of Singapore, Singapore, Singapore

Rüdiger Dillmann, Humanoids and Intelligent Systems Laboratory, Karlsruhe Institute for Technology, Karlsruhe, Germany

Haibin Duan, Beijing University of Aeronautics and Astronautics, Beijing, China

Gianluigi Ferrari, Università di Parma, Parma, Italy

Manuel Ferre, Centre for Automation and Robotics CAR (UPM-CSIC), Universidad Politécnica de Madrid, Madrid, Spain

Sandra Hirche, Department of Electrical Engineering and Information Science, Technische Universität München, Munich, Germany

Faryar Jabbari, Department of Mechanical and Aerospace Engineering, University of California, Irvine, CA, USA

Limin Jia, State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing, China

Janusz Kacprzyk, Systems Research Institute, Polish Academy of Sciences, Warsaw, Poland

Alaa Khamis, German University in Egypt El Tagamoa El Khames, New Cairo City, Egypt

Torsten Kroeger, Stanford University, Stanford, CA, USA

Yong Li, Hunan University, Changsha, Hunan, China

Qilian Liang, Department of Electrical Engineering, University of Texas at Arlington, Arlington, TX, USA

Ferran Martín, Departament d'Enginyeria Electrònica, Universitat Autònoma de Barcelona, Bellaterra, Barcelona, Spain

Tan Cher Ming, College of Engineering, Nanyang Technological University, Singapore, Singapore

Wolfgang Minker, Institute of Information Technology, University of Ulm, Ulm, Germany

Pradeep Misra, Department of Electrical Engineering, Wright State University, Dayton, OH, USA

Sebastian Möller, Quality and Usability Laboratory, TU Berlin, Berlin, Germany

Subhas Mukhopadhyay, School of Engineering & Advanced Technology, Massey University, Palmerston North, Manawatu-Wanganui, New Zealand

Cun-Zheng Ning, Electrical Engineering, Arizona State University, Tempe, AZ, USA

Toyoaki Nishida, Graduate School of Informatics, Kyoto University, Kyoto, Japan

Federica Pascucci, Dipartimento di Ingegneria, Università degli Studi "Roma Tre", Rome, Italy

Yong Qin, State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing, China

Gan Woon Seng, School of Electrical & Electronic Engineering, Nanyang Technological University, Singapore, Singapore

Joachim Speidel, Institut of Telecommunications, Universität Stuttgart, Stuttgart, Germany

Germano Veiga, Campus da FEUP, INESC Porto, Porto, Portugal

Haitao Wu, Academy of Opto-electronics, Chinese Academy of Sciences, Beijing, China

Walter Zamboni, DIEM - Università degli studi di Salerno, Fisciano, Salerno, Italy

Junjie James Zhang, Charlotte, NC, USA

The book series *Lecture Notes in Electrical Engineering* (LNEE) publishes the latest developments in Electrical Engineering - quickly, informally and in high quality. While original research reported in proceedings and monographs has traditionally formed the core of LNEE, we also encourage authors to submit books devoted to supporting student education and professional training in the various fields and applications areas of electrical engineering. The series cover classical and emerging topics concerning:

- Communication Engineering, Information Theory and Networks
- Electronics Engineering and Microelectronics
- Signal, Image and Speech Processing
- Wireless and Mobile Communication
- Circuits and Systems
- Energy Systems, Power Electronics and Electrical Machines
- Electro-optical Engineering
- Instrumentation Engineering
- Avionics Engineering
- Control Systems
- Internet-of-Things and Cybersecurity
- Biomedical Devices, MEMS and NEMS

For general information about this book series, comments or suggestions, please contact [leontina.dicecco@springer.com](mailto:leontina.dicecco@springer.com).

To submit a proposal or request further information, please contact the Publishing Editor in your country:

#### **China**

Jasmine Dou, Editor ([jasmine.dou@springer.com](mailto:jasmine.dou@springer.com))

#### **India, Japan, Rest of Asia**

Swati Meherishi, Editorial Director ([Swati.Meherishi@springer.com](mailto:Swati.Meherishi@springer.com))

#### **Southeast Asia, Australia, New Zealand**

Ramesh Nath Premnath, Editor ([ramesh.premnath@springernature.com](mailto:ramesh.premnath@springernature.com))

#### **USA, Canada:**

Michael Luby, Senior Editor ([michael.luby@springer.com](mailto:michael.luby@springer.com))

#### **All other Countries:**

Leontina Di Cecco, Senior Editor ([leontina.dicecco@springer.com](mailto:leontina.dicecco@springer.com))

**\*\* This series is indexed by EI Compendex and Scopus databases. \*\***

More information about this series at <https://link.springer.com/bookseries/7818>

Thangaprakash Sengodan · M. Murugappan ·  
Sanjay Misra  
Editors

# Advances in Electrical and Computer Technologies


Select Proceedings of ICAECT 2021

 Springer

*Editors*

Thangaprakash Sengodan  
Department of Electrical and Electronics  
Engineering  
SVS College of Engineering  
Coimbatore, India

M. Murugappan  
Department of Electronics  
and Communication Engineering  
Kuwait College of Science and Technology  
Safat, Kuwait

Sanjay Misra   
Department of Computer and Information  
Sciences  
Covenant University  
Ota, Nigeria

ISSN 1876-1100

ISSN 1876-1119 (electronic)

Lecture Notes in Electrical Engineering

ISBN 978-981-19-1110-1

ISBN 978-981-19-1111-8 (eBook)

<https://doi.org/10.1007/978-981-19-1111-8>

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2022, corrected publication 2022

This work is subject to copyright. All rights are solely and exclusively licensed 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

# Contents

<b>Towards Area-Delay Efficient Reverse Conversion of Higher-Radix Signed-Digit Number Systems</b> .....	1
Madhu Sudan Chakraborty, Ganti Sreelakshmi, Anirban Chakraborty, Sandip Kumar Sao, and Dulal Chandra Sahana	
<b>Multilayer Perceptron Mode and ANN to Assess the Economic Impact and Human Health Due to Alcoholism and Its Effect in Rural Areas</b> .....	13
Sendilvelan Subramanian, Sujatha Kesavan, Geetha Soman, Kanimozhi Natanam, Kanya Nataraj, Rajeswary Hari, Gomathi Kannayiram, and Jayalatsumi Umapathy	
<b>A Novel Approach to Predict Success of Online Games Using Random Forest Regressor for Time Series Data</b> .....	27
Rohit Renne Varghese, D. R. Aiswarya, Athulya Roy, Vighnesh Muraly, and Shini Renjith	
<b>On Strong Proper Connection Number of Some Interconnection Networks</b> .....	41
I. Annammal Arputhamary, D. Angel, and A. Shenbaga Priya	
<b>A Virtual Assistant for the Visually Impaired</b> .....	49
Alex Roy, Amal Saji, M. S. Gokul, and Subu Surendran	
<b>Automated Ultrasound Ovarian Tumour Segmentation and Classification Based on Deep Learning Techniques</b> .....	59
K. Srilatha, F. V. Jayasudha, M. Sumathi, and P. Chitra	
<b>Magnitude Comparison in Canonical Signed-Digit Number System</b> .....	71
Madhu Sudan Chakraborty	

<b>Region-Based Random Color Highlighting in Artistic Style Transfer Using CNN</b> .....	81
Katharotiya Krutarth and Manu Madhavan	
<b>A Novel Enhancement of Vigenere Cipher Using Message Digest</b> .....	91
Rajalaxmi Mishra and Jibendu Kumar Mantri	
<b>A Multi-criteria Decision-Making Approach to Analyze Python Code Smells</b> .....	103
Aakashshi Gupta, Deepanshu Sharma, and Kritika Phulli	
<b>The Importance of Validation Metrics in Chaotic Image Encryption</b> .....	119
M. Y. Mohamed Parvees and T. Vijayakumar	
<b>Application of Support Vector Machine and Convolutional Neural Network for Sentence-Level Sentiment Analysis of Companies Products Review</b> .....	133
Oluwatobi Noah Akande, Joyce Ayoola, Sanjay Misra, Ravin Ahuja, Akshat Agrawal, and Jonathan Oluranti	
<b>Application of XGBoost Algorithm for Sales Forecasting Using Walmart Dataset</b> .....	147
Yetunde Faith Akande, Joyce Idowu, Abhavya Misra, Sanjay Misra, Oluwatobi Noah Akande, and Ravin Ahuja	
<b>Comparison of Selected Algorithms on Breast Cancer Classification</b> .....	161
Olabiyisi Stephen Olatunde, Olalere Mofiyinfoluwa, Oluwatobi Noah Akande, Sanjay Misra, Ravin Ahuja, Akashat Agrawal, and Jonathan Oluranti	
<b>PCA-Based Feature Extraction for Classification of Heart Disease</b> .....	173
Roseline Oluwaseun Ogundokun, Sanjay Misra, Joseph Bamidele Awotunde, Akshat Agrawal, and Ravin Ahuja	
<b>Software Fault Prediction Using Machine Learning Algorithms</b> .....	185
M. S. Pavana, M. N. Pushpalatha, and A. Parkavi	
<b>Secret Key Generation: Single Edge-Triggered Flip-Flop PUF for IoT Environment</b> .....	199
S. Hemavathy, C. Renju Raju, Akshara Kairali, B. G. Hari Lavanya, and V. S. Kanchana Bhaaskaran	
<b>A Concise Survey on Solving Feature Selection Problems with Metaheuristic Algorithms</b> .....	207
Rama Krishna Eluri and Nagaraju Devarakonda	

**A SUMO Simulation Study on VANET-Based Adaptive Traffic Light Control System** ..... 225  
 Malose Mathiane, Chunling Tu, Pius Adewale Owola, and Mukatshung Claude Nawej

**FINNger: To Ease Math Learning for Children using Hand Gestures by Applying Artificial Intelligence** ..... 239  
 Pallavi Malavath, Nagaraju Devarakonda, and Zdzislaw Polkowski

**Control of Rice Blast Pathogens Using Back Propagation Fuzzy Neural Network** ..... 257  
 S. Anita Shanthi, G. Sathiyapriya, and L. Darwin Christdhas Henry

**Comparison of Machine Learning Algorithms for Bearing Failures Classification and Prediction** ..... 269  
 Yasser N. Aldeoes, Prasad Ghockle, and Shilpa Y. Sondkar

**A Parallelized Algorithm for Finding 3-D Mesh Illumination** ..... 283  
 Varun M. Khachane, V. Harshitha, Parthivi Khullar, S. V. Rishikesh, C. H. Mani Kumar, and K. V. Vineetha

**An Useful Survey on Supervised Machine Learning Algorithms: Comparisons and Classifications** ..... 293  
 B. Abhishek and Amit Kumar Tyagi

**Image Encryption and Decryption Using Chaotic Bimodal Quadratic Map** ..... 309  
 H. Soumya Babu, K. Gopakumar, and N. Vijayakumar

**A Novel March XR Algorithm, Design, and Test Architecture for Memories** ..... 321  
 I. G. Matri, Aishwaraya, N. Shreya, Saroja V. Siddamal, and Suneeta V. Budihal

**Analysis of Heterogeneous Queuing Model with Unreliable Server and Working Vacation** ..... 331  
 M. Seenivasan, H. Manikandan, and K. S. Subasri

**Performance Analysis of Single Server Low-priority Queue Based on Electronic Transmitter** ..... 347  
 R. Ramesh and M. Seenivasan

**Improved HELBP in Harsh Lightning Variations for Face Recognition** ..... 361  
 Shekhar Karanwal

**Determination of the TOE Factors Influencing the Adoption of Internet Banking Services on SMEs in Yemen: A Moderated Mediation Approach** ..... 371  
 Nabil Hussein Al-Fahim, Rawad Abdulghafor, and Sherzod Turaev



<b>Technological and Organizational Factors Influencing the Internet Banking Use Among SMES in Yemen: The Mediating Role of Attitude</b> .....	389
Nabil Hussein Al-Fahim and Rawad Abdulghafor	
<b>Markovian Queueing Model with Server Breakdown, Single Working Vacation, and Catastrophe</b> .....	409
M. Seenivasan, V. J. Chakravarthy, and R. Abinaya	
<b>Lessening Spectrum Sensing Data Falsification Attack by Weighted Fuzzy Clustering Means Using Simulation Annealing in Cognitive Radio Networks</b> .....	423
L. Thulasimani and A. Hyils Sharon Magdalene	
<b>Performance Analysis of Thinning Algorithms for Offline-Handwritten Devanagari Words</b> .....	437
Sukhjinder Singh and Naresh Kumar Garg	
<b>Application of Digital Image Correlation Technique to Study Strain Localization in Fe–Ni–W Alloy</b> .....	455
Kavita Tewari and Ramesh Kulkani	
<b>Design and Construction of Analogue Signal Processing System for Distributed Fibre Optic Sensors</b> .....	469
Kadambari Sharma, J. M. Nair, and P. P. Vaidya	
<b>Design of Versatile Reconfigurable ADC for Wide Range of Resolution and Conversion Time</b> .....	483
Jayamala Adsul, P. P. Vaidya, and J. M. Nair	
<b>Mini Interconnect IC Protocol and Camera Command Set Controller for Camera Communication in Mobile Phones</b> .....	505
Utsav Malviya, Gopal Kumar, Champalal Lalani, Sampad Barik, and P. S. Shrivastava	
<b>Automatic Estimation of Multiplicity in Partial Discharge Sources Using Machine Learning Techniques</b> .....	519
Lekshmi Kaimal and Ramesh Kulkarni	
<b>Comparison of Encryption Techniques to Encrypt Private Parts of an Image</b> .....	535
Nisha P. Shetty, Balachandra Muniyal, Rithish Reddy Kaithi, and Sarath Chandra Reddy Yemma	
<b>Speaker Identification Using Multiple Features and Models</b> .....	559
A. Revathi, G. Gayathri, and C. Jeyalakshmi	
<b>Performance Analysis on Flexible Modified Koch Fractal Patch Antenna for Wearable Healthcare Application</b> .....	575
K. A. Malar and R. S. Ganesh	

**Intelligent ICH Detection Using K-Nearest Neighbourhood, Support Vector Machine, and a PCA Enhanced Convolutional Neural Network** ..... 585  
 Shanu Nizarudeen and Ganesh R. Shunmugavel

**Feature Selection and Diagnosis Performance Evaluation of Breast Cancer** ..... 597  
 Anu Babu and S. Albert Jerome

**Grading and Classification of Retinal Images for Detecting Diabetic Retinopathy Using Convolutional Neural Network** ..... 607  
 Neetha Merin Thomas and S. Albert Jerome

**Ensured Configuration Security of FPGAs against CAD Attacks** ..... 615  
 Daliya A. John and M. Nirmala Devi

**Configuration Security of FPGA in IoT Using Logic Resource Protection** ..... 625  
 S. Meenakshi and M. Nirmala Devi

**A Novel Construction of Multiband Compact Microstrip Patch Antenna and Its Applications in WiMAX, HiPERLAN** ..... 635  
 Pritam Singha Roy and Moumita Guha

**Integrated Smart Alert System for Industrial Applications using Transceiver Module Analysis** ..... 647  
 Vaithiyanathan Dhandapani, Joel Jacob Thomas, and Y. Durga Sravanthi

**Analyzing Beamforming for Secure Transmission of MIMO-NOMA-Based CRN using Power Allocation Based on Second-Order Perturbation** ..... 665  
 L. Thulasimani and A. Hylis Sharon Magdalene

**A High Merit Factor Preamble Sequence for Better Synchronization in WiMAX** ..... 683  
 Suma Sekhar, Sakuntala S. Pillai, and S. Santhoshkumar

**Deep Learning Technique-Based Pulmonary Embolism (PE) Diagnosis** ..... 695  
 S. Vijayachitra, K. Prabhu, M. Abarana, A. Deepa, and L. Loga Priya

**Analysis of Daubechies 2 Wavelet in WPM System for Adhoc Network** ..... 703  
 M. B. Chakole and S. S. Dorle

**Adiabatic Multiplexer and Delay Flip-Flop** ..... 715  
 Y. Syamala and K. Srilakshmi

<b>A MobileNet-V2 COVID-19: Multi-class Classification of the COVID-19 by Using CT/CXR Images</b> .....	727
N. Mahendran and S. Kavitha	
<b>A Novel Method to Improve the Resolution of FLASH ADC for High-Resolution and High-Speed Applications</b> .....	739
Asma Parveen I. Siddavatam, J. M. Nair, and P. P. Vaidya	
<b>FPGA Implementation of Radix-2 Pipelined FFT Algorithm for High-throughput Applications</b> .....	753
Rajasekhar Turaka, S. Ravi Chand, Tavanam Venkata Rao, and V. Kumara Swamy	
<b>Design of High Speed Approximate Carry Select Adders Using RCPFA</b> .....	761
T. Phaneendra, A. Anitha, and Rajasekhar Turaka	
<b>Data Driven Approach to Achieve Coordinated Charging Among Electric Vehicles</b> .....	771
Venkata Nikitha Machineni, Korada Sri Vardhana, and Vaithiyathan Dhandapani	
<b>A 4-element Dual Band MIMO Antenna with L-Shaped Decoupling Stub for 5G Applications</b> .....	781
R. Seetharaman, B. Varsha Poorani, Kavya Santha Kumar, M. Tharun, S. S. Sreeja Mole, and K. Anandan	
<b>An Efficient Hybrid Genetic Algorithm and Whale Optimization Algorithm for Benefit Maximization of Optimal Sizing and Sizing Problem of DG in Distribution Systems</b> .....	791
K. Banumalar and B. V. Manikandan	
<b>The Tensor of Conductivity with the External Fields in Rectangular Quantum Wire for the Case of Electrons—Acoustic Phonon Scattering</b> .....	807
Hoang Van Ngoc	
<b>An Implementation of Hybrid CNN-LSTM Model for Human Activity Recognition</b> .....	813
Kolla Sai Krishna and Surekha Paneerselvam	
<b>Sliding Mode Control-Based Standalone Wind Energy System</b> .....	827
R. M. Meenakshi and K. Selvi	
<b>Instantaneous Power Estimation Algorithm in Hybrid Power Filter for Power Quality Improvement</b> .....	839
S. Sindhu and M. R. Sindhu	

**Multi-objective Optimal Power Flow Using Whale Optimization Algorithm Consists of Static VAR Compensator** ..... 849  
 G. Karthik Varma and Bathina Venkateswara Rao

**Standalone PMSG-Based Wind Energy Conversion System Under Unbalanced Load Conditions** ..... 859  
 P. N. Koteswara Rao, R. Mahalakshmi, and K. Sudarsana Reddy

**Application of Pythagorean Fuzzy Rough Distance Method in MCDM Problem for Buying Induction Motor** ..... 871  
 Revathi and Radhamani

**Real-Time Monitoring of Buses Utilizing Phasor Measurement Units for a 24-Bus System** ..... 881  
 G. Babu Naik

**State-Space Modelling for Two-Area Multiple-Source AGC Using PI, PID, PI-PD Controllers** ..... 891  
 P. Vidya Sagar and M. S. Krishna Rayalu

**Optimization of Power Generation Costs Through Soft Computing Techniques** ..... 901  
 M. V. Suganyadevi and A. R. Danila Shirly

**Implementation of Binary Particle Swarm Optimization for Image Thresholding using Memristor Crossbar Array** ..... 915  
 Priyanka B. Ganganaik, Omkar Mukul Gowaikar, V. Jeffry Louis, Rajesh K. Tripathy, Venkateswaran Rajagopalan, B. V. V. S. N. Prabhakar Rao, and Souvik Kundu

**Performance Analysis of Solar PV-fed BLDC Motor Under Partial Shading Condition Using Various PSO MPPT Algorithms** ..... 937  
 Pakki Pavan Kumar, V. Hemant Kumar, and R. N. Patel

**Design and Analysis of Fuzzy-Based Hybrid PV-Wind Power Quality Improvement for Local Nonlinear Loads Using MLMS** ..... 953  
 Pasala Gopi, M. Padma Lalitha, and S. Ayisha Jabeen Bhanu

**Small-Signal Analysis of Inverter to Maximise Power from Two PV Subarrays Under Different Environmental Conditions** ..... 969  
 S. Sneha and P. B. Savitha

**Some Properties of Silicene Nanoribbons When Doped with Two Lead Atoms in an Electric Field** ..... 987  
 Hoang Van Ngoc

**Optimal Solution of Economic Emission Load Dispatch Using a Hybrid Technique** ..... 995  
 Gutta Prem Kumar, G. V. Nagesh Kumar, and Vempalle Rafi

**Solution of Unit Commitment Using Genetic Algorithm with Population Refreshment** ..... 1007  
 Gundra Bhargavi, G. V. Nagesh Kumar, and Vempalle Rafi

**Modelling Series RLC Circuit with Discrete Fractional Operator** ..... 1019  
 George E. Chatzarakis, A. George Maria Selvam, R. Janagaraj, and D. Vignesh

**Design of Cascaded Multilevel Inverter-based STATCOM for Reactive Power Control with Different Novel PWM Algorithms** ..... 1033  
 Ch. Lokeshwar Reddy and G. Sree Lakshmi

**Outlier Detection Using Linear Regression in Wind and Solar Integrated Power Systems** ..... 1055  
 Priyanka Khirwadkar Shukla and R. Mahalakshmi

**Photovoltaic Systems Incorporated with Energy Storage System for Agricultural Implementation** ..... 1065  
 V. Hemant Kumar, Pakki Pavan Kumar, R. N. Patel, and Vivek Bargate

**Residential Feeder Energy Audit Analysis and Recommendation with Aid of Software** ..... 1081  
 K. Anitha, Shailesh, L. Ramesh, and Muruganath Gopal Raj

**An Investigation into the Applications of Real-Time Simulator in Experimental Validation of PMSM-Based Electric Drive System** ..... 1091  
 Anjaly Mohan, Meera Khalid, and A. C. Binojkumar

**A Genetic Algorithm Approach for the Optimal Allocation of On-Site Generation in Radial Distribution Networks** ..... 1109  
 Kasala Rajesh and J. Viswanatharao

**Enhanced Thermophysical Properties and Productive Yield of Pyramid Solar Still Combined with Shallow Solar Pond by Incorporating ZnO/Al<sub>2</sub>O<sub>3</sub> Nanocomposites** ..... 1121  
 B. Selvakumar, V. Shanmugapriya, K. Amudha, and P. Periasamy

**Correction to: Design of High Speed Approximate Carry Select Adders Using RCPEA** ..... C1  
 T. Phaneendra, A. Anitha, and Rajasekhar Turaka

# About the Editors

**Thangaprakash Sengodan** received his Bachelor of Engineering in Electrical Engineering from Bharathiar University, India, and Master of Engineering in Power Electronics and Drives from Anna University, Chennai, in 2002 and 2014, respectively. He received a Ph.D. degree in Electrical Engineering from Anna University, Chennai, India, in 2011. He is currently a visiting professor at SVS College of Engineering and the full-time head of the research and development wing of Diligentec Solutions, Coimbatore. His current research interests include power electronics circuits, renewable power conversion systems, and solid-state control of electric drives. He has authored over thirty papers in peer-reviewed international journals and conferences. Dr. Thangaprakash is a senior member of IEEE, IEEE-Power Electronics Society, IEEE-Communications Society, and a life member of the Indian Society for Technical Education (ISTE). Dr. Thangaprakash is currently Vice-Chair of Electronic Communication and Information Management, IEEE India Council (for the period 2021–2022), and also he is a member of the IEEE India Council EXCOM 2021–2022 (Executive Committee) and IEEE Madras Section EXCOM 2020–2021.

**M. Murugappan** received B.E. in Electrical and Electronics Engineering from Madras University, M.E. in Applied Electronics from Anna University, and a Ph.D. in Mechatronic Engineering from Universiti Malaysia Perlis, in 2002, 2006, and 2010, respectively. He is currently working as an associate professor in the Department of Electronics and Communication Engineering, Kuwait College of Science and Technology (KCST) (Private University), Kuwait. Before this, he worked as a senior lecturer in the School of Mechatronics Engineering, University Malaysia Perlis, Malaysia, from 2010 to 2016. He has been working in bio-signal processing applications for the past 12 years and, is cited as an expert in WHO IS WHO in the world. He has published over 120 research articles in peer-reviewed conference proceedings and journals. He has successfully guided 12 postgraduate students (eight Ph.D. and four M.Sc.). He is also serving as an editorial board member in several international peer-reviewed journals (SCI), a committee member in international conferences, and the chair of Educational Activities in IEEE Kuwait Section. His main research interest

is bio-signal/image processing applications (affective computing), neuroscience, machine learning, human-machine interaction, and artificial intelligence.

**Sanjay Misra** is a professor of Computer Engineering at Covenant University, Ota, Nigeria. He has 25 years of wide experience in academic administration and research in various universities in Asia, Europe, and Africa. He obtained his Ph.D. in Information and Know. Engineering (Software Engineering) from the University of Alcala, Spain, and M.Tech. (Software Engineering) from Motilal Nehru National Institute of Technology, India. He has authored/co-authored around 300 papers. He is the editor in chief of the book series on Advances in IT Personals and Project Management (IGI Global), the author of 1 book, and the editor (one of) in 31 Lecture Notes in Computer Science (Springer), six IEEE conference proceedings. He was the general chair of many international conferences. Presently, Dr. Sanjay is editor-in-chief of *International Journal of Physical Sciences* (SCOPUS Indexed), the founder-editor in chief of *Covenant Journal of ICT* and *International Journal of Computer Science and Software Technology*, and also serving as the editor, the associate editor, and the editorial board member of more than 20 journals (Including three SCIE) of international repute. His current researches cover the areas of software quality assurance, software process improvement, software project management, object-oriented technologies, XML, SOA, Web services, cognitive informatics, artificial intelligence, neural network, health Informatics, e-learning, cloud computing, and cyber security.

# Towards Area-Delay Efficient Reverse Conversion of Higher-Radix Signed-Digit Number Systems



Madhu Sudan Chakraborty, Ganti Sreelakshmi, Anirban Chakraborty, Sandip Kumar Sao, and Dulal Chandra Sahana

## 1 Introduction

Computer arithmetic (CA) or VLSI arithmetic is an interdisciplinary field of computer science and electrical engineering [1]. Arithmetic circuits (ACs) for addition/subtraction are the basic building block of every standard arithmetic unit (AU). Arithmetic addition and subtraction can be carried out at constant time using signed-digit number systems (SDNSs) [1]. The constant-time addition/subtraction may further ensure higher-speed in performing some more complex arithmetic operations (Aos) [1]. Apart from offering higher-speed, SDNSs support on-line arithmetic [2], fault tolerance [3], cryptography [4] and unified solutions for some apparently distinct CA-problems [5] too. Notable progresses have been achieved during the last two decades to address some seemingly problematic areas of SDNSs; particularly floating-point arithmetic [6], low-power computations [7, 8] and reducing the larger chip area commonly required in SDNS platform [9].

---

M. S. Chakraborty (✉)

CS Department, Indas Mahavidyalaya, Bankura Pin 722205, India  
e-mail: [madhusudhanchakraborty27@gmail.com](mailto:madhusudhanchakraborty27@gmail.com)

G. Sreelakshmi

ECE Department, GCET, Hyderabad Pin 501301, India

A. Chakraborty

ECE Department, Sastra University, Thanjavur Pin 613401, India

S. K. Sao

CS Department, J K College, Purulia Pin 723101, India

D. C. Sahana

CS Department, B C College, Bankura Pin 722101, India



SDNS with radix = 2, having digit-set (DS) as  $\{\bar{1}, 0, 1\}$ , is called binary SDNS (BSDNS) and all other classes of SDNSs with radix  $> 2$  are collectively called higher-radix SDNSs (HRSDNSs). Although BSDNS has been investigated more rigorously [2–9], some features of HRSDNSs have attracted the CA-community too [10–12].

Reverse conversion (RC) means expressing the signed-digit (SD) output of an AO in the standard/conventional form. RC appears as a troublesome area of SDNS [1]. In the existing literature, RC is projected as a necessary evil and RC algorithms (RCAs) for BSDNS have been extensively studied [13]. However, no significant work on RC of HRSDNS has been reported yet, excluding a typical RCA which was proposed for a particular SDNS in [14]. In this paper, RC of HRSDNS is investigated in details.

The rest of the paper will be organized with five sections. In The Background section, the arithmetic attributes of SDNSs, SD-encodings, RC and the philosophy of article [14] will be discussed. In The Proposed Algorithms: Schematic Specifications section, different alternative schemes for RC of HRSDNSs will be projected, on the basis of viewing the conversion control network (CCN) for HRSDNSs in terms of BSDNS. In The Proposed Algorithms: Realizing Arithmetic Expressions section, the CA expressions (CAEs) for various conversion control units (CCUs) will be developed. In the Results and Discussion section, simulation-based comparative performance study will be reported to determine the most suitable algorithm amongst different alternatives for RC of HRSDNSs. Finally, the proposed study will be ended in the Conclusion section, providing some future directions too.

## 2 The Background

At arithmetic level, the background of the RC problem for HRSDNSs rests with SDNSs, classification of SDNSs, encodings of SDs and obviously RC itself.

### 2.1 Signed-Digit Number Systems

The generalized SDNS of radix- $r$  ( $r \geq 2$ ) is defined as a positional number system (NS) that works on the DS  $\{\bar{\alpha}, \bar{\alpha} + 1, \dots, \bar{1}, 0, 1, \dots, \beta - 1, \beta\}$  where  $\alpha \geq 0$ ,  $\beta \geq 0$  and  $\alpha + \beta + 1 > r$  [15]. The primitive form of SDNSs is called ordinary SDNS (OSDNS). The OSDNS is defined on the DS  $\{\bar{\alpha}, \bar{\alpha} + 1, \dots, \bar{1}, 0, 1, \dots, \alpha - 1, \alpha\}$  where  $\frac{r}{2} < \alpha < r$  [16]. Actually the redundancy inherently acquired by the SDNSs using some unconventional DS, makes constant-time addition/subtraction admissible [1, 2].

**Table 1** BSD-encodings

Scheme	Encoding for BSD		
	$\bar{1}$	0	1
Positive-Negative Encoding (PNE) [1]	(0,1)	(0,0) or (1,1)	(1,0)
Two's Complement Encoding (TCE) [1]	(1,1)	(0,0)	(0,1)
Tripathy Encoding (TE) [17]	(1,0)	(1,1)	(0,0)

## 2.2 Encodings

As at the digital hardware level, all non-binary values mandatorily require interpretation as binary strings, for implementing AUs using a SDNS, binary encoding(s) of every digit of its DS is a pre-requisite. In the literature, some popular encodings for binary signed-digit (BSD)  $X_i$  as  $(X_{i,1}, X_{i,0})$  are shown in Table 1.

## 2.3 Reverse Conversion: Computer Arithmetic Interpretation

Let  $F = F_{n-1}F_{n-2} \dots F_0$  be the given radix- $r$  SD number (SDN) and  $Z = Z_nZ_{n-1} \dots Z_0$  be the equivalent radix- $r$  radix-complement (RCM) representation where  $Z_n = \{0, 1\}$ . Here in order to guarantee non-occurrence of any overflow-related errors,  $Z$  is formed with one more digit, compared to  $F$ . Then numerical agreement of  $F$  and  $Z$  yields:

$$\sum_{i=0}^{n-1} F_i.r^i = -Z_n.r^n + \sum_{i=0}^{n-1} Z_i.r^i \tag{1}$$

The primitive RC-rules [1] for a radix- $r$  Generic SDNS (GRSDNS) are as follows:

$$Z_i = \Delta(F_i - S_i), S_{i+1} = \Omega(F_i - S_i) \tag{2}$$

where  $\Delta(F_i - S_i) = (F_i - S_i)$  if  $F_i - S_i \geq 0$  and  $(r + F_i - S_i)$  otherwise. Also  $\Omega(F_i - S_i) = 0$  if  $F_i - S_i \geq 0$  and 1 otherwise. Here  $S_i$  is the conversion control variable (CCV) and  $S_0 = 0$ . CCVs may acquire values exploiting a network, called CCN. Obviously CCN provides conversion control information (CCI) for driving the RC.

As implied by (2), a simple linear version of the RCA for the  $n$ -digit BSDNS may be presented as Algorithm 1 as follows [1]:

**Algorithm 1:**

- a. Input at level 0:  $F = F_{n-1} \dots F_0$  where

$$\forall i \in [0, n - 1] F_i = (F_{i,1}, F_{i,0})$$

- b. Initialize:  $S_0 = 0$
- c. For  $i = 0$  to  $n-1$  repeat as follows:

$$Z_i = \overline{F_{i,0}} \cdot \overline{S_i} + F_{i,0} \cdot S_i, S_{i+1} = F_{i,1} \cdot \overline{F_{i,0}} + F_{i,1} \cdot S_i$$

- d. Finally set:  $Z_n = S_n$

However, better resolution on RC problem is still required as all prospects of SDNSs may be dampened otherwise [13, 18, 19].

## 2.4 On the Available Scheme for Higher-Radix

The existing on-the-fly RCA for radix- $r$  OSDNS [14] is presented as Algorithm 2 as:

### Algorithm 2:

- a. Input at level 0:  $F$
- b. Generate a linear CCN directly based on  $F$  having CCVs as  $C_i$
- c. Compute  $Z$  in that order

For Algorithm 2 step (b) and step (c) are driven by (3) and (4) respectively.

$$\forall i \geq 0 \ C_{i+1} = \frac{C_i - r \cdot F_1}{r^2 + C_i} \times (r - 1), C_0 = 1 \quad (3)$$

$$Z_i = (F_i + C_i) - \left( r \times \frac{F_i + C_i}{r} \right) \quad (4)$$

In [14] synthesis of associated ACs for Algorithm 2 was not discussed at all and still the problem is that in Algorithm 2 the arithmetic equations (Aes) involve a series of complex AOs containing multiplication/division, truncation/rounding-off, requiring different algorithms at their own capacities [1]. Obviously, Algorithm 2 even cannot be simulated instantaneously and primarily it rests with mathematical interests only.

## 3 The Proposed Algorithms: Schematic Specifications

For a given SDN,  $X = X_{n-1}X_{n-2} \cdots X_0, \forall i \in [0, n-1]$  define:  $P(X_i) = \bar{1}, 0, 1$ , if  $X_{i-1}X_{i-2} \cdots X_0$  holds negative, zero or positive sign, respectively. Then the methods proposed in this paper are based on utilizing Lemma 1, which is as follows:

**Lemma 1:** For a given higher-radix SDN,  $F = F_{n-1}F_{n-2} \cdots F_0$ , there exists a binary SDN,  $G = G_{n-1}G_{n-2} \cdots G_0$ , such that  $\forall i \in [0, n-1] \ P(F_i) = P(G_i)$  where  $G_i = \bar{1}, 0, 1$ , if  $F_i < 0, = 0$  or  $> 0$ , respectively.

**Proof:** If possible assume that Lemma 1 does not hold. It means, for some given  $F$  and  $G$ , there exists a least significant position,  $j$ , such that  $P(F_j) \neq P(G_j)$  whereas  $P(F_{j-1}) = P(G_{j-1})$ . Then obviously the sign of  $F_{j-1}$  and  $G_{j-1}$  must mismatch. However, it contradicts the definition of  $G_j$  as a derivative of  $F_j$  without altering its sign. Hence, the Lemma 1 holds true.

Therefore it may be possible to design the CCN for  $F$  in terms of  $G$ . Let  $G$  be called the binary sign equivalent (BSE) of  $F$ .

### 3.1 Proposed Scheme 1

#### Algorithm 3:

- a. Input at level 0:  $F$
- b. Determine at level 1:  $G$  as BSE of  $F$
- c. Generate a linear CCN with reference to  $G$ , yielding the CCI
- d. Compute  $Z$  in that order in terms of 1-digit blocks only, using the CCI in (2)

Algorithm 3 is actually a simplified, generic variant of Algorithm 2. For a 16-digit input, the CCN for Algorithm 3 is shown in Fig. 1a. In Fig. 1 all shaded circular-nodes represent terminal nodes and levels are prefixed with  $L$ . Moreover in Fig. 1a, the shaded rectangular node represents the initialization with  $C_0 = 0$ .

### 3.2 Proposed Scheme 2

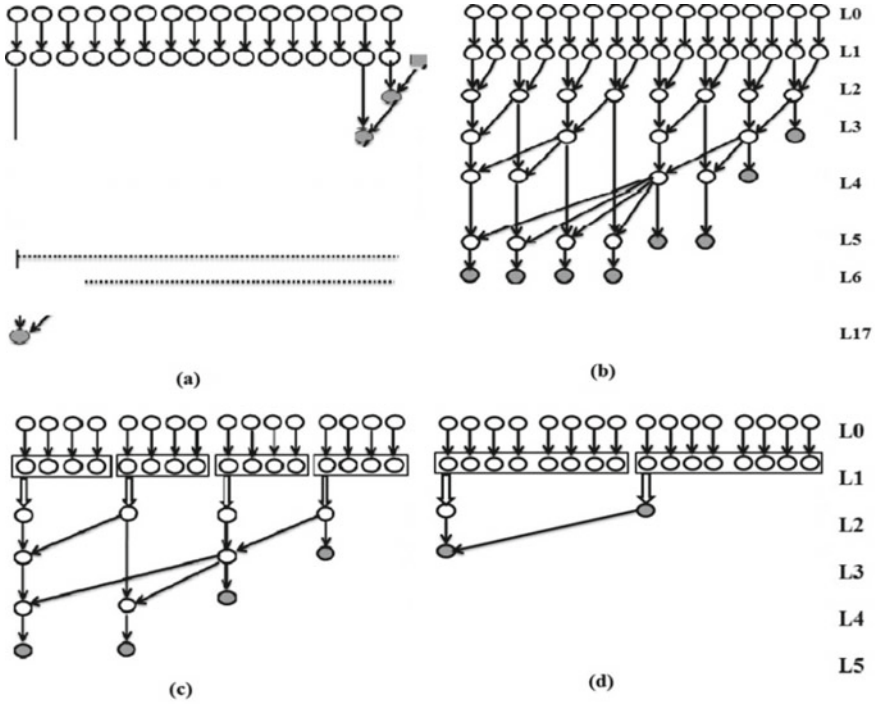
#### Algorithm 4:

- a. Input at level 0:  $F$
- b. Determine at level 1:  $G$  as BSE of  $F$
- c. Generate the CCN in terms of reverse tree structures (RTSs) [20] with uniform block factor (BF) = 2, yielding the CCI.
- d. Compute  $Z$  in terms of 2-digit blocks in parallel, using the CCI in (2), whilst performing the intra-block computations digit serially.

### 3.3 Proposed Scheme 3

#### Algorithm 5:

- a. Input at level 0:  $F$
- b. Determine at level 1:  $G$  as BSE of  $F$



**Fig. 1** CCNs for the proposed schemes for a 16-digit input

- c. Generate the CCN in terms of RTSs with  $BF = 4$  at level 2 and  $BF = 2$  for all higher levels, yielding the CCI.
- d. Compute  $Z$  in terms of 4-digit blocks in parallel, using the CCI in (2), whilst performing the intra-block computations digit serially.

### 3.4 Proposed Scheme 4

#### Algorithm 6:

- a. Input at level 0:  $F$
- b. Determine at level 1:  $G$  as BSE of  $F$
- c. Generate the CCN in terms of RTSs with  $BF = 8$  at level 2 and  $BF = 2$  for all higher levels, yielding the CCI.
- d. Compute  $Z$  in terms of 8-digit blocks in parallel, using the CCI in (2), whilst performing the intra-block computations digit serially.

For a 16-digit input, the CCNs for the proposed schemes 2, 3, 4 are shown in Fig. 1b, Fig. 1c and Fig. 1d, respectively. It is notable that all of Algorithms 4, Algorithm 5 and Algorithm 6 are the non-linear variants of the Algorithm 3.

## 4 The Proposed Algorithms: Realizing Arithmetic Expressions

Amongst various classes of HRSDNSs, the radix-4 SDNS having DS  $\{\{\bar{2}, \bar{1}, 0, 1, 2\}$ , known as the minimally redundant radix-4 SDNS (MRR4SDNS), is fairly popular in CA community [2] and it is the subject matter of case study presented in this paper. For encoding higher-radix SDs the extended-TCE is considered, where the DS of MRR4SDNS may be encoded as shown in Table 2.

On the other hand, for encoding all BSD information TE [17] is considered. The developing of CAEs for 16-digit reverse converters (RCRs) will be studied next:

For all proposed schemes, at step (a): For  $i = 0$  to 15: Input:  $F_{i,2}, F_{i,1}, F_{i,0}$

For all proposed schemes, at step (b) the transformation is done as per Table 3, yielding (5). From Table 3 onwards for any table the entry of hyphen (-) at any input cell is supposed to indicate formation of same output for its both 0 and 1 inputs.

$$G_{i,0} = \overline{F_{i,1} \cdot F_{i,0}}, G_{i,1} = F_{i,2} + G_{i,0} \quad (5)$$

For step (c) of the proposed schemes, six types of CCUs may be involved as a whole as shown in Fig. 2.

For type-1 CCU, involved in the proposed scheme 1, as shown in Fig. 2a, the computations are at par with Table 4, yielding (6).

$$S_{i+1,0} = G_{i,1} \cdot \overline{G_{i,0}} + G_{i,1} \cdot S_{i,0} \quad \text{Where } S_{0,0} = 0 \quad (6)$$

For type-2 CCU, involved in the proposed scheme 2, as shown in Fig. 2b, the computations are done at par with Table 5, yielding (7).

$$S_{r,1} = G_{p,1} \cdot (\overline{G_{p,0}} + G_{q,1}), S_{r,0} = G_{p,0} \cdot G_{q,0} \quad (7)$$

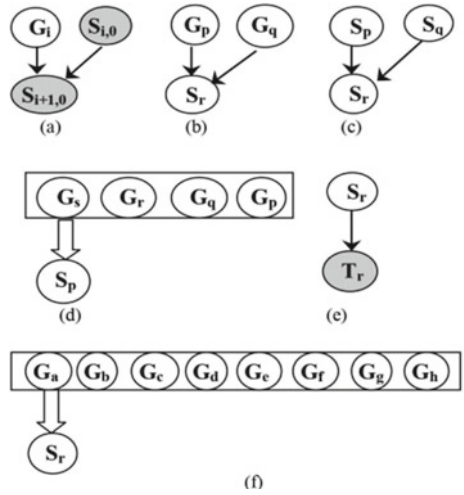
**Table 2** Higher-radix encodings

SD	$\bar{2}$	$\bar{1}$	0	1	2
Encoding ( $F_{i,2}, F_{i,1}, F_{i,0}$ )	(1, 1, 0)	(1, 1, 1)	(0, 0, 0)	(0, 0, 1)	(0, 1, 0)

**Table 3** BSE generation

Input ( $F_i$ )			Output ( $G_i$ )	
$F_{i,2}$	$F_{i,1}$	$F_{i,0}$	$G_{i,1}$	$G_{i,0}$
1	1	-	1	0
0	0	0	1	1
0	0	1	0	0
0	1	0	0	0

**Fig. 2** Different Types of CCUs used in step (c) of the Proposed Schemes



**Table 4** Computations by type-1 CCU

Input 1 ( $G_i$ )		Input 2 ( $S_{i,0}$ )	Output ( $S_{i+1,0}$ )
$G_{i,1}$	$G_{i,0}$		
1	0	–	1
1	1	0	0
1	1	1	1
0	0	–	0

**Table 5** Computations by type-2 CCU

Input 1 ( $G_p$ )		Input 2 ( $G_q$ )		Output ( $S_r$ )	
$G_{p,1}$	$G_{p,0}$	$G_{q,1}$	$G_{q,0}$	$S_{r,1}$	$S_{r,0}$
1	0	1	–	1	0
1	0	0	0	1	0
1	1	1	0	1	0
1	1	1	1	1	1
1	1	0	0	0	0
0	0	1	–	0	0
0	0	0	0	0	0

For type-3 CCU, involved in the proposed scheme 2, scheme 3 and scheme 4, as shown in Fig. 2c, computations similar to Table 5, yields (8)

$$S_{r,1} = S_{p,1} \cdot (\overline{S_{p,0}} + S_{q,1}), \quad S_{r,0} = S_{p,0} \cdot S_{q,0} \tag{8}$$

For type-4 CCU, involved in the proposed scheme 2, proposed scheme 3 and proposed scheme 4, as shown in Fig. 2e, it is observed that  $T_r = 1$  if  $S_r$  is negative and  $T_r = 0$ , otherwise. So type-4 CCU works at par with (9).

$$T_r = S_{r,1} \cdot \overline{S_{r,0}} \quad (9)$$

For type-5 CCU, involved in the proposed scheme 3, as shown in Fig. 2d, the computations can be viewed as an extension of type-2 CCUs and realized as in (10).

$$\begin{aligned} S_{r,1} &= G_{q,1} \cdot (\overline{G_{q,0}} + G_{p,1}), S_{r,0} = G_{q,0} \cdot G_{p,0} \\ S_{s,1} &= G_{r,1} \cdot (\overline{G_{r,0}} + S_{r,1}), S_{s,0} = G_{r,0} \cdot S_{r,0} \\ S_{t,1} &= G_{s,1} \cdot (\overline{G_{s,0}} + S_{s,1}), S_{t,0} = G_{s,0} \cdot S_{s,0} \end{aligned} \quad (10)$$

For type-6 CCU, involved in the proposed scheme 4, as shown in Fig. 2f, the computations may be viewed as an extension of type-5 CCU for 8 cells subsequently.

For step (d) of the proposed scheme 1, the transformations may be presented as in Table 6 and can be realized as in (11) where  $S_i = S_{i,0}$ .

$$Z_{i,1} = F_{i,1} \cdot F_{i,0} + F_{i,1} \cdot \overline{S_i} + \overline{F_{i,1}} \cdot \overline{F_{i,0}} \cdot S_i, Z_{i,0} = \overline{F_{i,0}} \cdot S_i + F_{i,0} \cdot \overline{S_i} \quad (11)$$

For step (d) of the proposed scheme 2, scheme 3 and scheme 4, subsequent instances of (11) may be applied in 2, 4 and 8 numbers, respectively where  $S_i = T_i$ .

**Table 6** Final output generation

Inputs			Output		
$F_i$			$S_i$	$Z_{i,1}$	$Z_{i,0}$
$F_{i,2}$	$F_{i,1}$	$F_{i,0}$			
–	1	0	0	1	0
1	1	1	0	1	1
0	0	0	0	0	0
0	0	1	0	0	1
–	1	0	1	0	1
1	1	1	1	1	0
0	0	0	1	1	1
0	0	1	1	0	0



**Table 7** Comparison of Alternative Schemes on FPGA Platform

Segment A				Segment B
Cases (Alternatives)	Delay (ns)	Area (Attributes in Number)		NADP
		Slice LUTs	Bonded-IOBs	
Scheme 1	6.50	29	60	1.4
Scheme 2	5.74	61	90	2.59
Scheme 3	5.79	63	92	2.69
Scheme 4	7.78	61	94	3.50
Reference (Algorithm 1)	5.21	26	49	1

## 5 Results and Discussion

On successful verification of Verilog codes, once written using text editors in the design entry stage of Xilinx 14.2 simulator for the 16-digit MRR4SDNS version of all alternative schemes as well as Algorithm 1, the codes are implemented on FPGA platform for the target device xc6vlx75tl-1Lff484 and the immediate outcomes are reported as segment A of Table 7. As bonded-IOBs consume minor area compared to slice-LUTs, in this study chip area is considered in terms of the latter attribute only.

For an unbiased comparative study, the slice-LUTs column and delay column of Table 7 are normalized [19] in form of multiplication by scaling factors  $(26)^{-1}$  and  $(5.21)^{-1}$ , respectively, and subsequently normalized area-delay products (NADPs) are determined. The NADP-column, which is presented as segment B of Table 7, projects the proposed scheme 1 as the best candidate for RCR design for MRR4SDNS, by outperforming its most potential contender, the scheme 2, with more than 45% comparative reduction in NADP. Table 7 also reveals that the RC of the 16-digit MRR4SDNS, which is arithmetically equivalent to the 32-digit BSDNS, attracts merely 40% increase in NADP overhead, compared to the RC of 16-digit BSDNS.

## 6 Conclusion

On the basis of exploiting the CCN for RC of HRS DNSs in terms of BSDNS, some alternative designs have been proposed in this paper. The alternative designs have been investigated for the 16-digit MRR4SDNS. In the environment of Xilinx 14.2 simulator, the possibly best radix-4 converter has demonstrated more than 45% reduction in NADP over its nearest contender. Another interesting feature of this report lies in the observation that RC of the 16-digit MRR4SDNS, which is arithmetically equivalent to the 32-digit BSDNS, attracts merely 40% increase in NADP

overhead, compared to the RC of 16-digit BSDNS. In other words, the performance loss caused by high overheads of RC ([1, 13]) may be, at least, partially compensated by employing HRSDNSs.

However, the report presented in this paper lacks in considering power/energy factor and radices higher than 4. These issues will be explored by the authors in future.

## References

1. Koren I (2001) Computer arithmetic algorithms 2. CRC Press, London
2. Ercegovac MD, Lang T (2004) Digital arithmetic, 1st edn. Morgan Kaufmann Publishers (An Imprint of Elsevier), San Francisco
3. Armand A, Timarchi S (2019) Efficient error detection and correction method for 1-out-of-3 binary signed digit adders. *Int J Electro* 106:1427–1440
4. Shah YA, Javeed K, Azmat S, Wang X (2019) Redundant-signed-digit-based high speed elliptic curve cryptographic processor. *J Circuits Syst Comput* 28:120–126
5. Chakraborty MS, Sao SK, Mondal AC (2018) Equivalence of reverse conversion of binary signed-digit number system and two's-complement to canonical signed-digit recording. In: IEEE international conference on recent advances in information technology. Dhanbad pp 662–666
6. Kaivani A, Ko S (2016) Floating point butterfly architecture based on binary signed digit representation. *IEEE Trans Very Large Scale Integr VLSI Syst* 24:1208–1211
7. Crookes D, Jiang M (2007) Using signed digit arithmetic for low power multiplication. *Electron Lett* 43:613–614
8. Smitha KG, Fahmy AH, Vinod AP (2006) Redundant adders consume less energy. In: IEEE Asia Pacific conference on circuits and systems. Singapore pp 422–425
9. Juang T-B, Wei C-C, Chang C-H (2009) Area-saving technique for low-error redundant binary fixed-width multiplier implementation. In: IEEE international symposium on integrated circuits. Singapore pp 550–553
10. Jaberipur G, Ghodsi M (2003) High radix signed digit number systems: representation paradigms. *Sci Iran* 10:383–391
11. Sudan SS, Singhal M (2017) A review on basics of QSD number addition & subtraction. In: IEEE international conference on computing and communication technologies for smart nation. Gurgaon pp 13–17
12. Chakraborty MS (2020) Designing an on-line magnitude comparator for higher-radix. *Int J Inf Tech Electr Eng* 9:92–98
13. Chakraborty MS (2016) Reverse conversion schemes for signed-digit number systems: a survey. *J Inst Eng (India): B* 97:589–593
14. Charoensiri V, Surarerks A (2006) On-the-fly conversion from signed-digit number system into complement representation. In: IEEE international symposium on communications and information technologies. Bangkok pp 1056–1061
15. Parhami B (1990) Generalized signed-digit number systems: a unifying framework for redundant number representations. *IEEE Trans Comput* 39:89–98
16. Avizienis A (1961) Signed-digit number representations for fast parallel arithmetic. *IRE Trans Electron Comput* EC-10 389–400
17. Tripathy SS, Barik RK, Pradhan M (2017) An improved conversion circuit for redundant binary to conventional binary representation. In: International conference on computational intelligence, communications and business analytics. Kolkata pp 363–371
18. He Y, Chang C-H (2008) A power-delay efficient hybrid carry-lookahead/carry-select based redundant binary to two's-complement converter. *IEEE Trans Circuits Syst I Regul Pap* 55:336–346

19. Sahoo SK, Gupta A, Asati AR, Shekhar C (2010) A novel redundant binary number to natural binary number converter. *J Signal Process Syst* 59:297–307
20. Srikanthan T, Lam SK, Suman M (2004) Area-time efficient sign-detection technique for binary signed-digit number system. *IEEE Trans Comput* 53:69–72

# Multilayer Perceptron Mode and ANN to Assess the Economic Impact and Human Health Due to Alcoholism and Its Effect in Rural Areas



Sendilvelan Subramanian , Sujatha Kesavan, Geetha Soman, Kanimozhi Natanam, Kanya Nataraj, Rajeswary Hari, Gomathi Kannayiram, and Jayalatsumi Umopathy

## 1 Introduction

The hardware developed for systems in IoT includes devices for an isolated console, controllers, servers, routers, bridging devices and transducers. These devices accomplish main responsibilities and enables activation of the system with required specification, safety, communication and recognition of precise objectives and activities. The sensors form the first and foremost part of the hardware in IoT. Such devices entail energy efficient units, power management units, RF components, and sensing elements. RF components accomplish communications by using signal processing techniques, WiFi, ZigBee, Bluetooth, radio transceiver and duplexer. Embedded

---

S. Subramanian (✉) · S. Kesavan · G. Soman · K. Nataraj · R. Hari · G. Kannayiram · J. Umopathy  
Dr.M.G.R Educational and Research Institute, Chennai 600095, India  
e-mail: [Sendilvelan.mech@drmgrdu.ac.in](mailto:Sendilvelan.mech@drmgrdu.ac.in)

S. Kesavan  
e-mail: [ksujatha.eee@drmgrdu.ac.in](mailto:ksujatha.eee@drmgrdu.ac.in)

G. Soman  
e-mail: [geethasoman@drmgrdu.ac.in](mailto:geethasoman@drmgrdu.ac.in)

K. Nataraj  
e-mail: [kanya.v@drmgrdu.ac.in](mailto:kanya.v@drmgrdu.ac.in)

G. Kannayiram  
e-mail: [gomathi.ibt@drmgrdu.ac.in](mailto:gomathi.ibt@drmgrdu.ac.in)

J. Umopathy  
e-mail: [jayalatsumi.ece@drmgrdu.ac.in](mailto:jayalatsumi.ece@drmgrdu.ac.in)

K. Natanam  
A.V.C. College of Engineering, Mayiladuthurai 609305, India  
e-mail: [kanimozhimca@avccengg.net](mailto:kanimozhimca@avccengg.net)

systems, middleware, partner computers are networked together which forms an integral part of IoT software. These discrete and chief applications are accountable for data gathering, device integration, and real-time data analytics, application-oriented process variables measurement systems networked using the IoT.

The data collection includes orientation of the sensor measurements, data filtering for noise removal, data safety and accumulation of data. It deploys some well-defined set of rules called protocols to facilitate the connection and communication of the sensors in real time so that they can interact with the computer machines thereby providing path for machine to machine (M2M) interaction. Thereafter it gathers all the data from several sensors which are heterogeneous and allocates it depending upon the reference values as mentioned in the specification for various devices. The scheme finally transfers all the data gathered related to alcoholism to the principal server. Dedicated software is used for integrating the heterogeneous device so as to form the skeletal structure of the IoT system. It safeguards the stability and ensures a strong cooperation between the devices that are networked together. This is the main feature of IoT. They manage the various applications, protocols, and limitations of each device to allow communication. The data about the alcohol addicts is the input obtained from various devices and transforms it to a pattern which enables human analysis [1, 2]. They investigate the information related to alcoholism based on various parameters so as to achieve automation and provide necessary information to de-addiction centres.

### ***1.1 Medical Application and Process Extension***

This application facilitates the integration of some specific special devices and special types of sensors. They also allow integration and access of few instruments related to the field of engineering. This scheme will increase the productivity and accuracy of data collection. IoT systems in health care has become the state of the art. They magnify the stretch out and cover a wide range of specialists. This scheme increases the accuracy, precision and the dimension of medical data related to the alcohol addicts in large scale from diversified real cases worldwide. It uses well-ordered environs, helps to even monitor the leftover examinations in the field of alcohol de-addiction. IoT-based de-addiction paves way for a wealthy real-time information with respect to the data from de-addiction centres followed by analysis and testing. IoT can distribute related data of good quality with typical analytics through various sensors that are integrated to perform research in the area of alcoholism [3, 4]. IoT-based integration provides a lot of important information relating to alcoholism and de-addiction. This technology supports de-addiction in health care domain by offering a highly consistent and real-world data thereby providing solution to unforeseen issues that may arise in future. It also permits researchers and physicians to avoid risks by gathering data without appropriate testing on human race. It then discloses the pattern and lost data in de-addiction-related health care applications. Conceivably the utmost development IoT brings to de-addiction in the field of health care is in

the real practice of medicine since it permits the better usage of the knowledge by health care professionals to recover a patient who is addicted to alcohol. They utilize good quality sensors which provides the relevant data in a more precise manner [5]. The decision-making is not affected by the noisy or irrelevant data that is to affect the quality of the entire scheme.

One of the challenges in de-addiction is the circulation of correct and recent information to the addicts. De-addiction care also tackles the complex nature involved in the guidance. IoT devices stretch a straight, 24 × 7 admission to the alcohol addicts in a non-invasive method as compared with the existing methods. They take de-addiction facilities to home, office, or social space. They motivate the addicts to take care of their health and get relieved from the influence of alcohol [6]. This kind of motivation reduces road accidents resulting from rash driving under the influence of alcohol, deaths due to liver cirrhosis and healthier precautionary care [7].

The innovative computerization and analytics of IoT permits more dominant emergency support services to de-addiction centres, which classically suffer from disconnection due to limited resources and fundamental facility. It offers a method to examine an emergency in a more comprehensive manner from long distance. It also gives more access to the alcohol addicts. IoT gives critical facts for providing indispensable care for alcohol addicts. It also increases the care given to the alcohol addicts when there is an emergency situation taking place [8, 9]. This decreases the related losses, and recovers emergency health care to de-addiction centres.

## ***1.2 Alcohol Addiction and the Implementation of IoT***

Human most commonly used for intoxication chemical substances is alcohol since time immemorial. Originally, alcohol is not a liquid, it is a powder [10]. In Arabian term a powder named 'al-kuhul', meaning 'the kohl' for the eyes that means 'finely divided spirit' [11, 12]. Alcohol has its own distinct properties with many different chemical compounds. Methanol and Isopropanol are the varieties of alcohol used in chemical laboratories and industries [13, 14]. For cleaning the household equipment this Isopropanol is generally used in industries [15, 16]. It is known as commonly 'rubbing alcohol' [17, 18]. Cleaning solvents and paint removers also contain alcohol [19, 20]. The photocopier developer and anti-freeze solutions also have considerable amount of alcohol. In the manufacturing of alcohol, the ethanol is extracted and the formaldehyde obtained as the end product is poisonous [21, 22]. Consuming even small amount of Methanol, which is another type of alcohol, leads to blindness and death [23, 24]. Ethanol is a type of alcohol [25]. Alcohol is unless otherwise referred to as ethanol or ethyl alcohol.

In most of the states, the important source of revenue generation is alcohol and also it has been an important international trade [26, 27]. It is the responsibility of researchers to provide necessary information to policy provider programme planners to make necessary decisions about the allocation of resources. A comparison of health care problems could be studied using this indicator of magnitude. There exists an

economics in alcohol addiction, reported nationally and internationally. It is a multi-dimensional technique to study the economics behind alcohol consumption [28]. Without assessing the social, psychological problems, it is not an easy task to predict the alcohol addicts in a particular area. If the prediction rate is not up to the mark, then it affects the economy of the nation starting from the individual and his family [29]. This study could be an eye-opener for people, make them aware of the consequences of alcoholism, and facilitate the statutory bodies and health organizations to orient their focus on the eradication of alcoholism [30].

## 2 Data Analysis

Background variables were calculated using descriptive statistics together with the data relating to structured population and their social and economic attributes, alcohol consumption information, its usage and dependence of alcoholism. Dependence of alcoholism was calculated with 95% Confidence Interval and also drinking methods, problematic drinkers and social problem creators. Outcome of alcoholism creates dispersion of data viability and economics. The association with marital status, education and occupation, along with living standard, with that of alcohol addiction is reported. Statistical significance investigation with Chi-square was used. For comparison purpose, t-test was conducted to identify the expenses related to alcohol addicts and occasional drinkers. This study was done at a primary health centre Poonamallee near Chennai, Tamil Nadu.

Questionnaires were used to study alcohol prevalence. Alcohol use disorder identification test was done by using questionnaire. Michigan alcohol screening test was used to assess psychosocial problems. Males of 19 years or more than 19 years living in that area were considered for study, for a minimum period of one year. Voter lists were used to draw sampling frame randomly. The population was identified using simple random methods. The size of the sample selected was 545. All gave consent form signed and were incorporated in the process study. Demographic characteristics are reported in Table 1.

In the study reported in Table 1, 23.9% of men have been found in the age group of 34–45 years, 19.1% men have been found in 25–34 age group, 18.7% men have been found in 45–54 age group, 15.8% men have been found in 18–24 age group, 10.8% men have been found in 55–64 age group and 11.7% men have been found in 64 and above age group. As per the study, majority of men were Hindus with 80.7%, then the Christian men with 15.6% and Muslims with 3.7%. In the statistics, 27.2% of men were unmarried, 67.9% of men were married, 4.6% of them were widower and 0.3% of them were separated/Divorcee.

**Table 1** Statistics

Statistics	No. of. men (Total = 545)	Percentage of men (%)
<i>Age (in yrs)</i>		
18–24	86	15.8
25–34	104	19.1
35–44	130	23.9
45–54	102	18.7
55–64	59	10.8
65 and above	64	11.7
<i>Religion of men</i>		
Hindu	440	80.7
Muslim	20	3.7
Christian	85	15.6
Others	0	0
<i>Marital status of men</i>		
Unmarried	148	27.2
Married	370	67.9
Widower	25	4.6
Separated/Divorcee	2	0.3
<i>Family Type</i>		
Nuclear	305	56
Extended nuclear	155	28.4
Joint	85	15.6

## ***2.1 Normal Drinkers' Prevalence with Respect to Occupation and Family***

Depending on occupation, it was found that 8% of normal drinkers were unemployed and 7% of them were unskilled. 42% of the normal drinkers seem to be semi-skilled, 53% of them seem to be skilled, 8% of them seem to be semi-professional and 2% of them are professional as shown in Table 2. The statistical significance has a difference which is about  $p < 0.0001$ . Based on family type, it was found that majority of normal drinkers belong to nuclear family, 13% of them belong to joint family and 26% of them belong to extended nuclear family as shown in Table 3. Here also, the difference is statistically significant ( $p < 0.0001$ ).



**Table 2** Normal drinkers' prevalence based on occupation

Occupation	Normal drinkers	Percentage (%)
Unemployed	8	6.7
Unskilled	7	5.8
Semi-skilled	42	35
Skilled	53	44.1
Semi-professional	8	6.7
Professional	2	1.7
Total	120	100.0

**Table 3** Normal drinkers' prevalence based on family type

Type of family	Normal drinkers	Percentage (%)
Nuclear	81	67.5
Joint	13	10.8
Extended nuclear	26	21.7
Total	120	100.0

## 2.2 Current Drinkers' with Respect to Age Group, Education

As mentioned earlier, current drinkers are those who have consumed any type of alcohol in the past 12 months. Normal drinkers and alcoholics are also included in this category. Current drinkers' prevalence is found to be highest in 25–34 age group with 24.7%, followed by 22.1% in 33–45 age group as shown in Table 4. Statistical significance is about  $p < 0.0001$ . Based on education, majority of the current drinkers were found to be higher secondary/diploma holders with 22.4%. Illiterates in current drinkers were 12.5% of them, 15.6% of them have finished primary schooling, 19.5% of them have finished middle schooling and 21.2% of them finished high schooling as shown in Table 5. Statistical significance is nearly ( $p < 0.0001$ ).

**Table 4** Current drinkers' prevalence based on age-group

Age in years	Current drinkers	Percentage (%)
18–24	75	23
25–34	81	24.7
35–44	72	22.1
45–54	50	15.3
55–64	28	8.8
65 and above	20	6.1
Total	326	100.0

**Table 5** Prevalence of current drinkers based on Education

Educational qualification	Current drinkers	Percentage (%)
Uneducated	41	12.5
Schooling(Primary)	51	15.6
Schooling(Middle)	64	19.5
Schooling (High)	69	21.2
Diploma/Higher secondary	73	22.4
Graduation	24	7.6
Post graduate	4	1.2
Total	326	100.0

### 3 Multilayer Perceptron (MLP) for Prevalence of Alcoholism

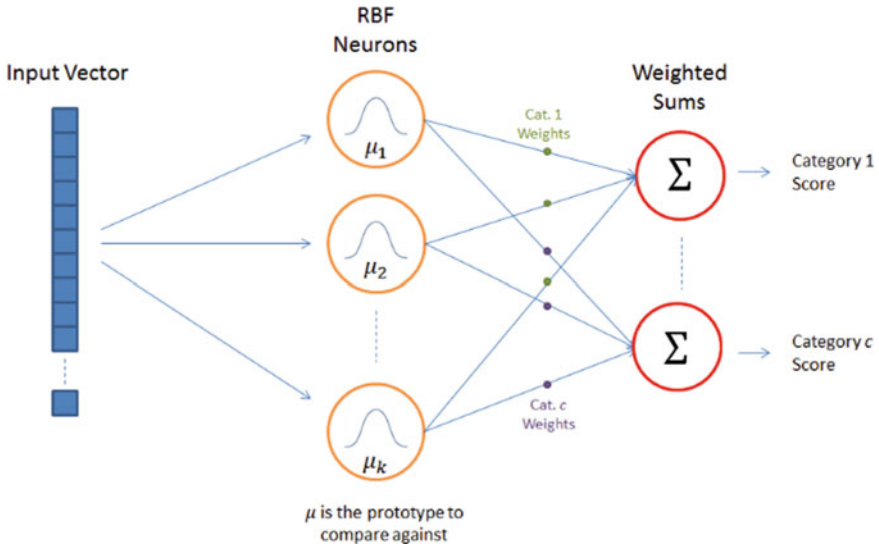
Investigation reveals that intelligent techniques like Multilayer perceptron (MLP) trained with RBF will be capable to predict the problems associated with various types of alcoholism [31, 32]. The data set is gathered by conducting survey at various places like rehabilitation centres, hospitals and de-addiction centres. This data is normalized before it is fed as inputs to the MLP. Normalization reduces computational complexity, hence this procedure is carried out.

The next stage is about the choice of the network architecture for MLP and choice of the training algorithm [33]. Since, these intelligent techniques are dependent on optimization of the objective function; they come under the class of optimization algorithms.

On analysis, it is understood that the MLP is a two-layered structure as shown in Fig. 1. Though, it is named as a two-layered model, physically it has three layers [34]. The input layer is the first layer, hidden layer is the second one and the output layer is the third layer. Input layer simply acts like a buffer, thereby transferring the input values. The inner product of these inputs are calculated and given as inputs to the hidden layer nodes. These nodes in the hidden layer possess, Gaussian activation function, which will process the values of the inner product by substituting them over the Gaussian activation function [35]. This function will facilitate the convergence of the MLP. Thirdly, the output layer uses sigmoidal activation function which will facilitate the faster convergence so that the actual output matches the target output [36].

#### 3.1 Steps to Implement Multilayer Perceptron

1. Random centres are chosen from the training set.
2. By applying the normalization method, computation of the spread for the RBF function is done.



**Fig. 1** Multilayer Perceptron (MLP) for Prevalence of Alcoholism

3. Infer the weights by pseudo-inverse method.

### 3.2 Current Drinkers' Health Seeking Pattern

The pattern of health seeking together with the health care visit reason were found out for alcoholics (Current drinkers). As per current drinkers' analysis, health care service sought 105 persons with alcohol-related health problems. In that analysis 59.0% took treatment in government hospitals but 41.0% went to private hospitals or nursing homes for their drinking problems as shown in Table 6. Among the current drinkers who took health care, 57.1% of them went as outpatient whereas 42.9% of them got admitted for their health problems as shown in Table 7.

Based on the evaluation, 105 drinkers took health care services mainly due to alcohol. Among the 105 persons, 35.2% were due to cirrhosis and liver disease, in the road due to accidents were 23 and 20% injury following violence, main injury concern is physical violence as shown in Table 8.

**Table 6** Health-caring places

Places	Users	Percentage
Hospital(Govt.)	62	59
Hospital(Pvt.)	43	41
Sum	105	100

**Table 7** Type of services for health

Type of services for health	Users ever $N = 105$	(%)
Outpatient	60	57.1
Inpatient	45	42.9
Sum	105	100

**Table 8** Health care seeking reasons

Health care seeking reasons	No. of participants $N = 105$	Percentage (%)
Cirrhosis/Liver disease	37	35.2
Accidents (Road)	24	23
Physical assault with others	21	20
Headache and gastritis	18	17.1
Symptoms withdrawal symptoms	5	4.8

### 3.3 Health Seeking Pattern

In this present study, majority of the person who drink sough health care for their drinking problems in the government hospitals and primary care centres (59.1%). Among those who sought health care (105), 42.7% got admitted in hospitals for their treatment, thus portraying the magnitude of the problem we are dealing with. In this present study, 17.1% of the alcoholics suffer from liver disease or cirrhosis, which is 35.2% of those who sought health care. WHO had mentioned 8.2% of alcoholics having cirrhosis and Lhachimi 9.7% had different findings [37]. This alarming increase in the liver diseases indicates the shift in the pattern of drinking towards harmful and dependence drinking. With increasing prevalence and changing pattern of drinking, more and more alcoholics will suffer from the wrath of liver disease and this problem will become a major health concern in the coming years. The health-related expenses on an average for a drinker was 3273 INR. The relatively lesser health-related expenses to drinkers are due to the fact that majority of the drinkers sought their health care from the government health delivery systems, which provide them services for free or with minimal charge. In this study, the current users of alcohol are found to be spending 13.3% from their yearly income for expenses related to alcohol. However, while categorizing the current drinkers depending on their present living standard, it shows drastic variation on the annual income proportion spent on alcohol and alcohol-related expenses. Current drinkers' stays like high standard of living spent 6.2% from their yearly income for alcohol-related expenditures whereas people having low living standard seems to spend a whopping of 36.1% of their annual income. This shows that alcohol not only has influences but

in diverse nature on people’s economic status but also the economic status of the people in turn surely affects alcohol use in many ways as described above.

### 4 Prediction of Alcohol Addiction by Feed Forward Neural Network Trained with Back Propagation Algorithm

By using the RBF and BPA, a classification is done to identify the addiction cases as high, medium and low as shown in Figs. 2 and 3.

Fig. 2 Prediction of alcohol addiction by FFNN trained with BPA

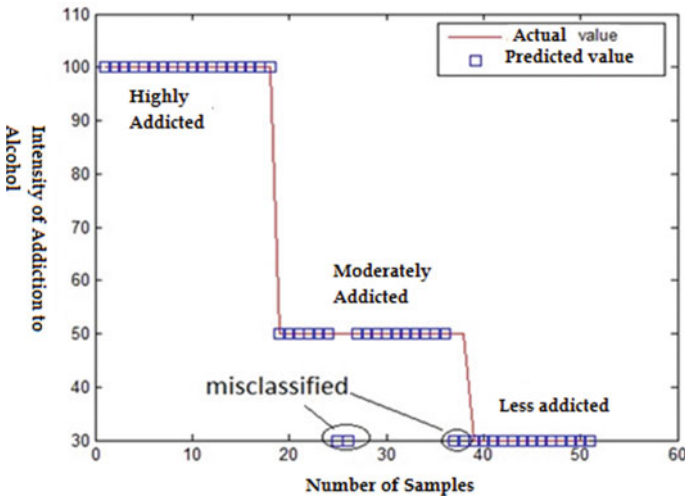
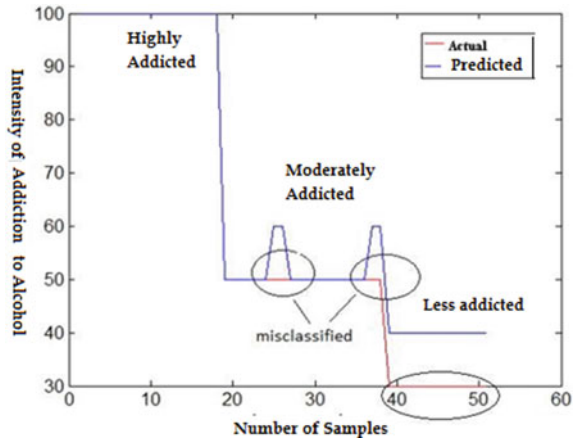


Fig. 3 Prediction of alcohol addiction by FFNN trained with RBF

In connection to the database obtained, a total of 51 data sets were collected pertaining to three categories of alcohol addiction. The first category denotes samples which correspond to less addiction to alcohol, the second category indicates moderate addiction to alcohol and thirdly, high addiction to alcohol. The inputs to FFNN include age, gender, height, weight, and blood group, level of alcohol content in blood and Blood pressure levels. Totally, 7 inputs are used to train the FFNN.

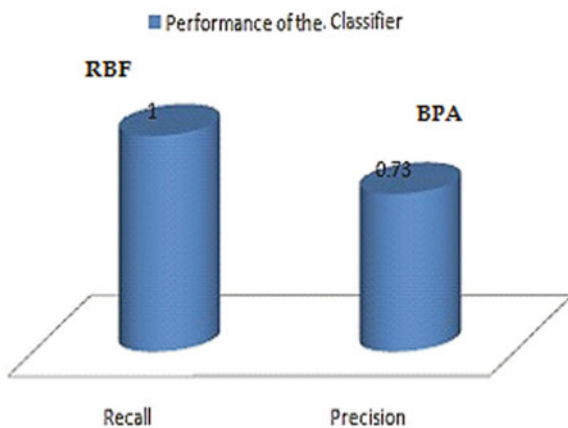
All these 7 inputs denote the normalized values. There are 3 outputs for this application which include less addiction, moderate addiction and high addiction to alcohol. The no. of nodes in layer (hidden) is finalized using this formula, No. of hidden layer nodes = (No. of nodes in input layer/2) rounded off to nearest possible integer. On substitution, it is found that

$$\begin{aligned}
 \text{No. of hidden layer node} &= (\text{No. of nodes in input layer}/2) \\
 &= (7/2) \\
 &= 3.5 \approx 4
 \end{aligned}$$

During the training process, it is also observed that only for 4 numbers of nodes in the hidden layer, the network attained convergence, with the lowest value of mean squared error (MSE) of 0.0001. Further during the training process, optimal values of the learning factor, rate of momentum and bias values are chosen.

In order to differentiate the representations with respect to alcohol addiction and their associated parameters using 7 attributes, Fig. 4 signifies the implementation of the BPA and RBF classifier. The efficiency metric used for this application is recall and precision, with a score closer to 1. Apart from using the conventional methods for prediction of alcohol addiction, an intelligent technique for predicting the level of alcohol addiction was developed. The generalized probabilistic methods with the capacity to analyse the data analysis with less computational complexity were found to be used for small-scale analysis in recent study. The proposed intelligent

**Fig. 4** Performance of the classifier



algorithms were especially dedicated for detecting the alcohol addicts and offer an appropriate remedy so that they get rid of their addiction.

## 5 Conclusion

The investigation of alcohol addiction from the data gathered from de-addiction centre is of significant importance in the field of data mining. The key goal of the work is the identification, interpretation and understanding of the level of alcohol addiction. The proposed method facilitates to properly monitor and regulate the alcohol addicts to de-addiction centres. This kind of data mining approach plays a vital role in reforming the lives of the alcohol addicts and saves their lives and future which is a huge challenge. Therefore, our aim is to focus on tracking and managing the alcohol addicts by using intelligent data mining algorithms with related artificial intelligence methods for analysis.

**Acknowledgements** The authors wish to thank the management of Dr. M.G.R. Educational and Research Institute, Chennai and A.V.C. College of Engineering, Mayiladuthurai, for their continuous encouragement and support rendered then and there to do the research inside the campus.

## References

1. Fathima F et al (2015) Alcohol consumption among adult males in a village in Bangalore Urban District : prevalence, harmful use and dependence. *Nat J Res Com Med*
2. Filov I et al (2014) Relationship between child maltreatment and alcohol abuse -findings from adverse childhood experience study in Republic of Macedonia. *Open Access Maced J Med Sci*. <https://doi.org/10.3889/oamjms.2014.067>
3. Dasgupta A et al (2013) Alcohol consumption by workers in automobile repair shops of a slum of Kolkata: an assessment with AUDIT instrument. *Nepal J Epidemiol*. <https://doi.org/10.3126/nje.v3i3.9188>
4. Dutta R et al (2014) 'A population based study on alcoholism among adult males in a rural area Tamil Nadu, India. *J Clin Diagn Res*. <https://doi.org/10.7860/JCDR/2014/6308.4411>
5. Zhao Q et al (2013) Validation study of 131I-RRL: assessment of biodistribution, SPECT imaging and radiation dosimetry in mice. *Mol Med Rep*. <https://doi.org/10.3892/mmr.2013.1338>
6. World Health Organization (2001) World health report : mental disorders affect one in four people. World Health Organization, Geneva
7. Thị Tuy ết Vân P (2018) Education as a breaker of poverty: a critical perspective. *Papers of Social Pedagogy*. <https://doi.org/10.5604/01.3001.0010.8049>
8. Steinebach P et al (2017) Factores de riesgo asociados al consumo problemático de alcohol en la adolescencia: El rol preventivo de mindfulness. *Psychology, Society, & Education*
9. Tan A et al (2014) Flaming alcoholic drinks: flirting with danger. *J Burn Care Res*. <https://doi.org/10.1097/BCR.0b013e3182a366de>
10. Apostoli P et al (2006) Elemental speciation in human health risk assessment. *Environmental Health Criteria*


11. Jackson KM, Chung T (2011) 'Alcohol use', in encyclopedia of adolescence. <https://doi.org/10.1016/B978-0-12-373951-3.00098-3>
12. Brányik T et al (2012) A review of methods of low alcohol and alcohol-free beer production. *J Food Eng.* <https://doi.org/10.1016/j.jfoodeng.2011.09.020>
13. Matar S, Mirbach MJ, Tayim HA (1989) Catalysis in petrochemical processes. *Catal Petrochemical Proc.* <https://doi.org/10.1007/978-94-009-1177-2>
14. Arifin S, Chien IL (2007) Combined preconcentrator/recovery column design for isopropyl alcohol dehydration process. *Ind Eng Chem Res.* <https://doi.org/10.1021/ie061446c>
15. Malekpour A, Mostajeran B, Koozhmareh GA (2017) Pervaporation dehydration of binary and ternary mixtures of acetone, isopropanol and water using polyvinyl alcohol/zeolite membranes. *Chem Eng Process.* <https://doi.org/10.1016/j.ccep.2017.04.019>
16. Setzler BP et al (2012) Further development of the isopropanol-acetone chemical heat engine. In: World Renewable Energy Forum, WREF 2012, Including World Renewable Energy Congress XII and Colorado Renewable Energy Society (CRES) annual conference
17. Jabbar U et al (2010) Effectiveness of alcohol-based hand rubs for removal of clostridium difficile Spores from hands. *Infect Control Hosp Epidemiol.* <https://doi.org/10.1086/652772>
18. Pires D et al (2017) Hand hygiene with alcohol-based hand rub: how long is long enough? *Infect Control Hosp Epidemiol.* <https://doi.org/10.1017/ice.2017.25>
19. Huang J et al (2017) Green preparation of a cellulose nanocrystals/polyvinyl alcohol composite superhydrophobic coating. *RSC Adv.* <https://doi.org/10.1039/c6ra27663f>
20. Buck S, Buck SL (1993) Three case studies in the treatment of painted furniture. In: Papers presented at the Wooden Artifacts Group: specialty session, June 4, 1993, AIC annual meeting, Denver, Colorado NV—13 bwills, 25 refs
21. Santasalo-Aarnio A et al (2011) Comparison of methanol, ethanol and iso-propanol oxidation on Pt and Pd electrodes in alkaline media studied by HPLC. *Electrochem Commun.* <https://doi.org/10.1016/j.elecom.2011.02.022>
22. Yip L et al (2020) Serious adverse health events, including death, associated with ingesting alcohol-based hand sanitizers containing methanol—Arizona and New Mexico, May–June 2020. *MMWR Morb Mortal Wkly Rep.* <https://doi.org/10.15585/mmwr.mm6932e1>
23. Moral AR et al (2015) Neuromuscular functions on experimental acute methanol intoxication. *Turk Anesteziyoloji ve Reanimasyon Dernegi Dergisi.* <https://doi.org/10.5152/TJAR.2015.13471>
24. Nsiewe NM et al (2018) Unusual deaths in two health districts in the east region-Cameroon: a case control study, 2016. *Am J Trop Med Hyg*
25. Hernández-Tobías A et al (2011) Natural alcohol exposure: is ethanol the main substrate for alcohol dehydrogenases in animals? *Chemico-Biol Interact.* <https://doi.org/10.1016/j.cbi.2011.02.008>
26. Voas RB, Johnson MB, Miller BA (2013) Alcohol and drug use among young adults driving to a drinking location. *Drug Alcohol Depend.* <https://doi.org/10.1016/j.drugalcdep.2013.01.014>
27. Hasin DS, Grant BF (2015) The National epidemiologic survey on alcohol and related conditions (NESARC) waves 1 and 2: review and summary of findings. *Soc Psychiatry Psychiatr Epidemiol.* <https://doi.org/10.1007/s00127-015-1088-0>
28. Moodie R et al (2013) Profits and pandemics: prevention of harmful effects of tobacco, alcohol, and ultra-processed food and drink industries. *Lancet.* [https://doi.org/10.1016/S0140-6736\(12\)62089-3](https://doi.org/10.1016/S0140-6736(12)62089-3)
29. Srinath R, Sendilvelan S (2017) A study on the prevalence of alcoholism among males in rural areas and its impact. *Indian J Public Health Res Develop.* <https://doi.org/10.5958/0976-5506.2017.00190.5>
30. Srinath R, Sendilvelan S (2017) Behavioral problems perceived by the alcoholic and his family—a study among males in rural areas. *Int J Psychosoc Rehabil* 21(2)
31. Tavana M et al (2016) A hybrid intelligent fuzzy predictive model with simulation for supplier evaluation and selection. *Expert Syst Appl.* <https://doi.org/10.1016/j.eswa.2016.05.027>
32. Hamdani M et al (2020) Prediction the inside variables of even-span glass greenhouse with special structure by artificial neural network (MLP-RBF) models. *J Agric Mach*



33. Mohd Amiruddin AAA et al (2020) Neural network applications in fault diagnosis and detection: an overview of implementations in engineering-related systems. *Neural Comput Appl.* <https://doi.org/10.1007/s00521-018-3911-5>
34. Alsmadi S, Omar KB, Noah SA (2009) Back propagation algorithm : the best algorithm among the multilayer perceptron algorithm. *Int J Comput Sci Netw Secur*
35. Chu Y, Fei J, Hou S (2020) Adaptive global sliding-mode control for dynamic systems using double hidden layer recurrent neural network structure. *IEEE Trans Neural Netw Learn Syst.* <https://doi.org/10.1109/TNNLS.2019.2919676>
36. Qian Y et al (2014) On the training aspects of deep neural network (DNN) for parametric TTS synthesis. In: *ICASSP, IEEE international conference on acoustics, speech and signal processing—proceedings.* <https://doi.org/10.1109/ICASSP.2014.6854318>
37. Lhachimi SK et al (2012) Health impacts of increasing alcohol prices in the European Union: a dynamic projection. *Prev Med.* <https://doi.org/10.1016/j.ypmed.2012.06.006>

# A Novel Approach to Predict Success of Online Games Using Random Forest Regressor for Time Series Data



Rohit Renne Varghese, D. R. Aiswarya, Athulya Roy, Vighnesh Muraly, and Shini Renjith 

## 1 Introduction

Online gaming is a highly profitable market, which is forecasted to reach \$180.1 billion by the end of 2021 [1]. There are different types of online games in the market, from which the PC games market saw an increase in digital sales and a high number of releases after valve launched steam [2]. This store saw a huge growth after 2012 with the help of a program called greenlight [3], which allowed the developers to easily release their games on this platform without the need of a publisher, and the users can download or buy the game from there. If the success of the game to be deployed can be predicted prior to the release, then it would be helpful for the creators to restructure the development to attract a more audience or even cast aside their endeavor to develop a game if it would not succeed. The previous attempts on the success prediction of video games assumed that the game was already released, and predictions were made based on that knowledge. In addition, they dealt with pre-2010 console games when there was no incentive to study the modern game market mostly dominated by FPS games. For this work, we considered an online FPS game called Counter-Strike: Global Offensive (CS: GO) [4] which currently has over 24 million monthly active users worldwide.

This work aims to propose a new system that is capable of determining the success of a game upon its release based on the game data collected from steam. The data are collected over the course of 2 months or more, and it includes various attributes like date, time, Twitch viewers, viewers ratio, etc. Once the data are collected, a random

---

R. R. Varghese (✉) · D. R. Aiswarya · A. Roy · V. Muraly · S. Renjith  
Department of Computer Science and Engineering, Mar Baselios College of Engineering and Technology, Thiruvananthapuram, Kerala 695015, India  
e-mail: [rohitrenne.varghese2021@vitstudent.ac.in](mailto:rohitrenne.varghese2021@vitstudent.ac.in)

S. Renjith  
e-mail: [shini.renjith@mbcet.ac.in](mailto:shini.renjith@mbcet.ac.in)

forest regressor model is used to predict the percentage increase in players which tells the developer whether the game would become a success or not in the coming months.

## ***1.1 Problem Definition***

As stated above, the online gaming domain has been growing rapidly. As this domain grows, the need for better success prediction systems grows too. The previous attempts at this gave unsatisfactory results. Moreover, they dealt with simple video games. So this work aims to design a novel system for predicting whether an online game, when released, would become successful, by utilizing time series data provided by the developers.

## **2 Background**

### ***2.1 Churn Prediction***

Everyone has free access to online games due to the easy availability of the Internet. The online gaming companies make profit with the help of in- game purchases, purchases, advertisements played in between, etc. Thus, as the number of players increases the profit, they make will also substantially increase. In addition to attracting new players, companies give very high focus on retaining their current players. They do this with the help of various marketing plans and pay-offs also known as the “churners.” The churn prediction involves various features like playtime, session time, purchase frequency, etc.

Playtime is accepted as a universal churn prediction feature, as it is more reliable than other features like the login data and time, etc., which may vary. Since short-term players do not yield much profit for the companies, they are more focused on getting the attributes of the long-term players. With the long-term players’ attributes, the empirical distributions and entropies can be calculated [5]. If the entropy of a long-term player is high, then it means that he/she is having an even distribution of attributes, i.e., having a more stable playtime. If it is not the case, it means that the player is having a less or uneven distribution of the, i.e., irregular/casual playtime. The attributes which were selected to constitute the dataset for the churn prediction include the following:

1. raw playtime data
2. total time spent
3. combined feature of last day of login and number of time slots played
4. daily and hourly standard deviation of playtime distribution.

For churn prediction, multiple classifiers like the following can be used.

1. Logistic regression (LR) [6]
2. Support vector machine (SVM) [7]
3. Random forests (RFs) [8]
4. Long short-term memory networks (LSTMs) [9]

## ***2.2 Predicting Success of Video Games***

Ever since valve launched the steam store, PC game sales have increased significantly, following the release of a program called greenlight, which allows developers to easily release games onto the steam store. Due to project greenlight, a massive number of games were added to the steam library, over 10,000 games between 2012 and 2016 [10], which is exponentially growing. Due to the massive number of games in the steam store, developers struggle to reach the limelight. Thus, to have a concept where the success of a game can be predicted, pre-release would be a vital tool in the game development industry. In the process to predict the success of the game, the first step is to analyze and identify the factors that affect the success of video game, generate a PC game database, and predict the success based on price, genre, game features, etc. [11]. The prime factors that influence a game's success include the number of clicks in search engine results, reviews given by players, and video/streaming services. The data collection process involved analyzing different sources and ordering the data. Steam spy is used for receiving the count of owners of the game at each point of time, steam chart for the concurrent player count, and steam database for information about the game [12]. As these alone are not sufficient for the learning processing, a database is created, with pre-release game data including genre, price, published, developer, languages, etc.

The price of the game is found to be a significant feature for successful games, where higher-priced games tend to be more successful, as they are usually developed from large studios, following bigger budgets. Similarly, games with better graphics tend to be more successful, as they are more visually appealing to the audience. But genre does not define whether a game will be successful or not, but user tags do. Games with tags such as open-world and third-person shooter tend to succeed. Added to this, multi-player, steam achievements, etc., also impose a significant effect on the success of the game. The data are analyzed with the help of many algorithms, such as binary classifiers and regressors with different settings. Algorithms such as baseline, linear, RPART, random forest, Gaussian process, and SVM are used for carrying out regression analysis whereas RPART, GLM, random forest, SVM, and Naive Bayes are used for binary classification modeling and analysis. When these algorithms are applied to smaller subsets of the dataset, it yielded better and more accurate results. A particular subset of 33% of the test data could yield the best result with a correlation of 0.82. The SVM model for regression is trained with a dataset including a minimum of 2 games recorded previously. The output given to the developer is an estimate of the average number of concurrent predicted players after releasing the games for

the first two months. Thus, a correlation is developed with features that were known before the release and concurrent average player count post-release for 2 months.

### ***2.3 Predicting Retention in Sandbox Games***

Sifa et al. [13] proposed a model for predicting the player retention for major commercial (AAA) games. They collected and processed spatio-temporal data about the gameplay times of players. It also collected their game actions. They aimed to learn over 14 days of activity and predict if the players will keep playing after the first 7 days. This work used mainly ensemble methods for evaluation and achieved a precision of 81% and a recall of 75% in the best case.

### ***2.4 Predicting Video Game Sales in the European Market***

This work focused on game and console sales in Europe from March 12, 2005 to December 31, 2011. The data were collected from about 2,450 games. The dataset had 9 attributes and attempted to predict “sales.” Simple regression models were fitted to predict weekly sales based on the first 2–6 weeks of sales. A prediction method for total sales was manually crafted and tested on all the data [14].

### ***2.5 Predicting Video Game Sales Using an Analysis of Internet Message Board Discussions***

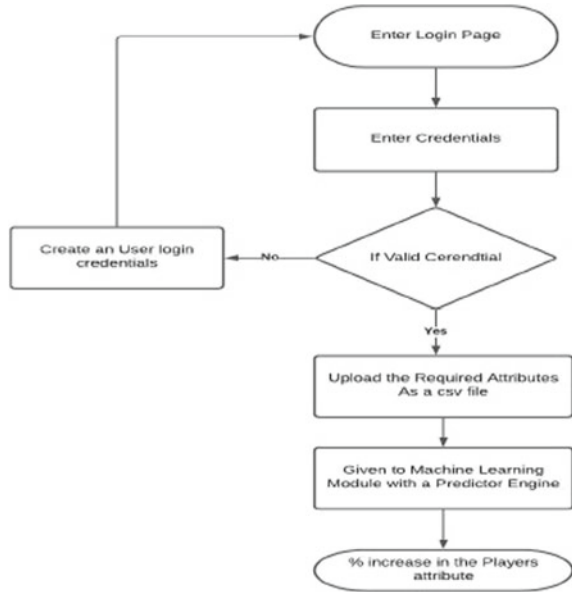
The aim of this work was to collect gaming forum posts and use this data to predict sales of video games. The data were collected from 2008 and 2009 from a major gaming message board. The developers [15] extracted mentions of each game and used the number of these mentions as well as sales from the previous two weeks to predict sales in the upcoming weeks. The only evaluation metric used is mean absolute error, making any conclusion of the results difficult.

## **3 Proposed System**

### ***3.1 Flowchart***

This system is intended for the use by a company that develops and launches games on the steam store. They are the ones who collect and give the data required. This

**Fig. 1** Flowchart for the model



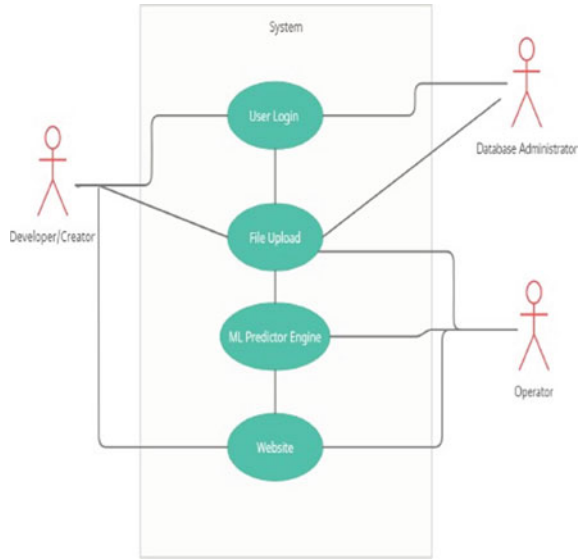
data should have the required attributes which are collected during the pre-release or during the beta testing of the game, and the model predicts the future of the game with these given values. The proposed system has a Web-based user interface where the gaming company has to first create an account and login into the main page, where they will be asked to upload the required file containing the attributes for the model. This Web page is connected to our predictor engine (random forest regressor) in the backend. The file is then processed by the model which gives the “percentage increase in the number of players (i.e., the audience)” for the game in the near future. Based on the predictions, the developers can adopt smart marketing strategies to boost their sales. They could also modify parts of the newly released game which they feel could please their users even more, thus making the game more popular and pushing it to the top. All these criteria are shown in the Fig. 1.

### 3.2 Design Diagrams

## 4 Dataset

For demonstrating the basic working of the model, we chose Counter-Strike : Global Offensive aka CS:GO, which is one of the most popular online FPS games. The dataset is collected from sullygnome.com and valve’s steam engine platform [16]. It has 2134 rows/instances, which has values from 17 July 2015 to 19 May 2021. It also has 6 attributes which are (Fig. 2)

Fig. 2 Use case diagram



1. Date Time: Date at which rest of the attributes were acquired. It is in the format of DD-MM-YYYY.
2. Day: Numerical data which are of the range of [1, 7] where 1 being Sunday and being Saturday.
3. Holiday: A binary attribute, whose value will be 1 if that particular day was a holiday and 0 if it was not.
4. Twitch Viewers: A numerical value which shows the number of viewers who were watching the game on Twitch at that particular day.
5. Players: A numerical value that shows the number of players who are playing this game on Twitch on that particular day.
6. Viewer Ratio: A float value, which is the maximum number of viewers who were watching the game on Twitch in the peak hour.

Figure 3 plots the correlation between the attributes used for predictions.

## 5 Methodology and Experiments

### 5.1 Time Series Models

**ARIMA:** Autoregressive integrated moving average (ARIMA) [17] is a type of model which explains a given time series model based on its historical data, i.e., its own lags and the lagged forecast errors, so that the equation can be used to forecast/predict the future values. The time in this model can be daily, monthly, or even per year.

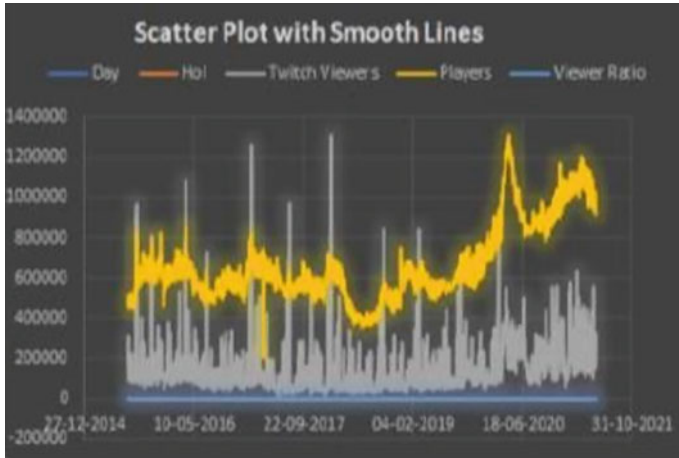


Fig. 3 Correlation between the attributes

During data preprocessing, the Date Time attribute is converted from object type to DateTime type. Date Time field is indexed as ARIMA is a time series model. All the rows other than the last 30 rows were used for training, while testing used the first 30 rows of the dataset. The RMSE value was found to be 0.47152 shown in Fig. 4. It also predicted the values for the next 30 days [18]. Figure 5 shows the plotting of the predictions got using ARIMA.

```
from sklearn.metrics import mean_squared_error
from math import sqrt
rmse=sqrt(mean_squared_error(pred,test['Players']))
print(rmse)

0.4715234227037225
```

Fig. 4 RMSE value of ARIMA

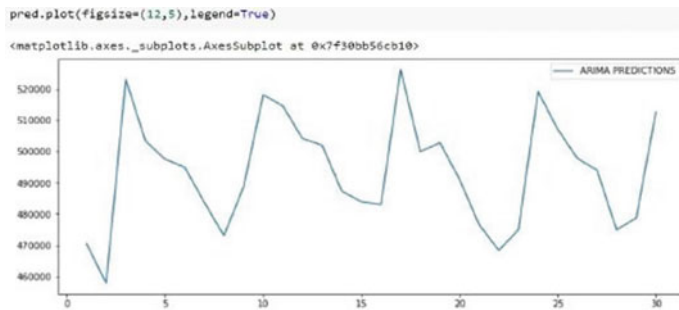


Fig. 5 Plotting of ARIMA predictions



```

from sklearn.metrics import mean_squared_error
from math import sqrt
rmse=sqrt(mean_squared_error(future_df['pred'],future_df['Players']))
print(rmse)
0.44783785744343774

```

**Fig. 6** RMSE score of SARIMA

**SARIMA:** SARIMA is the short term for seasonal autoregressive integrated moving average, [19] which is seasonal ARIMA. SARIMA is the extension of ARIMA that can model univariate time series data with seasonal component. The values were normalized using the MinMaxScaler during data preprocessing. The “Date Time” attribute was converted from object type to datetime type. The Date Time field was also indexed. For training the SARIMAX(1, 1, 1) × (1, 1, 1, 12) model, all the instances other than first 13 instances were used. The RMSE value was found to be 0.44783 [20]. Figure 6 shows the RMSE value of SARIMA developed.

## 5.2 Regressor Models

**SVR:** When support vector machine is used for regression purposes, it is then called support vector regression (SVR) [21]. It has the same principle as SVM. The difference is that the hyperplane is used to predict continuous numerical values based on the historical data. Once analyzed, SVR generates an appropriate hyperplane which is then used to predict the numerical values. Consider the distance of the decision boundaries from the hyperplane is “*a*” and the equation of hyperplane that satisfies SVR is

$$-a < y - wx + b + a \tag{1}$$

The dataset is split into training and testing sets using train-test-split with test-size as 50%. We then preprocess the data using StandardScaler for standardization and MinMaxScaler to perform normalization. This brings the values in the dataset between 0 and 1. Wu et al. [22] Once the model is set up, it predicts the “player” values. The r2\_score was found to be -0.1045 or -10.45% accuracy. Figure 7 is the r2\_score of SVR prediction.

```

from sklearn.metrics import classification_report,confusion_matrix,accuracy_score,r2_score, max_error
print(r2_score(Y_testlist,Y_predlist))
-0.10455388705740032

```

**Fig. 7** R2\_score of SVR

**Bayesian Regression:** The aim of Bayesian linear regression [23] is not to find the single best value of the model parameters like linear regression, but to determine the posterior distribution for the model parameters. Not only is the dependent variable generated from a probability distribution, but the model parameters are assumed to come from a distribution as well. The advantages of Bayesian regression over conventional regression are that: It is very effective when the size of the dataset is small. It is particularly well-suited for real-time-based learning, it gives much more compact confidence intervals [23].

For this model, the train-test-split is set as 0.7 with test-size as 30%. The Bayesian regressor model is imported and fitted with the data. A scatter plot is used to plot the predicted values and true values. The `r2_score` achieved using this model was 0.2089 which is an accuracy of 20.89%. Figure 8 shows the `r2_score` of this model, and Fig. 9 shows the scatter plot of predictions got using this method.

**Random Forest Regression:** Random forest regression [24] is a machine learning model which is capable of both regression and classification. It operates by constructing several decision trees during training time and outputting the mean of the classes as the prediction of all the trees. It usually performs smoothly on problems which may include features with non-linear relationships. The train-test ratio is 0.6, and the split index is 180. The random forest regressor functions are then used and fitted to the model. The predicted value and the true value are compared, and the error is calculated based on the difference between the predicted value and true value. The percentage increase from 0 to 10th day is 13.04%. The percentage

```
# Evaluation of r2 score of the model against the test set
print(f"r2 Score Of Test Set : {r2_score(y_test, prediction)}")

r2 Score Of Test Set : 0.20897615539910153
```

Fig. 8 R2\_score of Bayesian regression

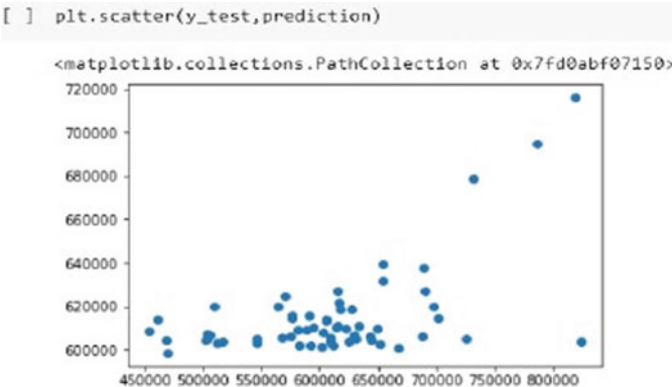


Fig. 9 Scatter plot of Bayesian regressor prediction

```
# print(confusion_matrix(y_test[0],Y_pred))
print(r2_score(Y_testlist,Y_predlist))

0.5530838650149592
```

**Fig. 10** R2\_score of random forest regression

```
diff=preds[10]-preds[0]
increase=(diff/preds[0])*100
print("Percentage increase: "+str(round(increase,2))+"%")

Percentage increase: 13.04%
```

**Fig. 11** Percentage increase in first 10 days of prediction

```
diff=preds[30]-preds[0]
increase=(diff/preds[0])*100
print("Percentage increase: "+str(round(increase,2))+"%")

Percentage increase: 2.57%
```

**Fig. 12** Percentage increase in first 30 days of prediction

increase for one month is 2.57 %. The r2\_score achieved is 0.553, or an accuracy of 55% [25]. Figure 10 shows r2\_score of random forest regression, and Figs. 11 and 12 show the percentage increase in first 10 and 30 days of prediction.

## 6 Results

This work involved trying various models for comparison. Both RMSE and r2\_score were used to evaluate the performance of the models.

### 6.1 Time Series Models

The RMSE value achieved for ARIMA was 0.47152, and the RMSE value for SARIMA was found to be 0.44783. There is no much difference between the values, or we cannot compare both values and use. This is because the dataset has few attributes and instances. If more data were available, the model would have become more accurate. So we cannot use ARIMA or SARIMA for our prediction engine model. Table 1 shows the RMSE values of ARIMA and SARIMA.

**Table 1** RMSE values for time series models

Models	ARIMA	SARIMA
RMSE value	0.4715	0.4478

**Table 2** r2\_score of machine learning models

Models	SVR	Bayesian regressor	Random forest regressor
r2_score	-0.014	0.208	0.55

## 6.2 Regressor Models

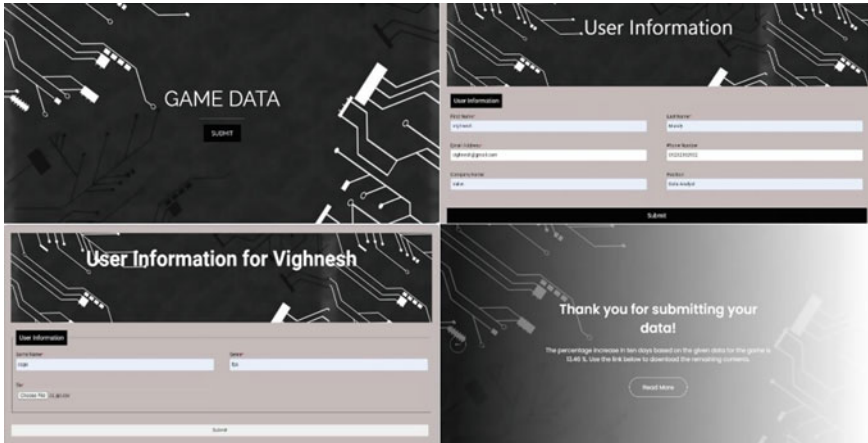
Table 2 shows the r2\_score of all the machine learning models which were tried. All three machine learning models tried were regressor models. The value of r2\_score ranges from -2 to 1 where 1 is perfect and -2 is the worst score possible. The r2\_score for SVR was found to be -0.1045 or -10.45% accuracy and that achieved using Bayesian regression model was 0.2089. The r2\_score of random forest regressor is 0.55, which shows the model is having less error and has the best accuracy of 55%. Thus, we choose random forest regressor as the predictor engine model for our system.

## 7 User Interface

### 7.1 Django

Django [26] is a popular and open-source Web framework in Python that enables quick development of maintainable Websites. The core of Django is a views.py Python file. This file integrates all the other Python files and returns HTTP responses to the client. The other Python files include models.py, urls.py, HTML template files, and forms.py Python file. Each of which has a very specific purpose. The models.py defines the database inbuilt with Django which is primarily SQLite but can be integrated with many other databases such as MongoDB and MySQL. Whenever a particular Website URL is encountered, the urls.py file is checked for the matching pattern and calls the corresponding view function for further action. When a view function has finished execution, it calls an HTML file in the templates directory and passes the data to be displayed or used as parameters.

Added to the default working of Django, in our Website, there exists a processing.py file which contains the function to run the random tree regression model explained in above. The views.py function output() called after submitting the user data calls the protry() function in processing.py to run the test dataset, and the returned output is given as argument to the html template, displaying the data as output. Figure 13 shows the user interface which was created using Django.



**Fig. 13** GUI using Django

## 8 Discussion

To develop this system, various models were tried including time series and machine learning models. A deep learning model (LSTM) [27] was also attempted. This type of RNN model is mainly designed to avoid the long-term dependency problem, hence it finds it easy to remember information of the data points over a long period of time. All recurrent neural networks have the form of a chain of repeating modules of neural network. Once creating and fitting the model, the loss calculated was very high since the number of instances and attributes was less. So the LSTM model failed in accuracy aspect.

As stated earlier, this system is novel as it provides a way to predict the success of PC/online FPS games compared to the previous systems which could handle only normal video games. It could help developers to increase their sales and/or to get an edge over the competition. Future enhancements are planned like adding more features to the dataset to improve the accuracy of prediction. Moreover, as the field of machine learning is expanding each day, a better model could be used for more accuracy.

## 9 Conclusion

Online gaming helps provide a fun and exciting way to pass time especially during this pandemic and also develops teamwork skills. A new player or a new developer coming to the field of online game production might not be aware of which game might be successful in future. This system aims at predicting the percentage increase

in the number of players within a time span, to know whether that game will have success or not in future.

Once a game developer releases a game onto the steam library, gamers will start downloading and trying the game out. The steam engine records this data automatically. Streamers on Twitch start streaming the game simultaneously, and the Twitch view count is recorded for the particular game. This data are easily accessible to the developer, and so they collect this data. The longer the game runs, the more data gets collected, the better the accuracy of the system. This data can be submitted through the Web portal of the proposed system. Once received, the model is trained with the dataset, and the necessary values are predicted. Out of all the five models tried, random forest regressor was the most accurate, and hence, it was selected as the predictor engine for this system, where it achieved an accuracy of 55%. In the system's Web UI, the prospective developers log in using their credentials and uploads the game's dataset in the form of a csv file. The system then displays the percentage increase in number of players on a new Web page.

## References

1. How Much Is the Gaming Industry Worth in 2021?—TechJury. <https://techjury.net/blog/gaming-industry-worth/>
2. Steam: Upping the game in 2020—Digital Innovation and Transformation. <https://digital.hbs.edu/platform-digit/submission/steam-upping-the-game-in-2020/>
3. News- ValveLaunchesSteamGreenlight. <https://store.steampowered.com/oldnews/8761#:~:text=August%2030%2C%202012%20%2D%2D%20Valve,community's%20help%20in%20selecting%20some>
4. Why CS: GO Is The Best FPS Game Of All Time. <https://www.gamebyte.com/csgo-best-fps-game-time/>
5. Yang W, Huang T, Zeng J, Chen L, Mishra S, Liu Y (2020) Utilizing players playtime records for Churn prediction: mining playtime regularity. *IEEE Trans Games* 1–1. <https://doi.org/10.1109/tg.2020.3024829>
6. Application of Logistic Regression with Different Sampling Models (2005) In: *Applied logistic regression*. John Wiley & Sons Inc., pp 203–222. <https://doi.org/10.1002/0471722146.ch6>
7. Yan Q (2020) SOC prediction of power battery based on SVM. In: *2020 Chinese control and decision conference (CCDC)*. IEEE. <https://doi.org/10.1109/ccdc49329.2020.9164245>
8. Tato JH, Brito MC (2018) Using smart persistence and random forests to predict photovoltaic energy production. *Energies* 12:100. <https://doi.org/10.3390/en12010100>
9. Yao W, Huang P, Jia Z (2018) Multidimensional LSTM networks to predict wind speed. In: *2018 37th Chinese control conference (CCC)*. IEEE. <https://doi.org/10.23919/chicc.2018.8484017>
10. Steam is turning into the App Store and that's OK | Engadget. <https://www.engadget.com/2016-09-27-steam-turning-into-the-app-store.html>
11. *Digital Games Assessment, and Learning* (2013) In: *Stealth assessment*. The MIT Press. <https://doi.org/10.7551/mitpress/9589.003.0005>
12. Steam: Game and Player Statistics. <https://store.steampowered.com/stats/>
13. Sifa R, Srikanth S, Drachen A, Ojeda C, Bauckhage C (2016) Predicting retention in sandbox games with tensor factorization-based representation learning. In: *2016 IEEE conference on computational intelligence and games (CIG)*. IEEE. <https://doi.org/10.1109/cig.2016.7860405>

14. The market for video games (2016) In: Video game marketing. Routledge, pp 21–34. <https://doi.org/10.4324/9781315748900-7>
15. Huang L-Y, Hsieh Y-J (2011) Predicting online game loyalty based on need gratification and experiential motives. *Internet Res* 21:581–598. <https://doi.org/10.1108/10662241111176380>
16. Steam, The Ultimate Online Game Platform. <https://store.steampowered.com/about/>
17. Ayaz Z, Fiaidhi J, Sabah A, Ansari MA (2020) Bitcoin price prediction using ARIMA model. <https://doi.org/10.36227/techrxiv.12098067.v1>
18. Leiner B (1985) Prediction with Arima filters. In: Contributions to econometrics and statistics today. Springer Berlin Heidelberg, pp 177–191. [https://doi.org/10.1007/978-3-642-70189-4\\_16](https://doi.org/10.1007/978-3-642-70189-4_16)
19. Chikkakrishna NK, Hardik C, Deepika K, Sparsha N (2019) Short-term traffic prediction using Sarima and FbPROPHET. In: 2019 IEEE 16th India council international conference (INDICON). IEEE. <https://doi.org/10.1109/indicon47234.2019.9028937>
20. Nikmatillah VM, Anggraeni D, Hadi AF (2018) Prediction interval in seasonal autoregressive integrated moving average (SARIMA) model for rainfall forecasting and drought. In: Proceedings of the international conference on mathematics and Islam. SCITEPRESS—science and technology publications. <https://doi.org/10.5220/0008517801010107>
21. Regression with Support Vector Machines (2009) In: Knowledge discovery with support vector machines. John Wiley & Sons Inc., pp 193–208. <https://doi.org/10.1002/9780470503065.ch12>
22. Wu Y, Huang S, Zhao H, Cao K, Gan J, Yang C, Xu Z, Li S, Su B (2021) Zebrafish minichromosome maintenance protein 5 Gene regulates the development and migration of facial motor neurons via fibroblast growth factor signaling. *Dev Neurosci* 1–11
23. Wang X, Yue Y, Faraway JJ (2018) Bayesian linear regression. In: Bayesian regression modeling with Inla. Chapman and Hall/CRC, pp 39–69. <https://doi.org/10.1201/9781351165761-3>
24. Nakashima H, Arai I, Fujikawa K (2019) Passenger counter based on random forest regressor using drive recorder and sensors in buses. In: 2019 IEEE international conference on pervasive computing and communications workshops (PerCom Workshops). IEEE. <https://doi.org/10.1109/percomw.2019.8730761>
25. Table 2: Modeling verification, and prediction accuracy evaluation index using random forest regression. [10.7717/peerj.8243/table-2](https://doi.org/10.7717/peerj.8243/table-2)
26. Welcome to Django. In: Practical Django Projects. Apress, pp 1–8. [https://doi.org/10.1007/978-1-4302-0868-6\\_1](https://doi.org/10.1007/978-1-4302-0868-6_1)
27. Li Y, Ye F, Liu Z, Wang Z, Mao Y (2021) A short-term photovoltaic power generation forecast method based on LSTM. *Math Probl Eng* 2021:1–11. <https://doi.org/10.1155/2021/6613123>

# On Strong Proper Connection Number of Some Interconnection Networks



I. Annammal Arputhamary, D. Angel, and A. Shenbaga Priya

## 1 Introduction

Consider a network of relay stations where each pair of relay stations has access to a certain frequency channel for message transmission. If in a relay station, a message is transmitted through a specific frequency, the same frequency cannot be reused in the same station as it would lead to interference. This means that we need two different frequency transmissions of message. The number of frequencies is required to allocate the connections among towers so that a path of shortest length is present between any two towers can be related to SPCN of the associated network [1, 2]. In this article, strong proper connection number (SPCN) of butterflies, honeycombs, and Benes networks is investigated. See Fig. 1.

## 2 An Overview of the Paper

For all basic definitions and notations related to graph, refer [1]. Andrews et al. [2] proposed the concept of proper-path colorings as a result of the inspiration of rainbow colorings. In 2011, proper colored cycles and paths were discussed [5]. In 2012, Borozan et al. [3] presented the theory of proper k-connection of graphs. Minimum degree conditions for proper connection number of graphs were investigated by Brause et al. [4] in 2017. In 2016, Laforge et al. [6] explored all graphs which are

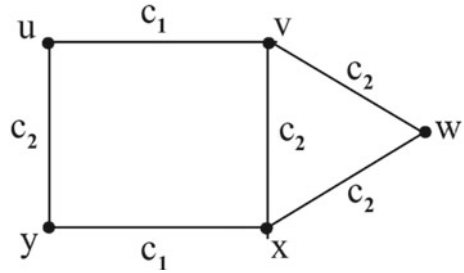
---

I. Annammal Arputhamary (✉) · A. S. Priya  
Department of Mathematics, St. Anne's Arts and Science College, Chennai 600 110, India  
e-mail: [annammalarputhamary@ssacollegechennai.com](mailto:annammalarputhamary@ssacollegechennai.com)

D. Angel  
Department of Mathematics, Sathyabama Institute of Science and Technology, Chennai 600 119, India



**Fig. 1** SPCN of the above graph is 2



connected having SPCN,  $n - p$  where  $p = 1, 2, 3$ . Directed proper connection of graphs was studied by Magnant et al. [8].

### 3 Strong Proper Connection Number of Some Interconnection Networks

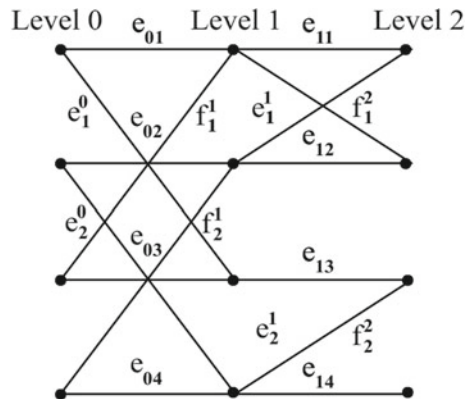
#### 3.1 Butterfly Network $\mathcal{BF}(n)$

$\mathcal{BF}(n)$  has  $(n + 1)2^n$  vertices and  $n(2^{n+1})$  edges,  $n + 1$  levels and  $2^n$  vertices in every level for all  $n \geq 1$  [9]. Vertices on level 0 and level  $n$  are of degree 2, and the vertices on level  $h$   $1 \leq h \leq n - 1$  are of degree 4. See Fig. 2.

**Theorem 3.1.1** SPCN for a  $\mathcal{BF}(n) = 4$ .

**Proof**  $\mathcal{BF}(n)$  has  $n + 1$  levels, namely level 0, level 1, ..., level  $n$  with  $2^n$  rows. Let  $e_{hi}$  be the straight edge (SE) between level  $h$  and level  $h + 1$ ,  $0 \leq h \leq n - 1$ ,  $1 \leq i \leq 2^n$ . Let the cross-edge (CE) from  $h$  level to  $h + 1$  level be  $e_i^h$  and CE from level  $h + 1$  to level  $h$  be  $f_i^{h+1}$ ,  $0 \leq h \leq n - 1$ ,  $1 \leq i \leq 2^{n-1}$ .

**Fig. 2** Labeling of  $n$ -dimensional butterfly



Denote  $c(e_{hi})$  as straight edge  $e_{hi}$  color,  $c(e_i^h)$  and  $c(f_i^{h+1})$  be the color of the cross-edges  $e_i^h$  and  $f_i^{h+1}$ , respectively.

**Algorithm for Strong Proper Connected Coloring.**

**Input:**  $\mathcal{BF}(n)$  of dimension  $n$ ,  $n \geq 2$ .

**Algorithm:**

- (i) For  $SE$  betwixt level  $h$  and  $h + 1$ ,  $0 \leq h \leq n - 1$  and  $1 \leq i \leq 2^n$ .

$$c(e_{hi}) = \begin{cases} c_1 \text{ for } h \equiv 0 \pmod{2} \\ c_2 \text{ for } h \equiv 1 \pmod{2} \end{cases}$$

- (ii) For  $CE$  from level  $h$  to  $h + 1$ ,  $0 \leq h \leq n - 1$  and  $1 \leq i \leq 2^{n-1}$

$$c(e_i^h) = \begin{cases} c_3 \text{ for } h \equiv 0 \pmod{2} \\ c_4 \text{ for } h \equiv 1 \pmod{2} \end{cases}$$

And the  $CE$  from level  $h + 1$  to  $h$ ,  $0 \leq h \leq n - 1$  and  $1 \leq i \leq 2^{n-1}$

$$c(f_i^{h+1}) = \begin{cases} c_3 \text{ for } h \equiv 1 \pmod{2} \\ c_4 \text{ for } h \equiv 0 \pmod{2} \end{cases}$$

**Output:**

$$SPCN(\mathcal{BF}(n)) = 4$$

**Proof of Correctness:**

Let  $G$  be the butterfly network  $\mathcal{BF}(n)$ ,  $n \geq 2$ . Let us define the colors to  $SE$  and  $CE$  as follows. Color  $SE$  is in the middle of level 0 as well as level 1 with the color  $c_1$ . That is  $c(e_{0i}) = c_1$ ,  $1 \leq i \leq 2^n$ . All the  $2^n$   $SE$ s in the middle of level 0, as well as level 1, acquire  $c_1$  color. Next, the  $SE$ s in the middle of level 1 as well as level 2 have  $(e_{1i}) = c_2$  color,  $1 \leq i \leq 2^n$ . That is all the  $2^n$   $SE$ s in the middle of level 1 as well as level 2 can be given the  $c_2$  color.

In general, color  $SE$ s in the middle of level  $h$  as well as level  $h + 1$ ,  $0 \leq h \leq n - 1$ , as

$$c(e_{hi}) = \begin{cases} c_1 \text{ for } h \equiv 0 \pmod{2} \\ c_2 \text{ for } h \equiv 1 \pmod{2} \end{cases}$$

That is all the  $2^n$  straight edges between level  $h$  and level  $h + 1$  get the color  $c_1$ ,  $h \equiv 0 \pmod{2}$  and with the color  $c_2$ ,  $h \equiv 1 \pmod{2}$ .

From the above discussion, we see that only two colors are required for straight edges.

Next, color the cross-edges as follows. Color the cross-edges from level 0 to level 1 with  $c_3$ , level 1 to level 2 with  $c_4$ , level 3 to level 4 with  $c_3$ , and so on. Similarly, color the cross-edges from level 1 to level 0 with  $c_3$ , level 2 to level 1 with  $c_4$ , level 4 to level 3 with  $c_3$ , and so on.

In general, color the cross-edges from level  $h$  to level  $h + 1$  as  $c_3, h \equiv 0(mod2)$  and with the color  $c_4, h \equiv 1(mod2)$ . Similarly, color the cross-edges from level  $h + 1$  to level  $h$  as  $c_3, h \equiv 1(mod2)$  and with the color  $c_4, h \equiv 0(mod2)$ . Thus, two colors are required for cross-edges.

We note that totally four colors are needed for strong proper connected coloring. If we reduce the color by 3, then there exists at least one path in which two adjacent edges receive the same color.

Thus,  $SPCN$  of  $\mathcal{BF}(n) = 2 + 2 = 4$ . See Fig. 3.

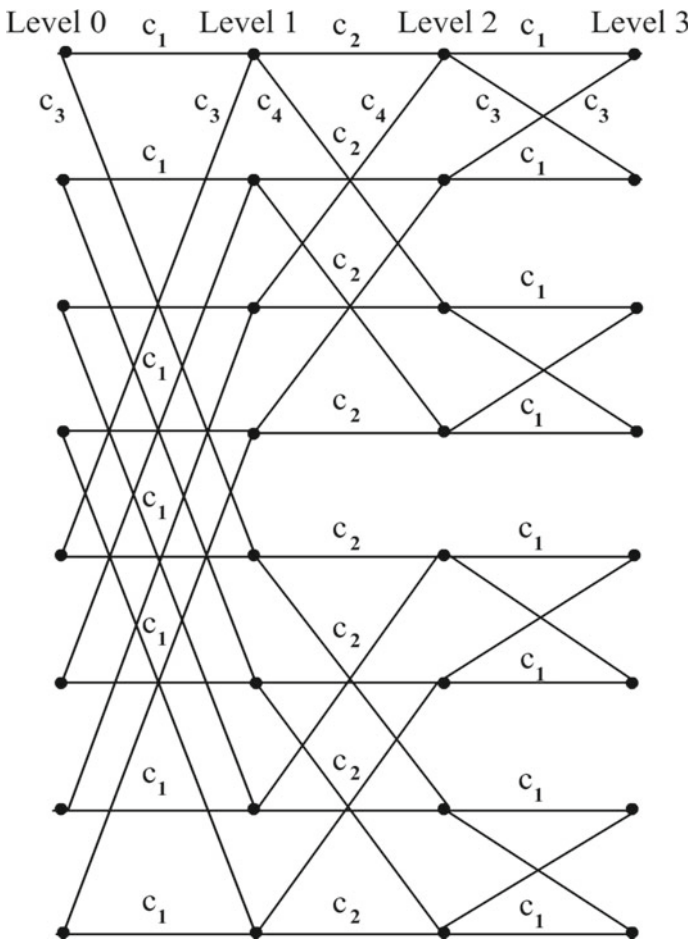


Fig. 3  $SPCN$  of  $\mathcal{BF}(3) = 4$

The above algorithm allocates colors such that every two vertices are joined by the shortest proper path.

### 3.2 SPCN for Benes Networks $B(r)$

$B(r)$  has  $2r + 1$  levels together with  $2^r$  nodes in every level.

**Theorem 3.2.1** SPCN of  $B(n)$  is 4 for all  $n \geq 2$ .

**Proof** Let  $e_{hi}$  be the SE between level  $h$  and level  $h + 1$ ,  $0 \leq h \leq 2n - 1$ ,  $1 \leq i \leq 2^n$ . Let CE from level  $h$  to level  $h + 1$  be  $e_i^h$  and the CE from level  $h + 1$  to level  $h$  be  $f_i^{h+1}$ ,  $0 \leq h \leq 2n - 1$ ,  $1 \leq i \leq 2^n$ .

We define  $c(e_{hi})$  as the color of the straight edge  $e_{hi}$ ,  $c(e_i^h)$  and  $c(f_i^{h+1})$  as the color of the cross-edges  $e_i^h$  and  $f_i^{h+1}$ , respectively.

**Algorithm Proper Connected Coloring:**

**Input:** A Benes network  $B(n)$  of dimension  $n$ .

**Algorithm:**

- (i) Color the SE between level  $h$  and level  $h + 1$ ,  $0 \leq h \leq 2n - 1$  as,  
For  $1 \leq i \leq 2^n$

$$c(e_{hi}) = \begin{cases} c_1 & \text{for } h \equiv 0 \pmod{2} \\ c_2 & \text{for } h \equiv 1 \pmod{2} \end{cases}$$

- (ii) Color the CE from level  $h$  to  $h + 1$ ,  $0 \leq h \leq 2n - 1$  as,  
For  $1 \leq i \leq 2^{n-1}$ ,

$$c(e_i^h) = \begin{cases} c_3 & \text{for } h \equiv 0 \pmod{2} \\ c_4 & \text{for } h \equiv 1 \pmod{2} \end{cases}$$

Color the CE from level  $h$  to  $h + 1$ ,  $0 \leq h \leq 2n - 1$  as,  
For  $1 \leq i \leq 2^{n-1}$

$$c(f_i^{h+1}) = \begin{cases} c_3 & \text{for } h \equiv 0 \pmod{2} \\ c_4 & \text{for } h \equiv 1 \pmod{2} \end{cases}$$

**Output:**

SPCN of  $B(n) = 4$ .

**Proof of Correctness:**

The above coloring gives properly connected graph, and shortest proper path is present betwixt two vertices. Thus, SPCN ( $B(n)$ ) = 4. See Fig. 4.

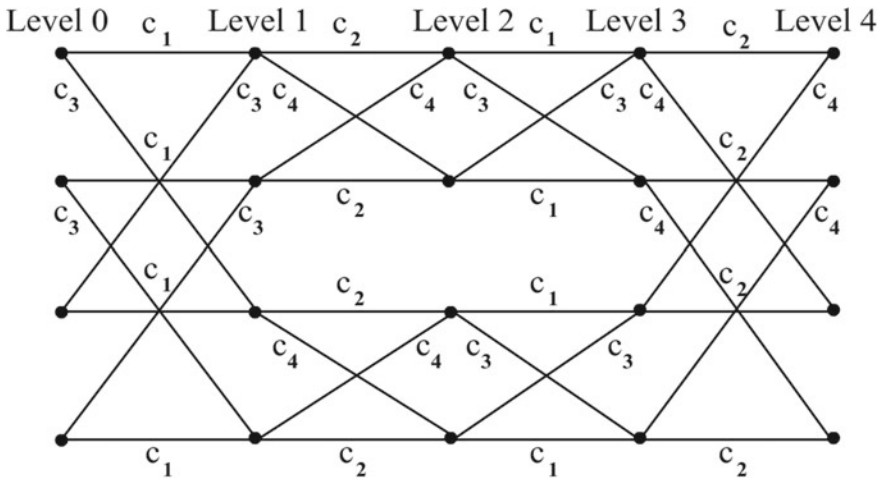


Fig. 4 SPCN of  $B(2) = 4$

### 3.3 Strong Proper Connection Number of Honeycomb Network $HC(n)$

Honeycomb network  $HC(n)$  is suitable to design wireless network base stations. Applications can be found in [7, 11]. Its structural properties are discussed by Manuel et al. [10].

The labeling structure in Fig. 5 is used to label the row edges (RE) and vertical edges (VE). Denote the rows in  $HC(n)$  by  $R_i, 1 \leq i \leq 2n$ . The row edges in  $HC(n)$  are represented by  $e_{i,l}; l$  is its position in the  $i$ th row,  $1 \leq l \leq 4n - 2$ .

The vertical edges between row  $R_i$  and  $R_{i+1}$  are indicated as  $e_{i,i+1}^{k,j}$  where  $k$  indicates its position between row  $R_i$  and  $R_{i+1}, 1 \leq i \leq 2n - 1$ , and  $1 \leq k \leq 2n$ . See Fig. 5.

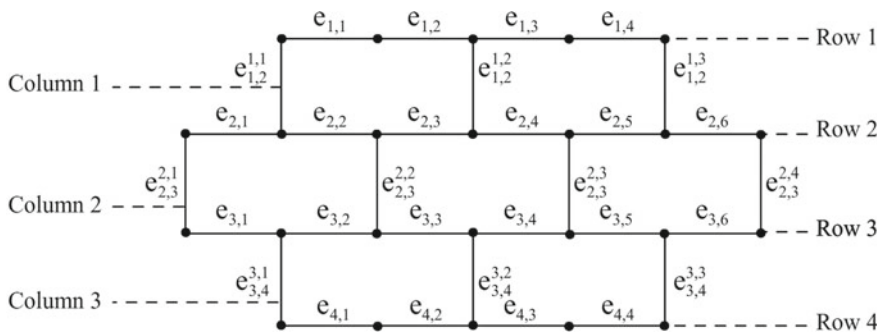


Fig. 5 Brick representation of  $HC(2)$

**Theorem 3.3.1**  $\text{SPCN}(\text{HC}(n)) = 3, n \geq 2$ .

**Proof** Let  $G$  be the graph  $(n), n \geq 2$ .

**Input:**  $\text{HC}(n), n \geq 2$ .

**Algorithm:**

**Step 1:** Consignment of colors to RE.

- (i) When  $1 \leq i \leq n$  and  $1 \leq l \leq 2(n + i - 1)$

$$c(e_{i,l}) = \begin{cases} c_1 \text{ if } l \equiv 1 \pmod{2} \\ c_2 \text{ if } l \equiv 0 \pmod{2} \end{cases}$$

- (ii) When  $n + 1 \leq i \leq 2n$  and  $1 \leq l \leq 6n - 2i$

$$c(e_{i,l}) = \begin{cases} c_1 \text{ if } l \equiv 1 \pmod{2} \\ c_2 \text{ if } l \equiv 0 \pmod{2} \end{cases}$$

**Step 2:** Consignment of colors to VE.

- (i) When  $1 \leq i \leq n, 1 \leq k \leq n$  and  $1 \leq j \leq n + k$ .

$$c(e_{i,i+1}^{k,j}) = c_3.$$

- (ii) When  $n + 1 \leq i \leq 2n, n + 1 \leq k \leq 2n - 1$  and  $1 \leq j \leq n + k - 2(k - n)$ .

$$c(e_{i,i+1}^{k,j}) = c_3.$$

**Output:**

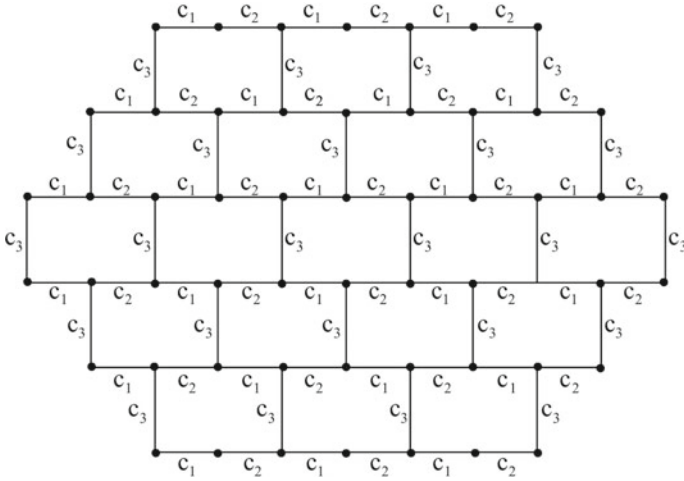
$$\text{SPCN}(\text{HC}(n)) = 3$$

**Proof of Correctness:**

Using the above algorithm, the edges in each row have been consigned with  $c_1$  and  $c_2$  alternately, and the edges in each column are consigned with  $c_3$ , respectively. Since there is the shortest path which is proper that connects each of the two nodes,  $\text{SPCN}$  of  $\text{HC}(n) = 3$ . See Fig. 6.

## 4 Conclusion

Exact values of strong proper connection of the  $n$ -dimensional butterflies, Benes, and  $\text{HC}(n)$  have been found out using the strong proper connected coloring. A secure communication protocol has been established using the strong proper connection model. These results are helpful in establishing a protected information exchanging system. The strong proper connection number for star and pancake networks can be explored in the near future.



**Fig. 6** SPCN of  $HC(3) = 3$

### References

1. Bondy JA, Murty USR (2008) Graph theory. Springer Edition
2. Andrews E, Laforge E, Lumduanhom C, Zhang P (2006) On proper-path colorings in graphs. *J Comb Math Comb Comput* 97:189–207
3. Borozan V, Fujita S, Gerek A, Magnant C, Manoussakis Y, Montero L, Tuza Z (2012) Proper connection of graphs. *Discrete Math* 312(17):2550–2560
4. Brause C, Doan TD, Schiermeyer I (2017) Minimum degree conditions for the proper connection number of graphs 33(1):833–843
5. Fujita S, Magnant C (2011) Properly colored paths and cycles. *Discrete Appl Math* 159:1391–1397
6. Laforge E, Lumduanhom C, Zhang P (2016) Characterizations of graphs having large proper connection numbers. *Discuss Math Graph Theory* 36(2):439–454
7. Lester LN, Sandor J (1984) Computer graphics on hexagonal grid. *Comput Graphics* 8:401–409
8. Magnant C, Morley PR, Porter SA, Nowbandegani PS, Wang H (2016) Directed proper connection of graphs. *Matematiqki Vesnik* 68(1):58–65
9. Manuel P, Abd-El-Barr MI, Rajasingh I, Rajan B (2008) An efficient representation of Benes network and its applications. *J Discrete Algorithms* 6:11–19
10. Manuel P, Rajan B, Rajasingh I, Monica C (2008) On minimum metric dimension of honeycomb networks. *J Discrete Algorithms* 6:20–27
11. Stojmenovic I (1997) Honeycomb networks: Topological properties and communication algorithms. *IEEE Trans Parallel Distrib Syst* 8:1036–1042

# A Virtual Assistant for the Visually Impaired



Alex Roy, Amal Saji, M. S. Gokul, and Subu Surendran

## 1 Introduction

Visual Impairment is a medical condition affecting many people around the world. According to Hindustan Times, “India currently has around 12 million blind people against 39 million globally.” Dealing with blindness is a challenge already in itself, yet they strive to be self-sufficient. It is in human nature to be forgetful. Even when we have two eyes, we find it difficult to locate items that we have misplaced. So it is easy to imagine how much worse it would be for a visually impaired person. They will have to seek help from others constantly. This dependent nature makes them susceptible to feel like a burden to others, and lose confidence in themselves, even when they try hard to be self-sufficient.

Some of the major challenges include difficulty in finding an object without the assistance of someone. Some devices available in the market help them to overcome a few of these challenges. There is always a huge number of researches involved with the sole aim of building devices to help these visually challenged people.

---

A. Roy · A. Saji (✉) · M. S. Gokul · S. Surendran  
Department of Computer Science and Engineering, Sree Chitra Thirunal College of Engineering,  
Pappanamcode, Thiruvananthapuram, Kerala 695018, India  
e-mail: [amalsaji25@gmail.com](mailto:amalsaji25@gmail.com)

A. Roy  
e-mail: [alexroy4773h@gmail.com](mailto:alexroy4773h@gmail.com)

M. S. Gokul  
e-mail: [gokulmsudha@gmail.com](mailto:gokulmsudha@gmail.com)

S. Surendran  
e-mail: [subusurendran@sctce.ac.in](mailto:subusurendran@sctce.ac.in)



## 2 Related Works

Some of the existing technologies to help the visually impaired include a smart eye system [1] that detects obstacles in front of them and identifies people as mentioned by Joe Louis Paul. Another work done in this area is the creation of smart guiding glasses [2] that help them to overcome traveling difficulty. Research works [3, 4] done in this field are based on obstacle detection which helps them to move from one place to another. One limitation of such works was that they used less accurate sensors. A better working sensor is generally more expensive and might not be economically feasible for the end user. Also, some of the object detection algorithms do not specify what the obstacles are, they just detect the presence of some object. This led this paper to focus on object detection integrated with a voice-enabled system.

## 3 Methodology

A system developed to help the visually impaired should provide the most convenient user interface. The virtual assistant in this paper is Google Assistant, which performs automatic speech-to-text [5] and text-to-speech [6] conversions. It is integrated with the most commonly used smart home device, which helps the user to convey his/her needs without much physical interaction with a device and helps the solution to reach a wide range of users. The various technologies used are Natural Language

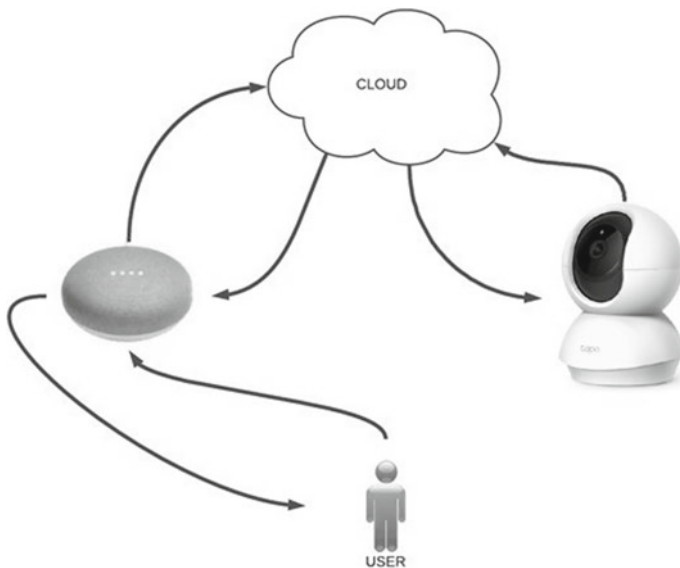


Fig. 1 Proposed model

Processing, Image Processing, and Cloud Computing. Figure 1 shows the proposed model:

### ***3.1 Integrating Smart Devices with Google Cloud***

The Google mini uses the “Actions on Google” [7] and “Dialogflow” [8] platforms to integrate cloud computing with the conversational user interface. Tp-link Tapo C200 IP camera is used with this experiment. This ip camera uses the RTSP (Real Time Streaming Protocol) to stream its feed to the cloud or media server using mp4 file format and codec H.264. A python function with OpenCV library is used to get an image frame from the camera feed/stream using VideoCapture function. It is sent to the cloud storage using google.cloud.storage library, as an image (jpg) input for further processing.

### ***3.2 Natural Language Processing with Dialogflow***

The Natural Language Processing functionalities presented in this paper are performed by the “Actions on Google” [7] and “Dialogflow” [8] platforms. Actions on Google is an interaction built for Google Assistant that performs specific tasks based on user input. DialogFlow is a module that uses Natural Language Understanding (NLU) and Machine Learning (ML) to transform user input into actionable data to be used by an Assistant application. The three main parts of DialogFlow are, Intent matching, which recognizes what the user wants. Entity Extraction, is used to identify the things the user mentioned. Fulfillment is the dialog control that invokes specific functionalities (here, cloud function) based on the needs of the user.

### ***3.3 Object Detection***

Object detection is more complex than classification, as it requires to recognize objects, as well as indicate where the object is located within the image. YOLOv3 [9] is more common in the field of computer vision, but in terms of accuracy YOLOv4 [10] shows much more promising results. YOLO [11] which stands for You Only Look Once, performs object detection in real time. A single neural network is applied to the entire image, which then divides the image into regions and predicts bounding boxes and probabilities for each region. The overview of Object Detection is shown in Fig. 2.

### ***3.4 Reference Object Detection***

The algorithm for calculating the reference objects for the queried objects is given below. Two reference objects are detected—one near the queried object and one on which the queried object is situated. In Algorithm 1,  $x$  represents the  $x$ -coordinate,  $y$  represents the  $y$ -coordinate of the top-left corner of the bounding box of each item,  $w$  represents the width, and  $h$  represents the height of the bounding box.

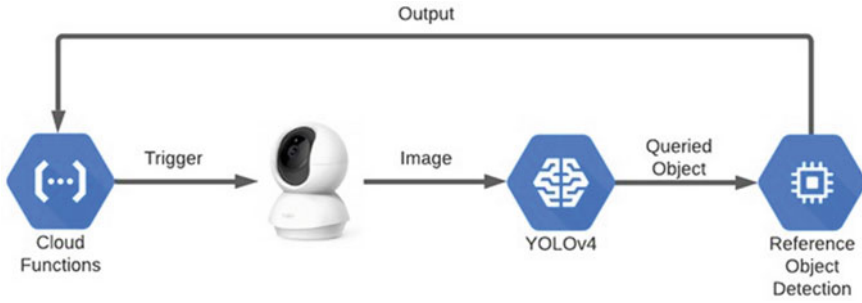


Fig. 2 Object detection overview

**Data:** The queried object, list of possible reference objects (list 01) and the output of object detection (index,label,x,y,w,h).

**Result:** Reference objects detected.

Initialize a list containing all the objects got from object detection as list 02;

```

for each object in list _02 do
    Calculate the centroid using the formula :
    Centroid (x,y) = ((2x+w)/2,(2y+h)/2) and append
    it to corresponding object in list 02 ;
end

for each object in list _02 do
    if object label = queried object then
        append the object to a new list required obj with elements
    end (index,label,x,y,w,h,near index,near name,on index,on name);
end

for each object 'A' in required obj do
    for each object 'B' in list _02 do
        Calculate the distance between 'A' and 'B' using the Euclidean distance formula
        if 'B' has least distance from 'A' then
            'B's index is stored in near index;
            'B's label is stored in near name;
        end
        if bottom edge of 'A' fully overlaps 'B' and 'B' is in list 01 then
            'B's index is stored in on index;
            'B's label is stored in on name;
        end
    end
end

```

Algorithm 1: Reference object detection algorithm

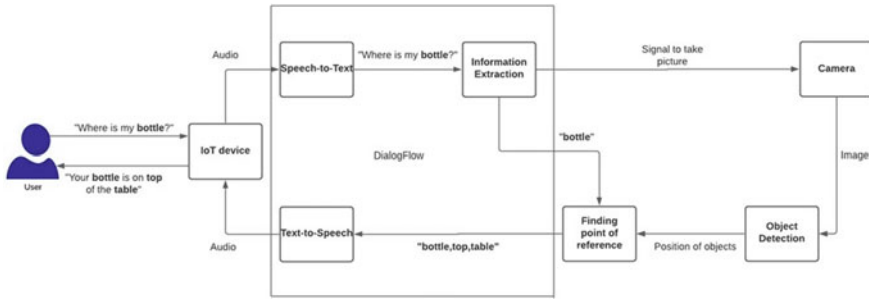


Fig. 3 Block diagram of the proposed model

### 4 Implementation

For the experiment, the object under consideration was a water bottle. It can be placed anywhere in a room. A camera was placed inside the room. The voice command given by the user, such as “find the water bottle” was first converted to text by Google’s speech-to-text API [5], the object to be found was extracted using the entity extraction module of DialogFlow. The next point in execution is to trigger a fulfillment that executes a cloud function that triggers the camera module to capture a real-time image of the room. The real-time image is passed as the input to YOLOv4 [10], which performs object detection.

Once object detection is performed successfully, information about detected objects like x and y coordinates of the top-left corner of the objects detected, height and width of the bounding box, confidence score for all the classes of objects are retrieved. This information is used to perform Reference Object Detection. The extracted entity from the DialogFlow module along with the results obtained from YOLOv4 [10] is passed to the reference object detection module, to find two reference objects, one near to the queried object and one on which the queried object is situated. The output obtained from this module is sent back to the DialogFlow which marks the completion of fulfillment and the text response is output to the user via Google Assistant. A Block Diagram showing the various interconnections of modules is shown in Fig. 3.

### 5 Results and Discussion

The entire dynamic conversation between the user and the system, starting from the user request till returning the result to the user is shown in Fig. 4. A dynamic conversation is implemented to overcome the ten-second execution time limit of fulfillment, which causes the conversation to end if it exceeds this time limit. It also helps in confirming that the system extracted the indented object by the user.

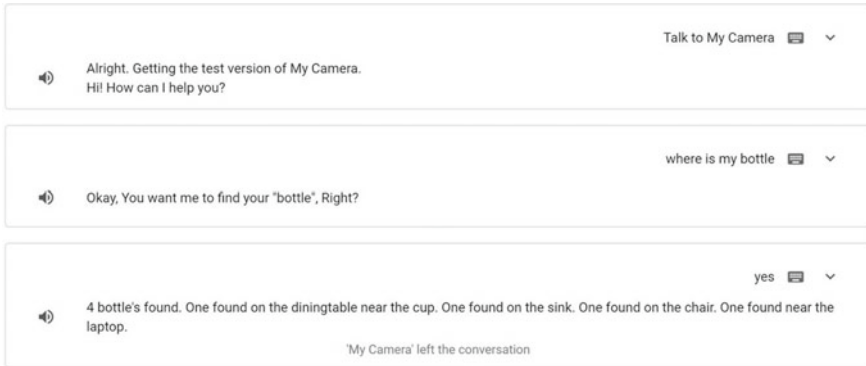


Fig. 4 Overview of the dynamic conversation

An example for an output image in which object detection has been performed is given below in Fig. 5.

The index, label, bounding-box coordinates, and the confidence values are passed to the reference object detection module, which then calculates the reference objects for the queried object. For this experiment, “bottle” was taken as the queried object. Table 1 represents the relative positions of objects found near the queried object. The “near name” stores the closest object while “on name” stores the object on which it is situated. The tabular representation of the result (Table 1) is converted into a meaningful sentence and output to the user in speech format.

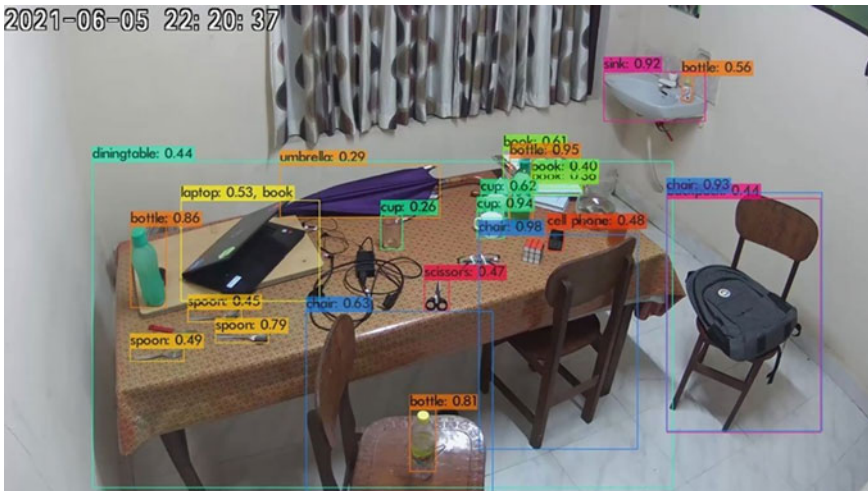


Fig. 5 YOLOv4 processed image

**Table 1** Reference object result

Index	Label	x	y	Width	Height	Near index	Near name	On index	On name
2	Bottle	278	487	79	180	4	Spoon	0	Diningtable
9	Bottle	893	890	58	139	7	Chair	7	Chair
15	Bottle	1114	337	51	159	13	Cup	0	Diningtable
22	Bottle	1492	154	24	58	19	Sink	19	Sink

**Table 2** Result comparison

Model	Mean accuracy (%)
YOLOv3-tiny	39.2
YOLOv3	59.6
YOLOv4	75.6

For this experiment, object detection was performed using three models. The YOLOv3-tiny model showed the least accuracy but fast prediction. The YOLOv3 model showed improved results with higher accuracy but the accuracy dropped

when the number of objects was very high. To further improve the accuracy of results, YOLOv4 model was used which showed even higher accuracy and the results were satisfactory. The comparison of mean accuracy of the three models is given in Table 2. The mean accuracy is calculated by Eq. (1)

$$\text{Mean Accuracy} = \frac{1}{n} \sum_j \sum_i \frac{TP_{ji}}{TP_{ji} + FP_{ji} + FN_{ji}} \quad (1)$$

where TP is True Positive (actual object matches predicted object), FP is False Positive (object does not match predicted object), FN is False Negative (actual object is not detected),  $i$  is the object, and  $n$  is the number of images.

To make this comparison, a total of twenty images with about 25 objects were given as input for each of the three models mentioned above.

## 6 Conclusion and Future Works

### 6.1 Conclusion

There are many challenges faced by visually impaired people. In most cases, they require constant support in almost all scenarios especially in their day-to-day activities. The aim of this work is to reduce their burden at least a little bit.

The paper introduces a voice-enabled system that helps to direct the visually challenged person in their day-to-day lives. The virtual assistant helps them to be

independent, than having to rely on another person. In order to find an object within a room, a visually impaired person just needs to ask the assistant where it is or to find it for them. The request gets converted to text and is processed to extract the required object. An IP camera captures an image of the room and uploads it to the cloud storage which is processed along with the extracted information to find the object. Making use of Cloud Computing enables the application to be deployed on a wide range of devices without depending on the hardware specifications. Object detection is performed using YOLOv4 (You Only Look Once) deployed in a cloud virtual machine. The YOLOv4 algorithm was used as it showed high accuracy in detecting small objects.

## 6.2 Future Works

This sort of application can find its use in multiple domains where the visually impaired would have to seek the help of others. The application can be extended to help the visually impaired shop independently at grocery shops and shopping malls. It can also be used in a different way for detecting the faces of people who visit the house and determine whether they are family members, friends, acquaintances, or a stranger, or perhaps even an intruder! It can also be extended to help them at Conferences and other social gatherings by helping them navigate through the crowd and also identifying their acquaintances without the help of another person. The NLP applications can enable the model to work with a range of languages.

## References

1. Paul IJ, Sasirekha S, Mohanavalli S, Jayashree C, Priya PM, Monika K (2019) Smart eye for visually impaired-an aid to help the blind people. In: International conference on computational intelligence in data science (ICCIDS)
2. Bai J, Lian S, Liu Z, Wang K, Liu D (2017) Smart guiding glasses for visually impaired people in indoor environment. *IEEE Trans Consum Electron* 63(3):258–266
3. Patil K, Kharat A, Chaudhary P, Bidgar S, Gavhane R (2021) Guidance system for visually impaired people. In: 2021 International conference on artificial intelligence and smart systems (ICAIS), Mar 2021
4. Chaccour K, Badr G (2016) Computer vision guidance system for indoor navigation of visually impaired people. In: 2016 IEEE 8th international conference on intelligent systems (IS), Sept 2016
5. Cloud Speech-to-Text. <https://cloud.google.com/speech-to-text/docs>
6. Cloud Text-to-Speech. <https://cloud.google.com/text-to-speech/docs>
7. Actions on Google. <https://developers.google.com/assistant/conversational/dfasdk/overview>
8. DialogFlow Documentation. <https://cloud.google.com/dialogflow/docs>
9. Redmon J, Divvala S, Girshick R, Farhadi A (2016) You only look once: unified, real-time object detection. In: IEEE conference on computer vision and pattern recognition. University of Washington, Allen Institute for AI, Facebook AI Research

10. Bochkovskiy A, Wang CY, Liao HYM (2020) Yolov4: Optimal speed and accuracy of object detection. arXiv preprint [arXiv:2004.10934](https://arxiv.org/abs/2004.10934)
11. Liu C, Tao Y, Liang J, Li K, Chen Y (2018) Object detection based on YOLO network. In: 2018 IEEE 4th information technology and mechatronics engineering conference (ITOEC), pp 799–803



# Automated Ultrasound Ovarian Tumour Segmentation and Classification Based on Deep Learning Techniques



K. Srilatha, F. V. Jayasudha, M. Sumathi, and P. Chitra

## 1 Introduction

Ovarian Tumour ranks fifth in women's tumour deaths, considering more deaths than other female reproductive system diseases. A woman's uncertainty of becoming an ovarian tumour through her endurance is around 1 in 78. Her existence probability of dying from the ovarian disease is about 1 in 108. Ovarian tumour and endometrial tumour are the most common kinds of gynaecological malignancy among post-menopausal women. Notwithstanding advancements in medical diagnostics, the lack of distinct signs for this sort of tumour suggests that the ovarian tumour remnant rate is low. There is, however, a lack of reliable symptomatic labels and other systems permitting early discovery, which would enhance the survival rate—ovarian tumour estimates for two hundred thirty-nine thousand anticipated new claims and one hundred fifty-two thousand deaths per year worldwide. The highest rates (11.4% one hundred thousand one and 6.0% one hundred thousand sequentially) are observed for Central and Eastern Europe. While China produces a relatively small percentage rate (4.1% one hundred thousand), the large community changes to fifty-two thousand one hundred expected new claims and twenty-two thousand five hundred associated mortality in the year of 2015.

In combination, twenty-one thousand two hundred ninety causes and fourteen thousand one hundred eighty associated fatalities are estimated to occur in the United States throughout the same period. The probability of a woman contracting an ovarian tumour is a one out of seventy-five, and her risk of dying from infection is a one out of a one hundred. The syndrome generally works in an advanced stage, during the five-year comparable rate remaining is only twenty-nine percentage. Fifteen percent are diagnosed with localized tumour stage-1, whereas the residual rate of five years

---

K. Srilatha (✉) · F. V. Jayasudha · M. Sumathi · P. Chitra  
Sathyabama Institute of Science and Technology, Chennai, Tamil Nadu 600119, India  
e-mail: [srilatha169@gmail.com](mailto:srilatha169@gmail.com)

is ninety percent. It is striking that the overall rate of comparable residues over five years generally varies between thirty percentages and forty percentages worldwide and that there were only quiet increases from two to four percentages in 1995.

Ultrasound (US) imaging is one of the several modern imaging modalities with a variety of demonstrative diligence due to the following benefits: no radioactivity, safer imaging, greater accuracy and sensitivity and specificity, and costs lower than other imaging modalities such as computed tomography (CT) or magnetic resonance imaging (MRI). Although sonography is operator-dependent, examining ultrasound (US) images necessitates the expertise of very well and experienced radiologists. Computer-aided diagnosis (CAD) has been proposed to reduce inter-observer variation among various doctors and motivate them to create more efficient and reliable diagnostic results [1].

A bilateral filter softens the ROI model, which is subsequently contrast-enhanced via histogram equalisation. The magnified image is then filtered to improve homogeneity using the base pyramid shift. With the particle swarm optimisation (PSO) technique, the RGB segmentation procedure is used to segment the filtered model [2]. These features provide U-net with a distinct utility within the medical imaging similarity. They have been seen in a lot of U-net implementations as the primary tool for segmentation tasks in medical imaging. The U-Net conclusion net is the widespread use of all original picture modalities, ranging from CT scans and MRI to X-rays and microscopes. Furthermore, while U-net is primarily a segmentation tool, it is possible that it will be used in other treatments [3].

U-net++, U-net with Res-net as the resolution (U-net with Resnet), and CE-Net all do automated segmentation. To remove radiomic-features from the segmented target volumes, a python script and the Pyradiomics package were used. The Jaccard similarity coefficient (JSC), Dice similarity coefficient (DSC), and average surface distance (ASD) were used to determine the efficacy of automated segmentation [4].

For 2-Dimension medical image segmentation, this research [5] suggests using a context encoder network (similar to CE-Net) to capture additional high-level data and interpret spatial data. A feature encoder, a context extractor, and a feature decoder are the three main components of CE Net.

This research [6, 7] proposes a novel collect layer recognition network to use the value of context data in feature extraction alone. It use channel influences and mix several scales in the application tool, unlike the correct attention tool.

Completely convolutional networks are styled and feature the period of fully convolutional networks, demonstrating its usefulness to spatially dense forecast challenges and drawing connections to prior designs. Target identification, segmentation, and classification have all benefited from deep learning, a cutting-edge computer learning system [8]. To distinguish distinct maize from temporally Lidar, this study [9, 10] suggested integrating deep learning and regional development methods. The tumefactions were classified as benign or cancerous (Ovary-Dx1 standard) by the DNN [11–13] ensemble, or as benign, indeterminate, or malignant by the DNN [11–13] ensemble (Ovary-Dx2 type). The symptomatic performance of the DNN models is similar to that of SA for analysing ovarian tumours in the study set in terms of sensitivity and specificity. To improve Contrast Adaptation Ultrasound, tumour

models, noise previously indicated as Adaptive Anisotropic Diffusion Filter, Speckle Noise, and Contrast Limited Adaptive Histogram Equalization were employed in the pre-processing step in [14]. The Ovarian tumour was segmented using the FL-SNN Model and GrabCut algorithms in the segmentation step. Xing et al. [15] speckle-noise filtering of PolSAR models is a significant pretreatment. Non-local (NL) intends to calculate the measurements among two pixels with related neighbouring areas, preferably two different pixels. Recognising that scraps include structural data, the NL base filter processes repeated structures and works better than different filters.

Marques et al. [16] the U-design Net examines several ovarian model parameters, including ovarian follicle partitioning. On Beam-formed-radio-frequency (BRF) data, the results of various post-processing approaches are also examined.

Olofsson et al. [17] The BF parameters were optimised using an extended grasshopper optimisation algorithm. The Rician noises and impulse are used to recreate the medical MRI images (with a few changes). To get the best filter parameters, the extended grasshopper optimization algorithm is applied to the noisy image to assess window size, and spatial and intensity regions. For optimisation, the PSNR has been used as a fitness rate. The following sections make up this paper: Sect. 2 presents the proposed ultrasound ovarian tumour segmentation and classifications. Section 3 describes the outcome and discussion. The conclusion part is shown in Sect. 4.

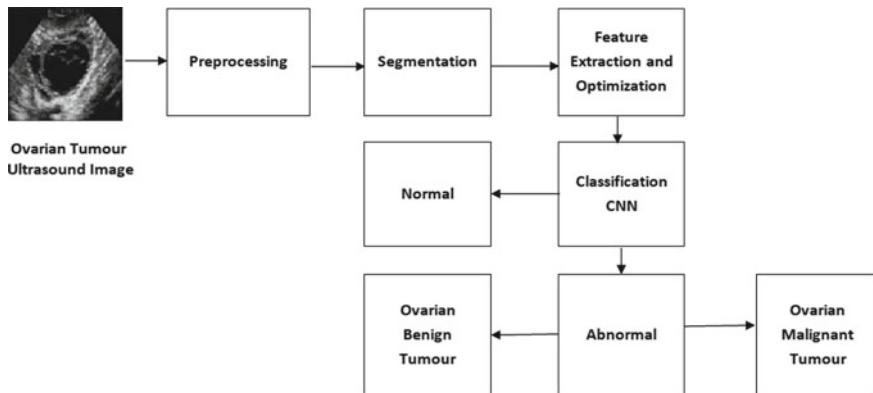
## 2 Materials and Methods

### 2.1 Database

In the present task, 50 US ovarian image processing denoising filtering techniques are utilised to prepare Entirety 150 Ovarian tumour ultrasound images as 80 benign and 70 malignant images obtained from a standard benchmark data set of US ovarian images available at [18, 19].

This proposed method can recognise and segment tumour areas automatically in US ovarian tumour images, which is of great benefit to dealing with several limitations of existing US ovarian tumour image segmentation methods (computational time, human intervention, contract problem, disappointing outcomes).

Figure 1 shows our model's four major steps, which incorporate many machine vision techniques.



**Fig. 1** Block diagram of proposed system of ultrasound ovarian tumour segmentation and classification

## 2.2 Pre-processing Module

This moment around due to the comparatively poor quality of clinical ovarian (US) images, a great segmentation process is performed with the use of a priori knowledge to improve the segmentation output. Furthermore, it is difficult to quantify the segmentation impact without any priori information; hence, constructing quantitative function(s) with no priori information is challenging. Consequently, it employs the priori information practised in [20–22], specifically, requiring the operator to irregularly select a comparatively inadequate four-sided ROI from the US Ovarian image. In this way, interventions from another separate area can be decreased as much as practicable, presenting the segmentation more comfortable and more effective. Also, it provides users a priori information for the objective function(s).

**Bilateral Filtering:** Because of different interventions in US Ovarian tumour images, speckle modification is needed to enhance the quality of ultrasound images. In the MOORGB, a bidirectional filter [1] is utilised, which has been proved to be an effective and efficient solution for speckle modification.

**Median Filter:** This filter is estimated as a spatial non-linear filter. Concerning decreasing speckle or pulse noise, that filter operates by determining the median rate of its acquaintances in the window can be performed based on the methods here Eq. 1:

$$f^{\wedge}(p, q) = \text{med}_{(m,n) \in S_{p,q}(p,q)} \{a(m, n)\} \quad (1)$$

where  $(m, n)$  gives the initial image,  $S_{p,q}$  denotes the set of coordinates in square image windows. This filter is able to eliminate impulse or short noise, but it is not quite satisfactory for speckle noise.

Boxcar filter is a single averaging filter that follows the pixel in the square centre by the mean amount of kernel pixels. This filter has an excellent achievement in overcoming the speckle noise in a similar area. Mathematically expressed as:

$$k(i, j) = \langle l(i, j_N) \rangle = \left( \frac{1}{N^2} \right) \sum_{x=-\frac{N}{2}}^{\frac{N}{2}} \sum_{y=-\frac{N}{2}}^{\frac{N}{2}} L_{i+x, j+y} \quad (2)$$

### 2.3 Segmentation Module

RGB Segmentation Method: The RGB approach [1] seeks to merge pixels of similar intensity into a minimum spanning tree (MST), which represents a pixel, based on an image first considered as a graph (i.e., region in the ovarian US image). As a result, the image is divided into several sub-regions. Obviously, the final segmentation results are determined by the stage of merging pixels into an MST. A pairwise comparison predicate was proposed using the RGB to assess whether a boundary between two subgraphs should be removed.  $A(B_1, B_2)$  contrasts within-subgraph differences with inter-subgraph differences, given the graph  $G = (V, E)$ .

$$A(B_1, B_2) = \begin{cases} \text{false, if } A(B_1, B_2) > \text{MInt}(B_1, B_2) \\ \text{true, other} \end{cases} \quad (3)$$

$$A(B_1, B_2) = |\mu(B_1) - \mu(B_2)| \quad (4)$$

$$\text{MInt}(B_1, B_2) = \min(\sigma(B_1) + \tau(B_1), \sigma(B_2) + \tau(B_2)) \quad (5)$$

$$\tau(B) = \frac{k}{|B|} \cdot \left( 1 + \frac{1}{\alpha \cdot \beta} \right), \beta = \frac{\mu(B)}{\sigma(B)} \quad (6)$$

where  $A(B_1, B_2)$  is the dissimilarity between two subgraphs,  $B_1$  and  $B_2 \in V$ ,  $\text{MInt}(B_1, B_2)$  signify the smallest number of internal dissimilarity of  $B_1$  and  $B_2$ ,  $\mu(B)$  indicates the typical intensity of  $B$ ,  $\sigma(B)$  is standard deviation (SD) of  $B$ , and  $\tau(B)$  is a threshold task of  $B$  at the same time as  $\alpha$  and  $k$  are positive limitations. While  $k$  raises,  $\tau$  raises as well as the area combine more without difficulty. In contrast, while  $\alpha$  raises,  $\tau$  diminishes and consequently the regions are amalgamated less with no trouble.

FGMAC: The snake technique, also known as the active contour method, is an area-based segmentation algorithm that recognises objects by minimising the energy differential. By employing image gradients, it improves on edge-based segmentation. Srilatha and Ulagamuthalvi [22] refined the Active-Contour approach by employing a

Global-Minimum for Active-Contour method to speed up curve change. Several well-known image models, including the snake model, Rudin-Osher-denoising Fatemi's approach, and Mumford-segmentation Shah's algorithm were proposed to be integrated. It minimises global Active Contour by combining traditional Geometric Active Contour dependent on edge detection with Active Contour No Edge based upon homogeneous areas. The following are the primary minimum processes:

$$\min_{x,y} \{E_3^r(x, y, u, v, \lambda, \alpha, \theta) = WU_g(x) + \frac{1}{2\theta} \|x - y\|_{L^2}^2 + \int_{\Omega} \lambda r_1(u, d_1, d_2)v + \alpha y(y)dm \quad (7)$$

$$v = \min\{\max\{x(u) - \theta\lambda r_1(u, d_1, d_2), 0\}, 1\} \quad (8)$$

## 2.4 Feature Extraction and ACO-PCO

Standard deviation (SD), Root Mean Square (RMS) Energy, Mean, Mode, Entropy, range, median, mode, variance, mean absolute deviation and kurtosis are all first-order statistics of an intensity distribution. The skewness and homogeneity of the data were used to calculate its intensity. According to [14–20], the grey-level co-occurrence matrix was used to describe the joint probability distribution of images. However, the sum variance was not included in the 25 GLCM. To set the GLCM, we applied the following principles: (1) The image input's dynamic range was equal to the number of greyscale; (2) the ranges were 1, 2, 3, and 4; and (3) the directions were  $0^\circ$ ,  $45^\circ$ ,  $90^\circ$ , and  $135^\circ$ . At each distance, the GLCM of texture features were averaged over four directions. A contourlet transform (CT) is a feature that is built using the Laplacian scale space and a directed filter banks. As a result, contourlet transformation can be used to extract spatial, temporal, and directional data. In our exercise, each image was separated into three stages, with each stage divided into four sub-bands. The aforementioned features were extracted by averaging the intensity and texture features of each reconstructed directional sub-band over four sub-bands. Shape features such as perimeter, short-axis length, area, orientation, long-axis length, diameters, convex area, mean, solidity, median, mean, max thickness, median, max width, and max width to max thickness ratio were extracted based on the segmentation results of BMUS images. TIC and perfusion patterns were studied in addition to quantitative features. The  $z$  scores were used to normalise the feature scores.

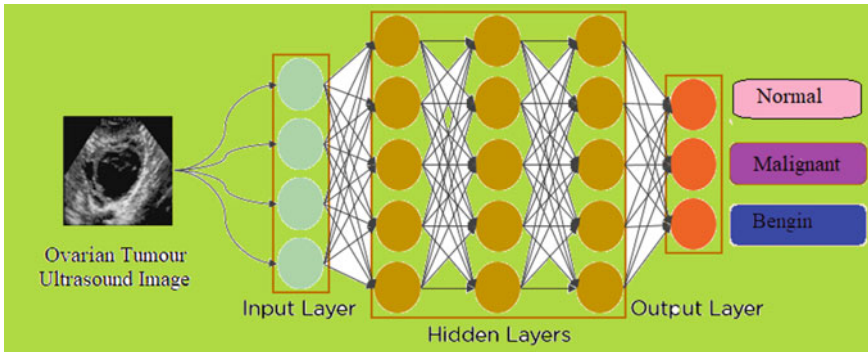
## 2.5 Classification

CNN-Deep Learning (DL) is a method for constructing multilayer models designed to solve classification, definition, prediction, and other problems. A DL model is an architecture for processing sound or images via a multilayered artificial neural network structure [23]. Choosing deep learning models means relying on them to discover features on raw data without expert guidance. In deep learning, the basic architecture for image categorisation is the Conventional Neural Network architecture. CNN made a major impression in the Image-Net-Large-Scale-Visual-Recognition-Challenge (ILSVRC) [23] for the first time in 2012. In the years that followed, ZfNet, VggNet, Google Net, and Microsoft ResNet were developed to classify images. As a result of the success of Convolutional Neural Networks in classification, Region-Based Convolutional Neural Networks (R-CNN), Yolo, Fast-RCNN, Faster-Region-Based Convolutional Neural Networks (F-R-CNN), and Mask Region-Based Convolutional Neural Networks (R-CNN) have been developed for tumour recognition [24]. A CNN's general architecture is depicted in Fig. 3. A construction for learning features and classifying them can be divided into two parts. The network is given an input image of a specific size at the start. The input layer can be supplied either an RGB three-channel image or a single-channel image. The convolution layer, which comes after the input layer, is the following phase. In feature discovery, the convolution layer is frequently referred to as the most significant layer. This layer discovers characteristics by applying the provided filters. Two-by-two, three-by-three, five-by-five, or seven-by-seven are the most common filters. As a result of the stride being applied, each of the filters circulating on the input image yields an output image. As a result, convolution is applied prior to the application of a convolution layer, and Ovarian US images are produced up to the amount of filters. The images that make up these maps are known as features maps. The next step is to classify the follicles according to their location.

The Convolutional Neural Networks Class is shown in Fig. 4 for the classification of ovarian cancers. Table 3 lists the layers and attributes of the Convolutional Neural Networks Class Section contain information on the Convolutional Neural Networks Class's training. U-Net and a CNN commonly used in biomedicine for segmentation tasks, are utilised to train the CNN (Fig. 2).

## 3 Results and Discussions

The proposed system results and discussion section show various experimental outcomes on real-time images of US ovarian tumours. It is estimated that there are 150 anomalous US ovarian tumour images from benign and malignant tumours included in the input data. Irregular US ovarian images are produced by ovarian images affected by injury to the ovary. The first step of reducing noise in this method is to combine the results of bilateral filtering and median filtering in Table 1. Following this, the



**Fig. 2** Convolutional Neural Networks (CNN) architecture

segmentation and classification steps were carried out by using RGB Segmentation and FGMAC, as shown in Fig. 3. According to FGMAC, the US ovarian images are divided into two categories: abnormality and normality.

Accordingly, the number of clusters in feature gap denotes the figure of module in the US ovarian image. The third phase involves extraction of texture features using GLCM and ACO-PCO. The homogeneity, energy, correlation, and contrast features are extracted in the proposed work. As a final step, CNN is used to classify the type of tumours in the US images shown in Fig. 3. Performance evaluation of the proposed method is based on accuracy, precision, sensitivity, specificity, F-score in Fig. 4 and Table 2 and lastly, performance analysis of accuracy and loss of training and validation of proposed method shown in Fig. 5.

## 4 Conclusion

The proposed system is a computerised process for segmenting and classifying different kinds of tumours using ACO-PCO and GLCM texture features extracted from ultrasound ovarian images. The system will be tested in relation to ultrasound images of the ovary which are caused by tumours in the ovary. According to the US Ovarian image, the ovarian area adjacent to a lesion was precisely broken up. In the proposed method, the general practitioner can determine an ovarian tumour's type, which helps further treatment. A comparison of the results obtained using the proposed method and the ground truth yields Precision, Sensitivity, Specificity, and F-Score values of 98.89%, 98.75%, 97.06%, 97.08%, and 98.11%, respectively. This classification was performed by Convolutional Neural Network (CNN) classifier, which was 98.11% accurate. Using only a small amount of variation, the proposed system can be extended to classifying other types of tumours as well.



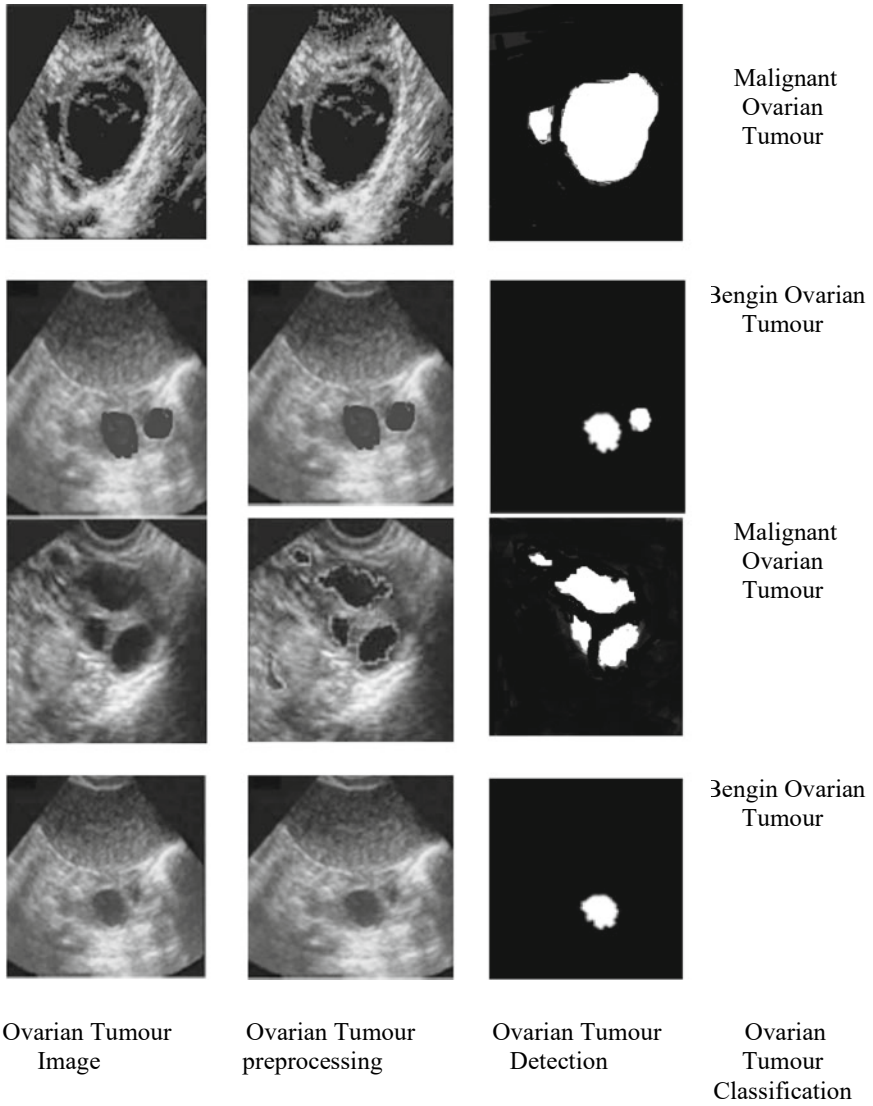


Fig. 3 Results of US ovarian tumour segmentation and classification

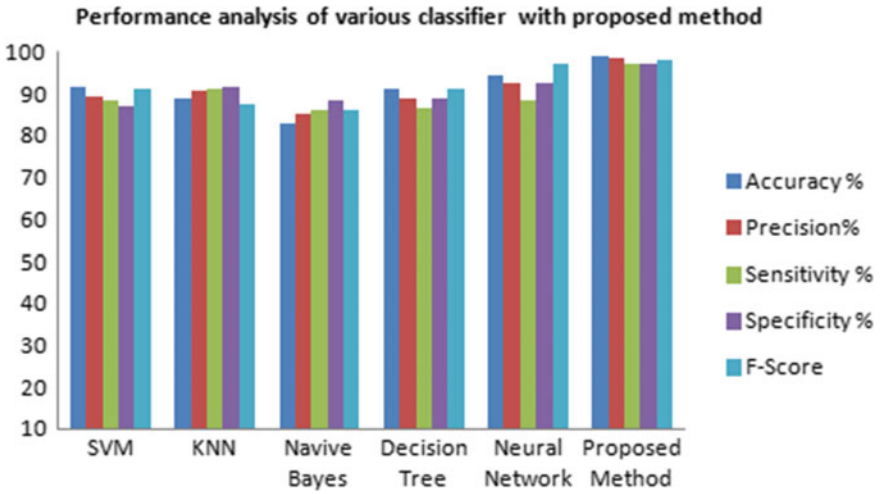


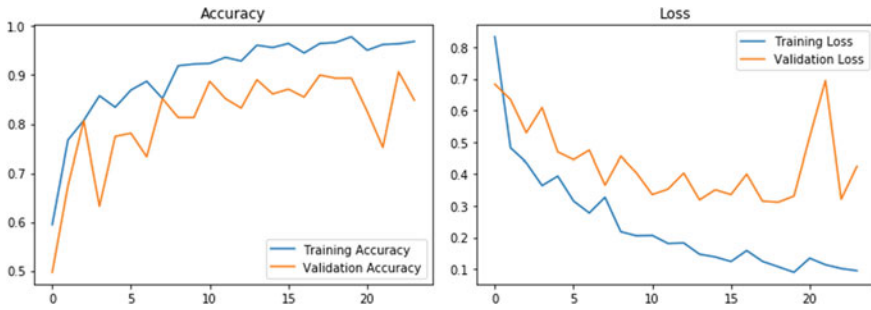
Fig. 4 Performance analysis of various classifiers with proposed method

Table 1 Preprocessing performance metrics

Performance metrics	Median filter	Boxcar
MSE	$2.018 \times 10^{-3}$	$5.1232 \times 10^{-4}$
SNR	38.46	51.13
PSNR	41.79	55.73

Table 2 Performance analysis of various classifier with proposed method

Classifier	Accuracy %	Precision%	Sensitivity %	Specificity %	F-score
SVM	91.75	89.36	88.37	87.18	91.37
KNN	88.96	90.76	91.29	91.47	87.59
Navive Bayes	82.96	85.09	85.98	88.36	86.34
Decision tree	90.98	89.06	86.46	88.86	91.28
Neural network	94.31	92.36	88.39	92.62	97.27
Proposed method	98.89	98.75	97.06	97.08	98.11



**Fig. 5** Performance analysis of accuracy and loss of training and validation of proposed method

**Table 3** The properties of the layers of the CNN used for ovarian tumour segmentation and classification

Layers	Input size	Kernel size	Stride length	Output size
Conv 1	128*128*1	7*7*36	1	122*122*36
Pool 1	122*122*36	2*2*36	2	61*61*36
Conv 2	61*61*38	5*5*36	1	56*56*36
Pool 2	56*56*36	2*2*64	2	28*28*64
Conv 3	28*28*64	3*3*64	1	24*24*64
Pool 3	24*24*64	3*3*64	3	8*8*64
FC	4096*1	–	–	1024*1
Softmax	1024*1	–	–	4*1

## References

1. Luo Y, Liu L, Huang Q, Li X (2017) A novel segmentation approach combining region- and edge-based information for ultrasound images. *Biomed Res Int* 27:2017
2. İnik Ö, Ceyhan A, Balcıoğlu E, Ülker E (2019) A new method for automatic counting of ovarian follicles on whole slide histological images based on convolutional neural network. *Comput Biol Med* 112:103350
3. Siddique N, Sidike P, Elkin C, Devabhaktuni V (2020) U-Net and its variants for medical image segmentation: theory and applications. *arXiv preprint arXiv:2011.01118*, 2 November 2020
4. Jin J, Zhu H, Zhang J, Ai Y, Zhang J, Teng Y, Xie C, Jin X (2021) Multiple U-net-based automatic segmentations and radiomics feature stability on ultrasound images for patients with ovarian cancer. *Front Oncol* 18(10):3428
5. Gu Z, Cheng J, Fu H, Zhou K, Hao H, Zhao Y, Zhang T, Gao S, Liu J (2019) Ce-net: context encoder network for 2D medical image segmentation. *IEEE Trans Med Imaging* 38(10):2281–2292
6. Jin B, Liu P, Wang P, Shi L, Zhao J (2020) Optic disc segmentation using attention-based U-net and the improved cross-entropy convolutional neural network. *Entropy* 22(8):844
7. Srilatha K, Ulagamuthalvi V (2019) A comparative study on tumour classification. *Res J Pharm Technol* 12(1):407–411
8. Long J, Shelhamer E, Darrell T (2015) Fully convolutional networks for semantic segmentation. In: *Proceedings of the IEEE conference on computer vision and pattern recognition 2015*, pp 3431–3440

9. Jin S, Su Y, Gao S, Wu F, Hu T, Liu J, Li W, Wang D, Chen S, Jiang Y, Pang S (2018) Deep learning: individual maize segmentation from terrestrial Lidar data using faster R-CNN and regional growth algorithms. *Front Plant Sci* 22(9):866
10. Xu Z, Wu Z, Feng J (2018) CFUN: combining faster R-CNN and U-net network for efficient whole heart segmentation. arXiv preprint [arXiv:1812.04914](https://arxiv.org/abs/1812.04914), 12 December 2018
11. Christiansen F, Epstein EL, Smedberg E, Åkerlund M, Smith K, Epstein E (2021) Ultrasound image analysis using deep neural networks for discriminating between benign and malignant ovarian tumors: comparison with expert subjective assessment. *Ultrasound Obstet Gynecol* 57(1):155–163
12. Guo P, Xue Z, Long LR, Antani S (2020) Cross-dataset evaluation of deep learning networks for uterine cervix segmentation. *Diagnostics* 10(1):44
13. Sergio Rodrigues P, Wachs-Lopes G, Morello Santos R, Coltri E, Antonio GG (2019) A q-extension of sigmoid functions and the application for enhancement of ultrasound images. *Entropy* 21(4):430
14. Srilatha K, Ulagamuthalvi V (2021) Performance analysis of ultrasound ovarian tumour segmentation using GrabCut and FL-SNNM. In: 2021 international conference on advances in electrical, computing, communication and sustainable technologies (ICAECT), 19 February 2021. IEEE, pp 1–7
15. Xing X, Chen Q, Xu Q, Yang S, Liu X (2017) Nonlocal means filtering for polarimetric SAR images based on heterogeneity. *J Rem Sens* 21(3):434–441
16. Marques S, Carvalho C, Peixoto C, Pignatelli D, Beires J, Silva J, Campilho A (2019) Segmentation of gynaecological ultrasound images using different U-net based approaches. In: 2019 IEEE international ultrasonics symposium (IUS), 6 October 2019. IEEE, pp 1485–1488
17. Olofsson K, Carannante V, Takai M, Önfelt B, Wiklund M (2021) Ultrasound-based scaffold-free core-shell multicellular tumor spheroid formation. *Micromachines* 12(3):329
18. <https://www.cancer.net/cancer-types>
19. [https://hms.harvard.edu/search-results?as\\_q=harvard%20ovarian%20cancer](https://hms.harvard.edu/search-results?as_q=harvard%20ovarian%20cancer)
20. Anoop V, Bipin PR (2019) Medical image enhancement by a bilateral filter using optimization technique. *J Med Syst* 43(8):1–2
21. Sahani A, Srilatha K (2014) Image forgery detection using SVM classifier. *Int J Adv Res Electr Electron Instrum Eng* 3(3)
22. Srilatha K, Ulagamuthalvi V (2019) Support vector machine and particle swarm optimization based classification of ovarian tumour. *Biosci Biotech Res Comm* 12(3):714–719
23. Alsinan AZ, Patel VM, Hacihaliloglu I (2019) Automatic segmentation of bone surfaces from ultrasound using a filter-layer-guided CNN. *Int J Comput Assist Radiol Surg* 14(5):775–783
24. Ma J, Wu F, Jiang TA, Zhao Q, Kong D (2017) Ultrasound image-based thyroid nodule automatic segmentation using convolutional neural networks. *Int J Comput Assist Radiol Surg* 12(11):1895–1910
25. Reid BM, Permuth JB, Sellers TA (2017) Epidemiology of ovarian cancer: a review. *Cancer Biol Med* 14(1):9

# Magnitude Comparison in Canonical Signed-Digit Number System



Madhu Sudan Chakraborty

## 1 Introduction

The basis of any number system (NS) is constituted by some well-defined digit-set (DS). Signed-digit number system (SDNS) is a NS which is unconventional by nature in the sense that every digit of its DS necessarily carries independent sign [1]. Although it was initially introduced to support constant-time addition/subtraction [2] to speed up more complex arithmetic operations (AOs) like multiplication [1], later some other promising features of SDNSs have been uncovered, including regularity in circuit design [1], fault tolerance [3] and energy efficiency [4, 5]. The simplest form of SDNS where the DS is fixed to  $\{\bar{1}, 0, 1\}$ , is called binary signed-digit number system (BSDNS) and owing to its relative closeness to the binary number system, BSDNS has become the most frequently investigated category of SDNSs [1, 6]. In order to implement an arithmetic unit using any SDNS, the unconventional digits of the DS of the SDNS need to be encoded into different binary strings. For BSDNS, two popular encodings are the positive–negative encoding (PNE) and two’s-complement encoding (TCE) [1]. In PNE  $\bar{1}, 0, 1$  are expressed as 01, 00 and 10, respectively. In TCE  $\bar{1}, 0, 1$  are expressed as 11, 00 and 01, respectively. SDNSs seem to be promising to process digital signals, images and cryptographic algorithms. In particular, a subclass of BSDNS, called canonical signed-digit number system (CSDNS) [1] has got a lot of attention in the recent years. In CSDNS, any value can be uniquely represented with least number of non-zero digits in such a way that non-zero digits always remain non-adjacent. CSDNS strives to reduce the required computations thanks to the larger number of zero entries in inputs where no processing is apparently required. Obviously compared to the other classes of SDNSs, for performing the same AOs, CSDNS may offer higher speed, involving smaller chip area and less energy [1, 6].

---

M. S. Chakraborty (✉)

Computer Science Department, Indas Mahavidyalaya, Bankura, WB 722205, India  
e-mail: [madhusudhanchakraborty27@gmail.com](mailto:madhusudhanchakraborty27@gmail.com)

For some AOs even it has been claimed that involvement of CSDNS might cause the chip area and powder consumption to be most economic [7].

The interesting areas of applications of canonical signed-digits (CSDs) or CSDNS include multiplier design [8], FIR filter design [9], computing trigonometric and hyperbolic functions [10], fault testing in IC [11] and image processing [12, 13]. Obviously CSDNS may be worthy for further investigations in wider context. Magnitude comparator (MCR) is an integral part of every general-purpose arithmetic processor and in this paper, CSDNS is considered for MCR design. The remaining portion of this paper is formatted with four sections. In the Problem Statement section, the problem and its backgrounds are uncovered. In the Proposed Algorithm section, a novel scheme is strived to develop by viewing magnitude comparison (MC) in CSDNS on the basis of comparative magnitudes and signs of some counterpart portions of inputs. In Results and Discussion section, the immediate outcomes as well as the future scopes of the proposed algorithm are discussed. Finally, in Conclusion section the proposed study ends, reminiscing its major contributions.

## 2 The Problem Statement

Although MC is a complex AO in general, ability to support MC with efficacy and efficiency is prerequisite for any standard arithmetic processor. In the recent years, several reports have been published on designing better-performing MCR using the conventional radix-complement NS [14, 15]. However, for MC in SDNS-platform, no significant progress has been attained yet, excluding the proposal presented in [16]. The algorithm [16] may still run in online mode only and for simplicity in addition, the algorithm [16] solely relies on higher-radix SDNDS instead of BSDNS as the latter involves more complex additive rules. It appears that the algorithm [16] becomes inapplicable in the context of designing traditional MCR over the DS  $\{\bar{1}, 0, 1\}$ .

Another feature of the problem is reflected by the inability to extend the basic arithmetic model behind the ordinary MC beyond the conventional NS. For MC of two ordinary unsigned numbers,  $A = A_{n-1}A_{n-2}\dots A_1A_0$  and  $B = B_{n-1}B_{n-2}\dots B_1B_0$ , an algorithm has been presented in the standard textbooks as follows:

### Algorithm 1

1. Set:  $i = n - 1$
2. If  $(A_i > B_i)$ 
  - a. Output:  $A > B$
  - b. Stop
3. If  $(A_i < B_i)$ 
  - a. Output:  $A < B$
  - b. Stop

4. If  $(A_i = B_i)$ 
  - a. If  $(i = 0)$ 
    - i. Output:  $A = B$
    - ii. Stop
  - b. Otherwise
    - i. Set:  $i = i - 1$
    - ii. Go to Step 2

Algorithm 1 cannot be directly and unconditionally scaled up for the DS  $\{\bar{1}, 0, 1\}$  as in that case  $|A_{n-1}| > |B_{n-1}|$  (or  $|A_{n-1}| < |B_{n-1}|$ ) does not necessarily imply  $|A| > |B|$  (or  $|A| < |B|$ ). For example, for two 4-digit binary signed-digit numbers (BSDNs),  $A = \bar{1}011$ ,  $B = 0111$ ,  $|A_3| > |B_3|$  holds true but  $|A| > |B|$  does not hold correct yet. Even  $(|A_{n-1}| = |B_{n-1}|)$  and  $(|A_{n-2}| > |B_{n-2}|$  (or  $|A_{n-2}| < |B_{n-2}|$ )) does not guarantee  $|A| > |B|$  (or  $|A| < |B|$ ). As an example, two 4-digit BSDNs,  $A = \bar{1}011$ ,  $B = 1\bar{1}\bar{1}\bar{1}$  may be considered to give an insight into the scenario just mentioned. Clearly further studies are required in this regard with newer approaches, including investigating more restricted input patterns.

### 3 The Proposed Algorithm

In this paper, an extension and mutation of Algorithm 1 is considered for possible adaptation by the CSDNS. The developmental procedure will be described next.

#### 3.1 Developing an Algorithm

For two n-digit canonical signed-digit numbers (CSDNs),  $A$  and  $B$ , where  $A = A_{n-1}A_{n-2}\dots A_1A_0$  and  $B = B_{n-1}B_{n-2}\dots B_1B_0$ , one pivotal part of the proposed algorithm is given by Theorem 1.

**Theorem 1** *If  $|A_{n-1}| > |B_{n-1}|$  then  $|A| > |B|$*

**Proof** Let  $MC(X, Y)$  denotes the outcome of MC of two equi-digit CSDNs,  $X$  and  $Y$  and  $\widehat{P}_i$  denotes  $P_{i-1}\dots P_1P_0 \forall$  n-digit CSDNs  $P = P_{n-1}\dots P_1P_0$ . If possible, assume that  $\exists$  an instance  $(X, Y)$  of  $(A, B)$  such that  $|X_{n-1}| > |Y_{n-1}|$  holds true and still  $|X| \leq |Y|$  occurs. Generate another instance of  $(A, B)$  as  $(U, V) = (U_{n-1} \widehat{U}_{n-1}, V_{n-1} \widehat{V}_{n-1})$  where  $|U_{n-1}| = |X_{n-1}|$ ,  $|V_{n-1}| = |Y_{n-1}|$ , both  $|\widehat{U}_{n-1}|$  and  $|\widehat{V}_{n-1}|$  are maximum, the signs of  $U_{n-1}$  and  $\widehat{U}_{n-1}$  mismatch and signs of  $V_{n-1}$  and  $\widehat{V}_{n-1}$  do not mismatch. So  $U = |U_{n-1}|2^{n-1} - |\widehat{U}_{n-1}|$  and  $V = |V_{n-1}|2^{n-1} + |\widehat{V}_{n-1}|$ . Clearly  $|U| \leq |X|$  and  $|V| \geq |Y|$  and so  $|U| \leq |X| \leq |Y| \leq |V|$  obviously holds true. As the number of input digits required for every standard arithmetic processor is typically a power of 2 (say 32-bit

or 64-bit), for CSDNS the maximum input values appear in the form 0101...0.01 and 1010...0.10, if the input starts with 0 and 1, respectively. It means, given that  $m$  is odd, for any  $m$ -digit CSD input starting with 0 and 1 (or  $\bar{1}$ ), the maximum values which can be represented are  $\frac{2}{3}(2^{m-1} - 1)$  and  $(\frac{2^{m+1}-1}{3})$ , respectively. Obviously here  $|U_{n-1}| = 1$ ,  $|V_{n-1}| = 0$ ,  $|\widehat{U}_{n-1}| = \frac{2}{3}(2^{n-2} - 1)$  and  $|\widehat{V}_{n-1}| = (\frac{2^n-1}{3})$ . Then  $|U| = 2^{n-1} - \frac{2}{3}(2^{n-2} - 1)$ , implying  $|U| - |V| = 1$  and accordingly  $|U| > |V|$ . It means  $\nexists (U, V)$  or  $(X, Y)$  instance of  $(A, B)$  such that both  $(|A_{n-1}| > |B_{n-1}|)$  and  $(|A| \leq |B|)$  hold true simultaneously. In other words,  $\forall$  CSDNs  $A, B$  if  $|A_{n-1}| > |B_{n-1}|$  then  $|A| > |B|$  and if  $|A_{n-1}| < |B_{n-1}|$  then  $|A| < |B|$ .

However, when  $(|A_{n-1}| = |B_{n-1}|)$  MC  $(A, B)$  needs to be determined on the basis of MC  $(\widehat{A}_{n-1}, \widehat{B}_{n-1})$  as well as the matching/mismatching of the signs of  $A_{n-1}$  and  $\widehat{A}_{n-1}$  as well as signs of  $B_{n-1}$  and  $\widehat{B}_{n-1}$ . This is another pivotal part of the proposed algorithm and can be resolved using two functions,  $h()$  and  $h'()$ , which are defined on  $A$  and  $B$ , respectively, as follows:  $h(A_{n-1}, \widehat{A}_{n-1}) = 1$ , if the signs of  $A_{n-1}$  and  $\widehat{A}_{n-1}$  match and 0, otherwise. Also  $h'(B_{n-1}, \widehat{B}_{n-1}) = 1$ , if the signs of  $B_{n-1}$  and  $\widehat{B}_{n-1}$  match and 0, otherwise. For computing both functions, zero is supposed to match to positive as well as negative signs. Thusly, it may be possible to generalize and interpret the computations for MC  $(\widehat{A}_{i+1}, \widehat{B}_{i+1})$  in terms of MC  $(A_i, B_i)$ , MC  $(\widehat{A}_i, \widehat{B}_i)$ ,  $h(A_i, \widehat{A}_i)$  and  $h'(B_i, \widehat{B}_i)$ . As the truth tables for switching circuits are basically the decision tables [17], in the present context, all possible scenarios may be presented concisely in a decision table, Table 1, as follows:

### 3.2 Realizing Arithmetic Equations for the Proposed Algorithm

Let the signs of  $\widehat{A}_i$  and  $\widehat{B}_i$  are represented by  $S_i$  and  $S'_i$ , respectively. Also let  $h(A_i, \widehat{A}_i)$  and  $h'(B_i, \widehat{B}_i)$  are abbreviated by  $h_i$  and  $h'_i$ , respectively. For realizing the algorithm presented in this paper in terms of arithmetic equations (AEs), the CSDN-inputs and signs are represented by PNE and TCE, respectively, where every stakeholder binary signed-digit variable  $X_i$  is realized in the form of a binary-pair as:  $(X_{i,1}, X_{i,0})$ . For comparator's output recording (COR), the occurrences of  $|\widehat{A}_i| > |\widehat{B}_i|$ ,  $|\widehat{A}_i| = |\widehat{B}_i|$  and  $|\widehat{A}_i| < |\widehat{B}_i|$  are encoded as 01, 11 and 10, respectively. In addition  $M_i = (M_{i,1}, M_{i,0})$  is supposed to denote the outcome of MC  $(\widehat{A}_{i+1}, \widehat{B}_{i+1})$  in COR representation. Then the AEs associated to the proposed algorithm, Algorithm 2, may be derived as follows:

#### Algorithm 2

Step 1:



**Table 1** MC ( $\widehat{A}_{i+1}, \widehat{B}_{i+1}$ ) in terms of MC ( $A_i, B_i$ ), MC ( $\widehat{A}_i, \widehat{B}_i$ ),  $h(A_i, \widehat{A}_i)$  and  $h'(B_i, \widehat{B}_i)$

		Rules									
		1	2	3	4	5	6	7	8	9	10
Conditions	$ A_i  >  B_i $	Y	-	-	-	-	-	-	-	-	-
	$ A_i  =  B_i $	-	-	Y	Y	Y	Y	Y	Y	Y	Y
	$ A_i  <  B_i $	-	Y	-	-	-	-	-	-	-	-
	$ \widehat{A}_i  >  \widehat{B}_i $	-	-	Y	Y	-	-	-	-	-	-
	$ \widehat{A}_i  =  \widehat{B}_i $	-	-	-	-	Y	Y	Y	Y	-	-
	$ \widehat{A}_i  <  \widehat{B}_i $	-	-	-	-	-	-	-	-	Y	Y
	$h(A_i, \widehat{A}_i) = 1$	-	-	Y	-	Y	-	Y	-	-	-
	$h(A_i, \widehat{A}_i) = 0$	-	-	-	Y	-	Y	-	Y	-	-
	$h'(B_i, \widehat{B}_i) = 1$	-	-	-	-	Y	-	-	Y	Y	-
	$h'(B_i, \widehat{B}_i) = 0$	-	-	-	-	-	Y	Y	-	-	Y
Actions	$ \widehat{A}_{i+1}  >  \widehat{B}_{i+1} $	×	-	×	-	-	-	×	-	-	×
	$ \widehat{A}_{i+1}  =  \widehat{B}_{i+1} $	-	-	-	-	×	×	-	-	-	-
	$ \widehat{A}_{i+1}  <  \widehat{B}_{i+1} $	-	×	-	×	-	-	-	×	×	-

At first, the magnitude of counterpart digits of  $A$  and  $B$  are compared and  $\forall \in [0, n - 1]$  the outcomes ( $A'_i, B'_i$ ) are shown in Table 2, yielding (1).

$$A'_i = B_{i,1} + B_{i,0} + \overline{A_{i,1}} \cdot \overline{A_{i,0}}, \quad B'_i = A_{i,1} + A_{i,0} + \overline{B_{i,1}} \cdot \overline{B_{i,0}} \quad (1)$$

**Table 2** MC of counterpart digits of  $A$  and  $B$

Inputs ( $A_i, B_i$ )				Outputs ( $A'_i, B'_i$ )	
$A_{i,1}$	$A_{i,0}$	$B_{i,1}$	$B_{i,0}$	$A'_i$	$B'_i$
0	1	0	1	1	1
0	1	0	0	0	1
0	1	1	0	1	1
0	0	0	1	1	0
0	0	0	0	1	1
0	0	1	0	1	0
1	0	0	1	1	1
1	0	0	0	0	1
1	0	1	0	1	1

**Table 3** Representing the sign of  $A_0$  in TCE

Inputs ( $A_0$ )		Outputs ( $S_1$ )	
$A_{0,1}$	$A_{0,0}$	$S_{1,1}$	$S_{1,0}$
0	1	1	1
0	0	0	0
1	0	0	1

For initialization of  $M_i$  as  $M_0$ :

$$M_{0,1} = A'_0, \quad M_{0,0} = B'_0 \tag{2}$$

Step 2:

$S_1$  is computed as per Table 3 and  $S'_1$  is computed likewise, yielding (3).

$$\begin{aligned} S_{1,1} &= A_{0,0}, \quad S_{1,0} = A_{0,1} + A_{0,0} \\ S'_{1,1} &= B_{0,0}, \quad S'_{1,0} = B_{0,1} + B_{0,0} \end{aligned} \tag{3}$$

Step 3:

$\forall i \in [1, n - 1]$  do as follows:

Sub-step 3.1:

The computations for  $h_i$  and  $S_{i+1}$  are performed as per Table 4. Also  $h'_i$  and  $S'_{i+1}$  can be computed likewise, yielding (4).

$$h_i = \overline{A_{i,1}} \cdot S_{i,1} + A_{i,1} \cdot \overline{S_{i,1}} + \overline{A_{i,0}} \cdot \overline{S_{i,1}} + A_{i,0} \cdot \overline{S_{i,0}}$$

**Table 4** Computing  $h_i$  and  $S_{i+1}$

Inputs ( $A_i, S_i$ )				Outputs ( $h_i, S_{i+1}$ )		
$A_i$		$S_i$		$h_i$	$S_{i+1}$	
$A_{i,1}$	$A_{i,0}$	$S_{i,1}$	$S_{i,0}$		$S_{i+1,1}$	$S_{i+1,0}$
0	1	1	1	1	1	1
0	1	0	0	1	1	1
0	1	0	1	0	1	1
0	0	1	1	1	1	1
0	0	0	0	1	0	0
0	0	0	1	1	0	1
1	0	1	1	0	0	1
1	0	0	0	1	0	1
1	0	0	1	1	0	1

**Table 5** MC at  $i$ th stage

Inputs ( $A'_i, B'_i, M_{i-1,1}, M_{i-1,0}$ )				Outputs ( $M_{i,1}, M_{i,0}$ )							
$A'_i$	$B'_i$	$M_{i-1,1}$	$M_{i-1,0}$	When $h_i = 0, h'_i = 0$		When $h_i = 0, h'_i = 1$		When $h_i = 1, h'_i = 0$		When $h_i = 1, h'_i = 1$	
				$M_{i,1}$	$M_{i,0}$	$M_{i,1}$	$M_{i,0}$	$M_{i,1}$	$M_{i,0}$	$M_{i,1}$	$M_{i,0}$
0	1	0	1	0	1	0	1	0	1	0	1
0	1	1	1	0	1	0	1	0	1	0	1
0	1	1	0	0	1	0	1	0	1	0	1
1	1	0	1	1	0	1	0	0	1	0	1
1	1	1	1	1	1	1	0	0	1	1	1
1	1	1	0	0	1	1	0	0	1	1	0
1	0	0	1	1	0	1	0	1	0	1	0
1	0	1	1	1	0	1	0	1	0	1	0
1	0	1	0	1	0	1	0	1	0	1	0

$$\begin{aligned}
 S_{i+1,1} &= A_{i,0} + \overline{A_{i,1}}, S_{i,1}, S_{i+1,0} = A_{i,1} + A_{i,0} + S_{i,0} \\
 h'_i &= \overline{B_{i,1}} \cdot S'_{i,1} + B_{i,1} \cdot \overline{S'_{i,1}} + \overline{B_{i,0}} \cdot \overline{S'_{i,1}} + B_{i,0} \cdot \overline{S'_{i,0}} \\
 S'_{i+1,1} &= B_{i,0} + \overline{B_{i,1}}, S'_{i,1}, S'_{i+1,0} = B_{i,1} + B_{i,0} + S'_{i,0}
 \end{aligned} \tag{4}$$

Sub-step 3.2:

Then at  $i$ th stage, MC ( $\widehat{A}_{i+1}, \widehat{B}_{i+1}$ ) is determined as per Table 5, yielding (5).

$$\begin{aligned}
 M_{i,1} &= P \cdot \overline{h_i} \cdot \overline{h'_i} + Q \cdot \overline{h_i} \cdot h'_i + R \cdot h_i \cdot \overline{h'_i} + T \cdot h_i \cdot h'_i \\
 M_{i,0} &= U \cdot \overline{h_i} \cdot \overline{h'_i} + V \cdot \overline{h_i} \cdot h'_i + W \cdot h_i \cdot \overline{h'_i} + X \cdot h_i \cdot h'_i
 \end{aligned} \tag{5}$$

where

$$\begin{aligned}
 P &= \overline{B'_i} + A'_i, M_{i-1,0}, \quad Q = A'_i, \quad R = \overline{B'_i}, \quad T = \overline{B'_i} + A'_i, M_{i-1,1} \\
 U &= \overline{A'_i} + B'_i, M_{i-1,1}, \quad V = \overline{A'_i}, \quad W = B'_i, \quad X = \overline{A'_i} + B'_i, M_{i-1,0}
 \end{aligned}$$

## 4 Results and Discussion

In this paper, an algorithm for MC of CSDNS is proposed on the basis of comparing the magnitudes and signs of some counterpart portions of the inputs. The author would like to project the proposed algorithm as the first ever proposal on designing

MCR over the DS  $\{\bar{1}, 0, 1\}$ , though non-adjacency of non-zero digits is another precondition for the applicability of proposed algorithm.

As an example, the trace of Algorithm 2 for some sample inputs is shown in Table 6 where the symbols  $\times$ ,  $\times$  and  $\checkmark$  stand for not required, not applicable and already computed entries for the corresponding steps. As every AE associated to the proposed algorithm can be mapped into digital hardware using some ordinary logic gates only, the proposed algorithm can be implemented at ease as a whole.

It is obvious that owing to non-existence of any potential contender, the question of determining the comparative merit of the proposed algorithm is irrelevant. However, as reduction in delay is the basic motivation for employing SDNS-based arithmetic units, the achievements of the proposed algorithm may be better understood by analyzing its delay characteristics in terms of the AEs. At this point, it is worthy to note that although all AEs, (1)–(5), must undergo serial execution in that order, various sub-equations of each of (1)–(5) may still run in parallel. Now restricting the fan-ins of all logic gates to four, equating the delays of both AND gate and OR gate to  $\Delta_T$  and neglecting the delay caused by NOT gate [1], the delay of the proposed algorithm may be empirically estimated as follows: the delays which are caused by (1), (2), (3), (4) and (5) are  $2\Delta_T$ ,  $0$ ,  $\Delta_T$ ,  $2\Delta_T$  and  $3\Delta_T$ , respectively. It means, the delay due to step 1 and step 2 are  $2\Delta_T$  and  $\Delta_T$ , respectively. On the other hand, for every single instance of step 3, the delay is  $5\Delta_T$ . So, the overall delay is:  $3\Delta_T + 5(n - 1)\Delta_T = (5n - 3)\Delta_T$ .

It is also important to note that as no classes of SDNS, other than CSDNS, conform Theorem 1, the proposed Algorithm 2 cannot be ordinary extended beyond the CDSNS. Clearly, the field is open for exploration with new ideas and techniques. As already shown, the asymptotic characteristic of the execution time of the proposed

**Table 6** Trace of algorithm 2: an example

Position	$i$	7	6	5	4	3	2	1	0
Input	$A_i$	$\bar{1}$	0	1	0	$\bar{1}$	0	0	1
	$B_i$	$\bar{1}$	0	0	0	0	$\bar{1}$	0	1
Step 1	$A'_i$	1	1	0	1	0	1	1	1
	$B'_i$	1	1	1	1	1	0	1	1
	$M_i$	$\times$	$\times$	$\times$	$\times$	$\times$	$\times$	$\times$	(1, 1)
Step 2	$S_i$	$\times$	$\times$	$\times$	$\times$	$\times$	$\times$	(0, 1)	$\checkmark$
	$S'_i$	$\times$	$\times$	$\times$	$\times$	$\times$	$\times$	(0, 1)	$\checkmark$
Step 3.1	$h_i$	–	1	0	1	0	1	1	$\checkmark$
	$S_i$	(0, 1)	(0, 1)	(1, 1)	(1, 1)	(0, 1)	(0, 1)	$\checkmark$	$\checkmark$
	$h'_i$	–	1	1	1	1	0	1	$\checkmark$
	$S'_i$	(1, 1)	(1, 1)	(1, 1)	(1, 1)	(1, 1)	(0, 1)	$\checkmark$	$\checkmark$
Step 3.2	$M_i$	(1, 0)	(0, 1)	(0, 1)	(0, 1)	(0, 1)	(1, 0)	(1, 1)	$\checkmark$
Conclusion		$ A  <  B $							

Algorithm 2 is obviously linear. However, in the context of MC in the conventional, radix-complement NS platform, the emerging algorithms [14, 15] are the faster, non-linear variants of Algorithm 1, beyond the primitive linear version of Algorithm 1. As the design of Algorithm 2 is basically motivated by Algorithm 1, some variants of Algorithm 2 may even exist with log-depth or even better asymptotic characteristics. So also it may be intriguing to carry out incremental investigations for finding non-linear variants of Algorithm 2 for possibly superior performance.

## 5 Conclusion

SDNS is an emerging unconventional NS and CSDNS refers to an important category of SDNS. Although MCR is an indispensable part of every standard arithmetic processor, no significant report on designing MCR for CSDNS and other important classes of SDNS has been published yet. In this paper, an algorithm for MC of CSDNS is proposed on the basis of comparing the magnitudes and signs of some counterpart portions of inputs. The proposed algorithm can be implemented at ease by means of cascade connection which employs some basic logic gates only and for every stage of the proposed algorithm, all operations may be completed in a fairly small number of logic levels.

A significant offering of the proposed work is obviously demonstrating the ability of the CSDNS to handle MC in an effective and efficient manner, seemingly unlike that of the other classes of SDNS. Although the investigation report presented in this paper remains in the preliminary stage with theoretical studies only, it may act as the basis for accelerated investigations over this field in future.

## References

1. Koren I (2001) Computer arithmetic algorithms, 2nd ed. CRC Press, London
2. Avizienis A (1961) Signed-digit number representations for fast parallel arithmetic. IRE Trans Electron Comput EC-10:389–400
3. Cardarilli GC, Pontarelli S, Re M, Salsano A (2006) Fault tolerant design of signed-digit based FIR filters. In: IEEE international symposium on circuits and systems, Greece, pp 2809–2812
4. Crookes D, Jiang M (2007) Using signed digit arithmetic for low power multiplication. Electron Lett 43:613–614
5. Smitha KG, Fahmy AH, Vinod AP (2006) Redundant adders consume less energy. In: IEEE Asia Pacific conference on circuits and systems, Singapore, pp 422–425
6. Parhami B (2009) Computer arithmetic: algorithms and hardware design, 2nd edn. Oxford University Press, New York
7. Hartley R, Parhi KK (1995) Canonic signed digit arithmetic. In: Digit-serial computation. The springer international series in engineering and computer science. Boston
8. Wang Y, DeBrunner LS, Zhou D, DeBrunner VE (2007) A multiplier structure based on a novel real-time CSD recoding. In: IEEE international symposium on circuits and systems, New Orleans, pp 3195–3198

9. Manuel M, Elias M (2013) Design of frequency response masking FIR filter in the canonic signed digit space using modified artificial bee colony algorithm. *Eng Appl Artif Intell* 26:660–668
10. Park CS, Park S (2016) CSD-based CORDIC algorithm and its VLSI implementation. *IEICE Electron Exp* 13:1–9
11. Palchoudhuri A, Dhar AS (2019) Fault localization and testability approaches for FPGA fabric aware canonic signed digit recoding implementations. *J Electron Test* 35:779–796
12. Yang S, Lee BU (2009) Efficient transform using canonical signed digit in reversible color transforms. *J Electron Imaging* 18:033010
13. Parvin A, Ahmadi M, Muscedere R (2013) Application of neural networks with CSD coefficients for human face recognition. In: *IEEE international symposium on circuits and systems, Beijing*, pp 1628–1631
14. Veeramachaneni S, Krishna MK, Avinash L, Sreekanth RP, Srinivas MB (2007) Efficient design of 32-bit comparator using carry look-ahead logic. In: *IEEE northeast workshop on circuits and systems, Montreal*, pp 867–870
15. Chuang PI-J, Sachdev M, Gaudet VC (2014) A 167-ps 2.34 mW single-cycle 64-bit binary tree comparator with constant-delay logic in 65 nm CMOS. *IEEE Trans Circ Syst I Regul Pap* 61:160–171
16. Chakraborty MS (2020) Designing an on-line magnitude comparator for higher-radix. *Int J Inf Technol Electr Eng* 9:92–98
17. Pawlak Z (1991) *Switching circuits. rough sets. Theory and decision library (Series D: System theory, knowledge engineering and problem solving)*. Springer, Dordrecht, pp 188–204

# Region-Based Random Color Highlighting in Artistic Style Transfer Using CNN



Katharotiya Krutarth and Manu Madhavan

## 1 Introduction

Artwork is used to represent aesthetic values for centuries. Artwork creation is an arduous process, so only professionals used to create it. The term “digital art,” which represents computer-generated art, was introduced in the early 1980s. Simple operations on camera captured images like crop, rotate, flip, merge, or create simple collages were also considered digital art back in that time. Since then, it became much easier to manipulate images and create artwork because of the evolution of computer vision and image processing algorithms. In the last decade, the deep learning field has also emerged strongly because of the availability of advanced computational machines.

Convolutional neural networks (CNNs) are widely used in computer vision and have various applications [1]. Neural Style Transfer (NST) techniques use CNNs for feature extraction. These powerful CNNs, which are already trained on larger data sets, make it easy to extract features for a particular image. According to NST, content, style, and target images are used as input to CNN, and feature maps are obtained for each image. These obtained feature maps are used to represent loss terms, which are content loss and style loss. After total loss minimization, the target image is generated such that it has content from the content image and style from the style image. Color information of an image is represented by the style. So, color information is synthesized in the target image from the style image while minimizing

---

K. Krutarth · M. Madhavan (✉)  
Department of Computer Science and Engineering, Amrita School of Engineering, Amrita  
Vishwa Vidyapeetham, Coimbatore, India  
e-mail: [m\\_manu@cb.amrita.edu](mailto:m_manu@cb.amrita.edu)

K. Krutarth  
e-mail: [cb.en.p2aid19018@cb.students.amrita.edu](mailto:cb.en.p2aid19018@cb.students.amrita.edu)

total loss. This limited scope for color information in traditional style transfer makes results more static.

For human beings, color is an essential natural characteristic of any object. In artworks, color information plays a significant role in making an image more artistic. This research introduces a new approach to generate an artistic image, such that the scope of color information is not only bounded to style image. This study introduces a new image called a “region-based random color image,” which could be used in the style transfer process to make the results more dynamic. This study also proposes a new loss term, “region-based random color loss,” responsible for transferring color information from “region-based random color image” to the target image. The paper also discusses the required preprocessing and layer selection for obtaining feature maps for “region-based random color image.” The results are compared with [2, 3] qualitatively, and a user survey is conducted to determine the general preference for the proposed method.

The following are the paper’s primary technical contributions:

1. “Region-based random color image” generation and required preprocessing before total loss minimization.
2. Enhanced total loss with the additional term “region-based random color loss.”

The remainder of the paper is organized as follows. Section 2 describes the related works in artistic style transfer. The proposed methodology is described in Sect. 3. Section 4 contains qualitative comparison and user survey analysis. Finally, Sect. 5 wraps up the study with conclusions and potential future improvements.

## 2 Literature Survey

There is a plethora of research that has been done for generating artistic images since the early ’90s. For the artwork generation, the non-photorealistic approaches were famous. Earlier methods for artistic image generation considers Image-based Artistic Rendering (IB-AR). Stroke-based rendering [4] is a particular IB-AR algorithm that creates brushy stripes on an image. Region-based techniques [5, 6] were also proposed, which consider regions in the image while transferring style or control the levels of transfer based on regions. These approaches are not applicable to arbitrary styles. Example-based rendering (EBR) [7] is also one kind of IB-AR which takes an input image and a corresponding artistic rendered image as a pair and learns a mapping between these two images. Filters based on image processing techniques can also be used to create artistic images. For example, Gooch et al. [8] uses a Gaussian filter difference to generate cartoon-like effects. Filters-based techniques are more straightforward but limited in style diversity. These traditional non-deep learning techniques were used for artistic image generation before deep learning became popular in the last decade.

Recent advances in deep learning algorithms have resulted in the development of neural network-based style transfer methods. Deep learning-based methods offer



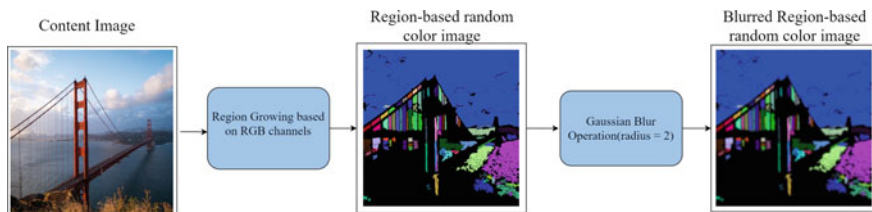
new possibilities and improvements in style transfer when compared to traditional approaches. There are two types of style transfer algorithms—image optimization-based and model optimization-based algorithms. Image representation is an essential step in style transfer algorithms. Image can be represented as a feature, and there are several methods for extracting image features [9–11]. NST [2] is the first image optimization-based technique, and it uses a CNN for image feature extraction. Gatys et al. [12] improved NST by allowing for more control over color and scale in style transfer by controlling perceptual factors. Hung et al. [3] achieves real-time artistic image generation capabilities, and it also works with arbitrary styles. Liu et al. [13] proposes a method for utilizing depth information of a content image, which helps preserve the perspective of an original image. Further, it improved by applying instance normalization by Kitov et al. [14]. Semantic style transfer [15] is also inspired by NST, which makes correspondence in style regions and gives spatial control. Doodle style transfer [16] is an extension of NST, which discards content loss while transferring style and uses doodle for mapping to achieve semantic style transfer. The shape stylization approach [17] is inspired by NST, which helps to generate logos. The abovementioned methods generate non-photorealistic results. To achieve photorealism, Luan et al. [18] uses semantic segmentation and photorealism regularization. Apart from these, there are many other exciting applications inspired by NST [19–21].

### 3 Methodology

This section explains the artistic image generation process, which includes two phases. The first phase discusses the need and process for generating a region-based random color image. It also covers additional preprocessing steps. The second phase introduces a new loss function for the adapted approach [2]. The second phase also discusses layer selection for obtaining feature maps for each image.

#### 3.1 *Blurred Region-Based Random Color Image Generation*

The region-based random color image is an image where neighboring pixels with the same color values get maps to a specific region, and each region initializes with a random color. The proposed method uses a region growing technique [22] to identify regions in the content image. The primary purpose of using region-based random color image is to transfer color to the target image and provide random color highlighting. The proposed method chooses the content image to generate a region-based random color image. The structural information and symmetry of the region-based random color image remain the same as the content image, which helps to transfer only color information. Generated region-based random color images may have distortions and some textures, which may conflict with the texture of the



**Fig. 1** Blurred region-based random color image generation

style image. Gaussian blur helps to reduce distortions and get better results in the artistic image generation process. Figure 1 depicts the process of creating a blurred region-based random color image.

### 3.2 Style Transfer Using Enhanced Total Loss

Figure 2 depicts the proposed architecture for artistic image generation. The proposed approach uses a pre-trained VGG19 deep neural network for extracting image feature maps. The proposed approach takes the content image, blurred region-based random color image, and target image as input. The target image is initialized with random pixel values before feature extraction. Feature maps are obtained for all four images from information-specific layers. Extracted feature maps are then used to define loss terms—content loss, style loss, and region-based random color loss.

According to NST [2], the content loss can be defined as the mean squared distance between the feature maps of content image and target image. VGG19’s initial layers are used to obtain content information. This study uses conv1\_1, conv1\_2, and conv2\_1 layers for extracting content-specific feature maps. FC in Eq. (1) represents the content image feature maps, whereas FT depicts feature maps of the target image.

$$L_{\text{content}} = \frac{1}{2} \sum_{i,j}^l (\text{FC}_{ij}^l - \text{FT}_{ij}^l)^2 \quad (1)$$

Style loss can be defined as the mean squared distance between correlated feature maps of the style and target image [2]. The correlation of feature maps is calculated by taking the dot product of feature maps with transposed feature maps. Depth layers of VGG19 contain style information which includes texture and color. The proposed approach utilizes conv3\_1, conv3\_4, conv4\_3, conv5\_1, and conv5\_4 for obtaining style information. GC and GT in Eq. (2) represents correlated feature maps of style and target image, respectively.

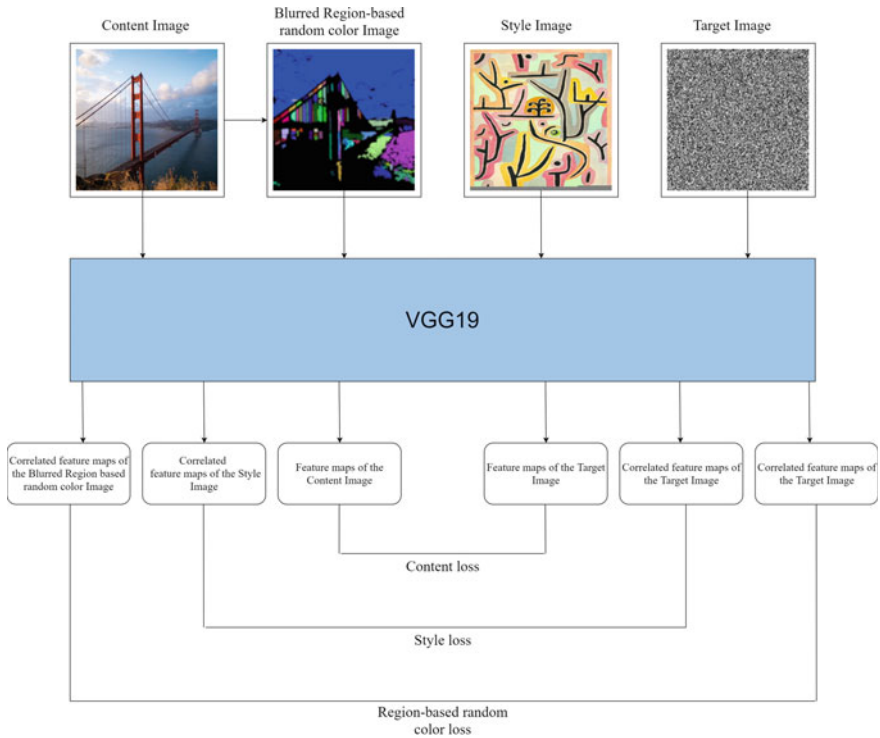


Fig. 2 Proposed architecture for artistic image generation

$$L_{\text{style}} = \frac{1}{2} \sum_{i,j}^l (GC_{ij}^l - GT_{ij}^l)^2 \tag{2}$$

This paper introduces a new loss term, “region-based random color loss.” This research considers color information as style information. So, the proposed loss term “region-based random color loss” can be calculated by taking the mean squared distance between correlated feature maps of the blurred region-based random color image and correlated feature maps of the target image. We found that correlated feature maps from early layers of VGG19 focus on color information and help to reduce texture. Thus, color information is extracted as correlated feature maps using conv1\_1, conv1\_2, and conv2\_1 layers. GR and GT in Eq. (3) represent correlated feature maps of the region-based random color image and target image, respectively.

$$L_{\text{region}} = \frac{1}{2} \sum_{i,j}^l (GR_{ij}^l - GT_{ij}^l)^2 \tag{3}$$

$$L_{\text{total}} = \alpha.L_{\text{content}} + \beta.L_{\text{style}} + \varepsilon.L_{\text{region}} \tag{4}$$

This study combines content loss, region-based random color loss, and style loss to formulate total loss as given in Eq. (4).  $\alpha$ ,  $\beta$ , and  $\varepsilon$  are weights associated with content, style, and region-based random color loss terms. After minimizing total loss for multiple iterations using backpropagation, the proposed method recreates the target image such that it has content from the content image, style from the style image, and color from the blurred region-based random color image. The results for the proposed method are generated after minimizing total loss for 5000 epochs with a learning rate set to 0.003. This study uses Adam optimizer for total loss minimization. The results shown in the paper are generated by keeping  $\alpha$ ,  $\beta$ , and  $\varepsilon$  values at 0.5, 0.5, and 1.0, respectively.

## 4 Results and Discussion

To validate the proposed method, we conducted multiple experiments. This study makes qualitative comparison and user evaluation study to get a general preference for the proposed method. For experiments, this study uses both realistic and artistic images, which are retrieved from [23, 24].

As shown in Fig. 3, the generated results are qualitatively compared to NST [2] and AdaIN [3]. For some image pairs of content and style, AdaIN results are distorted, and structural detail is lost, for example, eyes in the portrait of a lady. However, the proposed approach symmetrically transfers style and preserves the structural details of the content image. Random color highlighting using the proposed approach makes results unique and more artistic than the results of NST and AdaIN.

NST and AdaIN have a fixed scope for color information. Whereas, the proposed method has the advantage of generating results with different color highlighting by providing different region-based random color images, as shown in Fig. 4. The proposed method generates results with style and random color highlighting, making results unique than NST and AdaIN.

The proposed method also allows users to control color highlighting. It is observable from Fig. 5 that the degree of color highlighting varies based on  $\varepsilon$ . When  $\varepsilon$  is zero, the proposed method produces results that are similar to NST. The strength of color highlighting will increase as the  $\varepsilon$  value rises. From the visual comparison, the proposed method allows for greater flexibility in transferring color information, resulting in unique and better results than previous techniques in terms of color information.

In order to get a general preference of the proposed method, two individual surveys are conducted comparing the proposed approach with Neural Style Transfer NST [2] and AdaIN [3]. As part of the survey analysis, 25 participants were chosen who did not have any prior knowledge about the given techniques for each survey. They were given output images of each technique in random order and asked to choose best according to their visual perception. Table 1 summarizes the collected data as part of user survey analysis. According to the data collected, the results generated by the

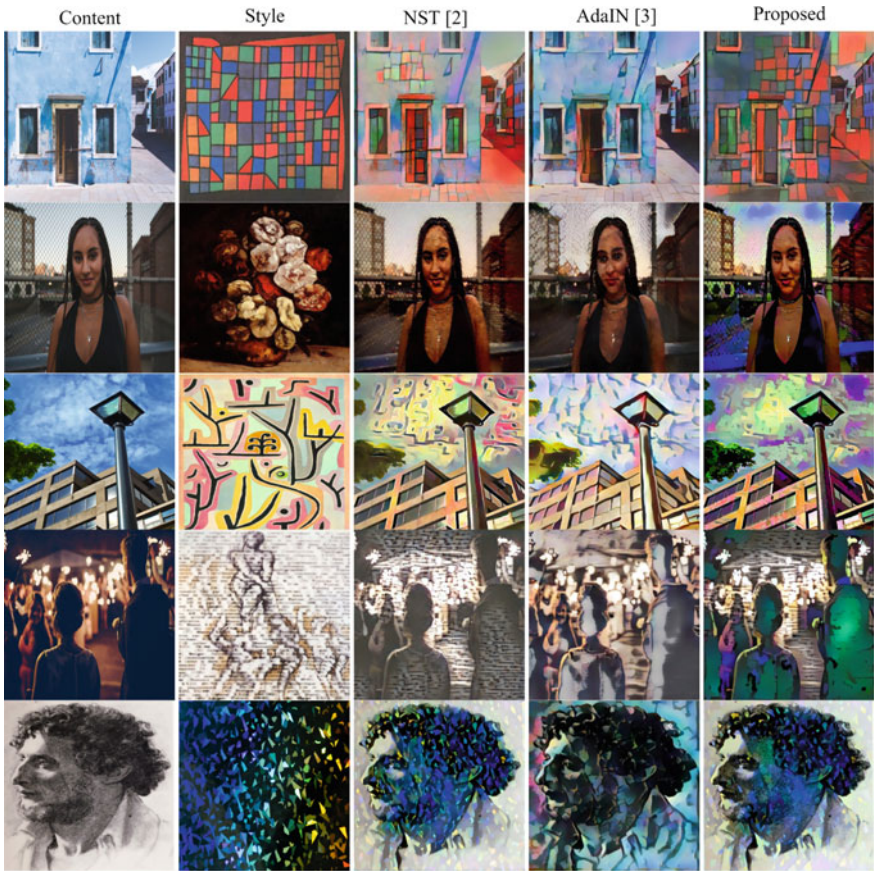


Fig. 3 Qualitative comparison of the proposed method with NST [2] and AdaIN [3]

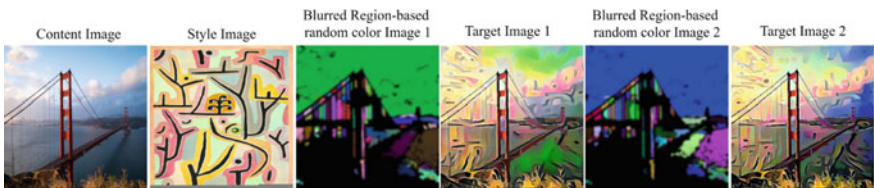
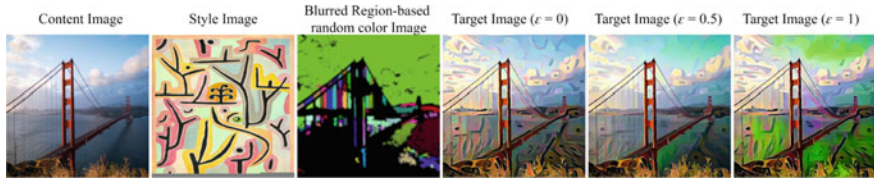


Fig. 4 Possible outcomes by different blurred region-based random color images

proposed method are moderately more preferable than other techniques with a 99% level of significance in the binomial test.



**Fig. 5** Effect of  $\varepsilon$  on the target image

**Table 1** User survey analysis

Experiments	Proposed versus NST [2]	Proposed versus AdaIN [3]
#Pairs of an images	10	10
#Participants	25	25
#Total Responses	250	250
#Responses in favor of the proposed method	178	159
<b>Favored Responses %</b>	<b>71.2%</b>	<b>63.6%</b>
P-value (Binomial test)	0.000001	0.000011

## 5 Conclusions and Future Scope

This work introduces a new deep learning-based approach to generate artistic images using a region-based random color image. We use region growing to generate a region-based random color image. Later after preprocessing, a region-based random color image is used in the style transfer process. The proposed approach also introduces a new loss term, “region-based random color loss,” which contributes to a total loss for transferring color information. The proposed method gives the advantage of generating different results using the same pair of content and style images using different region-based random color images each time. The proposed approach is more flexible and generates unique artworks. Qualitative evaluation and user survey analysis indicate that the proposed approach results are unique and preferable compared to NST and AdaIN methods.

However, the proposed approach is based on image optimization, so it cannot be applied to the use cases that require style transfer in real time. Hence, the proposed work can be enhanced by adapting the model optimization techniques.

## References

1. Harini S, Aarthi R (2018) A survey of deep convolutional neural network applications in image processing. *Int J Pure Appl Math* 118(7):185–189

2. Gatys LA, Ecker AS, Bethge M (2016b) Image style transfer using convolutional neural networks. In: 2016 IEEE conference on computer vision and pattern recognition (CVPR). IEEE, Las Vegas, NV, USA, pp 2414–2423
3. Huang X, Belongie S (2017) Arbitrary style transfer in real-time with adaptive instance normalization. In: IEEE international conference on computer vision (ICCV). IEEE, Venice, pp 1510–1519
4. Hertzmann A (1998) Painterly rendering with curved brush strokes of multiple sizes. In: Proceedings of the 25th annual conference on computer graphics and interactive techniques. Association for Computing Machinery, New York, NY, USA, pp 453–460
5. Song Y-Z, Rosin PL, Hall PM, Collomosse J (2008) Arty shapes. In: Proceedings of the fourth Eurographics conference on computational aesthetics in graphics, visualization and imaging. Eurographics Association, Lisbon, Portugal, pp 65–72
6. Gooch B, Coombe G, Shirley P (2002) Artistic vision: painterly rendering using computer vision techniques. In: Proceedings of the 2nd international symposium on non-photorealistic animation and rendering. Association for Computing Machinery, Annecy, France, pp 83–ff
7. Hertzmann A, Jacobs CE, Oliver N et al (2001) Image analogies. In: Proceedings of the 28th annual conference on computer graphics and interactive techniques. Association for Computing Machinery, New York, NY, USA, pp 327–340
8. Gooch B, Reinhard E, Gooch A (2004) Human facial illustrations: creation and psychophysical evaluation. *ACM Trans Graph* 23:27–44. <https://doi.org/10.1145/966131.966133>
9. Kanishka ND, Bhagavathi SP (2015) Learning of generic vision features using deep CNN. In: 2015 fifth international conference on advances in computing and communications (ICACC), Kochi
10. Reghunath A, Nair SV, Shah J (2019) Deep learning based customized model for features extraction. In: International conference on communication and electronics systems (ICCES), Coimbatore, India
11. Muthugnanambika M, Padmavathi S (2017) Feature detection for color images using SURF. In: 4th international conference on advanced computing and communication systems, ICACCS
12. Gatys LA, Ecker AS, Bethge M et al (2017) Controlling perceptual factors in neural style transfer. [arXiv:161107865](https://arxiv.org/abs/161107865) [cs]
13. Liu X-C, Cheng M-M, Lai Y-K, Rosin PL (2017) Depth-aware neural style transfer. In: Proceedings of the symposium on non-photorealistic animation and rendering. Association for Computing Machinery, Los Angeles, California, pp 1–10
14. Kitov V, Kozlovstev K, Mishustina M (2019) Depth-aware arbitrary style transfer using instance normalization. <http://arxiv.org/abs/1906.01123>
15. Lu M, Zhao H, Yao A et al (2017) Decoder network over lightweight reconstructed feature for fast semantic style transfer. In: 2017 IEEE international conference on computer vision (ICCV). IEEE, Venice, pp 2488–2496
16. Champandard AJ (2016) Semantic style transfer and turning two-bit doodles into fine artworks. [arXiv:160301768](https://arxiv.org/abs/160301768) [cs]
17. Atarsaikhan G, Iwana BK, Uchida S (2020) Guided neural style transfer for shape stylization. *PLoS One* 15:e0233489. <https://doi.org/10.1371/journal.pone.0233489>
18. Luan F, Paris S, Shechtman E, Bala K (2017) Deep photo style transfer. In: 2017 IEEE conference on computer vision and pattern recognition (CVPR). IEEE, Honolulu, HI, pp 6997–7005
19. Selim A, Elgharib M, Doyle L (2016) Painting style transfer for head portraits using convolutional neural networks. *ACM Trans Graph* 35:129:1–129:18. <https://doi.org/10.1145/2897824.2925968>
20. Atarsaikhan G, Iwana BK, Narusawa A et al (2017) Neural font style transfer. In: 2017 14th IAPR international conference on document analysis and recognition (ICDAR), pp 51–56
21. Jiang S, Fu Y (2017) Fashion style generator. In: Proceedings of the twenty-sixth international joint conference on artificial intelligence organization, Melbourne, Australia, pp 3721–3727

22. Gothwal R, Gupta S, Gupta D, Dahiya AK (2014) Color image segmentation algorithm based on RGB channels. In: Proceedings of 3rd international conference on reliability, Infocom technologies and optimization. IEEE, pp 1–5
23. <https://www.kaggle.com/ikarus777/best-artworks-of-all-time>. Last accessed 2021/06/27
24. <https://stocksnap.io/>. Last accessed 2021/06/27



# A Novel Enhancement of Vigenere Cipher Using Message Digest



Rajalaxmi Mishra  and Jibendu Kumar Mantri 

## 1 Introduction

Cryptographic methods are based on the theory of mathematics and involve the procedure of transformation of the ordinary plain text messages into unintelligible ciphered text messages and vice versa. These consist of the techniques of data storage and transmission in encrypted form which makes sure that only the intended user can read and process it. Many researchers have contributed in the area of cryptography and proposed quite secure and robust schemes like Data Encryption Standard (DES) [5, 10], 3DES [16, 18], AES [8, 10, 29], Blowfish [11], RSA [10, 29]. The complexity of these ciphers is too high and also requires more resources for implementation.

Blaise de Vigenère proposed the Vigenere cipher, a popular traditional cryptographic technique for encryption of alphabet text. Depending on the sequence of characters in the keyword, a sequence of Caesar cipher application is used. A number of researchers have worked on the improvement of Vigenere cipher and presented different modifications to it. To include small letters, capital letters and digits, a matrix of size  $62 \times 62$  [20] is constructed. Vigenere cipher is combined with Polybius cipher to present a new hybrid cipher [25]. To secure the key [2], the one-time pad cipher is used. The one-time pad cipher is used with a randomly generated key to generate the cipher text of the message. On the equal length key and the message, the XOR operation is applied. The transposition method is applied to the encrypted message and then some logical bits are added. Then, the one-time pad cipher followed by transposition cipher is used to provide the cipher text which is more secure. To avoid exchange of keys, a three-pass Protocol scheme [13] is proposed. It used separate keys for the process of encryption and decryption. The keys need not be shared by the communicating parties. The paper [15] presents an

---

R. Mishra · J. K. Mantri (✉)

Department of Computer Application, North Orissa University, Baripada, Orissa, India

e-mail: [jkmantri@gmail.com](mailto:jkmantri@gmail.com)

extension to Vigenère cipher; it is the Alpha-Qwerty cipher. It provides an encryption scheme for a set of 92 characters which include digits, some special symbols and case sensitivity to the original Vigenere cipher of 26 letters. It employs a modified mapping method to map the plain text to the cipher text. The paper [3] uses different keys and the Kasiski method of attack to analyze the modified Vigenere cipher. An enhanced Vigenère cipher is proposed in [19] by adding some random bits of padding to each byte. It employs a one-way function to find the accurate extent of pad to separate the message bits from the random padded bits. The size of the cipher text is increased reasonably in this method which significantly improves the security of the cipher. A strong hybrid cipher [17] is proposed by combining the characteristics of modern cipher with conventional ciphers like Caesar cipher and Vigenere cipher in their modified forms. The multilevel encryption algorithm [28] using modified Vigenere cipher enhances its security. This method is suitable for resources constrained light-weight applications as it has much lesser computational complexity as compared to most of the modern ciphers. The generalized implementation for Vigenere cipher [27] employs any two reversible square matrices having unique rows or columns, for encryption and decryption purposes. A poly alphabetic cipher is proposed by combining the features of the Vigenère cipher and the Affine cipher [26]. It incorporates the diffusion and confusion properties to improve the security level in storage and transmission of data in open public networks. An extension of the original Vigenere table into  $95 \times 95$  is presented in [12] which include all characters, digits, mathematical symbols and punctuations to enhance security and make the cryptanalysis procedure more difficult. Kester [21] presents a novel method of implementation of the Vigenère cipher algorithm where the cipher key is being changed automatically in every step of encryption. The subsequent key values are generated from the initial value of the key. The Vigenère cipher is reviewed [1] and also different existing modifications are discussed. A key generation procedure [4] generates a key whose length is dependent on the level of security of the message. The key length is dependent on the security level of the message which is determined by the sender on the basis of its importance. The paper [14] performs cryptanalysis and analyzes the classical encryption schemes. It also presents a modified algorithm by using simple computations to provide better security. By using simple shift and EXOR operation from a single key, it generates lots of keys.

The message digest algorithm (MD5) takes any string of arbitrary length as input data, and produces output of 128-bit (32-digit hexadecimal) as fingerprint [23]. Each message produces a different MD5 value. It has been widely used to check integrity of data. Rachmawati et al. [7] presents a comparative study between MD5 and SHA256 algorithm. A chaotic block image encryption scheme is presented in [9] by applying message digest algorithm. To enhance the level of security, the initial key value is generated by using the message digest algorithm MD5, which is used by 2D-STCM to generate a large key stream.

In this paper, we used the message digest MD5 algorithm applied on plain text message to generate the initial key as in [9], and then employed a key generation procedure based on [14] to generate the keys as the length of the plain text message. Then Vigenère cipher is applied to encrypt the message in plain text. The arrangement

of the rest of the paper is given as: the theoretical background is included in Sects. 2, 3 contains our proposed scheme, Sect. 4 includes the result and analysis of the scheme and finally, Sect. 5 contains the conclusion and scope of future work.

## 2 Theoretical Background

In 1991, Ronald Rivest designed MD5 as a substitute of the existing hash function MD4. The message-digest algorithm, MD5 is a commonly used hash function producing a hash value of 128-bit. The MD5 hash value of 128-bit (16 byte) called as the message digests are generally expressed as a series of 32 hexadecimal numerals. Message digest values are used for checking the integrity of the messages.

The symmetric key cryptography is the cryptographic scheme where both the processes of encryption and decryption use the same key. The private or secret key needs to be shared between the sender and the receiver in a secured manner. P stands for the plain text, C stands for the ciphered text and K is the shared secret key. The plain text is converted to the ciphered text by using the encryption algorithm and the key. Then the ciphered text can be sent to the receiver with whom the secret key is shared. After receiving the ciphered text, the receiver can decrypt it by using the decryption algorithm along with the shared secret key.

Encryption Algorithm:

$$C = \text{Encryption}(P, K)$$

Decryption Algorithm:

$$P = \text{Decryption}(C, K) = \text{Decryption}(\text{Encryption}(P, K), K)$$

Applying a series of Caesar cipher applications depending on the letters of the keyword, the Vigenere cipher produces the cipher text. The method of substitution is utilized that shifts each plain text letter with a different amount according to the letter of the keyword and uses a table called “Vigenere Table” to achieve this purpose. The table consists of 26 rows and 26 columns each representing a letter from A to Z.

The encryption process of Vigenere cipher generates cipher text from the plain text

$$C = \text{Encryption}(P, K) = (P_i + K_i) \bmod 26.$$

The original message in plain text is generated from the ciphered message by the process of decryption of Vigenere cipher

$$P = \text{Decryption}(C, K) = (C_i - K_i) \bmod 26.$$

The values of plain text, cipher text and key, respectively, denote at the offset position  $i$  denoted as  $P_i$ ,  $C_i$  and  $K_i$ . The alphabet, A–Z are mapped with the numbers from 0 to 25, so that A is mapped with 0, B mapped with 1...Z mapped with 25. To encrypt the plain text using Vigenère cipher, the key length has to be equal with the message length. If the key length is less than the message length, then the letters of

the key need to be repeated to get a key of length equals to the length of message in plain text. For example, the plain text P = CRYPTOGRAPHY

The numeric values are 2, 17, 24, 15, 19, 14, 6, 17, 0, 15, 7, 24

And the keyword is TRUST, and then the key for encryption is TRUSTTRUSTTR

The numeric values of the key are 19, 17, 20, 18, 19, 19, 17, 20, 18, 19, 19, 17

The cipher text is generated as VISHMHXLSIAP (Table 1)

There are some drawbacks of Vigenere cipher. The mathematical symbols, punctuation and digits are not included in the Vigenere table. Hence, the encryption methods of these symbols are not described. The repetitive key is its main flaw. The key length can be predicted by the cryptanalyst by using the Kasiski test and Friedman test. After getting the key length, the content of the cipher textual may be handled as interwoven Caesar ciphers, which can easily be broken separately.

### 3 The Proposed Scheme

This paper proposes the encryption and decryption scheme based on Vigenere cipher.

The encryption and decryption processes:

The plain text is encrypted as per Vigenere cipher

For all the characters in the plain text

$$C[i] = (P[i] + K[i]) \text{ mod } 256$$

Similarly, the ciphered text is decrypted to plain text as

$$P[i] = (C[i] - K[i]) \text{ mod } 256$$

For all the characters present in the cipher text.

#### The Key Generation Process:

On the plain text, the MD5 (Message Digest) algorithm is applied which gives a sequence of 32 hexadecimal numbers.

The hexadecimal numbers are converted to decimal numbers.

Sum of the decimal numbers has to be evaluated.

This sum of 32 decimal numbers is the key value.

This encryption scheme generates a sequence of as many keys as the number of characters present in the plain text message by using bit-level manipulations.

k = initial key that is sum of 32 decimal numbers (output of MD5), represented as the binary string k1 and it is divided into left-hand side (LHS) part and right-hand side (RHS) part. The left circular shift (LCS) operation also performed to generate key values.

$$k[i] = \text{LHS} \parallel \text{RHS}$$

$$k[i + 1] = \text{RHS}(k[i]) \parallel \text{LHS}(k[i])$$

**Table 1** The original Vigenere table [1]

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

(continued)

Table 1 (continued)

V	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
W	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	0	P	Q	R	S	T	U
X	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
Y	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	0	P	Q	R	S	T	U	V	W
Z	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	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

$$k[i + 2] = \text{LCS}(k[i + 1])$$

$$k[i + 3] = \text{LCS}(\text{LHS}(k[i + 1])\|\text{RHS}(k[i + 1]))$$

$$k[i + 4] = \text{LHS}(k[i + 3])\|\text{LCS}(\text{RHS}(k[i + 3]))$$

Repeat this set of processes to generate keys equal to number of characters of plain text.

After getting the plain text at the receiver side, the MD5 (message digest algorithm) is applied again and the output of 32 hexadecimal numbers is converted to 32 decimal numbers and their sum is evaluated. This sum is compared with the shared secret key. The matching result ensures data integrity.

### 4 Results and Analysis

This encryption scheme is implemented by using the programming language Java.

Plain Text: secure message scheme

Numeric Values of the plain text:

115 101 99 117 114 101 32 109 101 115 115 97 103 101 32 115 99 104 101 109 101

MD5 value: 60df1f408d4b845642b9e8d0780701c1 (32 hexadecimal numbers)

The decimal values: 6 0 13 15 1 15 4 0 8 13 4 11 8 4 5 6 4 2 11 9 14 8 13 0 7 8 0 7 0 1 12 1

The sum of the 32 decimal numbers = 210

The numerical values of the individual key:

210 45 90 178 212 77 154 180 216 141 27 184 209 29 58 177 210 45 90 178 212

The numerical values of the characters of the encrypted string

71 19 62 41 72 51 59 35 63 2 15 27 58 3 90 38 55 22 64 33 59

Encrypted String: G!! >)H3;#? ☉☼ ← :♥Z&7■@!;

The numerical values of the characters of the decrypted string

115 101 99 117 114 101 32 109 101 115 115 97 103 101 32 115 99 104 101 109 101

Decrypted string: secure message scheme

To check integrity of the message, MD5 algorithm is applied on the plain text and the output is:

60df1f408d4b845642b9e8d0780701c1 (32 hexadecimal numbers)

The decimal values are:

6 0 13 15 1 15 4 0 8 13 4 11 8 4 5 6 4 2 11 9 14 8 13 0 7 8 0 7 0 1 12 1

The sum of the 32 decimal numbers = 210

Hence, the integrity of the communicated message is ensured.

## 4.1 Security Analysis

In this scheme, the key is generated from the output of MD5, which is a sequence of 32 hexadecimal numbers. These 32 hexadecimal numbers are converted to 32 decimal numbers and the summation of these numbers is the secret key. On the basis of the secret key, a sequence of keys are generated, one for each character of plain text. Value of the shared secret key is dependent on the plain text, for different plain text this value must be different. Hence, it will be quite difficult for the cryptanalyst to guess the key.

For brute-force attack, the cryptanalyst needs to try  $(256)^n$  number of different numbers as the key value, as  $n$  is the message length.

This scheme possesses two significant characteristics called confusion and diffusion. The concept of diffusion is to conceal the association of the plain text with the cipher text [6]. It will discourage the cryptanalyst to get the plain text by using the cipher text statistics. This scheme possesses the diffusion property as the key is based on the message digest of the plain text message. The confusion concept is to conceal the association among the key and the cipher text. It will discourage the cryptanalyst to get the key by using the cipher text. It has the property of confusion as for each plain text character, the key value is different. The cryptographic cipher scheme's success and competence depends on the fact that how difficult it is to be cracked or broken by a cryptanalyst.

The major limitation of the Vigenère cipher is the repetition of the key. Once the key length is guessed correctly, the cipher text would be regarded as inter-woven Caesar ciphers, which can be cracked very easily. The Kasiski and Friedman tests can be used to find the key length. In this scheme, the key length is same as the length of the plain text; hence, Kasiski and Friedman tests fail for this scheme.

## 4.2 Avalanche Effect Analysis

The avalanche effect is one among the requisite character of cryptographic algorithms which mean that with the slight change of input, the output changes significantly. Avalanche effect around 50% or more means the cryptosystem has diffusion property.

To test the avalanche effect, two different pairs of input strings are chosen [22]

Data1 (contains only alphabets) cryptography and cryptography

Data2 (contains alpha numeric characters) Data4 and Data2.

The following table shows the avalanche effect of AES, RSA and our proposed method (Table 2).



**Table 2** Analysis of avalanche effect

Input text	AES (%)	RSA (%)	Proposed scheme (%)
Cryptography Cryptography	47	51	54.16
Data4 Data2	53	56	55.0

### 4.3 Frequency Analysis

Plain Text: secret message

Ciphered Text: 3R < q\$QXk"0 k] a

Figure 1 represents the frequency analysis of the plain text and Fig. 2 represents the frequency analysis of the cipher text. It is evident from the frequency analysis that the proposed method eliminates repetition and no trace of it is provided in the cipher text.

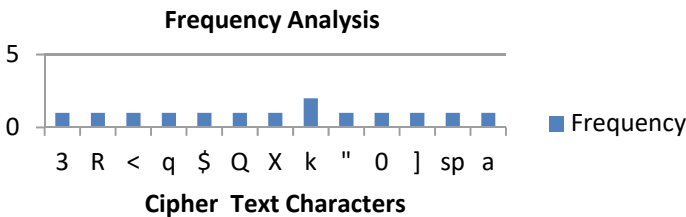
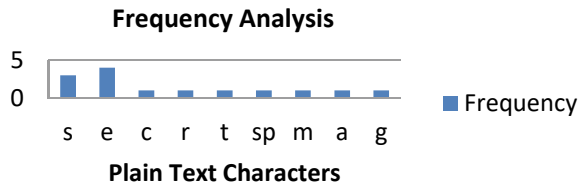
There is another test case.

Plain Text: aaabbbcccddeee

Cipher Text: E▶ > 6H1☺9Ms ● > H♦”

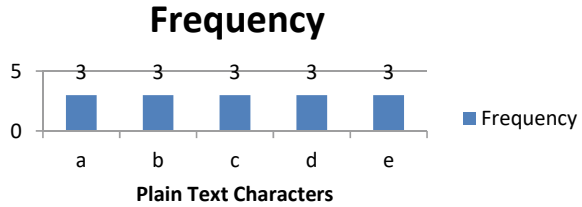
The frequency analysis of the message in plain text and in cipher text is shown in Figs. 3 and 4, respectively. From the frequency analysis, it is clear that the proposed method eliminates repetition and no trace of it is provided in the cipher text.

**Fig. 1** Frequency analysis of the plain text

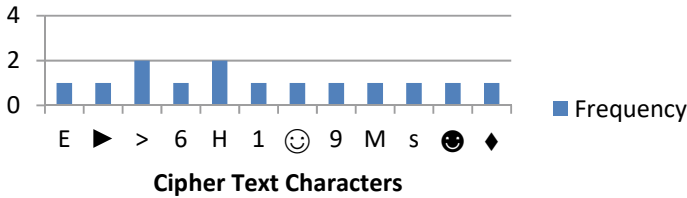


**Fig. 2** Frequency analysis of the cipher text

**Fig. 3** Frequency analysis of the plain text



### Frequency Analysis



**Fig. 4** Frequency analysis of the message in cipher text

## 5 Conclusions and Future Work

The proposed work is an improvement over Vigenère cipher as it defends the Kasiski and Friedman attacks, and it also possesses the properties like confusion and diffusion. The key value of the scheme is calculated from the message digest which is different for different plain texts; this makes it difficult for the cryptanalyst to guess the key. This scheme ensures privacy as well as message integrity. The secure way of sharing the key will be considered as our future work.

## References

1. Aliyu A-AM, Olaniyan A (2016) Vigenere cipher: trends, review and possible modifications. *Int J Comput Appl* 135(11). ISSN 0975-8887
2. Priyam A (2015) Extended Vigenère using double transposition cipher with one time pad cipher. *Int J Eng Sci Adv Res* 1(2):62–65
3. Hananto AL, Solehudin A, Irawan ASY, Priyatna B (2019) Analyzing the Kasiski method against Vigenere cipher. *Int J Comput Tech* 6(6)
4. Al-ahwal A, Farid S (2015) The effect of varying key length on a Vigenère cipher. *IOSR J Comput Eng* 17(2):18–23. Ver. VI, e-ISSN 2278-0661, p-ISSN 2278-8727
5. Krishna BM, Khan H, Madhumati GL, Kumar KP, Tejaswini G, Srikanth M, Ravali P (2017) FPGA implementation of DES algorithm using DNA cryptography. *J Theor Appl Inform Technol* 95(10). ISSN 1992-8645
6. Ferozan BA, Mukhopadhyay D (2013) *Cryptography and network security*, 2nd ed. McGraw Hill Education (India) Private Limited

7. Rachmawati D, Tarigan JT, Ginting ABC (2018) A comparative study of Message Digest 5(MD5) and SHA256 algorithm. In: 2nd international conference on computing and applied informatics. IOP Publishing. IOP Conf Ser J Phys Conf Ser 978:012116
8. Pandya D, Narayan KR, Thakkar S, Madhekar T, Thakare BS (2015) Brief history of encryption. Int J Comput Appl 131(9). ISSN 0975-8887
9. Ali FN, Ahmed MM (2019) Chaotic block image scheme using large key space and message digest algorithm. Int J Comput Sci Sec (IJCSS) 13(4)
10. Singh G (2013) A study of encryption algorithms (RSA, DES, 3DES and AES) for information security. Int J Comput Appl 67(19)
11. Singh G, Kumar A, Sandha KS (2011) A study of new trends in blowfish algorithm. Int J Eng Res Appl
12. Nahar K, Chakraborty P (2020) A modified version of Vigenere cipher using  $95 \times 95$  table. Int J Eng Adv Technol (IJEAT) 9(5). ISSN 2249-8958
13. Khasanah, Nguyen PT, Gunawan G, Rahim R (2020) Three-pass protocol scheme on Vigenere cipher to avoid key distribution. J Crit Rev 7(1). ISSN 2394-5125
14. Maya Mohan MK, Devi K, Jeevan Prakash V (2015) Security analysis and modification of Cassical encryption scheme. Indian J Sci Technol 8(S8):542–548
15. Rahmani MdKI, Wadhwa N, Malhotra V (2012) LPHA-QWERTY cipher: an extended Vigenere cipher. Adv Comput Int J (ACIJ) 3(3)
16. Abdelwahab MM (2018) A compact cryptosystem design of triple-DES, U of K E J. 8(1):25–29
17. Omolara OE, Oludare AI, Abdulahi SE (2014) Developing a modified hybrid Caesar cipher and Vigenere cipher for secure data communication. Comput Eng Intell Syst 5(5). ISSN 2222-1719
18. Abood OG, Guirguis SK (2018) A survey on cryptography algorithms. Int J Sci Res Publ 8(7):495. ISSN 2250-3153
19. Wilson PI, Garcia M (2006) A modified version of the Vigenère algorithm. IJCSNS Int J Comput Sci Netw Sec 6(3B)
20. Varghesea PE, Nair LR, Paul V (2017) Security enhancement of English text using Prathala table of alphanumericals. Procedia Comput Sci 115:667–673
21. Kester Q-A (2012) A cryptosystem based on Vigenère cipher with varying key. Int J Adv Res Comput Eng Technol (IJARCET) 1(10). ISSN 2278-1323
22. Verma R, Sharma AK (2020) Cryptography: avalanche effect of AES and RSA. Int J Sci Res Publ 10(4). ISSN 2250-3153, <https://doi.org/10.29322/IJSRP.10.04.2020.p10013>
23. Rivest RL (1991) The MD4 message digest algorithm. In: Advances in cryology—CRYPTO'90, LNCS 537, pp 303–311
24. Mandal SK, Deepti AR (2016) A cryptosystem based on Vigenere cipher by using multilevel encryption scheme. Int J Comput Sci Inform Technol 7(4):2096–2099. ISSN 0975-9646
25. Vatshayan S, Haidri RA, Verma JK (2020) Design of hybrid cryptography system based on Vigenère cipher and polybius cipher. In: International conference on computational performance evaluation (ComPE), North-Eastern Hill University, Shillong, Meghalaya, India
26. Aung TM, Naing HH, Hla NN (2019) A complex transformation of monoalphabetic cipher to polyalphabetic cipher: (Vigenère-affine cipher). Int J Mach Learn Comput 9(3)
27. Subhashini V, Geethanjali N (2017) An enhanced approach on Vigenere cipher by polyalphabets. Int J Latest Trends Eng Technol 8(1):372–379. e-ISSN 2278-621X
28. Mittal VK, Mukhija M (2019) Cryptosystem based on modified Vigenere cipher using encryption technique. Int J Trend Sci Res Dev (IJTSRD) 3(5). e-ISSN 2456-6470
29. Stallings W (2000) Cryptography and network security, 4th edn. Prentice Hall, Pearson
30. Singh YK (2012) Generalization of Vigenere cipher. ARPN J Eng Appl Sci 7(1). ISSN 1819-6608

# A Multi-criteria Decision-Making Approach to Analyze Python Code Smells



Aakanshi Gupta , Deepanshu Sharma , and Kritika Phulli 

## 1 Introduction

A software with less modifications is generally recommended in the current software industry. It is estimated that approximately 75% of the whole expense is exhausted to upkeep the maintenance cost [5]. Nonetheless, there still exist some code smells which hinder its maintenance in terms of design issues. It may result in compromising the software quality and pose an extreme danger to development and modifiability of the software (technical debt) [10]. Code smells happen to degrade design of the code leading to deviations from the expected execution and hamper the quality and maintainability of the application [3, 4]. Factors such as time constraints, market pressure and limited budget make it burdensome for developers to remove all the code smells. Refactoring seems to be a promising approach to eliminate the code smell so to prevent deterioration of performance and efficiency [3]. Emphasizing on Python language, it has been seldom investigated particularly for code smell domain. Besides, it is a dynamically typed language and has great support for varied libraries.

This proposed research involves the common Python-based code smells that impede the software's performance. It is worth marking that it is necessary to prioritize these distinguished code smells against the abovementioned odd factors and refactor those smells first which hamper the software quality [23]. Consequently, the refactoring strategies depend upon the qualitative and quantitative against the

---

A. Gupta

Department of Computer Science and Engineering, Amity School of Engineering and Technology, Amity University Uttar Pradesh, Noida, India

D. Sharma (✉)

Executive Branch-IT, Indian Navy, New Delhi, India

e-mail: [sharma.deepanshu97@gmail.com](mailto:sharma.deepanshu97@gmail.com)

K. Phulli

Application Development Associate, Accenture, India

nature of code [13]. This effort calculation to estimate the type of smell to be refactored defines the prioritization of code smells. In this study, it is followed by using a purely statistical approach, MCDM—Multi-Criteria Decision-Making [15]. MCDM approach assists in making decisions on the basis of its significance in the existence of other objectives. In relevance of resolving conflicts, the VIKOR method is applied which helps in determining the compromise solution closed to the ideal one and ranks the alternatives [12, 26].

Additionally, other MCDM approaches such as TOPSIS—Technique of Order Preference Similarity to the Ideal Solution—TOPSIS, and WASPAS—Weighted Aggregated Sum Product Assessment have also been used. Both these approaches have been explored in this study in comparison to VIKOR's order ranking solution for a better analyses of code smells induced in python language. These techniques utilize a weight estimation strategy which determines the criteria for decision-making. Hence, two types of objective weight estimators have been practiced in this study, namely ENTROPY and CRITIC method [2, 14].

The motivation to prioritize the dominant code smells for efficient refactoring rather than refactoring them all at once helps the authors devise an order ranking of the code smells for Python software and extract the vital software metrics affecting the maintenance of the software, employing the MCDM techniques referenced above alongside their weight estimators.

#### *Research Contributions*

1. *To extract the vital software metrics indispensable for order-ranking of code smells induced in Python software utilizing machine learning for enhanced software quality.*
2. *To evaluate the objective weights using weight estimations methods—Shannon Entropy and CRITIC Method.*
3. *To propose an order ranking of code smells induced in Python software through a MCDM approach which provides a compromised solution—VIKOR method.*
4. *To devise an order-ranking among the Python code smells on the basis of various MCDM techniques (VIKOR, TOPSIS, WASPAS) to be practiced for efficient refactoring needs.*

The research has been structured in the following manner: Sect. 2 aligns the background and related work while the research study and design has been illustrated in Sect. 3. Section 4 details the experimental setup performed for the work and Sect. 5 discusses the final results of the research. The threats to validity are depicted in Sects. 6 and 7 presents the conclusions of the work and briefs the future scope of this work.

## 2 Related Studies

This section explains the literature review and analyzes the updated published research work in the field of software engineering concerning code smells. The prioritization of code smells consistently stays a subject of ongoing examinations in the literature. Code smells seem to be demonstration of design issues which can be refactored using different strategies [4]. Fowler proposed 22 code smells in his study which suggest some refactoring strategies alongside [4]. A study by Gupta et. al. suggested a SLR—Systematic Literature Review on bad code smells induced in Java language [25]. The Python programming area has been less tested, possibly because of its dynamic nature.

Most previous researchers chose the data set based on the Java programming language, according to Kaur et al., who recently examined the subject of prioritising code smells. Even so, they haven't run many statistical tests [23]. A study by Mareschal discovered the rankings obtained in accordance with the changes in the weight criteria [8]. For prioritizing smelly instances, a ranking method was defined using heuristic-based detection of code smells [9]. Besides from configuration defects, there are numerous different examinations where the use of MCDM strategies has been noticed. The popular supplier selected problem has been resolved by Onder and Dag utilizing AHP and improved TOPSIS MCDM strategies [11, 19, 20]. Sehgal et al. used Fuzzy TOPSIS methodology whereas Yang Wu studied the construction of the evaluation model in health welfare [13, 16]. A recent study adopts the MCDM technique for the mobile application development using agile [22, 24]. A similar work recently published focusing on the prioritization of python code smell has been studied by Gupta et al. based on VIKOR methodology of MCDM approaches [26]. An integrated approach comprising of TOPSIS and AHP method has also been performed [17].

## 3 Research Study and Design

### 3.1 Research Design

The research work helps in studying the order-ranking of code smells induced in Python software for realizing their priority to attain sustainable development. A glimpse of this entire study has been illustrated in Fig. 1. This study supports the application of MCDM approaches which further helps in evaluating the conflicting criteria using the objective approach. It also depends upon the intensity of code smells induced in Python software at class level. The primary progress in this examination discovers the following:

- A set of crucial software metrics which leads to software degradation resulting in code smells for Python classes.

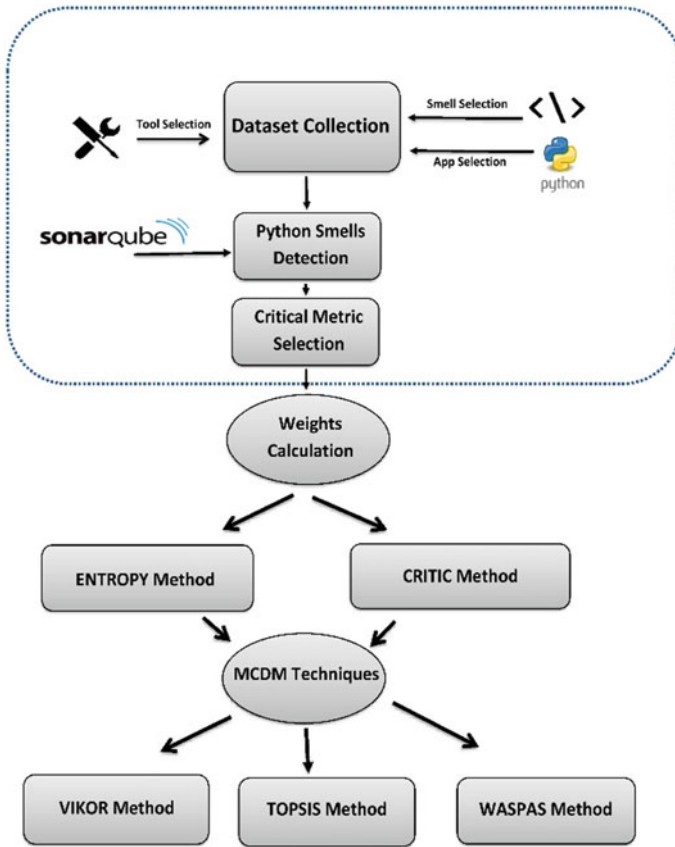


Fig. 1 Overview of research Study

- A comprised solution using the VIKOR method for code smells which are selected from a set of alternatives obtained from the order rankings in the existence of conflicting criteria (software metrics).
- The weight estimators required for the MCDM strategies using two weight determining criteria—*Shannon Entropy* and *CRITIC* method.
- The order ranking of the VIKOR technique further compared with other MCDM strategies—TOPSIS and WASPAS, utilizing both the weight criteria for order ranking calculations respectively.

This mathematical approach conforms the order-ranking of the Python code smells and represents the manner in which they should be refactored at the earliest and given utmost preference for significant maintainability enhancements.

**Table 1** Code smells induced in python software

Code smells	Description
Cognitive complexity	Cognitive complexity of a function should not be too high
Collapsible “IF”	Nested “IF” statements should be merged/collapsed in one if block
Many parameter	There should not be too many parameters in functions, methods and lambdas
Unused variable	The local variables which are unused should be removed
Naming conventions	The naming of class, function, method name should comply with the naming conventions

### 3.2 Context Selection

This section briefs about the data set collection and selection of Python software system along with its detection mechanism.

*Selection of Python Software:* This research has been conducted using open-source Python software which is extracted from the GITHUB repository. The following criteria has been applied for their cloning:

- Most starred filter of GITHUB with at least 15 K Line of Code (LOC)
- Python software having contrasting scopes and sizes and developed by different communities and are still in maintenance phase by Python developers.

Conclusively, 10,552 Python classes from a set of 20 Python applications alongside 6817 files were evaluated for determining the prioritization order using statistical techniques.

*Selection and Detection of Code Smell in Python Software:* The detection and prioritization of codes smells have been performed which contributes to the technical debts to the software. For detection of such code smells, SonarQube tool (<https://www.sonarqube.org/>) has been used. This study takes into account five code smells which were further detected by SonarQube, an open-source software inspection tool. It is worthy to highlight that the code smells detected using SonarQube does not align with Fowler’s known code smells [3]. The description of code smells has been listed in Table 1.

## 4 Experimental Setups

### 4.1 Data Pre-processing

A total of 10,552 classes have been extracted for five code smells diffused in 20 Python applications. A progressive approach has been practiced utilizing Machine



Learning Classification for evaluating the critical software metrics affected by code smells. Further, UNDERSTAND (<https://scitools.com/>), a static software analyzing tool has been employed to provide a set of 37 metrics. These metrics depict the conflicting criteria (multi-criteria) for deciding the prioritization of code smells. The next step accounts for selecting essential features to filter out the vital metrics against the detected code smell before classifying the data. The Rank searching approach used was Information Gain feature selector which minimizes the complexity of the detection rules.

Applying the feature selectors, a set of metrics were obtained which were later used in the rule-based approach of supervised learning. The metrics obtained after rule-based classifier approach empowered the criterion which determines the significant software metrics severely impacting the software systems. It is worth noting that machine learning rule-based classifiers have been applied. A paired t-test has been applied which resulted in JRIP classifier as the accurate one with an accuracy level of 89.81%.

## 4.2 Weight Estimation

The Multi-Criteria Decision-Making methods always require a weight as a part of its computations in association to each criterion (i.e., software metrics). Every metric is associated with a weight computed using different weight estimation techniques. In respective to this research, this procedure involves two weight estimators:

*Shannon Entropy* method and *CRITIC* method.

### *Shannon Entropy Method*

The Shannon Entropy is a fundamental quantity in information theory. It is based upon the amount of uncertainty in information determined by the probability theory [14].

When determining a proper weight for MCDM problem based on ranking, it is one of the known strategies [1, 6].

The formula for Computation of Shannon Entropy is as follows:

$$E_j = -h \sum_{i=1}^m r_{ij} \ln r_{ij} \quad j = 1, 2, \dots, n \quad (1)$$

where,

$m$  = number of code smells; ( $m = 5$ );

constant “ $h$ ” guarantees that  $E_j(j = 1, 2, \dots, n)$  lies in interval  $[0, 1]$ .

The formula for computation of weight vector is as follows:

$$w_j = \frac{1 - E_j}{\sum_{j=1}^n (1 - E_j)} \quad j = 1, 2, \dots, n \tag{2}$$

$1 - E_j$  = the degree of diversity of the information for the  $j$ th criterion.

*CRITIC Method*

(Criteria Importance Through Intercriteria Correlation)

This strategy establishes objective decision-making weights. For the MCDM problem, this method incorporates a conflict in the criteria and the severity of the contrast. It determines the information in the criterion by analyzing the variants in the decision matrix using analytical testing [2, 7]. Correlation analysis is used to look for differences between the criteria.

The formula for determining the quantity of the information concerning each criterion is as follows:

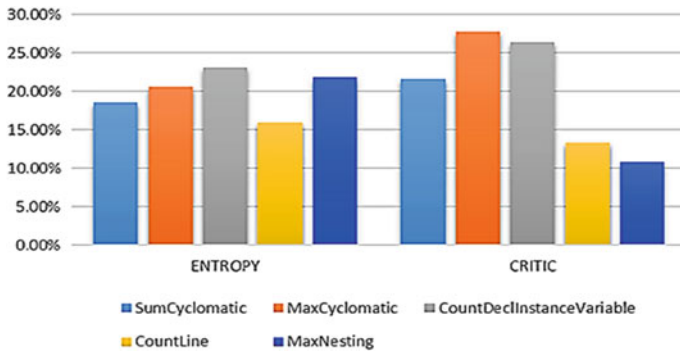
$$C_j = \sigma_j * \sum_{k=1}^m (1 - r_{jk}) \tag{3}$$

where  $C_j$  = quantity of information contained in  $j$ th criterion.

The formula for computation of weight is as follows:

$$W_j = \frac{C_j}{\sum_{k=1}^m C_j} \tag{4}$$

The above stated methods helped in computing weights for each software metric and defined below in Fig. 2.



**Fig. 2** Objective weights of software metrics

### 4.3 MCDM Techniques

Once the weights for the software metrics have been derived, MCDM techniques such as *VIKOR*, *TOPSIS* and *WASPAS* have been used in this study. These methods are elaboratively explained below:

#### *VIKOR Method*

The *VIKOR* methodology has proven to be an effective method for dealing with a wide range of judgment criteria. It proposes a multi-criterion optimal compromise ranking solution alongside an order ranking of code smells which resulted in an agreement between the code smells on mutual basis. The compromise ranking is obtained by evaluating its close alternative to the ideal one, using a criterion function to evaluate each alternative.

Applying the appropriate weights [12] has resulted in the following steps:

*Step 1:* Analyze the best and the worst value for each criterion. Due to the comparator of software metrics obtained using machine learning classifier, this study involves only beneficial criteria:

$$X_i^+ = \max(x_{ij}) - \text{Best Value}; X_i^- = -\min(x_{ij}) - \text{Worst Value}$$

*Step 2:* Using the Eq. (5), the values  $S_i$  and  $R_i$  are computed for every criterion for each code smell using weights of the criteria  $W_j$ .

$$S_i = \sum_{j=1}^m \left( W_j * \frac{X_i^+ - X_{ij}}{X_i^+ - X_i^-} \right) \quad (5)$$

$$R_i = \max_j \left( W_j * \frac{X_i^+ - X_{ij}}{X_i^+ - X_i^-} \right) \quad (6)$$

*Step 3:* Using Eq. (7), the value of  $Q_i$  is computed.

$$Q_i = \mu * \frac{S_i - S^*}{S^- - S^*} + (1 - \mu) * \frac{R_i - R^*}{R^- - R^*} \quad (7)$$

where,

$$S^* = \min_i S_i, S^- = \max_i S_i$$

$$R^* = \min_i R_i, R^- = \max_i R_i$$

$$\mu = 0.5 \text{ (considered).}$$

*Step 4:* Based on the  $Q$ ,  $R$  and  $S$  values, the code smells are ranked in descending order. Table 2 describes the ranking obtained.

**Table 2** Order preferencing of python-based code smells using VIKOR technique

Python code smells	$S_i$		$R_i$		$Q_i$		Order ranking
	ENTROPY	CRITIC	ENTROPY	CRITIC	ENTROPY	CRITIC	
Cognitive complexity	0.15	0.16	0.06	0.07	0.038	0.03	2nd
Collapsible 'IF'	0.08	0.1	0.06	0.07	-2.11E-07	1.14E-06	1st
Many parameter	0.27	0.26	0.07	0.1	0.12	0.15	3rd
Naming convention	1	1	0.23	0.27	1	1	5th
Unused variable	0.66	0.67	0.16	0.2	0.61	0.63	4th

*Step 5:* After rankings are obtained, following two conditions are checked to acquire a compromised solution.

$C_1$ : Acceptable advantage

$$Q(A^2) - Q(A^1) \geq DQ \tag{8}$$

where  $DQ = \frac{1}{j-1}$ ,  $j = \text{number of alternatives} = 5$ .

$C_2$ : Acceptable stability in decision-making:

It states that the alternative  $A^1$  must also be ranked by  $S$  or/and  $R$ .

The above conditions help in discovering a compromise solution and states that it is stable for the given decision-making process by majority rules through the voting.

The  $C_2$  conditions support the data set studied in this research but disagrees with  $C_1$  condition.

*TOPSIS Method*

TOPSIS (Technique of Order Preference Similar to the Ideal Solution) is used for comparing the set of considered code smells by recognizing the weights for each criterion (done using both ENTROPY and CRITIC). It is the second method used in accordance to VIKOR method.

*Principle:* The selected alternative should have the shortest geometric distance from the positive ideal solution and the longest geometric distance from the negative ideal solution.

The following procedure has been followed to evaluate the order preferencing using TOPSIS mechanism [12, 17, 20]:

*Step 1:* A normalized decision matrix is calculated using vector normalization for every element and then multiplied by the weights of each criterion (software metric) with the obtained normalized value.

*Step 2:* The ideal best ( $V_j^-$ ) and the ideal worst value ( $V_j^+$ ) are determined for each criterion.

*Step 3:* The separation measures (L2—norm distances) are evaluated using the Eqs. (9) and (10) from the ideal best ( $S_j^+$ ) and the ideal worst values ( $S_j^-$ ) using Euclidean distance.

$$S_j^+ = \left[ \sum_{j=1}^m (V_{ij} - V_j^+)^2 \right]^{0.5} \quad i = 1, 2, \dots, m. \tag{9}$$

$$S_j^- = \left[ \sum_{j=1}^m (V_{ij} - V_j^-)^2 \right]^{0.5} \quad i = 1, 2, \dots, m. \tag{10}$$

*Step 4:* The relative closeness to the ideal solution is calculated using the Eq. (11)

$$P_i = \frac{S_i^-}{S_i^+ + S_i^-} \quad 0 \leq P_i \leq 1 \tag{11}$$

*WASPAS Method*

The third approach used in this study is a combined method of WASPAS which proves to be an optimization method for multi-criteria decision-making approaches. This method is based on two criteria of optimality [18, 21]. ([Weighted Sum Method]; [WPM: Weighted Product Method]) technique,

*Step 1:* The weighted sum using WSM ( $S_i$ ) and weighted product using WPM ( $P_i$ ) from the decision matrix is calculated from the Eqs. (12) and (13).

$$S_i = \sum_{j=1}^n w_j x_{ij} \tag{12}$$

$$P_i = \prod_{j=1}^n x_{ij}^{w_j} \tag{13}$$

*Step 2:* The joint generalized criterion is calculated using WSM and WPM of weighted aggregation using the Eq. (14)

$$Q_i = \lambda S_i + (1 - \lambda) P_i; \quad \lambda = 0.5 \tag{14}$$

*Step 3:* Based on the  $Q_i$  values, rank the code smells.

**Table 3** Order preferencing of python code smells for TOPSIS and WASPAS technique

Python code smells	TOPSIS		WASPAS		Order ranking
	ENTROPY	CRITIC	ENTROPY	CRITIC	
Cognitive complexity	0.8	0.79	0.88	0.87	2nd
Collapsible ‘IF’	0.85	0.83	0.93	0.92	1st
Many parameter	0.72	0.73	0.79	0.8	3rd
Naming convention	0	3.52E-09	0.25	0.25	5th
Unused variable	0.33	0.31	0.49	0.49	4th

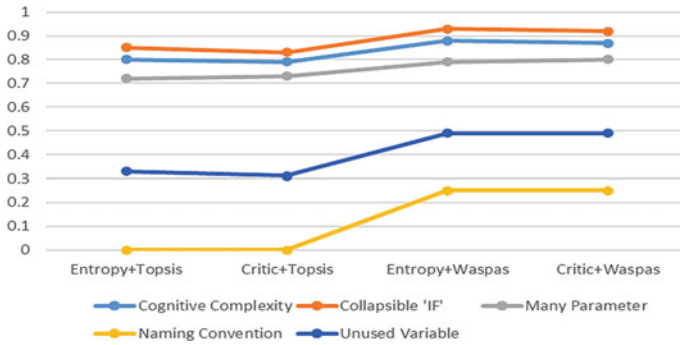
The combined preference scores for TOPSIS and WASPAS method are stated in Table 3.

## 5 Results and Discussion

This investigation devised an order-ranking for code smells induced in Python software by the application the Multi-Criteria Decision-Making (MCDM) strategies. The techniques using MCDM approaches help in better decision-making in various software domains. This research has explored techniques like VIKOR, WASPAS and TOPSIS alongside two weight estimators, i.e., *Shannon Entropy* and *CRITIC* method. A set of 20 Python software with a total of 10,552 classes were extracted from GITHUB repository. SonarQube tool has been preferred for code smell detection, detecting around 5 code smells in the name of technical debts.

Using a static analyzer tool, the software code has been analyzed for static software metrics. These metrics have been treated as a multi-criterion for procuring the vital software metrics (multi-criteria) and as a decision matrix for application of MCDM techniques. The vital software metrics has been selected using the machine learning rule-based classification method. Paired-t-test was applied to acquire an accurate algorithm. This test resulted into a highest obtained accuracy of 89.91% for JRIP algorithm, which generates detection rules. The obtained software metrics indicates the supreme existence of code smells labeled as: *MaxCyclomatic*, *CountDeclInstanceVariable*, *CountLine*, *SumCyclomatic*, *MaxNesting*.

It is significant that the acquired software metrics tend to pose a beneficial criterion for determining the order-ranking of the code smells. After the vital metric extraction, the subsequent stage helps in estimating the wrights of these metrics. Two methods namely, *Shannon Entropy* and *CRITIC*, were employed to ensure diversification, as discussed above in Sect. 4.2. Focusing on the statistics, the **Entropy** estimator reports the **highest** weight of **23.06%** for the metric, **CountDeclInstanceVariable** and the **least** weight of **15.92%** for the metric **CountLine**. In contrast, the **MaxCyclomatic** metric achieved the maximum weight of **27.78%**, whereas the **MaxNesting** metric weighed the least up to **10.82%** for the CRITIC estimator. Once the weights were



**Fig. 3** Visualization of order preference scores of TOPSIS and WASPAS method

calculated, the indicated MCDM techniques have been practiced. The methods have been explained in Sect. 4.3.

As discussed, a compromise solution is also devised in VIKOR method. For this study, the compromise solution sums up the following: A set of 3 code smells induced in Python software: Collapsible “IF” with Rank 1, Cognitive Complexity with Rank 2 and Many Parameters with Rank. This states that if the first ranked smell is non-existence, it is advised to refactor the second ranked smell. It is worth mentioning that both weight-determining methods emerged out to result in the identical order-preference rankings, despite having different performance score. The results from VIKOR method urged the authors to apply other MCDM techniques such as TOPSIS and WASPAS to acquire a firm ranking and proposing the finest order preferencing solution for the considered smells. Utilizing both the weight estimation methods, TOPSIS and WASPAS resulted in the indistinguishable (identical) order preference as acquired by the VIKOR method.

Nonetheless, the performance scores contrast for both the strategies just as the weight estimations.

It has been concluded that considered five code smells have identical order-ranking, though contrasting performance scores when evaluated through all the three MCDM methods utilizing both the weight estimators. The order preferencing scores are depicted in Fig. 3.

## 6 Threats to Validity

This section explains some of the limitations of the study that were encountered while examining the entire research study. The outcomes of the order-ranking acquired may vary when taking other smells into considerations. The deflections in the ranking may occur but some what it should continue to follow a similar ordering. The quality and quantity of the data set may pose some threats to validity. The enhancements in the mathematical formulas might affect the order-ranking of code smells achieved in

this study. Also, it is worth mentioning that the order-ranking achieved is currently based on the static code metrics. The results might vary when different nature of code metrics is considered.

## 7 Conclusion and Future Scope

This research accounts to order the code smells diffused in the popular language, Python to develop a sustainable software. The approach to prioritize the code smells aids the refactoring process to be more efficient and useful for further optimized coding solutions. The optimization in Python-based software is achievable by eliminating the design flaws, i.e., code smells which were earlier degrading the software quality. These mentioned flaws relate to the domain of “Technical Debt,” delaying the maintenance of the software in further developments. The detection of these technical debts has been accomplished by the SonarQube tool. This study intends to practice the MCDM technique to acquire the prioritization order for the considered five Python-based code smells. Along with the order preferencing, the authors were capable of extracting the vital software metrics which are harmful to the software and essential for code smells order preferencing measures.

This analysis involves 20 Python applications (10,552 classes) and 5 generic code smells. This examination extracts around 35 software metrics through UNDERSTAND (static code analyzer) and then discovers the most essential software metrics using the machine learning classifier based on propositional rule learner algorithm. This analysis yields a set of 5 vital software metrics along with a comparator which decides the beneficial or non-beneficial criteria. The weight estimators used are *Entropy* method and *CRITIC* method. The weight estimators triggered two contrasting percentages for each software metric. It was further utilized in MCDM techniques for their respective computations. All the MCDM techniques applied in this study were practiced in the combination of both weight estimation approaches. It was decided that the VIKOR method should be computed formerly to obtain the order preferencing ranks of the desired code smells and a set of compromise solutions. Subsequently, order preferencing of TOPSIS and WASPAS techniques were opted for prioritizing the code smells to ensure a proper order to realize the refactoring needs for smell rectification purposes. It was astonishing and encouraging that the other two methods resulted in the identical order-rankings for the code smell as obtained by the VIKOR method, although the performance scores differ. This consistent ranking of code smells using three MCDM techniques in combination with two objective weight estimators enhanced the efficiency of the proposed order preferencing.

The above devised order-ranking narrows down the expensive refactoring efforts, reducing time and saving efforts of the developer and maintenance team to deliver high quality software products. This ranking can be clearly observed through the Fig. 4. The conclusions obtained can be further analyzed for facilitating the Python developers for better decision-making to maintain better software quality software.





**Fig. 4** Order-ranking of python code smells

The similar approach can be practiced for different languages. Furthermore, this analysis can be carried out taking in account different code smells and realizing their intensity for empowering better decision-making.

## References

1. Choo EU, Wedley WC (1985) Optimal criterion weights in repetitive multicriteria decision-making. *J Oper Res Soc* 36(11):983–992
2. Diakoulaki D, Mavrotas G, Papayannakis L (1995) Determining objective weights in multiple criteria problems: the critic method. *Comput Oper Res* 22(7):763–770
3. Fowler M (1997) Refactoring: improving the design of existing code. In: 11th European conference, Jyväskylä, Finland
4. Fowler M (2018) Refactoring: improving the design of existing code. Addison-Wesley Professional, 20 November 2018
5. Lenarduzzi V, Saarimäki N, Taibi D (2020) Some SonarQube issues have a significant but small effect on faults and changes. A large-scale empirical study. *J Syst Softw* 170:110750
6. Ma J, Fan ZP, Huang LH (1999) A subjective and objective integrated approach to determine attribute weights. *Eur J Oper Res* 112(2):397–404
7. Madic M, Radovanović M (2015) Ranking of some most commonly used nontraditional machining processes using ROV and CRITIC methods. *UPB Sci Bull Ser D* 77(2):193–204
8. Mareschal B, Brans JP, Vincke P (1984) PROMETHEE: a new family of outranking methods in multicriteria analysis. ULB—Universite Libre de Bruxelles
9. Marinescu R (2012) Assessing technical debt by identifying design flaws in software systems. *IBM J Res Dev* 56(5):9–1
10. Narasimhan VL (2008) A risk management toolkit for integrated engineering asset maintenance. *Aust J Mech Eng* 5(2):105–114
11. Onder E, Sundus DA (2013) Combining analytical hierarchy process and TOPSIS approaches for supplier selection in a cable company. *J Bus Econ Finan* 2(2):56–74
12. Opricovic S (1998) Multicriteria optimization of civil engineering systems. 2(1):5–21. Faculty of Civil Engineering, Belgrade
13. Opricovic S, Tzeng GH (2004) Compromise solution by MCDM methods: a comparative analysis of VIKOR and TOPSIS. *Eur J Oper Res* 156(2):445–455
14. Sehgal R, Mehrotra D, Bala M (2018) Prioritizing the refactoring need for critical component using combined approach. *Decis Sci Lett* 7(3):257–272
15. Shannon CE (1948) A mathematical theory of communication. *Bell Syst Tech J* 27(3):379–423
16. Tsantalos N, Chatzigeorgiou A (2011) Ranking refactoring suggestions based on historical volatility. In: 2011 15th European conference on software maintenance and reengineering, 1 March 2011. IEEE, pp 25–34

17. Vidal SA, Marcos C, Díaz-Pace JA (2016) An approach to prioritize code smells for refactoring. *Autom Softw Eng* 23(3):501–532
18. Wei CC, Chien CF, Wang MJ (2005) An AHP-based approach to ERP system selection. *Int J Prod Econ* 96(1):47–62
19. Yilmaz B, Harmancioglu N (2010) Multi-criteria decision making for water resource management: a case study of the Gediz River Basin, Turkey. *Water SA* 36(5)
20. Zaidan AA, Zaidan BB, Al-Haiqi A, Kiah ML, Hussain M, Abdulnabi M (2015) Evaluation and selection of open-source EMR software packages based on integrated AHP and TOPSIS. *J Biomed Inform* 1(53):390–404
21. Zavadskas EK, Turskis Z, Antucheviciene J, Zakarevicius A (2012) Optimization of weighted aggregated sum product assessment. *Elektronika ir elektrotechnika* 122(6):3–6
22. Aljuhani A, Alhubaishy A (2020) incorporating a decision support approach within the agile mobile application development process. In: 2020 3rd international conference on computer applications and information security (ICCAIS), 10 March 2020. IEEE, pp 1–6
23. Kaur A, Jain S, Goel S, Dhiman G (2021) Prioritization of code smells in object-oriented software: a review. *Mater Today Proc*
24. Aljuhani A, Alhubaishy A (2020) Incorporating a decision support approach within the agile mobile application development process. In: 2020 3rd international conference on computer applications and information security (ICCAIS), pp 1–6
25. Gupta A, Suri B, Misra S (2017, July) A systematic literature review: code bad smells in Java source code. In: International conference on computational science and its applications. Springer, Cham, pp 665–682
26. Gupta A, Sharma D, Phulli K (2022) Prioritizing python code smells for efficient refactoring using multi-criteria decision-making approach. In: international conference on innovative computing and communications. Springer, Singapore, pp 105–122

# The Importance of Validation Metrics in Chaotic Image Encryption



M. Y. Mohamed Parvees and T. Vijayakumar

## 1 Introduction

The Confidentiality, Integrity and Availability (CIA) called CIA triad components in information security are to be preserved to achieve secure communication during this 'Information Age'. Particularly, the confidentiality is becoming one of the predominant components in secure communication and storage. The enormous development of communication technologies like cloud and IoT are communicating varieties of data to a large extent. Thus, the varieties of data are served and stored between various people, software and hardware. Among those data, the images need remarkable and difficult way to protect them. The images could be secured by providing the confidentiality to images. In order to provide confidentiality, the various crypto schemes are proposed by several researchers [1–8].

The review of literate evidence that the chaos-based block cipher cryptosystem is assessed by histogram, correlation coefficient and entropy analyses [9]. Bashir et al. [10] used correlation coefficient, histogram and data entropy to measure the randomness of the enciphered image which has been encrypted by Advanced Encryption Standard (AES). Wu et al. [11] also proposed a local entropy analysis for measuring the entropy of the variety of portions in the paired image and thereby, the local entropy becomes vivid measure for estimating the efficiency of an image encryption algorithm. Similarly, Praveenkumar et al. [12] encrypted the grey-scale images and inspected it with different new metrics such as chi-square analysis and local entropy analysis. Then, many researchers discuss the importance of various metrics

---

M. Y. Mohamed Parvees (✉) · T. Vijayakumar  
Department of Computer and Information Science, Faculty of Science, Annamalai University,  
Annamalai Nagar, Chidambaram 608002, India  
e-mail: [yparvees@gmail.com](mailto:yparvees@gmail.com)

T. Vijayakumar  
e-mail: [vijimca17@gmail.com](mailto:vijimca17@gmail.com)

such as Correlation, Coefficient, Net Pixel Change Rate (NPCR), Unified Average Changing Intensity (UACI), histogram analysis and chi-square analysis, etc. [13, 14]. All these metrics are useful for finding the efficiency of encryption algorithm in terms of randomness. At the same time, it is essential to consider few metrics like key sensitivity analysis, key space analysis and encryption time in terms of cipher key and complexity of the algorithm. Many researchers use these metrics as standard to evaluate their cryptosystem [15–17]. It is always essential to understand the roles of metrics in chaotic encryption.

Among the various crypto schemes, the chaos-based crypto algorithms are particularly robust towards images. The efficiency of the chaotic cryptosystem is measured by the ability to withstand different types of security attacks on encrypted images and randomness of image pixels. The withstanding ability and randomness of encrypted images are measured through the various metrics correlation coefficient, entropy, NPCR, UACI, key space, key sensitivity, randomness test suites, etc. [18, 19]. Therefore, the chaotic scrambling and diffusion are proposed with an improved logistic map [20] to encrypt a 24-bit colour image and discussed with few existing and proposed metrics. Ultimately, the research work is proposed to study the importance of various metrics and their roles in chaotic cryptosystem. The organisation of this article comprises of an introduction, mathematical background on chaotic cryptosystem, methodology adopted for chaotic cryptosystem, security analyses through validation of metrics and conclusion sections.

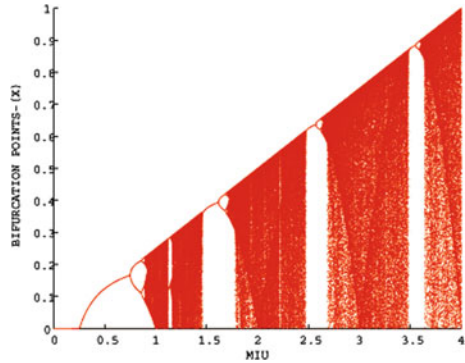
## 2 Mathematical Background

In this research, the chaotic equations are used as pseudorandom sequence generators. The chaotic Eq. (1) produces the chaotic random sequences. This is the one-dimensional improved logistic map equation. That is, the equation has been derived from the logistic map. The improved logistic map is producing better sequences in terms of randomness. From the literature, it could be observed that when the bifurcation is wider, then the equation could be used in encryption. Similarly, the improved logistic map has positive Lyapunov exponent (Figs. 1 and 2). Based on the above findings, this improved logistic map is used for this study to the role of metrics on cryptosystem.

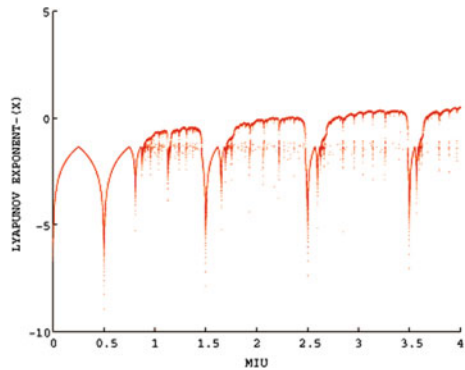
Equation (1) produces larger bifurcation range than logistic map. The key for the cryptosystem should be chosen from bifurcated range in order to provide better security. The key space is also higher due to the enhancement of logistic map. One more input parameter has been added to basic logistic equation which largely resists brute-force attack.

$$x_{n+1} = r \times \left[ \frac{\left[ \frac{x_n}{v} \right] \times v - x_n}{v} \right] \times \left\{ 1 - \left[ \frac{\left[ \frac{x_n}{v} \right] \times v - x_n}{v} \right] \right\} \quad (1)$$

**Fig. 1** The bifurcate diagram of improved logistic equation



**Fig. 2** The Lyapunov diagram for improved logistic equation



where,  $x_n$ —Initial parameter and  $r, v$ —Control parameters.

Equation (1) produces chaotic random sequences when  $r$  lies between 0 and 4,  $v$  lies between 0 and 0.3,  $x_n$  lies between 0 and 1. Figures 1 and 2 illustrate the larger bifurcation and positive Lyapunov exponent [20].

### 3 Methodology

In order to validate the various metrics involved in image encryption, the improved logistic map is employed to produce chaotic random sequences. The primitive operations used in the chaotic encryptions are permutation and diffusion. Using these random sequences, the permutation and diffusion are done to encrypt the images using the algorithm proposed in [9]. Equation (1) embodies the improved logistic map which is functional to form a non-linear dynamical structure. Sometimes, it is essential to validate the randomness of these permutation and diffusion sequences. That is, the cryptographically secure random sequences should be engaged in cryptography to achieve cent percent security [21]. This is also considered as a metric

to evaluate the randomness of the sequences produced or involved in the chaotic encryption. Further, the requirements for PRNGs are very high in various domains.

### **Algorithm 1 Steps involved in chaotic image encryption**

Step 01: Supply the source image with width  $M$  and height  $N$  and find  $l = M \times N$ .

Step 02: Read the pixels from the source image ( $256 \times 256$ ) and store them in one-dimensional array  $A$ .

Step 03: Split  $A$  pixels of length  $l$  into small  $n$  blocks.

Step 04: Store each block with indices.

Step 05: Produce a permutation sequence  $S_1$  for the length  $l = n$ .

Step 06: Produce a diffusion sequence  $S_2$  where  $S_i = \text{int} \left[ \left( \frac{\text{Seq}_i}{n_{\max}} \right) \times 255 \right]$ .

Step 07: Shuffle  $A$  pixels using the permutation sequence  $S_1$  and diffuse them using a bitwise XOR operations between  $A$  and  $S_2$ .

Step 08: Similarly, create permutation  $S_3$  and diffusion  $S_4$  sequences using the same Eq. (1).

Step 09: Shuffle the pixels again using permutation sequence  $S_3$  and execute a bitwise XOR operation between  $A$  and  $S_4$ .

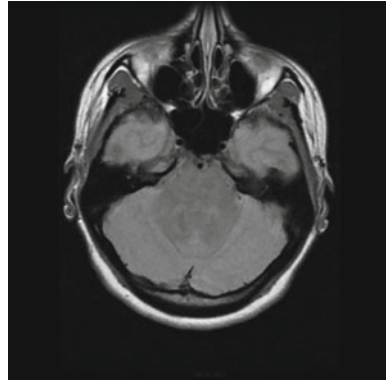
Step 10: To encrypt the image pixels completely,  $R$  rounds should be done between the steps 7 to 9 and get the cipher image  $I$  as the resulting pixels.

After the primitive operations, namely, permutation and diffusion, the pixel values of image are encrypted and the cryptosystem produces the cipher image from the plain image as given in Figures 3 and 4. The pixel locations are shuffled in permutation and pixel values are altered in diffusion using the chaotic random sequences. Hence, the act of shuffling and diffusion primarily depends on the chaotic random sequences produced by improved logistic map as mentioned in the Algorithm 1. Finally, the complete encrypted image is obtained as an output from the cryptosystem. The encrypted image does not reveal any useful information since it underwent chaotic scrambling and diffusion.

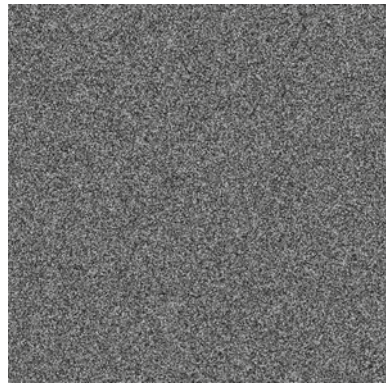
## **4 Security Analyses**

In experimental setup, this study is carried out on an Intel Core i5 2.6 GHz system with 8 GB RAM to encrypt the 8-bit grey-scale image with the size of  $256 \times 256$

**Fig. 3** Plain DICOM image



**Fig. 4** Encrypted DICOM image



(Fig. 3). During the encryption process, the idea is to break the relativity among the image pixels which are redundant in nature. The algorithm is coded using Java JDK 1.8.0 (64 bit). The symmetric keys for cryptosystem are selected from the control and initial parameters ( $x_n = 0.8000000000000005$ ,  $v = 0.2000000000000005$  and  $r = 3.9000000000000005$ ) of the improved logistic equation. The enciphered image is validated towards various metrics after chaotic scrambling and diffusion (Fig. 4). The most important portion of this work is to examine the aspect and roles of different metrics used to validate the chaotic cryptosystem. In this section, initially, the purviews of metrics involved for verifying encrypted pixel randomness have been studied. Secondly, the key strength of the cryptosystem is analysed through key space and key sensitivity analysis. Thirdly, the time complexity is analysed through the encryption time in milliseconds.

### 4.1 Correlation Coefficient Analysis

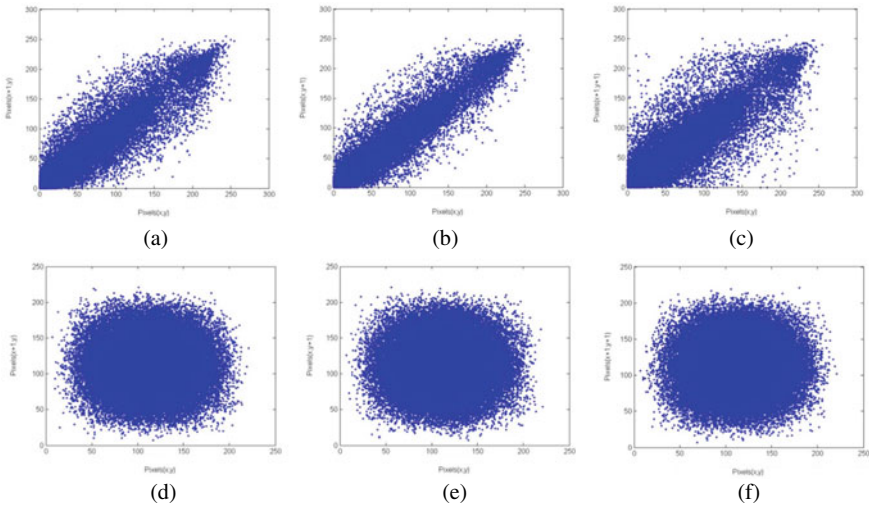
The correlation coefficient is used to examine the relations between the pixel pairs. The pixels are chosen randomly, at ten thousands adjust positions and the correlation value is calculated using the Eq. (2).

$$\gamma_{u,v} = \frac{\sum (u_i - \bar{u})(v_i - \bar{v})}{\sqrt{\sum (u_i - \bar{u})^2} \sqrt{\sum (v_i - \bar{v})^2}} \tag{2}$$

where,  $u$  and  $v$  are the pixel pair and  $\gamma$  is the correlation coefficient of the image pixels.

A good cryptosystem breaks the correlation among the pixel to get an encrypted image. Therefore, the value of the correlation coefficient should be nearing to zero (0) for an encrypted image. If the correlation value is nearing to one (1), the image pixels are not random which indicate that the cryptosystem has not adequately encrypted the image. The correlation coefficients are calculated for encrypted image in three different directions, that is, horizontal, vertical and diagonal correlations (Fig. 5).

Table 1 shows the correlation values of plain and encrypted image. The plain image has similarities among the pixels. Therefore, Table 1 indicates high correlation, that is, the values nearing to ‘1’. After chaotic encryption, the bonds between the pixels are broken using random shuffling and diffusion which lead to less correlation, that is, the values are nearing to ‘0’. This shows chaotic encryption makes the image pixels to random values. From, the encrypted image, the third person could not



**Fig. 5** The horizontal, vertical and diagonal correlation coefficient between neighbouring pixels representation of **a–c** for plain images and **d–f** for cipher images



**Table 1** The correlation coefficient of two adjacent pixels of plain image

Metrics	DICOM plain image (256 × 256)			DICOM Cipher image (256 × 256)		
	HC	VC	DC	HC	VC	DC
Correlation Coefficient	0.9656	0.9766	0.9472	0.0019	-0.0236	0.0799
Entropy	1.4999			7.9995		
Local entropy	1.5987			7.9977		
NPCR	99.6721					
UACI	33.3842					
MSE	121.2524					
PSNR	28.1257					
Mean–variance	27.1241			64.2120		
Chi-square	25,698.61			281.97		
Key space	$2^{384}$					
Encryption time	14.55 Mbit/s					

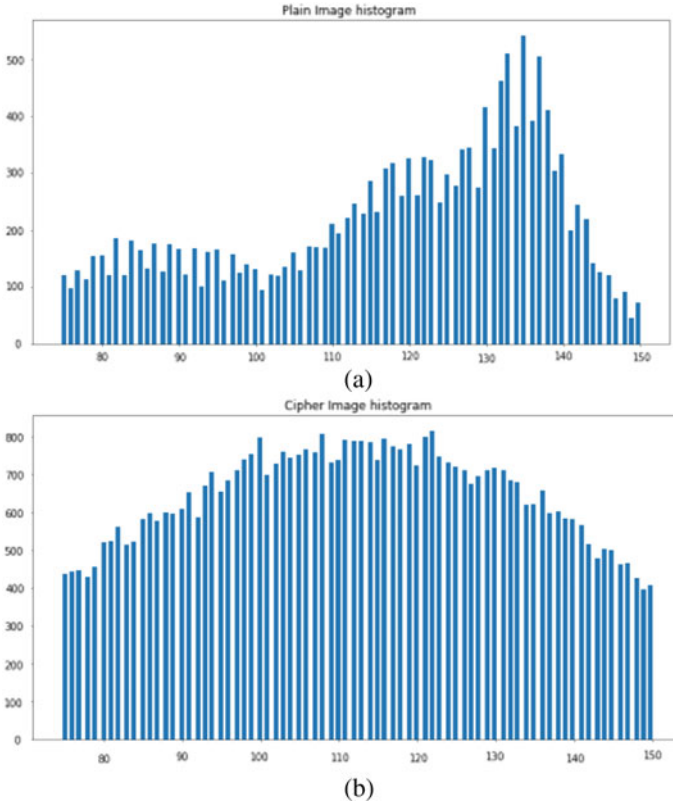
reveal any information. Similar results are found in the literatures also [12, 13, 17, 20]. The correlation coefficient is an effective metric to measure the cryptosystem with respect to the randomness of the image.

### 4.2 Histogram Analysis

Histogram depicts the graph-based representation of image pixel distribution. Figure 6a, b plot the pixel distribution of plain and cipher image, respectively. Usually, the adversary tries to attempt to infer information from the histogram of the image. If the chaotic encryption makes the distribution flat as shown in Fig. 6b, then the adversary could not infer any information from the encrypted image. That is, the chaotic encryption algorithm should flatten the pixel distribution. This clarifies that the histogram analysis is one of the measures used to evaluate the encrypted image in order to protect it from statistical attack.

### 4.3 Information Entropy

In 1949, Shannon proposed the information entropy. It is a statistical measure to find the degree of uncertainties in the system. That is, it tells the randomness. The entropy ( $H$ ) of a system  $m$  is calculated using the Eq. (3).



**Fig. 6** Histogram **a** original image, **b** cipher image

$$H(m) = \sum_{i=0}^{255} p(m_i) \log\left(\frac{1}{p(m_i)}\right) \quad (3)$$

where,  $m$  denotes pixels of an image  $m_i \in m$  and  $p(m_i)$  denotes the probability of occurrence ' $m_i$ '. In this case, the image has 8-bit information. Therefore, the ideal entropy value for encrypted image should be 8 nearing to 8 which informs that all the bits in each pixels have some value. The adversary will try to do entropy attack by looking at the varying values in each bit. In this case, the adversary could not predict the information from the encrypted image. Table 1 lists the entropy value of plain and cipher image and it is comparable with literature [12, 18–20]. The chaotic encryption makes the entropy value of cipher image higher than the plain image.

#### 4.4 Local Shannon Entropy Analysis

The local Shannon entropy is also an effective measure to check the randomness locally at various blocks of an image. It is considered as an effective analysis than global entropy analysis because the analysis has been carried out on cipher image pixels block by block. The idea of this metric is to test the randomness in local view-points. Further, the local entropy is calculated even among the overlapping blocks. It overcomes the demerits of global entropy in terms of inaccuracy, inconsistency and low efficiency. This local entropy has been calculated by selecting the random blocks in images and results listed in Table 1. The results coincide with the literatures [11, 12].

#### 4.5 Differential Attack Analysis

The adversary will try to reveal the information using differential attack. That is, the attacker change the pixel values, then encrypt and try to form the patterns among the changing pixel values thereby trying to guess the required plain image. Hence, the chaotic cryptosystems are tested towards this attack using the metrics, called Number of Pixels Change Rate (NPCR) and Unified Average Changing Intensity (UACI). NPCR actually finds that the percentage pixels differ between pair of encrypted images. UACI finds the ratio of changes among the pair of encrypted images. The NPCR value should be near to 98.6 and the UACI value should be near to 28. The NPCR and UACI are calculated using Eqs. (4) and (5).

$$\text{NPCR} = \frac{1}{W \times H} \sum_{i=1}^W \sum_{j=1}^H f(i, j) \times 100\% \quad (4)$$

$$\text{UACI} = \frac{1}{W \times H} \sum_{i=1}^W \sum_{j=1}^H \frac{|U_1(i, j) - U_2(i, j)|}{225} \times 100\% \quad (5)$$

where  $i, j$  are image pixel locations,  $C_1$  is cipher image one and  $C_2$  is another cipher image. If  $U_1(i, j) = U_2(i, j)$ , then  $f(i, j)$  will be 0. If  $U_1(i, j) \neq U_2(i, j)$ , then  $f(i, j)$  will be 1.

The NPCR and UACI values of encrypted image are 99.6721 and 33.3842, respectively. Table 1 lists the NPCR and UACI values of cipher image and it is identical with literature [1–3, 14]. The metrics NPCR and UACI prove that whenever there is a change even in single pixel will make larger effect on randomness. Whenever the image gets encrypted, it produces different cipher images for different plain images. Hence, NPCR and UACI are effective measurements towards image randomness.

#### 4.6 Mean Square Error (MSE) and Peak Signal to Noise Ratio (PSNR)

In contrary to the above metrics, the Mean Square Error (MSE) and Peak Signal to Noise Ratio (PSNR) are also used to prove the difference between the pixel values of plain and cipher images. MSE discusses the difference between the pixels of plain and cipher images. It is computed using the Eq. (6).

$$\text{MSE} = \frac{1}{W \times H} \times \sum_{i=1}^W \sum_{j=1}^H |\text{ob}(i, j) - \text{ex}(i, j)| \quad (6)$$

where,  $\text{ob}(i, j)$  denotes plain image;  $\text{ex}(i, j)$  denotes the cipher image;  $W \times H$  denotes the dimensions of the images.

Similarly, the ratio between the pixels of plain and cipher images is denoted by Peak Signal to Noise Ratio. It is inversely proportional to the MSE. The PSNR is calculated using the Eq. (7)

$$\text{PSNR} = 10 \cdot \log_{10} \left( \frac{\text{MAX}_I^2}{\text{MSE}} \right) \quad (7)$$

where  $\text{MAX}_I$  denotes the maximum value of the pixel.

The MSE value should be high which means that the chaotic encryption absolutely changes the pixel values. The PSNR value should be low. It is inversely proportion to MSE which states that the cryptosystem encrypts better. The best value for PSNR is near to 0. But, the MSE and PSNR are proposed to be around 125 and 27, respectively, as per the literature [8]. The MSE and PSNR values are 121.2524 and 28.1257, respectively. The results coincide with other literature [8]. Therefore, it is essential to consider MSE and PSNR metrics while applying chaotic encryption.

#### 4.7 Mean-Variance Analysis

The mean-variance analysis is performed on encrypted images to find the scrambling effect on the pixels. The cipher image should have high mean-variance value for its pixels than the plain image pixels. Equation (8) is used for calculating the mean-variance analysis.

$$C = \frac{1}{W \times H} \sum_{i=1}^W \sum_{j=1}^H |\text{MV}(i, j) - \overline{\text{MV}}| \quad (8)$$

where  $\overline{MV}$  is the mean value of pixels, and  $W \times H$  is the size of the image.

The mean–variance values for the plain and cipher image are 27.1241 and 64.2120, respectively. In mean–variance analysis, the difference between plain and cipher image pixels are measured and concluded that the cipher image always has higher value than the plain image. If the associations between the pixels are broken by chaotic cryptosystem, the mean–variance value of pixels will be higher in cipher image. The researcher supports this discussion [8].

#### 4.8 Chi-square Analysis

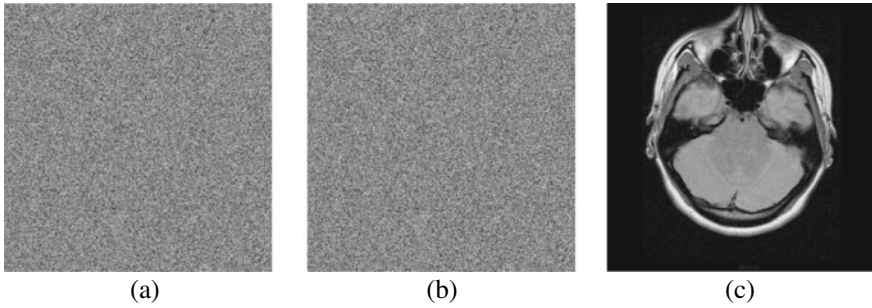
Chi-square is a statistical similarity test. This metric could be employed to validate the chaotic cryptosystem. It is used for measuring the similarities of pixels of plain and cipher images separately. The chi-square value of plain image pixels is high, whereas it is low for cipher image pixels. The chi-square test equation is given in Eq. (9).

$$X^2 = \sum_{i=1}^{256} \frac{(O(i) - E(i))^2}{E(i)} \quad (9)$$

where  $i$  denotes grey values,  $O(i)$  denotes histogram value of the observed cipher image and  $E(i)$  is the histogram value of the cipher image. Table 1 shows the chi-square value pixel distribution of cipher image is 281.97. This value is lower than the chi-square value of the plain image *i.e.* 25,698.61. Hence, the chaotic cryptosystem not only scrambles the pixels as well as alters the pixels better. The results are identical with the results of the literatures [12, 18]. Therefore, it is essential to employ the chi-square analysis for validating the chaotic cryptosystem.

#### 4.9 Key Space Analysis

The key space analysis is done to ensure that the cryptosystem withstands on brute-force attack. An adversary may attempt to search the key through key space to break the cryptosystem. Therefore, it is essential to analyse key space also. The chaotic encryption key space has been decided by the initial and control parameters supplied to the chaotic equation. The input parameters are represented in 64-bit. The key space is usually represented in the power of 2. This cryptosystem has the key space of 2384. It is very hard to do brute-force search through this key space since it is very large. The larger key space yields secure cryptosystem. The chaotic cryptosystem usually provides larger key space [5, 6, 12]. Hence, the key space is an essential metric to be studied while developing the chaotic cryptosystem.



**Fig. 7** Key sensitivity analysis **a** encrypted image, **b** decrypted with altered key, **c** decrypted with correct key

#### **4.10 Key Sensitivity Analysis**

In key sensitivity analysis, a part of the key has been altered and tried to break the cryptosystem with the altered key. The altered key may be the one which has a single bit or single digit difference from the original key. In this study, the original key value is  $q = 0.2000000000000005$  and altered one  $q = 0.2000000000000006$ . Now, the encrypted image is decrypted with altered key and observed the difference. The cryptosystem with altered key is not producing the original image as shown in Fig. 7c; instead, it produces the false image Fig. 7b. Hence, key sensitivity is an important metric to be considered while developing chaotic cryptosystem [15].

#### **4.11 Encryption Time**

The time complexity should be measured to check the efficiency of the proposed cryptosystem. Actually, the chaotic cryptosystem consists of two parts: (i) the generation of chaotic sequences and (ii) performing cryptographic primitive operations such as permutation and diffusion. It is always essential to consider both the parts in order to calculate the encryption time. Further, it is essential to represent the encryption time with respect to Mbit/s [17]. In this study, the cryptosystem took 14.55 Mbit/s to encrypt the image.

### **5 Conclusion**

At present, the chaotic cryptosystems are effective in enciphering the redundant pixels in images. This paper discusses the significance of various metrics employed for evaluating the chaotic cryptosystem. Therefore, this work is focused on employing the chaotic cryptosystem and its validating metrics. Hence, the chaotic encryption

algorithm is validated in terms of randomness of the cipher image by using correlation coefficient, histogram, global entropy and local entropy analyses. The same cryptosystem is validated towards the difference between the plain and cipher images using NPCR, UACI, MSE, PSNR, mean–variance, chi-square. Though the MSE and PSNR are not only involved in evaluating watermarking, they also perform important roles in evaluating chaotic cryptosystems. The chaotic cryptosystem is also evaluated for its security through key space and key sensitivity analyses in order to avoid brute-force attack. Further, the time complexity is a good metric to measure the efficiency of chaotic cryptosystem. Overall, this study analyses and explores the importance of the metrics involved in chaotic cryptosystem.

**Acknowledgements** We would like to thank Department of Computer and Information Science, Annamalai University authorities for providing RUSA-funded Computer Laboratory and Dr.C.P.Ramaswami Aiyar Library, Annamalai University.

## References

1. Liu Y, Zhang J, Han D, Wu P, Sun Y, Moon YS (2020) A multidimensional chaotic image encryption algorithm based on the region of interest. *Multimedia Tools Appl* 79(15). <https://doi.org/10.1007/s11042-020-08645-8>
2. Wei X, Guo L, Zhang Q, Zhang J, Lian S (2012) A novel color image encryption algorithm based on DNA sequence operation and hyperchaotic system. *J Syst Softw* 85:290. <https://doi.org/10.1016/j.jss.2011.08.017>
3. Kadir A, Hamdulla A, Guo W (2014) Color image encryption using skew tent map and hyper chaotic system of 6th order CNN. *Optik* 125:1671. <https://doi.org/10.1016/j.ijleo.2013.09.040>
4. Liu G, Li J, Liu H (2014) Chaos-based colour pathological image encryption scheme using one-time keys. *Comput Biol Med* 45:111. <https://doi.org/10.1016/j.combiomed.2013.11.010>
5. Parvees MYM, Samath JA, Bose BP (2017) Medical images are safe—an enhanced chaotic scrambling approach. *J Med Syst* 41(10). <https://doi.org/10.1007/s10916-017-0809-1>
6. Parvees MYM, Samath JA, Raj IK, Bose BP (2016) A colour byte scrambling technique for efficient image encryption based on combined chaotic map. In: *Proceedings of the international conference on electrical, electronics, and optimization techniques (ICEEOT'16)*, pp 1067–1072, March 2016. <https://doi.org/10.1109/iceeot.2016.7754851>
7. Parvees MYM, Samath JA, Bose BP (2016) Secured medical images—a chaotic pixel scrambling approach. *J Med Syst* 40(11):232:1–232:11. <https://doi.org/10.1007/s10916-016-0611-5>
8. Parvees MYM, Samath JA, Raj IK, Nirmal RM (2017) Chaos-based steganocryptic approach to protect medical images with text data of patients. *J Med Imaging Health Inf* 7:118–125. <https://doi.org/10.1166/jmih.2017.1993>
9. Ragab AHM, Alla OSF, Noaman AY (2014) Encryption quality analysis of the RCBC block cipher compared with RC6 and RC5 algorithms. In: *International association for cryptologic research cryptology ePrint archive*
10. Bashir AA, Basari H, Bin AS, Hamida A (2012) A novel image encryption using an integration technique of blocks rotation based on the magic cube and the AES algorithm. *Int J Comput Sci Issues (IJCSI)* 9(4):41
11. Wu Y, Zhou Y, Saveriades G, Agaian S, Noonan JP, Natarajan P (2013) Local shannon entropy measure with statistical tests for image randomness. *Inf Sci* 222(10):323–342. <https://doi.org/10.1016/j.ins.2012.07.049>

12. Praveenkumar P, Devi K, Ravichandran D, Avila J, Thenmozhi K, Rayappan JBB, Amirtharajan R (2017) Transreceiving of encrypted medical image—a cognitive approach. *Multimedia Tools Appl*. <https://doi.org/10.1007/s11042-017-4741-7>
13. Liu Y, Zhang J (2020) A multidimensional chaotic image encryption algorithm based on DNA coding. *Multimedia Tools Appl* 79(29–30). <https://doi.org/10.1007/s11042-020-08880-z>
14. Wang X, Chen S, Zhang Y (2021) A chaotic image encryption algorithm based on random dynamic mixing. *Opt Laser Technol* 138(12):106837. <https://doi.org/10.1016/j.optlastec.2020.106837>
15. Abbas AM, Alharbi AA, Ibrahim S (2021) A novel parallelizable chaotic image encryption scheme based on elliptic curves. *IEEE Access* 9:54978–54991. <https://doi.org/10.1109/ACCESS.2021.3068931>
16. El-Samie FEA, Ahmed HEH, Elashry IF, Shahieen MH, Faragallah OS, El-Rabaie EM, Alshebeili SA (2014) *Image encryption: a communication perspective*. CRC Press, Boca Raton, Florida
17. Parvees MYM, Samath JA, Bose BP (2017) Protecting large size medical images with logistic map using dynamic parameters and key image. *Int J Netw Sec* 19(6):984–994. [https://doi.org/10.6633/IJNS.201711.19\(6\).15](https://doi.org/10.6633/IJNS.201711.19(6).15)
18. Huang CK, Nien HH (2009) Multi chaotic systems based pixel shuffle for image encryption. *Opt Commun* 282:2123–2127. <https://doi.org/10.1016/j.optcom.2009.02.044>
19. Wu X, Kan H, Kurths J (2015) A new color image encryption scheme based on DNA sequences and multiple improved 1D chaotic maps. *Appl Soft Comput* 37:24. <https://doi.org/10.1016/j.asoc.2015.08.008>
20. Parvees MYM (2019) *Studies on the development and analysis of improved chaotic symmetric cryptographic solutions for image security*. Ph.D. thesis, Bharathiar University, Coimbatore. <http://hdl.handle.net/10603/301501>
21. Parvees MYM, Samath JA, Bose BP (2019) Chaotic sequences are cryptographically secure now—an improved chaotic approach. *J Comput Theor Nanosci* 16(4):1438–1443. <https://doi.org/10.1166/jctn.2019.8056.96>



# Application of Support Vector Machine and Convolutional Neural Network for Sentence-Level Sentiment Analysis of Companies Products Review



Oluwatobi Noah Akande, Joyce Ayoola, Sanjay Misra , Ravin Ahuja, Akshat Agrawal, and Jonathan Oluranti

## 1 Introduction

Product review is the most common approach and tool used to ascertain the level of satisfaction or dissatisfaction of products, especially goods and services purchased by or rendered to consumers. The sustainability of companies in the long run depends largely on their ability to meet the demand of their customers' needs [1]. However, customers' opinions provide feedback on products as this is essential for the success or failure to meet their satisfactions; hence, this determines the growth and acceptability of a company as positive reviews are important for the success of a product [2]. The traditional approach to product review involved people hearing views about products and deciding which to buy based on affordability, availability, and most

---

O. N. Akande · J. Ayoola

Computer Science Department, Landmark University, Omu-Aran, Kwara State, Nigeria

e-mail: [akande.noah@lmu.edu.ng](mailto:akande.noah@lmu.edu.ng)

J. Ayoola

e-mail: [ayoola.joyce@lmu.edu.ng](mailto:ayoola.joyce@lmu.edu.ng)

S. Misra (✉)

Department of Computer Science and Communication, Ostfold University College, Halden, Norway

e-mail: [sanjay.misra@hiof.no](mailto:sanjay.misra@hiof.no)

R. Ahuja

Delhi Skill and Entrepreneurship University, New Delhi, India

A. Agrawal

Amity University, Gurgaon, India

J. Oluranti

Center of ICT/ICE, CUCRID, Covenant University, Ota, Nigeria

e-mail: [jonathan.oluranti@covenantuniversity.edu.ng](mailto:jonathan.oluranti@covenantuniversity.edu.ng)

especially quality [3]. This method of research aids producers in not only determining what customers want or need, but also in designing new goods and improving existing ones [4]. Recently, the use of sentiment analysis and opinion mining for the analyzes and evaluation of text data from different platforms such as social media, e-commerce, and mobile trading applications have become important tools for product reviews, and this has been attributed to the massive amount of information related to individuals that are recorded in the digital forms such as texts, on daily basis [5]. Consumers also write reviews about items that have a lot of functionality or characteristics, and they have varying feelings or thoughts on each of them. While some shoppers might find some facets of goods enticing and worthwhile, others may be disappointing and unyielding. As a result, categorizing a single review as either favorable or negative may ignore useful details found within it where a customer receives comments on various features of a product [6]. For example, Amazon, the e-commerce behemoth, has a section where shoppers can read thousands of recommendations left by other consumers on the things they want. These reviews offer useful information about a product, such as its features, pricing, and feedback, allowing buyers to consider nearly any aspect of the product. This is valuable not only to customers, but also to sellers who manufacture their own goods, as it allows them to better consider consumers and their desires.

According to [7], sentiment analysis as a process of opinion mining in order to detect the opinion or mood of different sentences which are in the form of texts. Sentiment analysis provides an enabling environment for businesses to gain in depth insights on consumers' opinions and also to work actively on improving strategies and approaches of their products based on buyers' feedbacks [8]. Due to increase in the number of Internet users and platforms, sentiment analysis has been successfully used for online reviews, comments, and reactions in areas such as business review analysis, financial market survey, recommendation, and effectiveness analysis as individuals use the Internet to express their opinions and views using star rating, thumbs up, and thumbs down [9]. Consequently, the huge amount of brand-related information goes a long way in providing decisive potential business values for commercial purposes, thereby influence the brand image and positioning [10]. There are three primary characterization levels in sentiment analysis based on the structure of dataset, namely document-level, sentence-level, and aspect-level. Najma et al. [11] described document-level sentiment analysis as one which classifies opinions in term of positive or negative sentiments, while sentence-level sentiment analysis classifies sentiments based on the feelings clarified in each sentence, and aspect-level sentiment analysis groups the opinions for a particular aspect of substance to be considered. In addition to the aforementioned, [12] clarified that there is no difference among document and sentence-level orders, in that sentences are simply short archives. Another aim of sentiment analysis in product review is to determine the subjectivity or objectivity of a sentence. Subjective sentences consist of users' views, perspectives, thoughts, comments, and opinions about the sentence level which are considered for important procedures in sentiment analysis.

Several approaches have been employed in sentiment analysis of product reviews, especially in machine learning, these include supervised learning methods such as

support vector machine (SVM) and deep learning methods such as convoluted neural network. The application of support vector machine in classification task as illustrated by [12] is based on the idea of finding a hyperplane that separates features embedded in sentiments into different domains, which is highly essential for linearly separable and non-separable data with proper kernel functions. Furthermore, it has been reported that SVM achieved better classification than evaluation with respect to unigram model compared to other models. Convoluted neural network which is a deep learning approach for sentiment analysis has been described to better demonstrate competitive performance on sentiment classification as experimented by [13] with different convoluted kernels assisting in learning the local features of various texts. This study seeks to carry out a sentence-level sentiment analysis on customers' reviews about musical products purchased on Amazon using support vector machine and convoluted neural network. It considers the significance of product review analysis, literature reviews on sentiment analysis and concludes with recommendations on how to improve product reviews sentiment analytical tools for market monitoring, business survey, e-commercial activities, and existing competitions.

## 2 Related Works

Sentiment analysis, an aspect of natural language processing and computational linguistics, is designed to extract emotions, sentiment, or more generally opinions expressed in human texts in order to identify, extract, and classify the polarity embedded in each word or sentence [14]. According to [15], sentiment analysis has been an essential method for a variety of social media uses, including the analysis of consumer views on goods and services, political party funding, and even industry trending recognition. In the same vein, sentiment analysis has proven to be tremendously helpful in business review analysis, financial market prediction, and multimodal analysis; to assess prospective strategic strategies based on consumer opinion on new and current products, to monitor patterns in product sales that can be predicted by measuring consumer sentiment, and to allow marketing agencies to advise businesses on the best way to advertise a product based on public feedback derived from social media messages or product reviews [9]. Sentiment analysis has been carried out based on the level of sentiment in the collected data. These levels of sentiments include document/text, sentence or aspect related, although this depends on the discretion of the author, available data, and the unit of information considered. Each sentence is treated as a separate entity in the sentence-level grouping, which means that each sentence can only contain one point of view. In document sentiment analysis, a whole document or text is classified and analyzed in order to identify the sentiment orientation [16]. Word level analysis determines the polarity of an opinion in word or a phrase or on a single entity [17]. The challenge with this approach of sentiment classification comes up when a document has multiple product reviews [18]

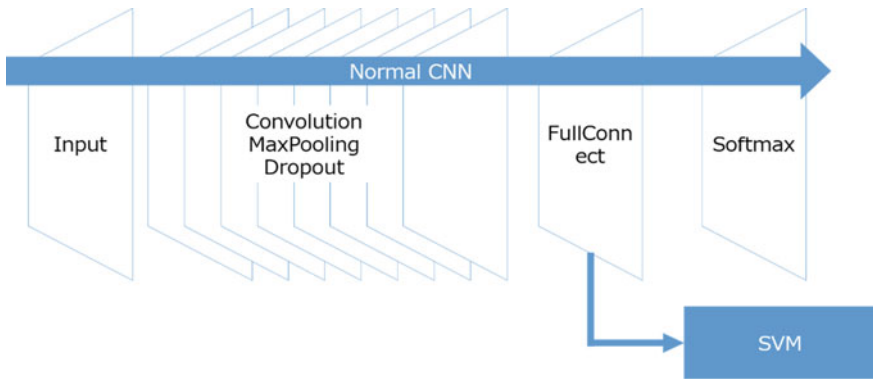
Sentence-level sentiment analysis has been considered in literature [19, 20]. Authors in [21] conducted a linguistic study on sentence-level sentiment analysis, focusing on classifying each sentence in a summary rather than the whole review. In addition, to pay greater attention to the language properties of the sentence and to comprehending the sentence structures for each sentence, so that the objects of the feeling and the aspect being represented can be identified. Since it includes feature-based opinion mining and summarization, aspect-level sentiment analysis is also known as feature-level sentiment analysis [22]. The aim of feature-level classification is to provide an opinion overview of various feedbacks focused on features [23], however, this analysis is faced with the challenge of words possibly carry different sentiments for different aspects [24]. Several tools and models have been used for sentiment analysis of product reviews, with varying degrees of accuracy when compared to other models [2, 8, 12]. Machine learning method has been identified as a major approach to sentiment analysis [5]. Authors in [25] developed a convolutional neural network (CNN) for emotion analysis that performed admirably over a variety of datasets. This network's input layer is made up of concatenated Word2vec word embeddings, preceded by a convolutional layer with several filters, a max pooling layer, and finally, a softmax layer. Instead of using low-dimensional word vectors as input to the convolutional neural network, authors in [26] used a version of the bag-of-words paradigm to generate function vectors (CNN). Machine learning techniques were seen to have obtained the best results to classify the products reviews, such as Naïve Bayes and SVM, which obtained accuracies of 98.17% and 93.54%, respectively, for camera reviews [8]. This study explores the use of SVM and CNN for a sentence-level sentiment analysis of product reviews.

### 3 Methodology

The following phases are activities involved in developing the proposed method: data collection, data pre-processing, model architecture definition, model architecture implementation, model execution, model performance evaluation, and result visualization.

#### 3.1 Model Architecture Definition

In this study, a hybrid CNN-SVM paradigm was presented for the interpretation and classification of sentiment polarity of product reviews. Figure 1 demonstrates the machine architecture. First, the structure of the CNN part allows the model to learn complex word embeddings from a large raw text array. This is a simple way to use the CNN as a feature extractor and eliminates the properties of the SVM convex problems. Finally, a distributed sentence function display is used by observing the distribution of probability over labels as features for SVM classificatory model



**Fig. 1** Architecture of the proposed model

training. Features learned by (convolutional) neural networks are powerful enough that they generalize to different kinds of objects and even completely different parts of the review sentences. As stated earlier, input to the CNN model is an embedding layer of 200 dimensions, 2 convolutional layers, a max pooling layer for each convolutional layer, a fully connected layer, and finally, and the SVM serves as the classification layer of the model. The first convolutional layer has a total of 128 filters and a kernel size of 2, whereas the second, 256 filters and a kernel size of 3. All layers used the “ReLU” activation function except the softmax layer with a sigmoid activation function. This softmax layer, however, is not used in the model classification process. In model execution, the CNN is first created and fitted to the training data. After the weights have been learned, these weights are extracted from the model and fitted to an SVM classifier to develop its hyperplane from the trained instances. Hence, the SVM is fitted to the weights of the CNN. Once SVM training is complete, the SVM classifier is used to classify the sentiments of the validation set of product reviews.

The proposed model architecture is not end-to-end training as it incorporates two machine learning algorithms, and both algorithms were implemented using separate libraries. The learned representations (weights) from the CNN layers serve as training inputs for the SVM classifier. To create an end-to-end trainable version of the proposed algorithm, a second implementation was necessary. This second implementation (M2) uses the hinge loss function of the Keras library. This loss function is mostly used for maximum-margin classification and its activations approximate the SVM kernel functions, that is, it behaves exactly like the first model, only that its end-to-end trainable, and it is implemented with the one library. Hence, the first model implementation would be hereafter referred to as M1. Table 1 describes the hyperparameters of both model variations, where the “hinge” keyword is used to denoted parameters specific to the hinge model (M2).

**Table 1** Parameters of M1 and M2 implementation of proposed model

Hyper parameter	Value
Total training samples	44,463
Test and validation samples	19,056
Sequence max length	116
Chosen training epochs	30
Training batch size	128 samples
Training callbacks	Early stopping and Tensor_board for visualization
SVM kernel	Radial basis function
SVM hyperparameter tuning algorithm	Grid search SVC sklearn implementation
Hinge classifier activation function	Linear
Hinge classifier optimizer function	Adams optimizer
Hinge loss function	Categorical hinge
Hinge classifier kernel regularizer	L2 regularizer at 0.01

## 4 Results and Discussion

The proposed model was built and developed using Amazon’s review dataset on musical instruments and office products. This combined dataset is to be divided and used for both training and validating the proposed model architecture. Figure 2 shows the structure of both datasets.

In evaluating the performance of the proposed method for sentiment classification, both implementations of the proposed model are tested and evaluated using the aforementioned metrics such as the accuracy, precision, recall, and F1 score on both datasets. Table 2 highlights the values obtained by the model for each metric.

Recall on the M2 model shows the proportion of positive reviews were correctly detected by the model, which means the model can perfectly distinguish positive reviews. This may be as a result of overfitting to positive class resulting from the drawback of having an imbalanced dataset.

	Total Samples	Number of Samples for each rating (n)					Positive Samples	Negative Samples	Train Samples		Test/ Validation Samples
		5	4	3	2	1			Positive	Negative	
Musical Dataset	10,261	6938	2084	772	250	217	9022	1239	Positive	6316	2706
									Negative	867	372
Office Product Datasets	53,258	30327	15015	5060	1726	1130	45342	7916	Positive	31739	13603
									Negative	5541	2375

**Fig. 2** Samples of the training and validation datasets

**Table 2** Performance evaluation results of the proposed model

Model/performance metric	ACC (%)	Precision (%)	Recall (%)	AUC	TE	Training time
CNN feature extractor model (M1)	85.38	90.56	95.14	0.836	3	408.6 s ~ 6.081 min
Hinge loss model (M2)	85.74	85.62	100	0.5	3	167.57 s ~ 2.73 min

TE = Training epoch, ACC = Accuracy

The area under curve metrics illustrated in Figs. 3 and 4 show how well (as a fraction between 0 and 1) the models can distinguish sequences belonging to a particular class. A value closer to 0 shows an inability to differentiate sentiment polarity and according to the training data, our model shows high proficiency in distinguishing sentiment polarity labels. From the table (Table 3) and figures (Figs. 2 and 3) above, the CNN feature extractor fits well with the training data. The hinge model (M2) seems to fit better, having a perfect recall value. This may be fitted to a particular classification label. Table 3 describes the performance of all the models on new test data and highlights the values obtained by the models for each metric. The CNN feature extracted SVM classifier has a similar performance with a state-of-the-art CNN model with the softmax classification for sentiment analysis on product reviews. The hinge classifier, however, performs poorly on the provided datasets. Figures 5, 6, and 7 describe the number of classifications and their classifications by the models. 0 represents the negative sentiment label, whereas 1 represents positive instances.

## 5 Conclusion

This research work applied a hybrid CNN-SVM algorithm model in carrying out the sentence-level sentiment analysis of companies' products review on Amazon, specifically musical and office products. It showed that 10,261 datasets and 53,258 datasets were collected on musical and office products, respectively. Based on sentiment polarity of the obtained datasets, musical products reviews contained 9022 positive sentiments and 1239 negative sentiments with 87.9% and 12.1%, respectively, whereas office products review comprised 45,342 positive sentiments and 7916 negative sentiments, representing 84.7% and 15.3%, respectively. This establishes the distribution of opinions and customers' reviews on musical and office products purchased on Amazon, such that, there are more positive sentiments and satisfaction obtained from these products as against the negative sentiments. Hence, the level of satisfaction on these products reflected on the sentiments analyzed. The work has showed that computational-based approaches can be used on business survey, marketing, sales, and opinions of customers on different products available or purchased from the market, in order to drive productive commercial activities.

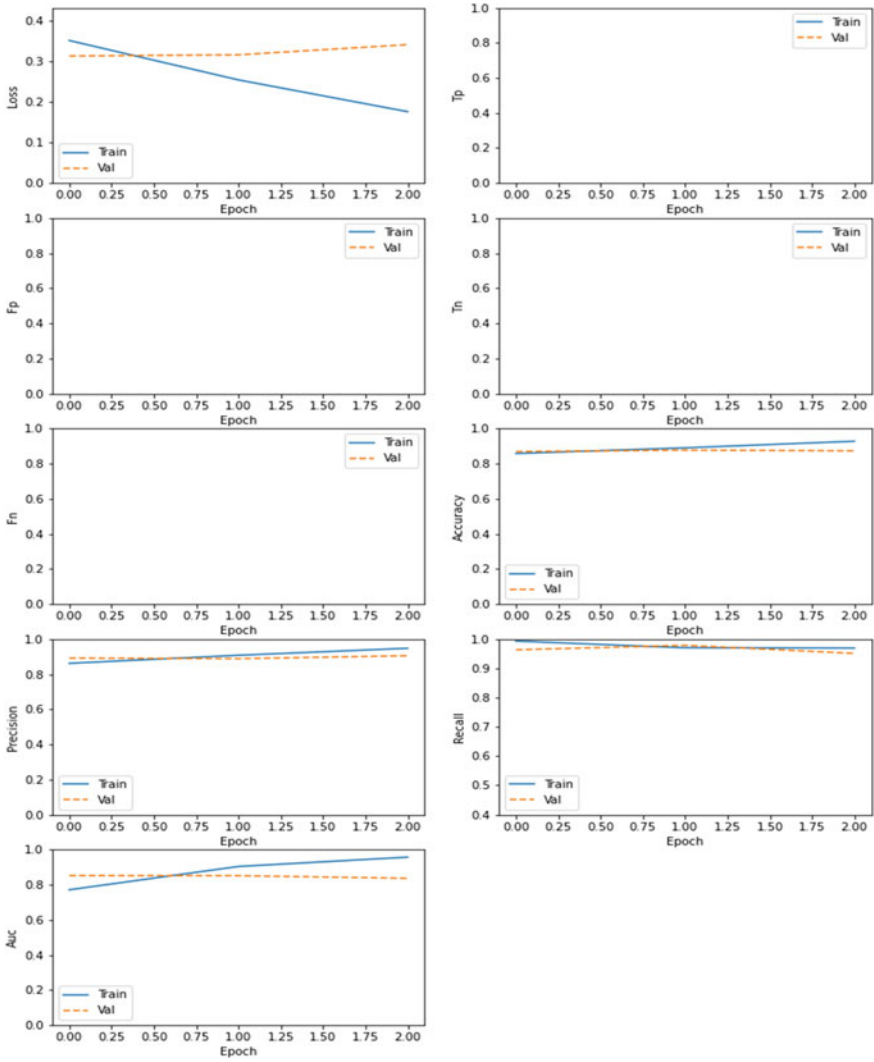


Fig. 3 CNN feature extractor performance on training

More so, it will serve as a feedback mechanism for companies and provide awareness of other competitors in the global market. It is therefore recommended that the science of opinion mining and sentiment analysis should be further incorporated and well-funded to serve as analytical tools.



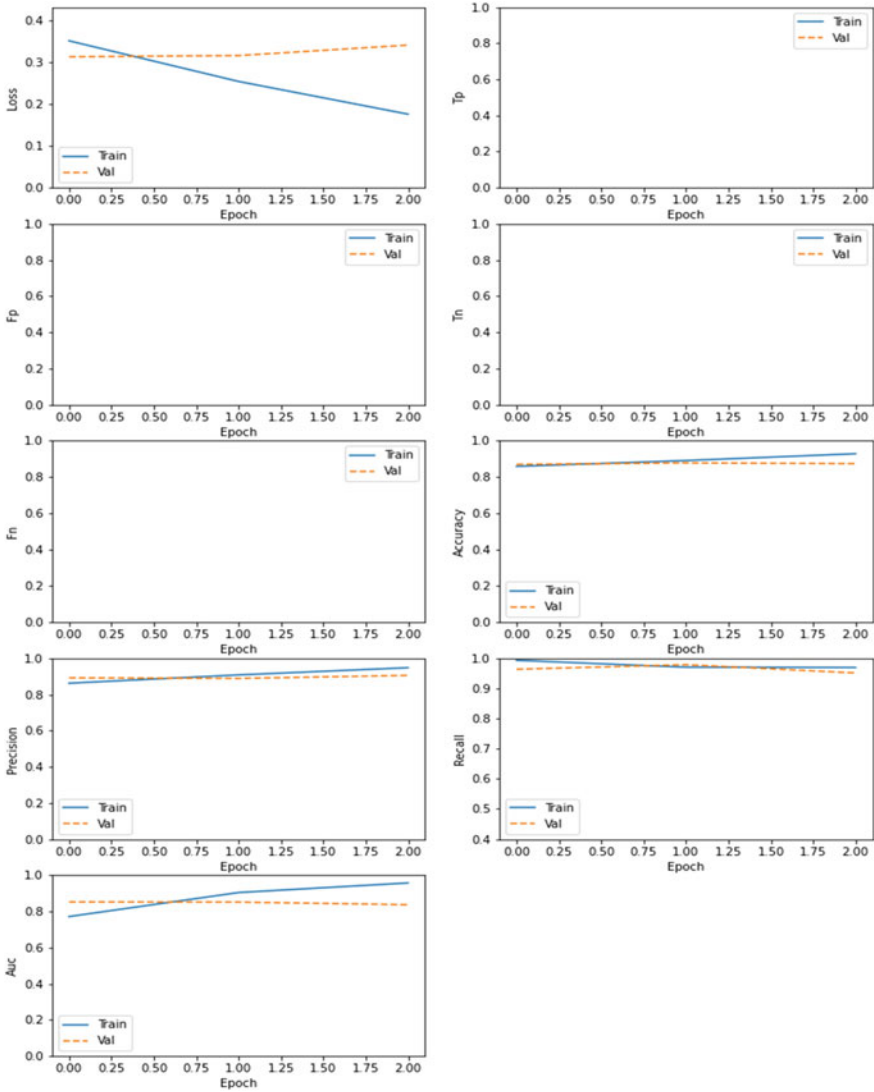
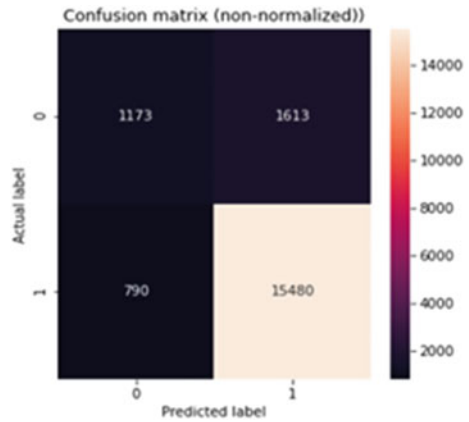


Fig. 4 Training performance of the hinge model

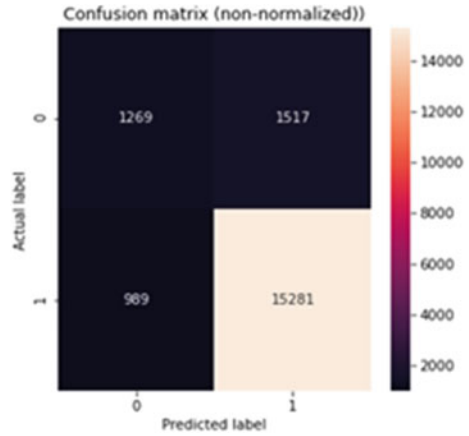
**Table 3** Performance metrics of the model on testing data

Model /performance metric	Accuracy	Precision (%)	Recall (%)	F score (%)
CNN model (softmax classifier)	Positive	87.4	91.1	93.7
	Negative		60.8	46.0
	Macro average		75.0	71.0
	Weighted average		86.0	86.0
CNN feature extractor model (SVM classifier) (M1)	Positive	87.5	91.5	92.0
	Negative		56.2	50.0
	Macro average		74.0	71.0
			86.0	86.0
Hinge loss model (M2)	Positive	86.7	86.3	92.0
	Negative		0.0	0.0
	Macro average		43.0	46.0
			74.0	79.0

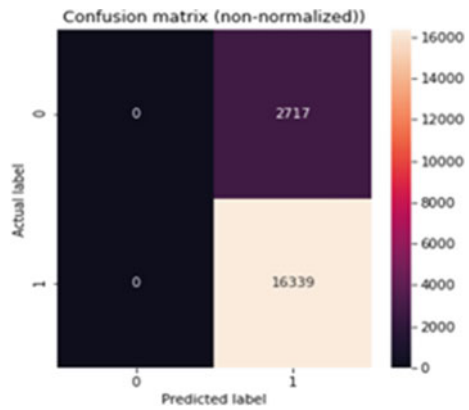
**Fig. 5** Performance confusion matrix of CNN-softmax



**Fig. 6** Performance confusion matrix of SVM with CNN extracted features



**Fig. 7** Performance confusion matrix of the hinge model



**Acknowledgements** The authors appreciate the sponsorship from Covenant University through its Center for Research, Innovation and Discovery, Covenant University, Ota Nigeria.

## References


1. Kang M, Choi Y, Choi J (2019) The effect of celebrity endorsement on sustainable firm value: evidence from the Korean telecommunications industry. *Int J Advert* 2019(38):563–576
2. Rajesh KP, Pandiaraja P (2019) An efficient sentiment analysis approach for product review using Turney algorithm. In: *International conference on recent trends in advanced computing 2019, ICRTAC 2019*. *Procedia Comput Sci* 165:356–362
3. Andrea E, Sebastiani F (2012) Determining the semantic orientation of terms through gloss classification. In: *Proceedings of the 14th ACM international conference on information and knowledge management (CIKM 2005)*, Bremen, DE, pp 617–624

4. Fan Z, Xi Y, Li Y (2018) Supporting the purchase decisions of consumers: a comprehensive method for selecting desirable online products. *Kybernetes* 2018(47):689–715
5. Gandomi A, Haider M (2015) Beyond the hype: big data concepts, methods, and analytics. *Int J Inf Manag* 2015(35):137–144
6. Pooja M, Pandya S (2020) A review on sentiment analysis methodologies, practices and applications. *Int J Sci Technol Res* 9(2). ISSN 2277-8616
7. Sailunaz K, Alhajj R (2018) Emotion and sentiment analysis from twitter text. *J Comput Sci* 36(101003):1–18
8. Jagdale RS, Shirsat VS, Deshmukh SN (2019) Sentiment analysis on product reviews using machine learning techniques. In: *Cognitive informatics and soft computing. Adv Intell Syst Comput* 768. [https://doi.org/10.1007/978-981-13-0617-4\\_61](https://doi.org/10.1007/978-981-13-0617-4_61)
9. Ashima Y, Vishwakarm DK (2020) A deep language-independent network to analyze the impact of COVID-19 on the world via sentiment analysis
10. Erick K, Peral J, Gil D, Ferrández A, Sellers R, Mora H (2019) Managing marketing decision-making with sentiment analysis: an evaluation of the main product features using text data mining. *Sustainability* 11:4235. <https://doi.org/10.3390/su11154235>
11. Najma S, Kumar P, Patra MR, Chandra S, Safikul Alam SK (2019) Sentiment analysis for product review. *ICTACT J Soft Comput* 9(3). ISSN 2229-6956 (Online). <https://doi.org/10.21917/Ijsc.2019.0266>
12. Dishi J, Vardhan BH, Kandasamy SK (2019) Sentiment analysis of product reviews—a survey. *Int J Sci Technol Res* 8(12). ISSN 2277-8616
13. Naz S, Sharan A, Malik N (2018, December) Sentiment classification on Twitter data using support vector machine. In: *2018 IEEE/WIC/ACM international conference on web intelligence (WI)*. IEEE, pp 676–679
14. Kalchbrenner N, Grefenstette E, Blunsom P (2014) A convolutional neural network for modelling sentences. *arXiv preprint [arXiv:1404.2188](https://arxiv.org/abs/1404.2188)*
15. Konstantinos K (2018) Sentiment analysis for Tweets. Athens University of Economics and Business
16. Araújo M, Pereira A, Benevenuto F (2020) A comparative study of machine translation for multilingual sentence-level sentiment analysis. *Inf Sci* 512:1078–1102
17. Pang B, Lee L, Vaithyanathan S (2002) Thumbs up? Sentiment classification using machine learning techniques. In: *Proceedings of the 2002 conference on empirical methods in natural language processing (EMNLP)*, Philadelphia, PA, USA, July 2002, pp 79–86
18. Takamura H, Inui T, Okumura M (2007) Extracting semantic orientations of phrases from dictionary. In: *Proceedings of human language technologies: the annual conference of the north American chapter of the ACL*, pp 292–299
19. Turney PD (2002) Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews. In: *Proceedings of the 40th annual meeting on association for computational linguistics*. Association for computational linguistics
20. Zhang C, Zeng D, Li J, Wang F-Y, Zuo W (2009) Sentiment analysis of Chinese documents: from sentence to document level. *J Am Soc Inform Sci Technol* 60(12):2474–2487
21. Aljuhani SA, Alghamdi NS (2019) A comparison of sentiment analysis methods on Amazon reviews of mobile phones. *Int J Adv Comput Sci Appl* 10(6):608–617
22. Arun M, Prabhakar TV (2007) Sentence level sentiment analysis in the presence of conjuncts using linguistic analysis. In: *ECIR 2007, LNCS vol 4425*, pp 573–580
23. Hu M, Liu B (2004) Mining and summarizing customer reviews. In: *Proceedings of the tenth ACM SIGKDD international conference on knowledge discovery and data mining*. ACM
24. Jagtap VS, Pawar K (2013) Analysis of different approaches to sentence-level sentiment Classification. *Int J Sci Eng Technol* 2(3):164–170. ISSN 2277-1581
25. Donatas M, Frasinca F (2020). ALDONAr: a hybrid solution for sentence-level aspect-based sentiment analysis using a lexicalized domain ontology and a regularized neural attention model. *Inform Process Manage* 57:102211
26. Fang X, Zhan J (2015) Sentiment analysis using product review data. *J Big Data* 2(1):5

27. Tian F et al (2015) A topic sentence-based instance transfer method for imbalanced sentiment classification of Chinese product reviews. *Electron Comm Res Appl.* <https://doi.org/10.1016/j.elelap.2015.10.003>
28. Gui L, Zhou Y, Xu R, He Y, Lu Q (2017) Learning representations from heterogeneous network for sentiment classification of product reviews. *Knowl-Based Syst* 124:34–45
29. Raheesa S, Sharmila KR, ShriSubangi TS, Vimal EA (2017) Sentiment analysis on online product review. *Int Res J Eng Technol (IRJET)* 4(4). e-ISSN 2395-0056
30. Kim Y (2014) Convolutional neural networks for sentence classification. In: *Conference on empirical methods in natural language processing*, Doha, Qatar, pp 1746–1751
31. Johnson R, Zhang T (2017, July) Deep pyramid convolutional neural networks for text categorization. In: *Proceedings of the 55th annual meeting of the association for computational linguistics*, vol 1: Long Papers. pp 562–570

# Application of XGBoost Algorithm for Sales Forecasting Using Walmart Dataset



Yetunde Faith Akande, Joyce Idowu, Abhavya Misra, Sanjay Misra , Oluwatobi Noah Akande, and Ravin Ahuja

## 1 Introduction

The importance of sales forecasting cut across all phases of a company's supply chain. Sales forecasting is critical for manufacturing, logistics, and decision-making [1]. Because of shorter lead times, the consumer demands, and the need to manage scarce resources, forecasts have become extremely relevant. Market forecasting that is accurate normally results in lower inventory costs and better customer service. Accurate forecasting is also critical in retail and industrial activities [2]. Forecasting is used in many processes and management decisions that influence a company's profitability [3]. For example, long-term forecasting is used in production preparation to assess a sufficient amount of personnel and as an input for business planning, such as planning

---

Y. F. Akande

Accounting Department, College of Business Sciences, Landmark University, Omu-Aran, Nigeria  
e-mail: [akande.faith@lmu.edu.ng](mailto:akande.faith@lmu.edu.ng)

J. Idowu

Department of Computer Science, Ladoke Akintola University of Technology, Ogbomosho, Nigeria

A. Misra

Brandan University, Brandan, Canada

S. Misra (✉)

Department of Computer Science and Communication, Ostfold University College, Halden, Norway  
e-mail: [sanjay.misra@hiof.no](mailto:sanjay.misra@hiof.no)

O. N. Akande

Computer Science Department, Landmark University, Omu-Aran, Kwara State, Nigeria  
e-mail: [akande.noah@lmu.edu.ng](mailto:akande.noah@lmu.edu.ng)

R. Ahuja

Center of ICT/ICE, Covenant University, Ota, Nigeria

for expansion or reduction of production units. When a company wants to incorporate a forecasting framework, it is common to copy ideas, goals, and values from other firms to speed up the process. But, since most forecasting approaches were designed for commercial goods, there is a danger of adopting unreasonable targets and unsuitable error measures when the world changes [4]. Inaccurate demand predictions, for example, have resulted in suboptimal levels of production workforce, resulting in workload imbalances and rises in the costs of recruiting, firing, and overtime labor activities; in many situations, outsourced commodity procurement and inaccurate forecasts may result in a shortage of production supplies. As a result, fast reactions to sales results along with forecasting error delivery help to reduce market risks while still maintaining the balance between manufacturing materials and inventory. In periods of crises, unexpected events, and volatile market conditions, certain responses become ever more important. According to authors in [5], long-term, exponentially emerging disasters, such as coronavirus (COVID-19), perplex policymakers and decision-makers, and profoundly disturb population activities and supply chains; COVID-19's novelty and uncertainty make it difficult to deal with, resulting in resource shortages and supply chain disturbances. Furthermore, lockdowns and restricted goods flow render delivery centers unavailable, disrupting the entire supply chain [6]. Traditionally, forecasting analysis aimed to find the best model for a given collection of results [7].

Sales forecasts and enhancement of business processes can be done in a variety of ways, depending on the situation [8, 9]. Traditional time series approaches, such as the ARIMA model and the Holt-Winters technique, have traditionally been used to forecast sales, but recent AI-based methods have piqued interest due to their potential to improve predictive accuracy and model nonlinear trends [10, 11]. Orthodox predictive forecasting techniques, according to [1], extrapolate past patterns and seasonal variations to estimate the future, and as a result, these methods are incapable of predicting environmental macroeconomic developments in the market, which usually have a major impact on production. Firms either manually changed their mathematical forecasts or relied on expert judgmental forecasts to deal with these changes. These tactics, though, are skewed since humans are slow at making certain changes, and the method is time consuming. However, machine learning algorithms have shown their prowess as a forecasting technique whose outputs can be used to make informed decision about sales prediction and forecasting. This work studied the application of extreme gradient boosting (XGBoost) algorithm for sales forecasting. Machine learning algorithms are probability-based. They provide the tendency of an event occurring in percentage. This probability is derived from data fed into the system. Machine learning is a subset of artificial intelligence which is employed in so many industries in developing; self-driven cars, speech recognition, and speech-to-text applications. Sales forecasting is one of the areas where machine learning is applied in predicting future sales and demand using learnt data. Probability occurs mainly in two ways either 0 or 1 (yes or no). The present and past data will be processed using machine learning putting all factors whether externally or internally into consideration in making the best decisions for the sales process. This is done through long short-term memory (LSTM); which uses neural network for making

predictions with provided sequential data. Also, regressive models can be used which include applications in linear regression, random forest regression, and XGBoost. Modeling and comparing models are used by looking into the mean absolute error (MAE) and root mean squared error (RMSE) which are used in comparing model performance. This type of probability-based decisions helps sales leaders in understanding how natural factors like rainfall, seasons, or man-made factors, e.g., low budget capital, economic recession limited staff, and resources affects sales process in business. However, this study explores how XGBoost can be used to forecast sales. Walmart stores sales data that were collected from 45 stores were used to train and test the model.

## 2 Related Works

Since top officials of organizations are the ones to give the final verdicts if the results of a sales forecast will be embraced or discarded, authors in [12] studied the perception of sale's managers to sales forecasting. The views of 400 business-to-business sales managers were retrieved and analyzed, and this were then evenly divided using gender, to find out their forecasting priorities, values, and wishes. Furthermore, gender differences were examined and recognize several key female attitudes that differ from those of their male counterparts were noted. The views of women's and men's, habits, and demographics were also retrieved and analyzed. The study revealed that when it comes to their forecasting roles, women and men have equal perceptions. The discrepancies that do exist, however, emphasize the importance of ongoing preparation and instruction, as well as the advancement of constructive organizational cultures and practices. Given the emphasis of the paper on forecasts and the outdated myth that work functions involving numbers are best fit for men, the survey items in the analysis that classify men and women of the same expertise identify a more even playing field, providing optimism and confidence that sales are progressing favorably in terms of gender equity. Authors in [13] carried out a two-level statistical model for big mart sales prediction. A two-level methodology was used to forecast product sales from a specific outlet, which produced better predictive results than any of the common single model predictive learning algorithms. The method was tested using 2013 revenue data from big marts. In order to forecast reliable performance, data discovery, data transformation, and feature engineering became crucial. The results showed that a two-level statistical approach outperformed a single model approach because the former had more detail, which resulted in better prediction.

Furthermore, artificial neural networks and analytical hierarchy process were employed for predicting car sales in [14]. In order to elicit factors impacting car sales in North America, a questionnaire was provided to car sellers. The questionnaire variables were then given weights and fed into the proposed neural network as data. To compare the degree of error and achieve a more suitable final output that is similar to truth, linear and exponential regression methods were chosen. The



obtained result shows the neural network's good efficiency as compared to other approaches, and it was discovered that it has a lower minimum square error (MSE) than other techniques. Moreover, a study aimed at forecasting sales of electric vehicles was carried out in [15]. To derive further value from the available observations, a new self-adaptive optimized gray model was used, which uses a complex weighted series. The proposed model's weighted coefficient and adjusted initial condition will change to different samples, increasing the model's applicability. The history value was then reconstructed using Simpson's algorithm to improve forecasting accuracy. The modification was discovered to improve the gray model's adaptability and prediction accuracy, making it ideal for predicting electric vehicles. A similar study was carried out in [16]. Furthermore, the results of the forecast show that from 2018 to 2020, China's new energy vehicle production and sales will rise at an average annual rate of 27.53% and 30.49%, respectively.

Similarly, meta-learning was employed for sales forecasting was explored in [17]. For the sales forecast, the meta-learning architecture used deep convolutional neural networks. Authors in [18] observed that the effects of marketing strategies over sales can change anytime resulting in structural change problem. As a result, more accurate approaches for forecasting retailer inventory prices that account for structural change have been suggested. The proposed approach outperforms traditional forecasting approaches that neglect the probability of such improvements, according to data from a well-known US retailer. A sales forecasting model for computer products manufacturers was proposed in [19]. In the meanwhile, sales forecasting's success is heavily influenced by the complex relationships between the major competitors. A research summarized as (1) seasonal and cyclic patterns in computer product demand are recorded, (2) sales forecasting consider product uncertainty changes, and (3) the complex relationships between the manufacturers and computer products are considered to evoke managerial insights. Experiments demonstrate that the proposed framework accomplishes the above objectives and has the ability to be applied to other manufacturing components. The application of deep neural networks for sales forecasting in the fashion industry was carried out in [20]. The models were created using a broad range of factors, including product physical characteristics and domain experts' opinions. Furthermore, this research contrasts the revenue forecasts obtained using a range of shallow techniques, with those obtained using a collection of deep learning techniques. The deep learning model was found to have strong results in predicting sales in the fashion retail industry, but it does not do substantially better than some of the shallow strategies, such as random forest, for some of the measurement metrics considered. However, this study explores how XGBoost can be used to forecast sales. Walmart stores sales data that were collected from 45 stores were used to train and test the model.

### 3 Methodology

The process flow involved in developing the model is shown in Fig. 1. It is necessary that each technique goes through the build up to generate accurate results. There are about six steps involved, and each one plays a crucial role to build up the proposed model. The pre-processing part is the first three part of the steps before the model was built. The single model was developed using XGBoost algorithm. The performance evaluation gives the performance of the model which was obtained using the testing set the performance metrics.

#### 3.1 Data Collection

The data used in this study were collected from 45 Walmart stores in the United States. They are the record of sales made in their 45 branches between from 2010-02-05 to 2012-11-01. Each store is made up of several departments, and our task is to predict the department-wide sales for each store. Walmart did not provide much information about what is the type of the store or the units of the store size. The datasets consist of four different files, namely stores.csv, train.csv, test.csv, and features.csv.

#### 3.2 Data Exploration

Given the dataset, key features of the dataset that will be used for the prediction need to be extracted. These are features about the stores: store unique ID, store type, store size, temperature of the region during the week, fuel price in the region during

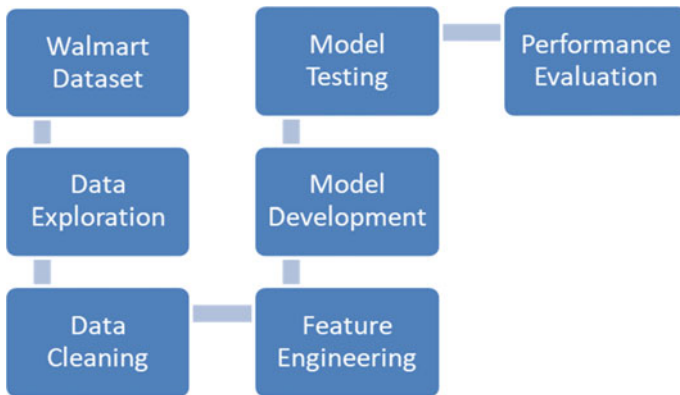


Fig. 1 Flow diagram of processed system

a particular week, unemployment rate during that week in the region of the store, date, Weekly\_sales, CPI, IsHoliday, and Markdowns. These extracted features will be used for the training and testing of the model.

### ***3.3 Data Cleaning***

In this phase of the study, the null values present in the dataset were removed. While exploring the dataset, it was observed that the markdown 1–5 contains some null values. To know if the features will be useful in the model development, a threshold of 60% was set as an acceptable percentage for null values. The number of missing values from the markdown columns and their percentage over the total of registers were counted. However, it was observed that 60% of the data from all markdown columns are missing; therefore, the columns were discarded while other columns whose null values does not exceed the set threshold were converted to 0. The IsHoliday column was also transformed to numeric values 0 for false and 1 for true.

### ***3.4 Feature Engineering***

Feature engineering refers to a process of selecting and transforming features when a predictive model is created by machine learning or statistical modeling. Feature engineering techniques used in this study are

- **One-Hot Encoding:** This technique was used to split the type column and create separate columns for each of the types and assign 0 or 1 to each column depending on the type column, and this was also used to convert the types column from char to integer and give the columns some form of hierarchy
- **Extracting Date:** Even though date columns provide valuable information on the model objective, the information is disregarded as an input. This may be because dates can be found in numerous formats that make algorithms difficult to understand, even in formats such as 01-01-2017. If the date columns are not re-arranged, it will be difficult for a machine learning algorithm to build an ordinal relationship between the values. Therefore, the different sections of the dates were separated into different columns: year, month, and day. With this, the date information becomes easy for the machine learning algorithm to manipulate.

### ***3.5 Model Building Using XGBoost Algorithm***

The next task is model building using the XGBoost algorithm. This was implemented using the XGBoost library. The library contains all the tools needed to build different

XGBoost model. It is an open-source implementation of the gradient boosting algorithm that is quite effective. As such, XGBoost is a Python library, an open-source project, and an algorithm. It is designed to be both computationally efficient (i.e., fast to execute) and highly effective, perhaps even more so than other open-source implementations. Execution speed and model consistency are the two key factors to use XGBoost. On classification and regression predictive modeling issues, the algorithm's main goal is to maximize the objective function's value. Rather than using feature vectors to measure the similarities between forecasting and history days, gradient boosting builds boosted trees to obtain feature scores intelligently, showing the value of each feature to the training model. The higher a feature's score becomes the more, and it is used to make crucial decisions with boosted trees. The algorithm divides the value into three categories: "benefit," "frequency," and "cover." The key criterion for determining the value of a function in the tree branches is gain. Frequency is a simplified form of gain. The dataset is then split into two sets: a training set and a testing set. The training set is used to train the model during the development process, while the testing set is used to assess the model's results. The dataset was split into two parts in this report: 70% for the training set and 30% for the research set.

### **XGBoost Algorithm**

XGBoost is a modular tree boosting machine learning framework. The software can be downloaded as an open-source kit. The technology has a big influence and is well-known in a variety of deep learning and data mining problems. The scalability of XGBoost in all situations is the most important explanation for its success. On a single processor, the technology is ten times faster than current common implementations, and it scales to billions of examples in distributed or memory-limited environments. XGBoost's scalability is attributed to many significant machine and algorithmic optimizations, including a novel tree learning algorithm for sparse data handling and a technically justified weighted quantile sketch technique for estimated tree learning instance weight handling. Parallel and distributed processing speed up research, allowing for more rapid model discovery. More specifically, the model takes advantage of out-of-core computing to enable data scientists to process a hundred million examples on a single computer. Finally, after integrating these methods to create an end-to-end method, it can expand to even more data by using the fewest cluster resources possible. We used the XGBRegressor from the XGBoost open-source kit available in Python. We set our parameter to the following after tuning the XGBoost model several times as proposed in:

- Set the learning rate parameter to 0.1.
- Set the number of estimators to 200.
- Set regularization parameters: reg lambda = 0.45, reg alpha = 0, gamma = 0.5, min child weight = 2, subsample = 1, colsample bytree = 0.8.

### 3.6 Model Evaluation

The performance of the proposed prediction model was carried out using the following metrics:

- The root mean square error (RMSE) is the residuals' standard deviation (prediction errors). It calculates the degree to which the data are concentrated in the best fit axis. It was determined using Eq. (1):

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (S_i - O_i)^2} \quad (1)$$

- Mean absolute error (MAE): Tests how reliable the estimated values are relative to the actual ones. It was calculated using Eq. (2):

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y(i) - \hat{y}(i)| \quad (2)$$

- Coefficient of determination or  $R^2$  uses two mean squared error calculations. The former is the mean squared error of the current and expected values, and the latter is the mean squared error of each real value and the sum of observations. It was determined using Eq. (3):

$$R_2 = \frac{\sum_{i=1}^m (y(i) - \hat{y}(i))^2}{\sum_{i=1}^m (y(i) - \bar{y})^2} \quad (3)$$

## 4 Results and Discussion

The results obtained from the model implementation are discussed in this section.

### 4.1 Results of the Data Exploration

Data exploration was carried out to understand the features present in the dataset and their purpose. The distribution of the train dataset is shown in Fig. 2. This shows the distribution of the data in each column such as the mean, standard deviation, the count, and other values. The next phase was to check the number of null values in the train data. However, there was no null value in the train data.

Afterward, the distribution of the features dataset was also explored as shown in Fig. 3, and the presence of null values were examined in the columns.

	Store	Dept	Weekly_Sales
count	421570.000000	421570.000000	421570.000000
mean	22.200546	44.260317	15981.258123
std	12.785297	30.492054	22711.183519
min	1.000000	1.000000	-4988.940000
25%	11.000000	18.000000	2079.650000
50%	22.000000	37.000000	7612.030000
75%	33.000000	74.000000	20205.852500
max	45.000000	99.000000	693099.360000

Fig. 2 Distribution of the train dataset

	Store	Temperature	Fuel_Price	Markdown1	Markdown2	Markdown3	Markdown4	Markdown5	CPI	Unemployme
count	8190.000000	8190.000000	8190.000000	4032.000000	2921.000000	3613.000000	3464.000000	4050.000000	7605.000000	7605.0000
mean	23.000000	59.356198	3.405992	7032.371786	3384.176594	1760.100180	3292.935886	4132.216422	172.460809	7.8268
std	12.987966	18.678607	0.431337	9262.747448	8793.583016	11276.462208	6792.329861	13086.690278	39.738346	1.8772
min	1.000000	-7.290000	2.472000	-2781.450000	-265.760000	-179.260000	0.220000	-185.170000	126.064000	3.6840
25%	12.000000	45.902500	3.041000	1577.532500	68.880000	6.600000	304.687500	1440.827500	132.364839	6.6340
50%	23.000000	60.710000	3.513000	4743.580000	364.570000	36.260000	1176.425000	2727.135000	182.764003	7.8060
75%	34.000000	73.880000	3.743000	8923.310000	2153.350000	163.150000	3310.007500	4832.555000	213.932412	8.5670
max	45.000000	101.950000	4.468000	103184.980000	104519.540000	149483.310000	67474.850000	771448.100000	228.976456	14.3130

Fig. 3 Distribution of the feature dataset

However, there are null values in the Markdown 1–Markdown 5 columns as well as the CPI and unemployment columns. The remaining datasets were also explored, and the null values were removed. The next stage of the exploration is the combining the datasets together to get one dataset that is going to be cleaned and be used to train the model that will built in the report. First, we will merge the train dataset and feature dataset on the columns that are common in both the datasets and that is the Store, Date, and IsHoliday columns and it going be an inner join after joining them, then we are going to join the store dataset to the new dataset that is the combination of the train dataset and the feature dataset, they will be combined on the store column and also an inner join after joining all the dataset an new dataset is created that contains all the columns from the train, feature, and the store datasets. After that, count plot of the number of sales per store was plotted as shown in Fig. 4. This was also done to get a better knowledge of the relationship among related features in the dataset.

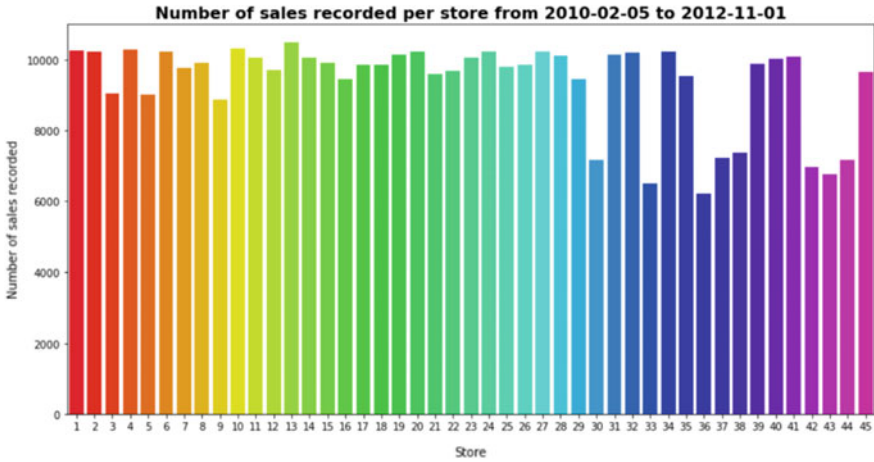


Fig. 4 Count plot of number of sales per store

The histogram of the input dataset was also generated so as to examine the frequency of sales during the time of the data collection. Figure 5 shows the total number of sales generated by each of the 45 stores during the time of the collection of the dataset. It was observed that more stores have between 8000 and 10,000 sales during this period.

This line plot shows the average sales of all the store is per week for the years, and the dataset was also explored as shown in Fig. 6. It was observed that in the latter

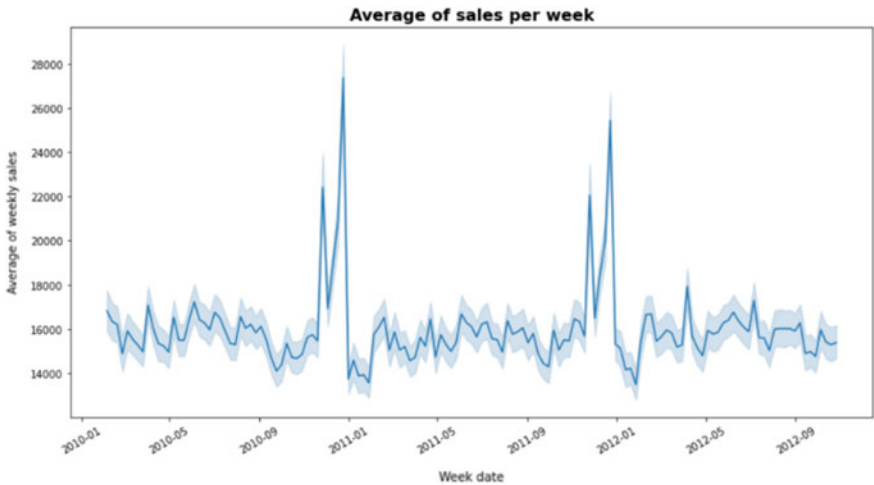


Fig. 5 Average sales per week

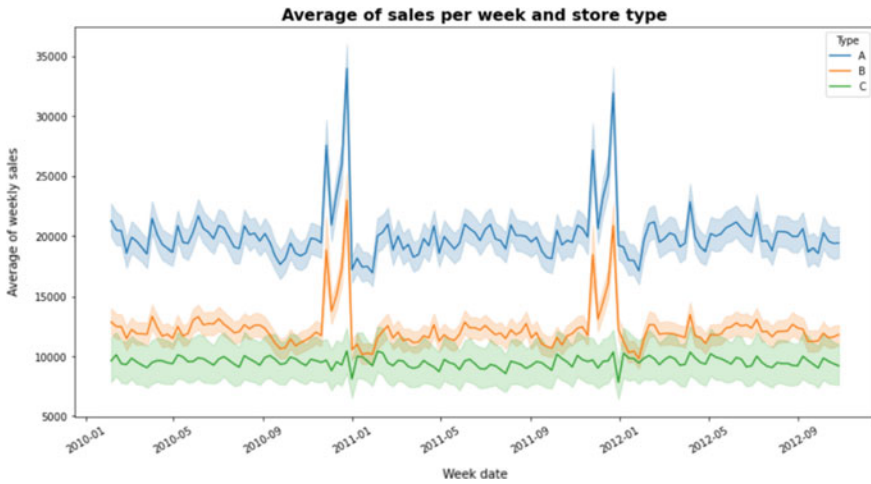


Fig. 6 Line plot of average sales per week and store types

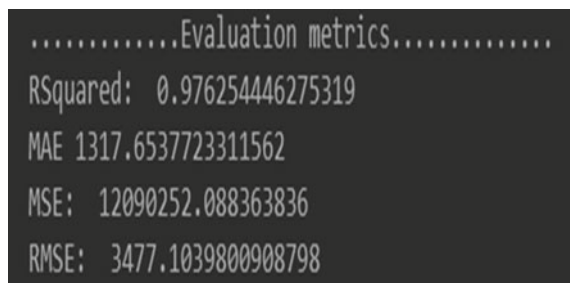
part of 2010, the stores made more sales than any other year followed by the latter part of 2011.

This line plot shows the average sales for each type of the stores which are type A, B, C. Type A made the most sales, and type C made the least sales for the duration of the dataset.

### 4.2 Performance Evaluation of the Proposed Model

In this stage after training and testing the model, mean absolute error (MAE),  $R^2$ , mean square error (MSE), and root mean square error (RMSE) were computed as shown in Fig. 7. The results obtained revealed that XGBoost achieved an appreciable prediction accuracy.

Fig. 7 Evaluation of XGBoost model





## 5 Conclusion

In supply chain, sales forecasting is a big concern and a critical role. This significance stems from the fact that it has an effect at all levels of the organization. A company's inaccuracy or lack of forecasting may result in inadequate inventory and material flow control, revenue losses or surplus inventory, and consumer frustration. When forecasting issues are combined with other supply chain flaws like the bullwhip effect, the effects may be disastrous. While research and practice favor historical sales forecasting techniques, causal methods have been shown to be more accurate and precise. For reliable revenue forecasting, extensive analysis is being conducted at the enterprise stage. Since a company's profit is directly proportional to its ability to forecast profits accurately, big marts are looking for a more reliable prediction algorithm to ensure that the company does not lose money. This study explores how extreme gradient boosting (XGBoost) can be used to forecast sales. Walmart stores sales data that were collected from 45 stores were used to train and test the model. Results obtained from the mean absolute error,  $R^2$ , mean square error, and root mean square error revealed that XGBoost machine learning technique can be effectively used to predict sales, and this could guide sales managers and other practitioners in making decisions about how to fix prices of goods.

**Acknowledgements** The authors appreciate the sponsorship from Covenant University through its Center for Research, Innovation and Discovery, Covenant University, Ota Nigeria.

## References

1. Verstraete G, Aghezzaf E-H, Desmet B (2020) A leading macroeconomic indicators' based framework to automatically generate tactical sales forecasts. *Comput Ind Eng* 139:106169. <https://doi.org/10.1016/j.cie.2019.106169>
2. Boone T, Ganeshan R, Jain A, Sanders NR (2019) Forecasting sales in the supply chain: consumer analytics in the big data era. *Int J Forecast. Special Section: Supply Chain Forecast* 35:170–180. <https://doi.org/10.1016/j.ijforecast.2018.09.003>
3. Hyndman RJ, Athanasopoulos G (2018) *Forecasting: principles and practice*. OTexts
4. Kerkkanen A, Korpela J, Huiskenen J (2009) Demand forecasting errors in industrial context: measurement and impacts. *Int J Prod Econ* 118:43–48. <https://doi.org/10.1016/j.ijpe.2008.08.008>
5. Govindan K, Mina H, Alavi B (2020) A decision support system for demand management in healthcare supply chains considering the epidemic outbreaks: a case study of coronavirus disease 2019 (COVID-19). *Transp Res Part E Logist Transp Rev* 138:101967. <https://doi.org/10.1016/j.tre.2020.101967>
6. Kumar MS, Raut DRD, Narwane DVS, Narkhede DBE (2020) Applications of industry 4.0 to overcome the COVID-19 operational challenges. *Diabetes Metab Syndr Clin Res Rev* 14:1283–1289. <https://doi.org/10.1016/j.dsx.2020.07.010>
7. De Gooijer JG, Hyndman RJ (2006) 25 years of time series forecasting. *Int J Forecast. Twenty five Years Forecast.* 22:443–473. <https://doi.org/10.1016/j.ijforecast.2006.01.001>

8. Arogundade O, Misra S, Odusami M, Adelaja E, Abayomi-Alli O (2020) An algorithm-centric approach to enhance business process compliance management. *Int J Eng Res Africa* 50:15–28. Trans Tech Publications Ltd
9. Mustapha AM, Arogundade OT, Misra S, Damasevicius R, Maskeliunas R (2020) A systematic literature review on compliance requirements management of business processes. *Int J Syst Assur Eng Manage* 11(3):561–576
10. Akande ON, Abikoye OC, Kayode AA, Lamari Y (2020) Implementation of a framework for healthy and diabetic retinopathy retinal image recognition. *Scientifica* 2020:1–14. Article ID 4972527
11. Karmy JP, Maldonado S (2019) Hierarchical time series forecasting via Support Vector Regression in the European Travel Retail Industry. *Exp Syst Appl* 137:59–73. <https://doi.org/10.1016/j.eswa.2019.06.060>
12. Wilson JH, Dingus R, Hoyle J (2020) Women count: perceptions of forecasting in sales. *Bus Horiz* 63(5):637–646. <https://doi.org/10.1016/j.bushor.2020.06.001>
13. Punam K, Pamula R, Jain PK (2018) A two-level statistical model for big mart sales prediction. In: 2018 international conference on computing, power and communication technologies (GUCON), September, pp 617–620. <https://doi.org/10.1109/GUCON.2018.8675060>
14. Farahani DS (2016) Car sales forecasting using artificial neural networks and analytical hierarchy process case study: Kia and Hyundai Corporations in the USA, pp 57–62
15. Ding S, Li R (2021) Forecasting the sales and stock of electric vehicles using a novel self-adaptive optimized grey model. *Eng Appl Artif Intell* 100:104148. <https://doi.org/10.1016/j.engappai.2020.104148>
16. He L, Pei L, Yang Y (2020) An optimised grey buffer operator for forecasting the production and sales of new energy vehicles in China. *Sci Total Environ* 704:135321. <https://doi.org/10.1016/j.scitotenv.2019.135321>
17. Ma S, Fildes R (2021) Retail sales forecasting with meta-learning. *Eur J Oper Res* 288(1):111–128. <https://doi.org/10.1016/j.ejor.2020.05.038>
18. Huang T, Fildes R, Soopramanien D (2019) Forecasting retailer product sales in the presence of structural change. *Eur J Oper Res* 279(2):459–470. <https://doi.org/10.1016/j.ejor.2019.06.011>
19. Wang C, Yun Y (2020) Demand planning and sales forecasting for motherboard manufacturers considering dynamic interactions of computer products. *Comput Ind Eng* 149:106788. <https://doi.org/10.1016/j.cie.2020.106788>
20. Loureiro ALD, Miguéis VL, Lucas FM (2018) Exploring the use of deep neural networks for sales forecasting in fashion retail. *Decis Support Syst* 114:81–93. <https://doi.org/10.1016/j.dss.2018.08.010>

# Comparison of Selected Algorithms on Breast Cancer Classification



Olabiyisi Stephen Olatunde, Olalere Mofiyinfoluwa,  
Oluwatobi Noah Akande, Sanjay Misra , Ravin Ahuja, Akashat Agrawal,  
and Jonathan Oluranti

## 1 Introduction

Cancer is a disease that is destructive and has a poor estimated survival rate. Owing to the high relapse and death rates, the recovery phase is lengthy and expensive. To improve the patient's survival rate, accurate early cancer detection and prognosis prediction are critical. Many scientists have been motivated to use quantitative techniques such as multivariate regression analysis to assess the prognosis of the disease as a result of advances in mathematics and data engineering over the years, and the precision of such analyzes is considerably higher than that of observational forecasts. Cancer is difficult to detect in the early stages and is prone to recurrence following surgery. Furthermore, predicting the prognosis of an illness with high accuracy is

---

O. S. Olatunde (✉) · O. Mofiyinfoluwa  
Ladoke Akintola University of Technology, Ogbomosho, Oyo State, Nigeria  
e-mail: [soolabiyisi@lutech.edu.ng](mailto:soolabiyisi@lutech.edu.ng)

O. N. Akande  
Computer Science Department, Landmark University, Omu-Aran, Kwara State, Nigeria  
e-mail: [akande.noah@lmu.edu.ng](mailto:akande.noah@lmu.edu.ng)

S. Misra (✉)  
Department of Computer Science and Communication, Ostfold University College, Halden,  
Norway  
e-mail: [sanjay.misra@hiof.no](mailto:sanjay.misra@hiof.no)

R. Ahuja  
Shri Viswakarma University, Gurgaon, India

A. Agrawal  
Amity University, Gurgaon, India

J. Oluranti  
Center of ICT/ICE, CUCRID, Covenant University, Ota, Nigeria  
e-mail: [jonathan.oluranti@covenantuniversity.edu.ng](mailto:jonathan.oluranti@covenantuniversity.edu.ng)

extremely difficult. Because of their subtle effects and indistinct tell-tale signals on mammograms and scans, certain tumors are impossible to spot throughout their early stages. As a result, improved statistical models based on multivariate data and high-resolution diagnostic methods are critical in clinical cancer research. A brief search of the literature reveals that the number of publications on cancer analysis has exploded, especially those concerning AI instruments and vast datasets containing historical clinical cases for AI model training [1].

Machine learning (ML) is applied in various domains in these days [2, 3]. ML is a subset of AI that is used to create predictive models that recognize logical patterns from large amounts of historical data in order to forecast a patient's survival rate. ML has been widely used to improve prognosis [4, 5]. When cells grow out of control and crowd out regular cells, cancer develops. This makes it difficult for normal cells to operate properly. It is a malignant tumor that begins in breast cells and spreads through breast tissue, accounting for about 25% of all cancer cases in women [6]. Cancer is a term used to describe any disease in which abnormal cells in the body destroy healthy cells, as well as fluid leaking from the nipple, an inverted nipple are all symptoms of distant illness dissemination.

Furthermore, age, formation control, childbearing antiquity, race, and being grossly overweight after menopause are all factors that contribute to breast cancer, which can be diagnosed by a breast exam, mammogram, or breast ultrasound. In various countries, women diagnosed with stage I/II breast cancer (small tumors or limited local dissemination to nodes under the arm) have a five-year survival rate of 80–90%. Various machine learning techniques have been explored for cancer tumor classification. However, this study explored the performance of SVM, DTC, NB, KNN, sequential model (SM), and discriminant analysis algorithm for breast cancer tumor diagnosis.

## 2 Related Works

This section provides an overview of existing works that have been done breast cancer detection and classification. Using wisconsin breast cancer records as train and test data, KNN, C4.5, and NB were employed for cancer classification [7]. The major goal is to evaluate the validity of each algorithm's classification accuracy, precision, sensitivity, and specificity in terms of efficiency and effectiveness. An accuracy of 97.13% was achieved with SVM. All tests were run in the WEKA data mining tool. In a similar vein, researchers in [8] used a large dataset to develop cancer prediction models based on Naive Bayes, RBF networks, and J48. The Naive Bayes algorithm had the global optimum solution of 97.36%, followed by RBF network with 96.77% accuracy and J48 with 93.41% accuracy, according to the data. In addition, the authors of [9] proposed the use of a computer-aided design technique for breast cancer diagnosis. When compared to the previous studies, the classifier complex has a 99.68% accuracy rate, indicating encouraging findings. The proposed method yields an effective breast cancer classification model. Breast cancer was classified

using NB and KNN classifiers in [10]. An accuracy of 97.51% was achieved with KNN classifier while NB recorded an accuracy of 96.19%.

A study to compare and contrast computational intelligence algorithms for detecting breast cancer in its early stages was carried out in [11]. Microarray-based and medical imaging-based data interpretations were also considered. In comparison with the previous studies, this one took into account the unbalanced existence of the evidence. When accuracy was used as a performance metric, the SMO algorithm performed better for the majority of the test data, particularly for microarray-based data. Given the data's unbalanced nature, the Naive Bayes algorithm was found to perform well in terms of true positive rate (TPR). Authors in [12] also studied the performance of NB, DT, CART, radial basis function (RBF), and SVM-RBF for cancer classification. Result obtained revealed that SVM-RBF performed better than the remaining classifiers as it achieved an accuracy of 96.84%. Summarily, the performance of several machine learning algorithms for cancer classification has been carried out with diverse accuracies recorded for respective algorithms. This study attempts to further examine the performance of SVM, DT, NB, KNN, sequential model (SM), and discriminant analysis algorithm for the detection and classification of cancer tumors in patients into benign and malignant. Wisconsin breast cancer dataset was used to train and test the machine learning techniques while accuracy, precision, and sensitivity were used as the performance evaluation metrics.

Also, authors in [13] proposed the use of a genetic algorithm for breast cancer categorization. This breast cancer classification system was able to optimize the artificial neural network parameter. The method of breast cancer classification becomes more complicated as the calculation becomes more intricate. A hybridize PSO-KDE models were proposed in [14]. The kernel bandwidth was determined using particle swarm optimization. The purpose function of PSO-KDE was designed using the concepts of classification performance and the number of features picked.

Authors in [15] proposed a decision tree-based ensemble learning technique for breast cancer classification. RF and ET algorithms make up the ensemble approach. In all cases, the developed RF models were proven to be 100% accurate, sensitive, and specific in forecasting the WBCD type. In 2016, writers in [16] suggested a two-phase approach for a breast cancer classifier. The gradient descent back-propagation algorithm, or LM, is utilized in the initial phase, along with an ANN. Following the pre-training phase, the unsupervised phase begins. According to the findings, the DBN-NN approach is more accurate than the RIW-BPNN method. Authors in [17] built a rule-based approach to breast cancer classification. Despite the fact that FL systems produce accurate outcomes in the vast majority of circumstances, the design process is difficult and needs multiple stages of implementation.

### 3 Methodology

Kaggle breast cancer wisconsin (diagnostic) dataset that has 569 instances and 10 attributes was used to train and test the selected algorithms considered in this study.

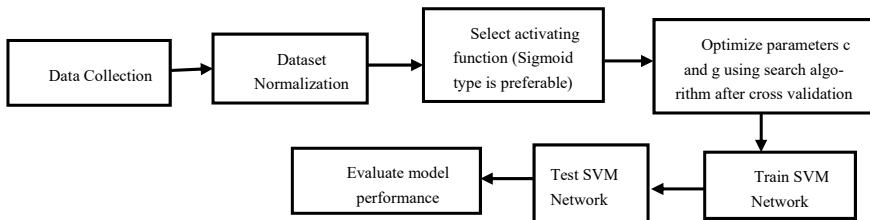
357 of these datasets are benign while 212 are malignant. This dataset's feature is derived from a digitized image of a breast tumor fine needle aspirate (FNA). The prognosis is recorded as M for malignant and B for benign in the target feature. Hence, the need for preprocessing the prognostic values indicated as malignant was converted to 1 and benign to 0 for smoother data processing. The dataset contained 32 float value attributes such as id, diagnosis, radius\_lean, texture\_lean, perileter\_lean, area\_lean, sloothness\_lean, colpactness\_lean, concavity\_lean, and concave points\_lean. Then, the 10 folds cross validation was applied. In this research, six algorithms were implemented which are support vector machine (SVM), decision tree classifier (DTC), discriminant analysis algorithm (DAL), K nearest neighbor algorithm (KNN), Naive Bayes, and sequential model (SM). This research utilized Google Colab with the use of Python programming language. The 10 folds cross validation was implemented with test size of 25% and train size of 75%, with 30 clinical variables. The data were pre-processed to eliminate empty rows of data, and then normalized in such a way that numerical values could be expressed in terms of 0–1 binary sets where 1 represents malignant and 0 represents benign tumors. The questions posed during the research are which algorithm is fastest? Which has to highest accuracy? Which has the least error? This has been illustrated.

### 3.1 Adopted Classifiers

This section explains how the selected classifiers work.

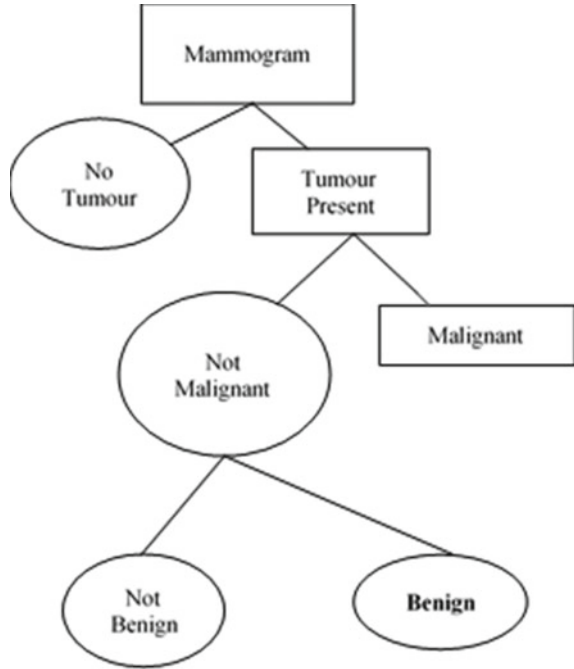
#### 3.1.1 Support Vector Machine.

SVM functions generate the right line or decision boundary that splits n-dimensional space into groups, making it easy to categorize fresh data points in future. The best judgment boundary is known as a hyperplane. Figure 1 illustrates how it works.



**Fig. 1** Support vector machine flowchart (Huang et al. 2015)

**Fig. 2** Decision tree classifier



### ***3.2 Decision Tree Classifier (DTC)***

This is a multilevel decision-making algorithm used to make predictions based on initial inferences. It can be viewed as a hierarchical classifier in which each node leads to the next based on the data being analyzed. Figure 2 illustrates how it works.

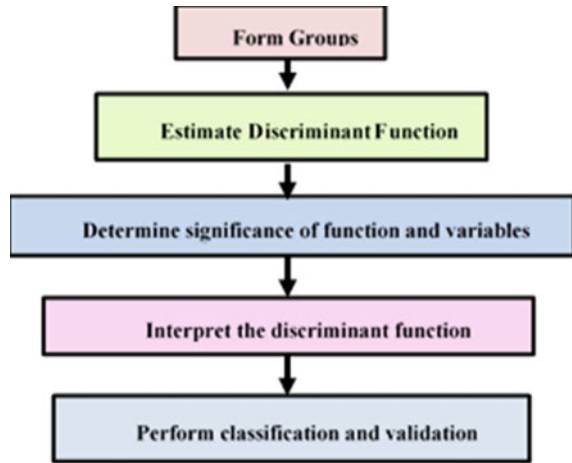
### ***3.3 The Discriminant Analysis Algorithm (DAL)***

This divides objects into classes, reducing within-class variance while optimizing variance between classes, and determining the linear combination of the initial variables (directions). These are known as discriminant functions, and the number of them is proportional to the number of groups minus one. Figure 3 illustrates how it works.

### ***3.4 K Nearest Neighbor (KNN) Algorithm***

This is a model for classifying data points based on their closest neighbors. It makes a “informed guess” on what an unclassified point can be classified as based on test

**Fig. 3** Discriminant analysis algorithm



results. The training dataset is all that is needed to build the model. Figure 4 illustrates how it works.

### 3.5 *Naive Bayes Classifiers*

These are a basic family of “probabilistic classifiers” based on Bayes’ theorem and strict independence assumptions between features. Figure 5 illustrates how it works.

### 3.6 *Sequential Model (SM)*

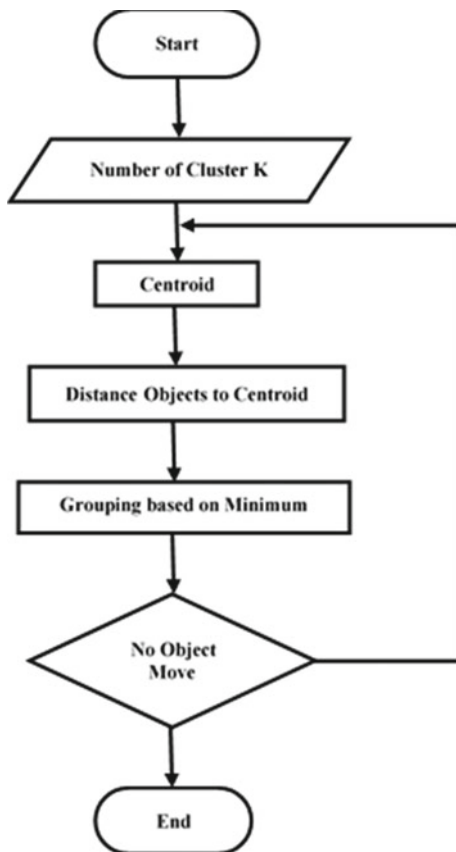
Sequence modeling is a method for generating a sequence of values by assessing a collection of input values. These input values could be time series data, which demonstrate how a specific variable varies over time, such as demand for a given product. Figure 6 illustrates how it works.

## 4 Results

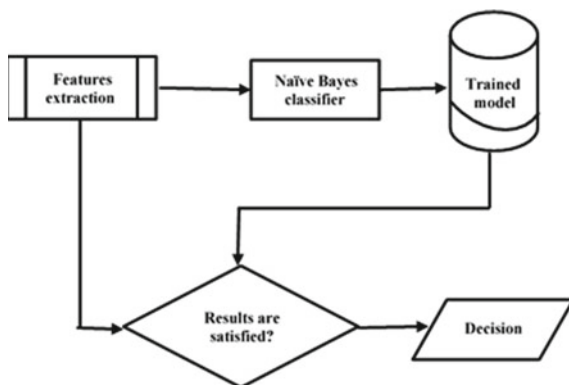
In comparison with discriminant analysis, which takes 4.77 s to construct the model, Naïve Bayes takes around 55.1 s. It can be explained by the fact that unlike other classifiers who construct models. The error rate for each algorithm is critical for determining the likelihood that patients will be misdiagnosed by the classifiers. This is shown through the average error rate where average error rate is calculated by



**Fig. 4** K nearest neighbor algorithm



**Fig. 5** Naïve Bayes classifier



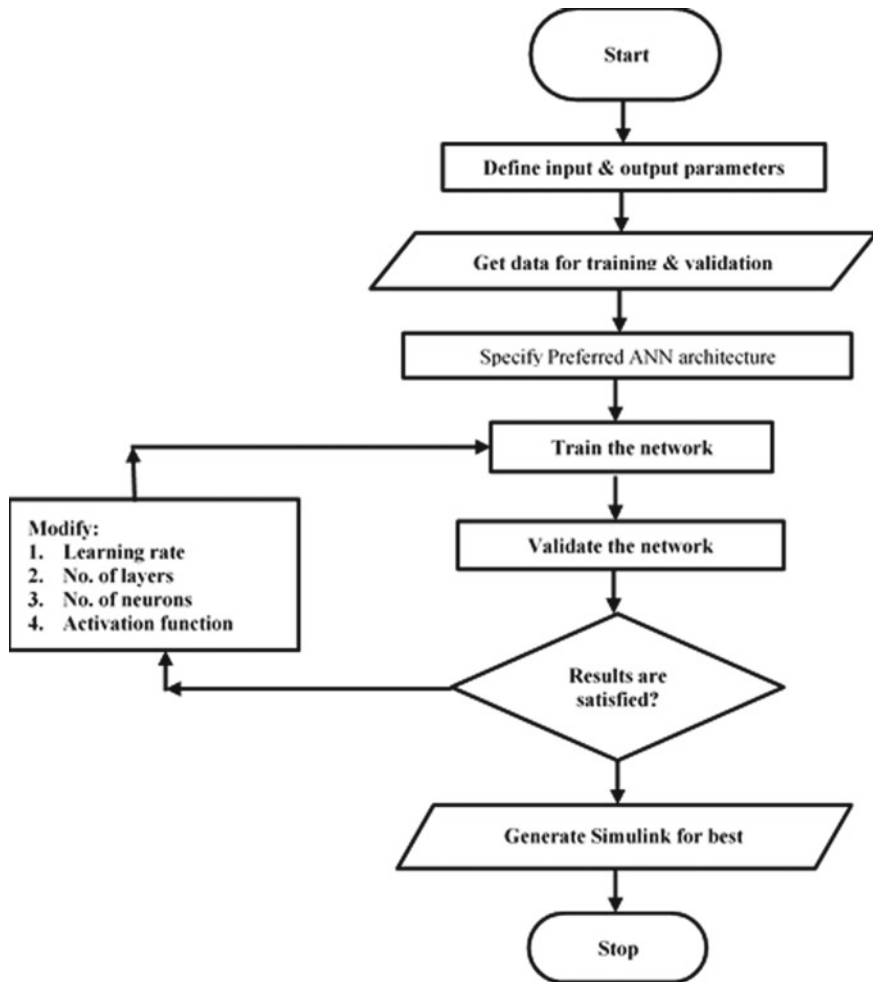


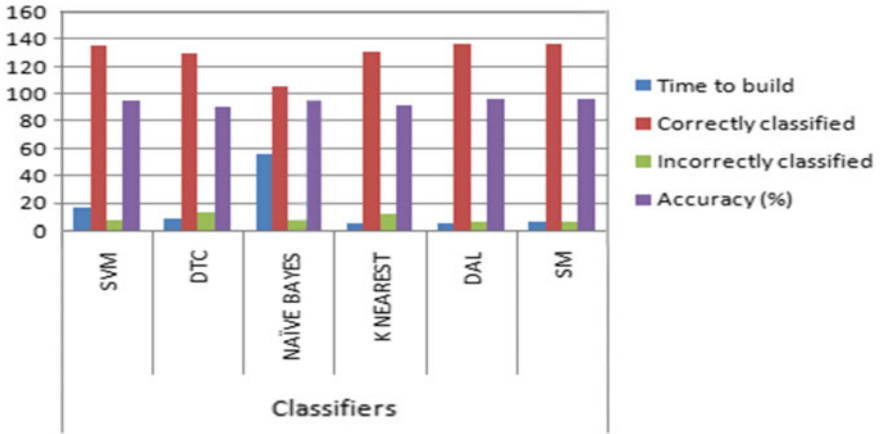
Fig. 6 Sequential modeling

$$\frac{\text{False Positive Rate} - \text{False Negative Rate}}{2}$$

From Table 1, discriminant analysis showed the optimal performance of 0.00 on false positive rate implying that 0 patients will be falsely diagnosed with breast cancer with an error rate of 0.0577 in the false negative category implies that 5.7% of patients categorized as negative have been falsely classified; from the experiment, the algorithm still displays the best average rate of 0.02855. This is the most consistent with the precision and reliability of the dataset. Due to the vast number of wrongly labeled instances for each algorithm, DTC and KNN have the highest error rates, as seen in Fig. 1 (13 for DTC and 12 for KNN) (Fig. 7 and Table 2).

**Table 1** Classifier(s) performance

Evaluation criteria	Classifiers					
	SVM	DTC	Naive Bayes	KNN	DAL	SM
False positive rate	0.025	0.1000	0.0476	0.0541	0.0000	0.0244
False negative rate	0.0583	0.0874	0.0495	0.0943	0.0577	0.0396
Average error rate	0.0417	0.0937	0.0486	0.0742	0.0289	0.0320



**Fig. 7** Performance evaluation

**Table 2** Comparison of classification accuracy

Classifiers	Symptoms	Precision	Recall	F1 score
SVM	Benign	0.94	0.99	0.97
	Malignant	0.97	0.87	0.92
DTC	Benign	0.91	0.96	0.94
	Malignant	0.90	0.80	0.85
NB	Benign	0.95	0.98	0.96
	Malignant	0.95	0.89	0.92
KNN	Benign	0.91	0.98	0.94
	Malignant	0.95	0.78	0.85
DAL	Benign	0.94	1.00	0.97
	Malignant	1.00	0.87	0.93
SM	Benign	0.95	0.99	0.97
	Malignant	0.98	0.89	0.93

## 5 Conclusion

A range of data processing and deep learning approaches is used to interpret medical data. We studied the classification prowess of six major algorithms employed for cancer classification: support vector mechanism, K nearest neighbor, decision tree classifier, Naive Bayes, sequential model, and discriminant analysis algorithm. Accuracy, precision, recall, and F1 score were used to evaluate the performance of the algorithms. Results obtained revealed that all the algorithms could be used for cancer tumor classification. However, DAL and SM achieved the highest classification accuracy of 95.80% which was closely followed by SVM and Naïve Bayes.

**Acknowledgements** The authors appreciate the sponsorship from Covenant University through its Center for Research, Innovation and Discovery, Covenant University, Ota Nigeria.




## References

1. Obermeyer Z, Emanuel EJ (2016) Predicting the future—big data, machine learning, and clinical medicine. *N Eng J Med* 375:1216–1219
2. Awotunde JB, Ogundokun RO, Jimoh RG, Misra S, Aro TO (2021) Machine learning algorithm for cryptocurrencies price prediction. In: Artificial intelligence for cyber security: methods, issues and possible horizons or opportunities. Springer, Cham, pp 421–447
3. Ogundokun RO, Awotunde JB, Misra S, Abikoye OC, Folarin O (2021) Application of machine learning for ransomware detection in IoT devices. In: Artificial intelligence for cyber security: methods, issues and possible horizons or opportunities. Springer, Cham, pp 393–420
4. Alaba AF, Oluwatobi AN, Akinkunmi AI (2017) Smallest univalue segment assimilating nucleus based brain MRI image segmentation using fuzzy C-means and fuzzy K-means algorithms. *Int J Comput Technol* 16(3):98–116
5. Kayode AA, Akande NO, Asani EO (2019) Implementation of a hybrid feature selection algorithm for improving classification of mammograms. *J Eng Appl Sci* 14:4419–4429
6. Federal Ministry of Health Nigeria National Cancer Control Plan 2018–2022 (2018). Available at [https://www.iccp-portal.org/system/files/plans/NCCP\\_Final%5B1%5D.pdf](https://www.iccp-portal.org/system/files/plans/NCCP_Final%5B1%5D.pdf)
7. Asri H, Mousannif H, Al Moatassime H, Noel T (2016) Using machine learning algorithms for breast cancer risk prediction and diagnosis. *Procedia Comput Sci* 83:1064–1069. ISSN 1877. <https://doi.org/10.1016/j.procs.2016.04.224>
8. Chaurasia V, Pal S, Tiwari BB (2014) Prediction of benign and malignant breast cancer using data mining techniques. *J Algorithms Comput Technol*
9. el Zaher M, Allah A, Eldeib A (2015) Breast cancer classification using deep belief networks. *Exp Syst Appl* 46:139–144. <https://doi.org/10.1016/j.eswa.2015.10.015>
10. Elbachiri M, Fatima S, Bouchbika Z et al (2017) Cancer du sein chez l’homme: à propos de 40 cas et revue de la littérature. *Pan Afr Med J* 28:287
11. Nahar J, Imam T, Tickler KS, Shawkat Ali ABM, Chen Y-PP (2012) Computational intelligence for microarray data and biomedical image analysis for the early diagnosis of breast cancer. *Exp Syst Appl* 39(16):12371–12377. ISSN 0957-4174, <https://doi.org/10.1016/j.eswa.2012.04.045>
12. Chaurasia V, Pal S (2014) Data mining techniques: to predict and resolve breast cancer survivability. 3(1):10–22
13. Bhardwaj A, Tiwari A, Bhardwaj H, Bhardwaj A (2016) A genetically optimized neural network model for multi-class classification. *Exp Syst Appl* 60:211–221

14. Sheikhpour R, Sarram MA, Sheikhpour R (2016) Particle swarm optimization for bandwidth determination and feature selection of kernel density estimation-based classifiers in diagnosis of breast cancer. *Appl Soft Comput* 40:113–131
15. Ghiasi MM, Zendehboudi S (2021) Application of decision tree-based ensemble learning in the classification of breast cancer. *Comput Biol Med* 128. <https://doi.org/10.1016/j.compbiomed.2020.104089>
16. Abdel-Zaher AM, Eldeib AM (2016) Breast cancer classification using deep belief networks. *Exp Syst Appl* 46:139–144
17. Pota M, Esposito M, De Pietro G (2017) Designing rule-based fuzzy systems for classification in medicine. *Knowl Base Syst* 124:105–132
18. Risk factors for breast cancer (n.d.) Havard health. Retrieved 2 March 2021, from <https://www.health.harvard.edu/cancer/risk-factors-for-breast-cancer>
19. Breast Cancer Survivors (2021). <https://www.wcrf.org/Dietandcancer/Breast-Cancer-Survivors>. Retrieved 18 January 2021
20. Pratheep K, Amala M, Nair GG (2021) An efficient classification framework for breast cancer using hyper parameter tuned Random Decision Forest Classifier and Bayesian Optimization. *Biomed Sig Process Control* 68:102682. <https://doi.org/10.1016/j.bspc.2021.102682>
21. Wang P, Wang J, Li Y, Li P, Li L, Jiang M (2021) Automatic classification of breast cancer histopathological images based on deep feature fusion and enhanced routing. *Biomedical Sig Process Control* 65. <https://doi.org/10.1016/j.bspc.2020.102341>

# PCA-Based Feature Extraction for Classification of Heart Disease



Roseline Oluwaseun Ogundokun , Sanjay Misra ,  
Joseph Bamidele Awotunde , Akshat Agrawal, and Ravin Ahuja

## Abbreviations

ML	Machine Learning
PCA	Principal Component Analysis
HD	Heart Disease
DR	Detection Rate
NB	Naïve Bayes
RF	Random Forest
SVM	Support Vector Machine
FPR	False-Positive Rate
DM	Data Mining

---

R. O. Ogundokun

Department of Computer Science, Landmark University Omu Aran, Omu Aran, Nigeria

e-mail: [ogundokun.roseline@lmu.edu.ng](mailto:ogundokun.roseline@lmu.edu.ng); [rosogu@ktu.lt](mailto:rosogu@ktu.lt)

Center of ICT/ICE, CUCRID, Covenant University, Ota, Nigeria

Department of Multimedia Engineering, Kaunas University of Technology, Kaunas, Lithuania

S. Misra (✉)

Department of Computer Science and Communication, Ostfold University College, Halden,

Norway

e-mail: [sanjay.misra@covenantuniversity.edu.ng](mailto:sanjay.misra@covenantuniversity.edu.ng)

J. B. Awotunde

Department of Computer Science, University of Ilorin, Ilorin, Nigeria

e-mail: [awotunde.jb@unilorin.edu.ng](mailto:awotunde.jb@unilorin.edu.ng)

A. Agrawal

Amity University, Gurgaon, India

R. Ahuja

Shri Vishwakarma University, Gurgaon, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2022

T. Sengodan et al. (eds.), *Advances in Electrical and Computer Technologies*,

Lecture Notes in Electrical Engineering 881,

[https://doi.org/10.1007/978-981-19-1111-8\\_15](https://doi.org/10.1007/978-981-19-1111-8_15)

AI Artificial Intelligence  
EC Ensemble Classifier

## 1 Introduction

Heart disease (HD) is the foremost originator of demise in both India and the rest of the world. Conferring to the World Health Organization (WHO), cardiac illnesses claim the existence of 17.7 million persons per annum, computing for 31% of the entire fatalities globally [1]. It was concluded that the proposed PCA+SVM classifier outperformed the other two classifiers with a precision of 88.44%, specificity of 67.31, f1-score of 65.49%, and FPR of 0.3269 while proposed PCA+NB outperformed the other two algorithms with an accuracy of 55.33% and detection rate of 52.88%. Data mining (DM), artificial intelligence (AI), or ML tools might be employed to extract knowledge from a dataset. The newly found information can be utilized by hospital administrators to improve service quality. Medical practitioners can utilize the new knowledge to decrease the number of adverse medication effects and to recommend less-priced therapeutically comparable alternatives [2, 3]. One of the most significant uses of DM, AI, or ML techniques in healthcare management is predicting a patient's future behavior based on their past behavior. The provision of high-quality services at reasonable prices is a key problem for healthcare institutions (hospitals, medical facilities) [4, 5]. Quality service includes accurately diagnosing patient's health challenges and providing effective therapies. Poor clinical judgments can have devastating effects, which are unacceptably dangerous. Clinical testing must also be kept to a minimum in hospitals. They can attain these outcomes by utilizing suitable computer-based information and/or decision-making technologies. The amount of data in the healthcare industry is enormous [6] that is it comprises patient information, resource management information, and data that have been modified. Organizations in the healthcare industry must be able to evaluate data. Millions of patient information may be saved, and computers and data mining techniques may aid in addressing numerous essential healthcare issues. Clinical choices are frequently made built on doctors' instinct and knowledge rather than the database's knowledge-rich data. This practice results in unintended biases, mistakes, and exorbitant medical expenditures, all of which affect the eminence of care delivered to the patient role. Integrating clinical decision support systems with computer-based patient records, according to Wu et al., might reduce medical mistakes, increase patient safety, eliminate undesired practice variation, and improve patient outcomes [6]. This study has promised to deliver data modeling and analysis technologies such as machine learning, thereby having the ability to create a knowledge-rich environment that may assist to enhance the quality of healthcare choices considerably.

This study, therefore, is projected to examine the discriminative power of numerous prognosticators in the research to intensify the efficiency of HD discovery

through their warning signs. A dimensionality reduction method PCA and three ML classifiers which are NB, RF, and SVM are employed in the study. The three ML classifiers were implemented with PCA DR. The techniques' performance was evaluated on a dataset attained from National Health Service (NHS) database. The system was evaluated using confusion matrix measures which include accuracy, precision, specificity,  $f1$ -score, detection rate, and FPR. A comparative analysis was conducted between the three ML classifiers with PCA DR. The techniques' performance was evaluated on a dataset attained from National Health Service (NHS) database. The study major contribution is as follows:

- a. Identification of well-known ML classifiers techniques employed for heart disease (HD) detection
- b. Computation of system results over the dataset gathered from NHS database
- c. Introduction of a dimensionality reduction called PCA
- d. Implementation of three classification techniques which are NB, RF, and SVM
- e. Comparative analysis of the results with performance metrics

The remaining part of the article is structured as thus: The literature review is presented in the second segment. Segment 3 discussed the materials and method employed for the implementation of the study, whereas the results and discussion are discussed in segment 4. Segment 5 lastly concludes the study and the impending study is suggested.

## 2 Literature Review

ML is a branch of AI that deals with the advancement and research of procedures that can acquire knowledge from datasets [1, 6]. Machine learning provides a set of rules and procedures that makes the system computationally intelligent [7]. These rubrics create systems based on the contribution, which are subsequently employed to take verdicts or forecast outcomes [8]. Because the health sector plays such a significant role in people's lives, extracting information from medical databases has become a genuine challenge. Medical data research is becoming increasingly popular. Treatment, screening, diagnosis, management, prognosis, and monitoring are just a few of the several sorts of medical solutions [9, 10]. Furthermore, as the healthcare industry becomes more computerized, machine learning methods are critical to assisting clinicians in recognizing and correcting problems at an early stage. One of the most significant challenges in the medical realm is a medical diagnosis. The accuracy of the diagnosis aids the physician in making the best selection for the best course of action and, ultimately, in illness treatment. Machine learning is frequently utilized in the diagnosis of diseases, for instance, HD, cancer, skin disease, and diabetes. ML procedures assist to improve the diagnostic procedure's accuracy, dependability, and rapidity. In the context of particular disorders such as heart disease, researchers have effectively applied machine and statistical learning approaches to create prediction models, much as they have in other fields. Because of the limits in current heart



disease diagnosis methods and the difficulties in predicting heart illness, researchers are looking for more acceptable approaches.

Machine learning can help forecast and diagnose the cardiac disease as a result of its capacity to excerpt information and uncover concealed designs. To increase prediction accuracy, several initiatives have been involved to forecast HD utilizing ML techniques. Few of the prevailing researches that have used ML classification techniques to forecast and diagnose HD are discussed as follows:

Bashir et al. [11] proposed employing an ensemble classifier dubbed BagMoov. The NB, linear regression (LR), quadratic discriminant analysis (QDA), instance-built learner, and SVM techniques were suggested in their study, and these techniques were centered on bagging and multi-objective optimal-weighted voting. To predict CAD, Arabasadi et al. [12] presented a correct combination method. By determining the initial weights of the neural network (NN) and applying a genetic algorithm (GA), their suggested technique enhanced the accuracy of the NN by 10%. Jabbar [13] enhanced the performance of the KNN algorithm to improve HD prognosis. In this technique, particle swarm optimization (PSO) was utilized to close the features during the preprocessing stage, and thereafter, KNN was trained on the data generated from the preprocessing stage. To forecast the risk of cardiac failure, Samuel et al. [14] employed an ANN technique and a fuzzy analytic hierarchy procedure (AHP) was utilized to calculate global weights of features established on their specific improvements in their techniques. The global weights (GWs), which indicated the contribution of each feature, are then added to train the ANN classifier to envisage the probability of HD in the affected role. Yekkala et al. [15] postulated a bagging ensemble technique, RF, and AdaBoost with PSO. PSO was employed as an FE technique to forecast HD. The findings of the study revealed that using PSO with bagging algorithm produces excellent accuracy. Mustafa et al. [16] presented an ensemble technique to improve prognosis by combining the prediction abilities of multi-classifiers. Ensemble learning was used in their study to forecast and identify the recurrence of cardiovascular illness by combining five classifiers which include SVM, ANN, NB, RF, and regression examination. Founded on ReliefF and rough set theory, Liu et al. [17] created a hybridizing classifier system called RFRS. An RFRS feature selection technique and a classification system with an ensemble classifier (EC) are the two subsystems employed for the proposed technique. The first model has three stages: discretization, FE with the ReliefF method, and feature reduction with a heuristic rough set reduction approach. An EC built on C4.5 was employed in the second model. Nguyen et al. [18] projected a health classification technique that combined a fuzzy typical improver method with a GA technique. The adaptive vector quantization clustering is in charge of rule initialization in this technique. A genetic algorithm and a gradient descent method, respectively, were used for rule optimization and constraint tuning. To excerpt discriminatory features from high-dimensional datasets, the wavelet transformation was used.

It was discovered from the literature reviewed that the use of feature extraction approaches is ignored by several of the prevailing research works to enhance the classification degree of state-of-the-art ML approaches. Hence, in this study, efficient ML classifiers were employed to enhance the classification outcomes that is a feature

extraction (FE) technique PCA, and three ML classifiers which are NB, RF, and SVM are employed in the study.

### 3 Materials and Method

In this segment, the dataset acquisition and the projected technique in this study are presented.

#### 3.1 Dataset Attainment

A heart disease dataset gathered from the National Health Service database is employed in this study. The dataset is assumed founded on factual hospital directorial data named Hospital Episodes Statistics. The assumed excerpt comprises an arbitrary sample of emergency admissions for HD (ICD10 code 150).

#### 3.2 Proposed Method

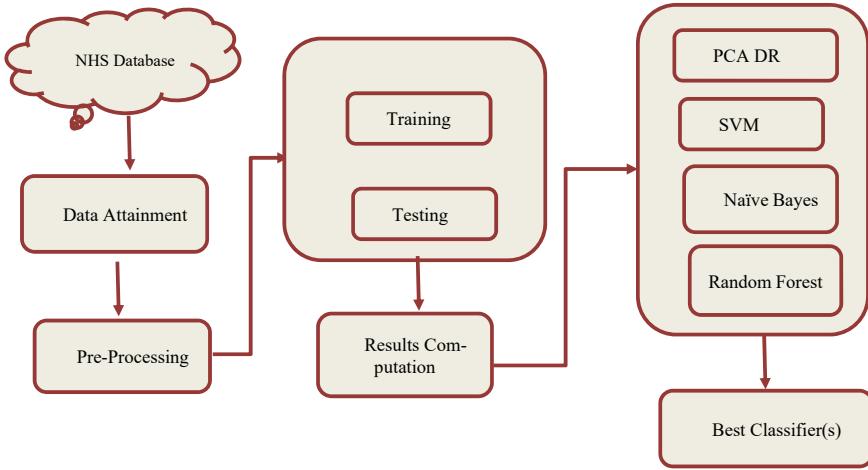
The proposed method starts with data attainment from the NHS database which is followed by the preprocessing stage. The classifiers are thereafter trained and tested on the HD dataset. The ML classifiers with PCA DR were implemented, and classification was obtained for each of the classifiers, and these were computed and evaluated to recognize the best ML classifier(s) for HD discovery. Figure 1 thereby displays the block movement for the proposed system.

#### 3.3 ML Classifiers

The choice of ML classifiers hinges on certain aspects like their extensive use in the context of disease identification. A brief description of the ML classifiers, namely PCA, SVM, NB, and RF.

**Naïve Bayes:** The Naive Bayes classifier is founded on the NB theorem and operates on the likelihood of occurrences. All characteristics are independent, according to NB [19].

**Support Vector Machine (SVM):** SVM is a classifier that uses the nearest data points to create a decision region beside a margin. They have supervised learning models that examine data for classification and regression analysis and have related learning algorithms [20].



**Fig. 1** Proposed system block flow

**Random Forest (RF):** Random forests are a type of ensemble learning approach for classification, regression, and further problems that works by building a large number of DT throughout training. For classification complications, the RF output is the class preferred by the great number of trees.

**Principal Component Analysis (PCA):** PCA is an approach for decreasing the dimensionality of datasets while maximizing interpretability minimizing information loss. This is achieved by engendering novel uncorrelated variables that optimize variance [21].

### 3.4 Performance Evaluation

Six distinct statistics, which include accuracy, precision, specificity,  $f1$ -score, DR, and FPR, are produced utilizing the values of the confusion matrix to assess the efficiency and competence of the ML classifiers approach. These metrics, whose formulation is shown in Eqs. (1–6), are equally important for comparing the different system methods.

$$\text{Accuracy} : \frac{TP + TN}{TP + TN + FN + FP} \quad (1)$$

$$\text{precision} : \frac{TP}{TP + FP} \quad (2)$$

$$\text{Specificity} : \frac{TN}{TN + FP} \quad (3)$$

**Table 1** Confusion matrix

	Predicted 0	Predicted 1
Actual 0	TN	FP
Actual 1	FN	TP

$$F1 - \text{score} : 2 \times \frac{\text{Precision} \times \text{Sensitivity}}{\text{Precision} + \text{Sensitivity}} \tag{4}$$

$$\text{Detection rate} : \frac{TP}{TP + FN} \tag{5}$$

$$\text{FPR} : \frac{FP}{FP + TN} \tag{6}$$

True-positive, true-negative, false-positive, and false-negative rates of classification techniques are denoted by the letters TP, TN, FP, and FN, respectively. The calculated results of the used ML methods are displayed in Table 1 in a comparison style using these equations. The likelihood of a false alarm being sounded that is a positive result being provided, while the real value is negative, is identified as the FPR. The detection rate, commonly identified as the true-positive rate (TPR), is the proportion of valid positive class predictions made to all predictions made. The accuracy rate is regarded as an excellent and acceptable result for HD detection.

## 4 Results and Discussion

### 4.1 Experimental Setup

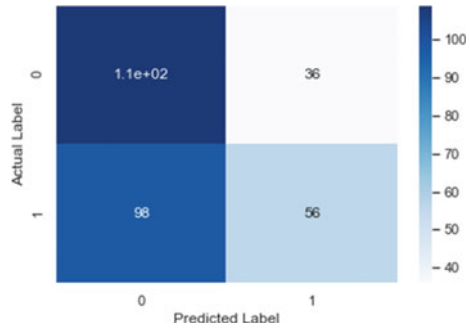
The study executed the postulated ML context for HD detection utilizing Python programming language and implemented the models in a DEL computer with Intel i7 3.40 GHz CPU and 8 GB memory.

### 4.2 Analysis of Result

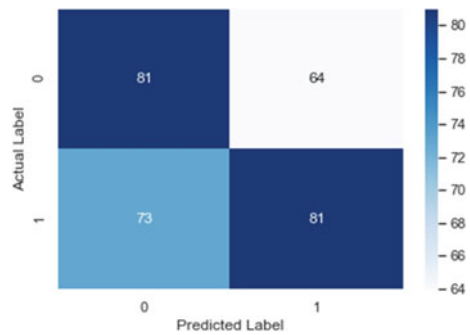
In this segment, the study directed different experiments to assess the two ML procedures for the detection of HD. The investigation of the two classifiers was evaluated for the exposure of the HD dataset. Table 1 shows the values obtained for the confusion matrix using different ML techniques.

TP: This means true positive, and it means that the patients diagnose have HD, and the test is positive.

**Fig. 2** Confusion matrix for PCA + NB



**Fig. 3** Confusion matrix for PCA + RF



FP: This means false positive, and this means that the patients diagnose do not have HD, but the test is positive.

TN: This means true negative meaning the patient diagnoses does not have HD, and the test is negative.

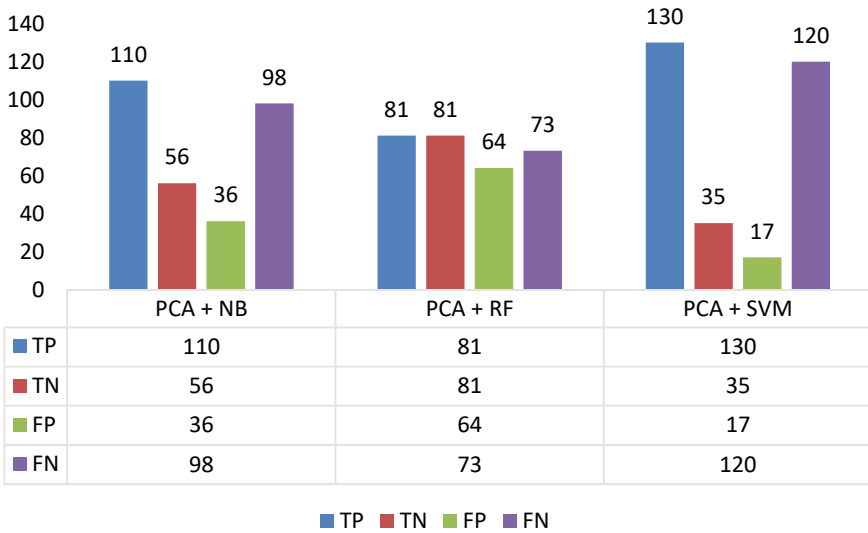
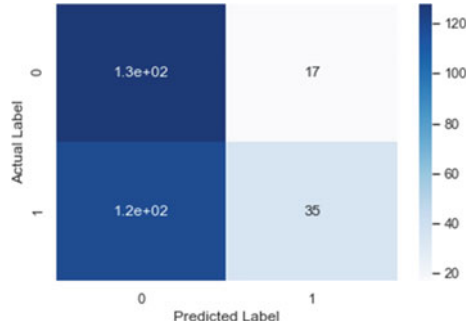
FN: This means false negative, and this means the patient has HD, but the test is negative.

In the study, the preprocessed dataset was utilized to execute the implementation, and the projected techniques were employed and applied. The mentioned performance metrics were obtained utilizing a confusion matrix as seen in. The confusion matrix demonstrated the performance of the system, and the confusion matrix obtained by the projected techniques is shown in Figs. 2, 3 and 4, and their values are shown in Fig. 5.

### 4.3 Discussion

The classification accuracy, precision, specificity, *f1*-score, detection rate, and FPR for the three classifiers that are PCA + NB, PCA + RF, and PCA + SVM are summarized in Table 2. Based on the derived results, PCA + NB outperformed PCA + RF and PCA + SVM concerning accuracy. While all three models tend to

**Fig. 4** Confusion matrix for PCA + SVM



**Fig. 5** Values obtained from confusion matrix using different classifiers

**Table 2** Analysis of the ML classifiers

Classifiers	Accuracy (%)	Precision (%)	Specificity (%)	F1-score (%)	DR (%)	FPR
PCA + NB	55.33	75.34	60.87	62.15	52.88	0.3913
PCA + RF	54.18	55.86	55.86	54.18	52.60	0.4414
PCA + SVM	54.64	88.44	67.31	65.49	52	0.3269

show an accuracy rate of more than 50%, accuracy alone can't be considered as the performance criterion for the underlying study. Out of the original 299 instances, only, 155 patients are said to have HD, whereas the remaining 144 individuals do not suffer from HD. It is seen that there is only a slight difference between the performance

of PCA + NB and PCA RF concerning detection rate and FPR measures. PCA + SVM classifier outperformed that of the other two classifiers in terms of precision of 88.44%, specificity 67.81%, *f1*-score 65.49%, and FPR value of 0.3269. It was also discovered that PCA + NB outperformed PCA + RF and PCA + SVM in terms of DR of 52.88%.

## 5 Conclusion

With the rising number of deaths due to heart disease, it has become necessary to build a system that can efficiently and precisely forecast heart illness. The study aimed to discover the most effective machine learning system for detecting cardiac problems. This research evaluates the accuracy, detection rate, and FPR scores of PCA with NB, RF, and SVM techniques for forecasting HD using NHS database dataset. The result of the research specified that the PCA + SVM classifier outperformed that of the other two classifiers in terms of precision of 88.44%, specificity 67.81%, *f1*-score 65.49%, and FPR value of 0.3269. It was also discovered that PCA + NB outperformed PCA + RF and PCA + SVM in terms of DR of 52.88 for the detection of HD.

The limitation of this study is that the dataset was small which led to the poor values and percentage for the accuracy, precision, specificity, *f1*-score, and detection rate with the high FPR value. Therefore, in the future, the study can be improved by using a larger HD dataset when evaluated to the one employed in this analysis, and this will assist to deliver better results and assist medical practitioners in diagnosing the HD efficiently and effectively. The study in the future can also be improved by the emergent of a Web or android application based on the best ML classifier.

**Acknowledgements** The authors appreciate the sponsorship from Covenant University through its Center for Research, Innovation, and Discovery, Covenant University, Ota Nigeria

## References

1. Kononenko I (2001) Machine learning for medical diagnosis: history, state of the art and perspective. *J Artif Intell Med* 1:89–109
2. Gultepe Y, Rashed S (2019) The use of data mining techniques in heart disease prediction. *Int J Comput Sci Mob Comput* 8(4):136–141
3. Pouriyyeh S, Vahid S, Sannino G, De Pietro G, Arabnia H, Gutierrez J (2017) A comprehensive investigation and comparison of machine learning techniques in the domain of heart disease. In: IEEE symposium on computers and communications (ISCC). Heraklion, pp 204–207. <https://doi.org/10.1109/ISCC.2017.8024530>
4. Oladele TO, Ogundokun RO, Misra S, Adeniyi JK, Jaglan V (2021) A joint neuro-fuzzy malaria diagnosis system. *J Phys: Conf Series* 1767(1):012038
5. Gokulnath CB, Shantharajah SP (2018) An optimized feature selection based on a genetic approach and support vector machine for heart disease. *Clust Comput* 22(S6):14777–14787

6. Abdulkadir SJ, Yong SP (2013) Unscented Kalman filter for noisy multivariate financial time-series data. In: International workshop on multi-disciplinary trends in artificial intelligence, pp 87–96
7. Abdulkadir SJ, Yong SP, Zakaria N (2016) Hybrid Neural Network Model for Metocean Data Analysis. *J. Informatics Math. Sci.* 8(4):245–251
8. Abdulkadir SJ, Yong SP, Marimuthu M, Lai FW (2014) Hybridization of ensemble Kalman filter and non-linear autoregressive neural network for financial forecasting. In: Mining intelligence and knowledge exploration. Springer, pp 72–81
9. Silver MA et al (2004) BNP consensus panel 2004: a clinical approach for the diagnostic, prognostic, screening, treatment monitoring, and therapeutic roles of natriuretic peptides in cardiovascular
10. Ogundokun RO, Sadiku PO, Misra S, Ogundokun OE, Awotunde JB, Jaglan V (2021) Diagnosis of long sightedness using neural network and decision tree algorithms. *J Phys: Conf Series* 1767(1):012021
11. Bashir S, Qamar U, Khan FH (2015) BagMOOV: a novel ensemble for heart disease prediction bootstrap aggregation with multi-objective optimized voting. *Australas Phys Eng Sci Med* 38(2):305–323
12. Arabasadi Z, Alizadehsani R, Roshanzamir M, Moosaei H, Yarifard AA (2017) Computer-aided decision-making for heart disease detection using a hybrid neural network-genetic algorithm. *Comput Methods Programs Biomed* 141:19–26
13. Jabbar MA (2017) Prediction of heart disease using k-nearest neighbor and particle swarm optimization. *Biomed Res (0970-938X)* 28(9)
14. Samuel OW, Asogbon GM, Sangaiyah AK, Fang P, Li G (2017) An integrated decision support system based on ANN and Fuzzy\_AHP for heart failure risk prediction. *Expert Syst Appl* 68:163–172
15. Yekkala I, Dixit S, Jabbar MA (2017) Prediction of heart disease using ensemble learning and particle swarm optimization. In 2017 International conference on smart technologies for smart nation (SmartTechCon). IEEE, Aug 2017, pp 691–698
16. Mustafa J, Awan AA, Khalid MS, Nisar S (2018) Ensemble approach for developing a smart heart disease prediction system using classification algorithms. *Res Rep Clin Cardiol* 9:33
17. Liu X, Wang X, Su Q, Zhang M, Zhu Y, Wang Q, Wang Q (2017) A hybrid classification system for heart disease diagnosis based on the RFRS method. *Comput Math Methods Med*
18. Nguyen T, Khosravi A, Creighton D, Nahavandi S (2015) Classification of healthcare data using the genetic fuzzy logic system and wavelets. *Expert Syst Appl* 42(4):2184–2197
19. Faisal MI, Bashir S, Khan ZS, Khan FH (2018) An evaluation of machine learning classifiers and ensembles for early-stage prediction of lung cancer. In 2018 3rd International conference on emerging trends in engineering, sciences, and technology (ICEEST). IEEE, Dec 2018, pp 1–4
20. Abdulsalam SO, Mohammed AA, Ajao JF, Babatunde RS, Ogundokun RO, Nnodim CT, Arowolo MO (2020) Performance evaluation of ANOVA and RFE algorithms for classifying microarray dataset using SVM. *Lecture Notes Bus Inform Process* 402:480–492
21. Rasheed J, Hameed AA, Djeddi C, Jamil A, Al-Turjman F (2021) A machine learning-based framework for the diagnosis of COVID-19 from chest X-ray images. *Interdiscip Sci: Comput Life Sci* 13:103–117. <https://doi.org/10.1007/s12539-020-00403-6>



# Software Fault Prediction Using Machine Learning Algorithms



M. S. Pavana, M. N. Pushpalatha, and A. Parkavi

## Abbreviations

ML	Machine learning
SFP	Software fault prediction
SVM	Support vector machine
LR	Logistic regression
NB	Naïve Bayes
RF	Random forest
FE	Feature extraction
FS	Feature selection

## 1 Introduction

In this current scenario, defects are the most serious issues, and predicting them is a challenging process or task. The existence of this faults/defect may cause high risk of project failure. In addition to, it may cause in decrease of the project quality

---

M. S. Pavana (✉) · M. N. Pushpalatha  
Department of Information Science and Engineering, M S Ramaiah Institute of Technology,  
Bangalore, India  
e-mail: [pavanashivanandms@gmail.com](mailto:pavanashivanandms@gmail.com)

M. N. Pushpalatha  
e-mail: [pushpalathmn@msrit.edu](mailto:pushpalathmn@msrit.edu)

A. Parkavi  
Department of Computer Science and Engineering, M S Ramaiah Institute of Technology,  
Bangalore, India  
e-mail: [parkavi.a@msrit.edu](mailto:parkavi.a@msrit.edu)

and increasing of the cost and time consumption. Therefore, detecting these errors at early stage of SDLC leads to decrease in time consumption and also cost of the overall project. As a result, defect/fault prediction plays crucial role in the developing and testing phases, and it adds to the entire projects success. Defects/faults should be predicted at early stages of SDLC. For this purpose, various software faults/defects prediction models have been built for practitioners to find the modules identifies as faulty in initial phases [7].

The classification model for predicting the faults is very important; otherwise, it will more hectic to developers because of high cost and effort, and time scheduled will be delayed. This inspires the development of defect-prone model to overcome these problem. Software fault prediction via machine learning techniques is most frequent use case among the researchers and developers. Many machine learning techniques have been applied to detect the defects.

In this paper, four machine learning algorithms were implemented to predict the faults. Software fault/defect prediction model consists of different stages to give better performance of the machine learning model in predicting the software faults. Different stages such as importing dataset, data preprocessing, feature selection, feature extraction, and performance evaluation. The dataset is collected from the NASA repository which consists of open-source datasets such are CM1, PC1, PC2, and JM1. To perform this experiment, we have used CM1 dataset which includes 22 attributes and 498 instances. The preprocessing stage consists of data cleaning, identifying missing values, and handling them. Feature selection steps include the reduction of input variables by applying correlation technique such as Spearman's and Kendall method. It is also a kind of dimensionality reduction. The feature extraction step includes reduction of input features. This step is called dimensionality reduction. Two reduction techniques were used they are principal component analysis (PCA) and linear discriminant analysis (LDA). Machine learning algorithms are used to distinguish classes either faulty or non-faulty. Many machine learning techniques were used to develop a high performance prediction model using software metrics for predicting problems. In this paper work, machine learning algorithm such as NB, RF, LR, and SVM was applied to the sampled dataset. After obtaining results from each of the machine learning algorithm are compared to identify the best algorithm with highest accuracy. This paper gives the comparative analysis of four different algorithms.

## 2 Related Works

In this section, we have included the previous related work done by researcher's based on different techniques was explained.

Ezgi Ertruk et al. [1], authors implemented adaptive neuro-fuzzy inference system (ANFIS) for the software fault/defect prediction model problem. Dataset

was collected from Promise repository, and McCabe's metric is selected, and accuracy is used as a performance evaluation. The result shows 0.77, 0.88 for SVM and ANN, respectively.

Wu [2] performed case study on empirical analysis to predict faults in the software. Author used KC1 dataset from NASA, which uses different metrics like WMC, NOC, LCOM, and CBO for defect/fault prediction.

Yang et al. [3] proposed defect prediction by applying NB evaluation method for the prediction and performance evaluations like ROC, TPR, and FPR were used. The usage of feature selection technique was enhanced to improve the results. During the data extraction, accuracy of model didn't achieve performance compared to selection method.

Anuradha chug et al. [6] implemented different algorithms for predicting defects. The datasets are collected from the NASA. Classification and clustering techniques were applied to build a defect prediction model. Several measures like ROC, precision, and MAE are used to perform the evaluation of the ML model.

In this paper [7], authors used feature selection technique like swarm optimization algorithm, and the bagging technique was used for data class imbalance. In this research, nine datasets were used from NASA. Authors used different classifier on them; comparison of results of both swarm and bagging techniques is done. Results show impressive improvement by using both methods.

In this research [13], authors proposed ensemble method for the feature selection, and dataset is collected from open-source dataset. This paper shows comprehensive study of different ensembles methods and filtering techniques to improve the accuracy. Ensemble methods show impressive improvement in the accuracy.

Ayse Tosum et al. [8] proposed a Naïve Bayes with optimized threshold which performs better default threshold. In this paper, author selected different 13 datasets for the evaluation purpose. For performance evaluation, ROC curve was considered.

In this paper [9], authors proposed multi-layered perceptron, Naive Bayes and Bayesian networks to build a predictive model. Ten datasets were downloaded from NASA repository. Comparisons of different classifiers were observed. Results show neural network classifier models are more superior when compared to other classifiers.

Ren et al. [11], in this paper, authors have proposed a model to solve the imbalance class problem which in turn reduces the performance of the predictive model. Authors have proposed Gaussian function that is asymmetric kernel partial least squares classifier (AKPLSC). NASA dataset is used in this paper.

Gray et al. [10] used SVM classifier for the static code metrics for NASA dataset for the automated software defect module detection. For the implementation of the classification problem includes preprocessing (cleaning dataset) removal of repeated data. In this paper, SVN shows 70% of accuracy.

Jaspreet kaur et al. [5], in this paper, k-means clustering application is used for fault-prone module of the object-oriented system, and they used metrics values of JEdit open-source software to generate criteria for categorizing software modules into the categories of faulty and not faulty. It also shows the high and low value of probability of detection.

Pushpalatha et al. [4], predicting the severity of NASA bug reports is performed using the different ensemble methods, and data reduction is done using chi-square and information gain.

### 3 Methodologies

This section describes the dataset collection, data preprocessing, feature selection methods, feature extraction methods, classification algorithms, and performance metrics. The below figure describes the proposed methodology of the model (Fig. 1).

#### 3.1 Dataset

Dataset is collected from the NASA metric data program (MDP) which was donated by Tim Menzies. There are different cleaned datasets are available in NASA repository. In this project, CM1 dataset is used as an input data. This dataset has 22 attributes which includes five line of code measures, four McCabe's metrics, and four Halstead metrics, and eight derived Halstead measures, a branch count and with known output class(defects) which represents faulty (True) and non-faulty (False) instances. In this project, dataset is divided into 75% of training data and remaining 25% of testing data.

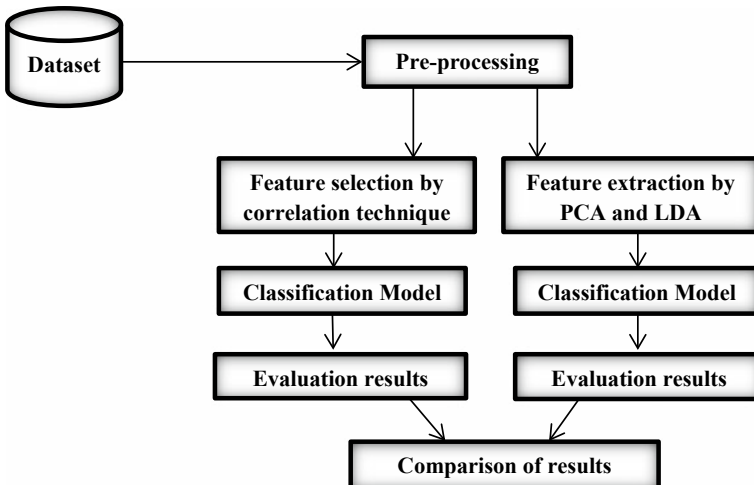


Fig. 1 Proposed methodology

### 3.2 Data Preprocessing

Our main approach aims to create a high-quality training dataset from the original dataset because it may contain redundant features and imbalanced classes in the dataset. The main goal is to increase performance of the machine learning (classification) models. In this preprocessing stage, we undergo different steps; they are importing required libraries and dataset, identifying missing values and handling missing values, data cleaning, encoding the categorical data, splitting the dataset etc. The splitting of dataset ratio is (75:25), i.e., 75% of training data and 25% of testing data.

### 3.3 Feature Extraction Using Dimensionality Reduction Methods

In the current situation, to get more accurate results, we tend to add more features. On basis of this, final prediction is done. At some point, this may lead to over fitting and reduce the performance of the model because of using more features. The higher the number of variables the more difficulty in visualizing and exploring the data. This is where dimensionality reduction comes into play. Dimensionality reduction seeks lower dimensional representation of numerical input data that preserves most important relationships in the data. This dimensionality reduction is a process of reducing the input features in the dataset. In this paper, two dimensionality reduction methods are used; they are LDA and PCA.

**Principal component analysis (PCA):** PCA is a dimensionality reduction technique used to reduce the high features, i.e., higher dimensionality to lower dimensionality. It reduces larger dataset to smaller dataset in order to give better accuracy. It is easier for data exploration and visualization. The algorithm to calculate PCA is as follows:

Step1: Standardize the data.

We need to find the mean and standard deviation for each features. Standardization formula is shown below:

$$x_{\text{new}} = \frac{x - \mu}{\sigma} \quad (1)$$

Step2: Calculate covariance matrix.

The formula to calculate covariance matrix is given below:

$$\text{cov}(x, y) = \frac{\sum (x_i - \bar{x}) * (y_i - \bar{y})}{(N - 1)} \quad (2)$$

Step3: Sort the eigenvalues in descending order and choose  $k$ -eigenvectors.

Step4: Build projection matrix from the selected  $k$ -eigenvectors.

Step5: Transform the original dataset via projection matrix to obtain a  $k$ -dimensional feature subspace.

**Linear discriminant analysis (LDA):** LDA is one of the feature extractions method of dimensionality reduction. LDA is a supervised machine learning method used to reduce the features (columns) in the given dataset. LDA is used to solve dimensionality reduction for data with higher attributes. LDA finds the linear relationship between the features that separates the different classes. The algorithm to calculate LDA is follows:

Step1: Computing the dimensional mean vectors.

Step2: Computing the scatter matrices.

Step3: Solving the generalized eigenvalue for the matrices.

Step4: Selecting the linear discriminants for new feature (variable) subspace.

Step5: Finally transform the samples to new subspaces.

### 3.4 Feature Selection Using Correlation Method

Feature selection is process of identifying and removing the redolent and irrelevant input variables. This helps in reduction of over fitting and helps to allow machine learning algorithms to perform faster and effectively. Feature selection steps play an import role in building a good model. There are different feature selection methods are available. One of the most prominent and simple feature selection method is filter-based correlation coefficient. In this paper, Spearman's rank coefficient technique is used for feature selection.

**Spearman's rank coefficient:** This is non-parametric test used to measure the connection between two variables. The Spearman's rank coefficient is calculated using the following formula:

$$\rho = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)} \quad (3)$$

$\rho$  = Spearman's rank correlation coefficient.

$d_i$  = Difference between the two variables of each observations.

$n$  = Number of observations.

In Spearman's rank correlation, R value which indicates the degree to which variables are correlated, i.e.,

$R = 1$  Strong degree of relationship between two variables.

$R = 0$  No correlation.

$R = -1$  Negative degree of correlation.

### 3.5 Classification

The aim of this project is to classify the software defects/faults as faulty or non-faulty. In this experiment, machine learning (ML) algorithms such as Naïve Bayes, support vector machine, logistic regression, and Random forest were applied to the feature extracted data. The result will be class 0 which indicates true (faulty) and class 1 indicates false (Non-faulty).

**Support vector machine:** SVM is a supervised machine learning technique that may be used to solve the classification and regression problems. It identifies the optimal hyperplane to classify between the two different classes.

**Logistic Regression (LR):** LR is a supervised machine learning algorithm which is to predict the probability of a target class, i.e., output class. Typically used to represent a discrete group of classes. This LR uses sigmoid function to predict the probability of an output.

**Naïve Bayes (NB):** NB is effective, simple, and widely used machine learning algorithm, and it works on Bayes' theorem. To classify the data, NB uses posterior decision rule.

**Random Forest (RF):** RF is a type of ensemble learning algorithm based on the trees. From randomly collected training data, this RF builds a group of decision trees (DTs). It combines all of the DT polls and selects the majority poll.

### 3.6 Performance Evaluation

In machine learning, models are able to give accurate prediction to solve the problem of classification. Once predictive model is built, the evaluation techniques were applied to testing dataset to show how well a machine learning model generalize to new results. This evaluation aims to estimate the accuracy of a model.

#### Accuracy

Accuracy is the simplest prediction metrics, which is the ratio of number of correct prediction with all prediction made. The accuracy equation is shown below.

$$\text{Accuracy (AC)} = \frac{\text{no. of correct predictions}}{\text{total no. of predictions}} \quad (4)$$

For binary classification, in terms of positive and negative values, accuracy can be evaluated.

$$\text{Accuracy (AC)} = \frac{(TN + TP)}{(TP + FP + TN + FN)} \quad (5)$$

Whereas,

TP is true positive.  
 TN is true negative.  
 FP is false positive.  
 FN is false negative.

### **Precision**

Precision (also called as positive predictive values) is the fraction or ratio between number of error instances for a particular class including faulty and non-faulty.

$$\text{Precision} = \frac{(\text{TP})}{(\text{TP} + \text{FP})} \quad (6)$$

### **Recall**

Recall is the ratio of correct predicted positive observation to all observation in actual class.

$$\text{Recall} = \frac{(\text{TP})}{(\text{TP} + \text{FN})} \quad (7)$$

### **F1-Score**

F1-score is the weighted average of precision and recall. Therefore, it takes both false positive and negative.

$$F1 - \text{Score} = \frac{2 * \text{Recall} * \text{Precision}}{(\text{Recall} + \text{Precision})} \quad (8)$$

## **4 Results and Discussion**

In this section, the experimental results of the experiment we conducted in this study are discussed. Different experiments were conducted to validate the results of our model. We performed experiment on CM1 dataset. This original dataset is preprocessed. To preprocessed data, feature selection and feature extraction methods have been implemented to reduce the input features. Machine learning algorithms were applied to this extracted data for further evaluation. The results of feature selection and feature extraction methods are shown in below tables.



### 4.1 Comparison Results of Feature Selection and Feature Extraction

The above Table 1 summarizes the prediction results of machine learning algorithms using Spearman’s rank correlation technique. SVM classifier has high accuracy rate of 94% when compared to other classification models.

The above Table 2 summarizes the prediction results of machine learning algorithms using PCA technique. Random forest classifier has high accuracy rate of 92.8% when compared to other classification models.

The above Table 3 summarizes the prediction results of machine learning algorithms using LDA technique. Logistic regression and Naïve Bayes classifiers have high accuracy rate of 92% when compared to other classification model.

The above Table 4 summarizes the prediction results of machine learning algorithms using LDA technique. SVM classifiers have high accuracy rate of 93.6% when compared to other classification model.

**Table 1** Performance report of ML techniques based on Spearman’s rank correlation

Algorithms	Accuracy Scores (%)
Naïve Bayes	87
SVM	94
Random forest	93
Logistic regression	92

**Table 2** Performance report of ML techniques based on PCA method

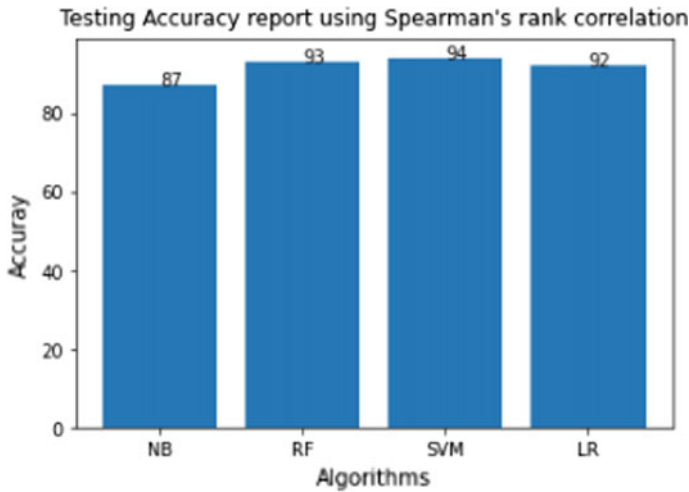
Algorithms	Accuracy Scores (%)
Naïve Bayes	88
SVM	90
Random forest	92.8
Logistic regression	90.6

**Table 3** Performance report of ML techniques based on LDA method

Algorithms	Accuracy Scores (%)
Naïve Bayes	92
SVM	90
Random forest	86.4
Logistic regression	92

**Table 4** Performance report of ML techniques without any methods

Algorithms	Accuracy Score (%)
Naïve Bayes	82.4
SVM	93.6
Random forest	92.8
Logistic regression	90.4



**Fig. 2** Graphical representation of testing accuracy report of Spearman’s rank correlation

#### 4.2 Graphical Representation of Classification Models Using Feature Selection and Feature Extraction Methods

The below figure shows graphical representation of testing accuracy scores of feature selection and feature extraction methods (Figs. 2, 3, 4 and 5).

### 5 Conclusions and Future Work

Software defect/fault prediction plays a vital role in the quality of software. To increase in the quality of software, it should be fault free. So, SFP helps in identifying the software which as faulty or non-faulty instances. To determine the prediction of faults in the software, many machine learning techniques were used such as NB, SVM, LR, and RF. Each of these classification model results were analyzed to show which algorithm gives better accuracy. The results show that using Spearman’s rank correlation and Kendall rank correlation method, SVM gives 94% of the highest accuracy compared to other algorithms, and using LDA method on the dataset, NB

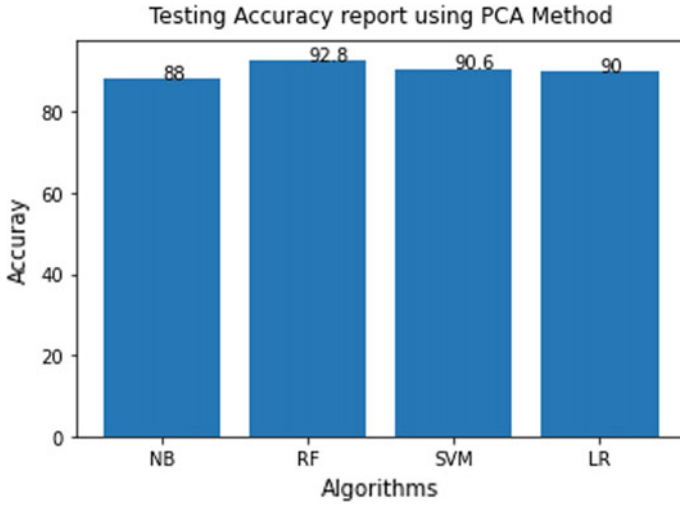


Fig. 3 Graphical representation of testing accuracy report of PCA method

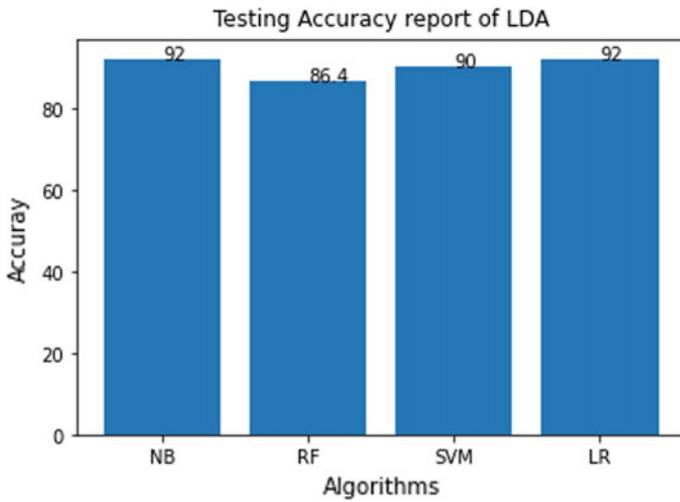
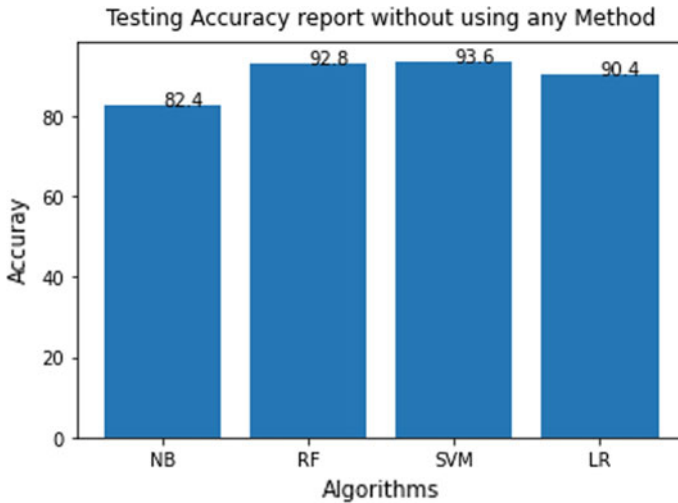


Fig. 4 Graphical representation of testing accuracy report of LDA method

and LR give highest accuracy of 92% over other techniques, and using PCA method, RF gives 92.8% of accuracy, which is highest among four classifiers.

For future works, we will implement more machine learning (ML) and deep learning algorithms (DL) to verify the prediction results. The same model can be implemented using GPU which will enhance the performance of the model.



**Fig. 5** Graphical representation of testing accuracy report without feature extraction and feature selection methods

## References

1. Erturk E, Sezer EA (2014) A comparison of some soft computing methods for software fault prediction. *Expert Systems with Applications*. Int J 42. <https://doi.org/10.1016/j.eswa.2014.10.025>
2. Fangjun Wu (2011) Empirical validation of object-oriented metrics on NASA for fault prediction. *Advances in information technology and education*, vol 2011. Springer, Berlin, pp 168–175
3. Yang C, Hou C, Kao W, Chen I (2012) An empirical study on improving severity prediction of defect reports using feature selection. In: 2012 19th Asia-Pacific software engineering conference (APSEC), vol 1, pp 350–355
4. Pushpalatha MN, Mrunalini M (2019) Predicting the severity of closed source bug reports using ensemble methods. In: Satapathy S, Bhateja V, Das S (eds) *Smart intelligent computing and applications*. smart innovation, systems and technologies, vol 105. Springer, Singapore. [https://doi.org/10.1007/978-981-13-1927-3\\_62](https://doi.org/10.1007/978-981-13-1927-3_62)
5. Kaur J, Sandhu PS (2011) A k-means based approach for prediction of level of severity of faults in software systems. In: *Proceedings of international conference on intelligent computational systems*
6. Chug A, Dhall S (2013) Software defect prediction using supervised learning algorithm and unsupervised learning algorithm, pp 5–01
7. Wahono RS, Suryana N (2013) Combining particle swarm optimization based feature selection and bagging technique for software defect prediction. *Int J Softw Eng Appl* 153–166
8. Tosun A, Bener A (2009) Reducing false alarms in software defect prediction by decision threshold optimization. In: *Third international symposium on empirical software engineering and measurement* ©2009 IEEE
9. Dhankhar S, Rastogi H, Kakkar M (2015) Software fault prediction performance in software engineering. *IEEE*, 04 May 2015
10. Gray D, Bowes D, Davey N, Sun Y, Christianson B (2009) Using the support vector machine as a classification method for software defect prediction with static code metrics. In: *Engineering applications of neural networks*. Springer Berlin Heidelberg, pp 223–234

11. Ren J, Qin K, Ma Y, Luo G (2014) On software defect prediction using machine learning. *J Appl Math*
12. Kim S, Zhang H, Wu R, Gong L (2011) Dealing with noise in defect prediction. In: 2011 33rd International conference on software engineering (ICSE). IEEE, pp 481–490
13. Wang H, Khoshgoftaar TM, Napolitano A (2012) Software measurement data reduction using ensemble techniques. Elsevier, pp 124–132
14. Challagulla VU, Bastani FB, Yen IL, Paul RA (2005) Empirical assessment of machine learning based software defect prediction techniques. In: Proceedings of the 10th IEEE international workshop on object-oriented real-time dependable systems (WORDS'05), pp 1530–1443

# Secret Key Generation: Single Edge-Triggered Flip-Flop PUF for IoT Environment



S. Hemavathy, C. Renju Raju, Akshara Kairali, B. G. Hari Lavanya, and V. S. Kanchana Bhaaskaran

## 1 Introduction

Physical Unclonable Function (PUF) is a physical device that generates a response based on the challenges or the input. The responses or the output of PUF are obtained based on the intrinsic variations that are inherently available in an integrated circuit. The output from every individual integrated circuit can act as a digital fingerprint to reproduce unique identity and non-resilient keys for highly secure crypto applications. In early periods, PUF gained attention in smartcard applications to create unique cryptographic keys for individual smartcards. Even they are used in many FPGAs to secure the secret keys for commercial applications.

The combination of input and output in a PUF is called as Challenge-Response pairs (CRPs). A PUF is a physical device integrated into an integrated chip. The semiconductor industry's continuous design of the digital circuit leads to some diverse changes in the oxide thickness, threshold voltage, and many other parameters. PUF utilizes these inherent manufacturing variations to generate a random response. Designing a PUF with the same Challenge-Response behavior is impracticable as it depends purely on the manufacturing process variations. These factors make the PUF unpredictable and uncontrollable structure.

Mathematically unclonable means it should be tough to compute CRPs for any PUF device with a given CRP. A complex challenge interaction with many or all of the unexpected properties of CMOS determines a response. In other words, PUF proves to be an unclonable device as it is physically and mathematically unclonable. These PUF features can be utilized as an inimitable and untampered device identifier.

---

S. Hemavathy · C. Renju Raju · A. Kairali · B. G. Hari Lavanya · V. S. Kanchana Bhaaskaran (✉)  
Vellore Institute of Technology, Chennai Campus, Chennai, India  
e-mail: [kanchana.vs@vit.ac.in](mailto:kanchana.vs@vit.ac.in)

## 2 Literature Survey

D flip-flop PUF architecture with symmetric cross-coupled inverters is proposed with two additional pass transistors to make the architecture symmetrical and give a high value of uniqueness [1]. The uniqueness is enhanced in the conventional D flip-flop by adding tristate logic instead of inverters [2]. The PUF metrics of flip-flop-based arbiter PUF is improved with a novel design of Feedback Oriented XORed Flip-Flop Arbiter PUF (FOXFFAPUF) [3]. The author of Coin-Flipping Physically Unclonable Function (CF-PUF) [4] has proposed a robust PUF design against the machine-learning attack. It uses the threshold voltage changes due to the nonlinearity of the convergence time of bistable rings. The author exemplifies the security of the PUF, along with less area overhead [5].

A new scheme [6] to protect the secret key against scan-based side-channel attacks is proposed to improve the uniqueness of PUF from every chip by using a lock and key scheme. Many flip-flop (FF) designs have been compared and evaluated for suitability as PUF generators by comparing their reliability, uniqueness, uniformity, and bit aliasing characteristics for PUF applications [7]. The Arbiter Physically Unclonable Function (APUF) is presented to give a low-cost and unique security solution over the expensive, standard cryptography system [8].

## 3 Design Methodology

### 3.1 Physical Unclonable Function

The PUFs can be broadly classified as Silicon and Non-Silicon based on the fabrication type as shown in Fig. 1. The various silicon PUFs can be broadly categorized based on logic as digital, analog, and adiabatic logic-based PUF. Adiabatic PUF [9] and quasi-adiabatic Tristate PUF [10, 11] are based on adiabatic logic, which can be used for low-energy applications.

The delay-based PUF generates a response based on the comparison of the delay between two paths (Arbiter PUF) or based on frequency as in Ring oscillator PUF. The memory-based PUF utilizes the cross-coupled latches to obtain the responses. Some of them are SRAM PUF, Latch PUF, and D flip-flop-based PUF.

Clock pulses are used in almost every digital circuit PUF to control the flow of responses utilizing the Finite state machine. Since the D flip-flops are spread in a wide range in an FPGA board, it is much easier to generate more CRPS. This feature also increases the robustness of the D flip-flop against side-channel attacks. The D flip-flop was earlier introduced as intrinsic PUF on the reconfigurable device to generate digital fingerprints. During power-up, the metastability in the D flip-flop generates randomness. This feature is exploited in PUF to generate a secret key. Similar to SRAM, upon power-up, all the D flip-flops are initialized to the specified

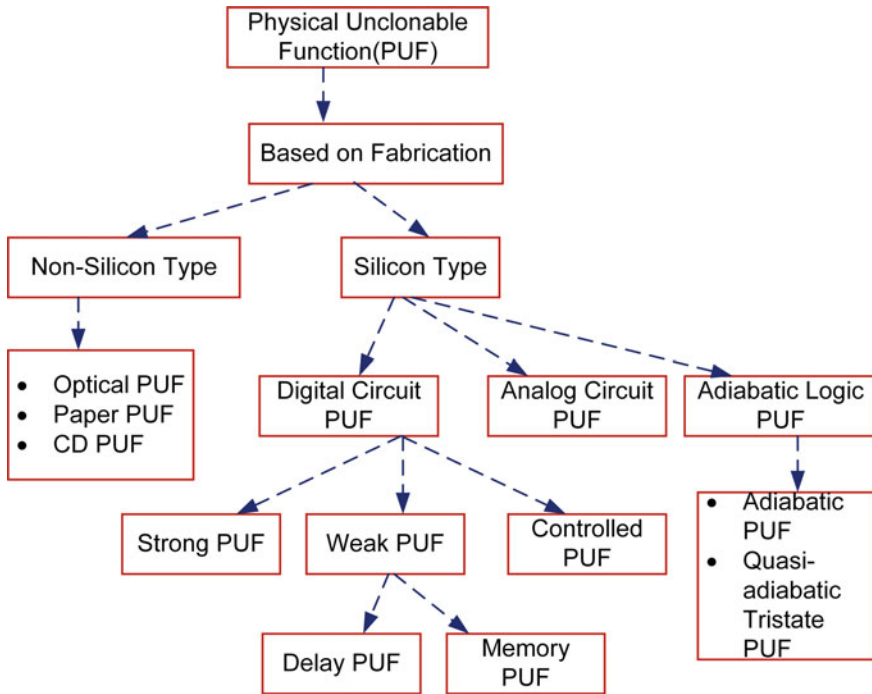


Fig. 1 Classification of PUFs [11]

initial value or ‘0’ if the user does not specify an initial value. In FPGA, the author has used the global restore line command to remove the initialization from the bit file.

### 3.2 PUF Metrics

The essential security metrics for the various forms of PUFs used to characterize their resilience are Uniqueness, Uniformity, and Reliability.

**Uniqueness** is the most important key feature of PUF. It elucidates distinctly the number of Challenge-response pairs (CRPs) obtained from any integrated chip. It characterizes how distinct is the CRP of a PUF instance from those of other PUF devices. The fact is, it should be nearly hard to clone two PUFs. Each chip should have a different input-output mapping. Environmental conditions like the voltage, temperature, and aging can affect the CRPs obtained from each PUF design. Hence, error-correcting codes are added to replicate the responses. Inter-chip hamming distance ( $HD_{inter}$ ) is employed to measure uniqueness. For the same set of challenges under nominal voltage and temperature,  $HD_{inter}$  is calculated for different PUF instances. With  $R_i(n)$  and  $R_j(n)$  be the  $n$ -bit responses obtained from chips  $i$  and  $j$  ( $i \neq j$ ) with



the same set of challenges, the uniqueness is evaluated by

$$HD_{INTER} = \frac{2}{k(k-1)} \sum_{i=1}^{k-1} \sum_{j=i+1}^k \frac{HD(R_i(n), R_j(n))}{n} \tag{1}$$

**Uniformity** is another metric of PUF where the percentage of 1's and 0's is calculated. Even distribution of bit 1 and 0 ensures a vital secret key that is hard to duplicate. The ideal value is 50%.

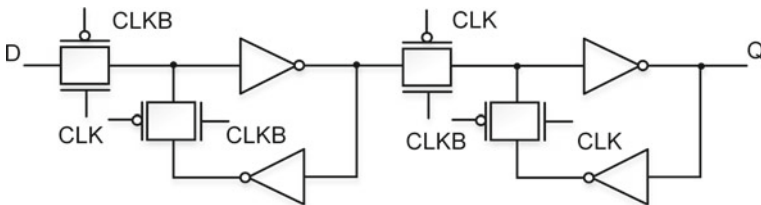
**Reliability** is a metric with which we evaluate a PUF circuit. It is an important attribute determining how repeatable or dependable a PUF's CRPs are under various environmental situations. Even at different voltages and temperatures, the reliability of the PUF needs to be ideally equal to 100%. Intra-chip hamming distance is used to evaluate Bit-Error Rate (BER). It is as given by

$$HD_{INTRA} = \frac{1}{k} \sum_{i=1}^k \frac{HD(R_i(n), R'_j(n))}{n} * 100\% \tag{2}$$

Here,  $k$  denotes the number of samples with the same set of challenges.  $R_i$  is the reference response obtained at normal working conditions, and  $R'_j$  is the response obtained by changing temperature and supply voltage.

### 3.3 Single Edge-Triggered Flip-Flop PUF (SETFF PUF)

Single edge-triggered flip-flops (Fig. 2) are the most commonly used flip-flops used in the semiconductor industry. Only one clock edge, either rising or falling, can be used to load data in this flip-flop. To maintain a desired logic 1 or a logic 0, all sequential components in a design must meet specific minimum data arrival timing requirements. For the data to latch at the exact clock edge (setup time), the data must arrive before the active edge of the clock (and be stable), and the data must also remain stable for a minimum specified duration after the active edge of the clock (hold time). Any breach of these time criteria may result in the latching of inaccurate results. This



**Fig. 2** Single edge-triggered flip-flop PUF (SETFF PUF)

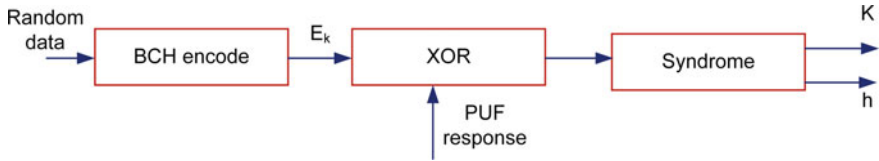


Fig. 3 Generation phase

Fig. 4 Reproduction phase



metastability condition is used to generate the random numbers. During the power-up state of the SETFF PUF, the cross-coupled inverters provide randomness in the output response of the circuit. This feature is utilized for the generation of secret keys from PUF.

### 3.4 Secret Key Storage and Generation

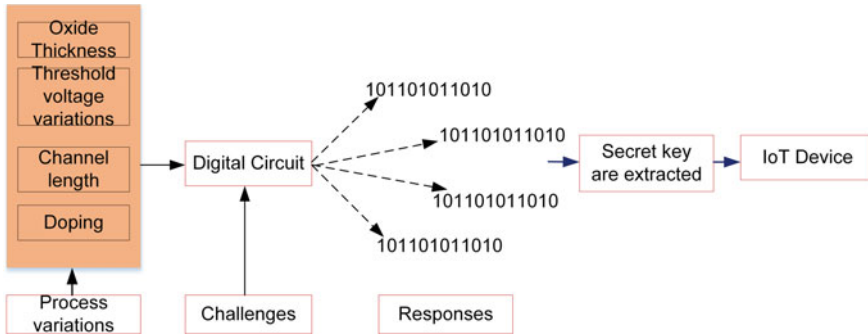
SETFF PUF can be used for key storage applications. The responses generated from the SETFF PUF cannot be used directly for any applications. A post-processing stage is usually required to verify that the device is working efficiently once it is deployed in the field. There are several methods for generating secure bit responses that are based on standard fault tolerance algorithms. In this paper, the noisy responses from the SETFF PUF can be recovered as a cryptographic key using error-correction codes. The BCH codes are generally used compared to other error-correcting code methods since they occupy less area overhead and are suitable for RFID and IoT applications.

**Generation phase** the noisy responses from the SETFF PUF is XORed along with the encoded key ( $E_k$ ) obtained from the BCH encoder. The BCH encoder generates the encoded key from the random data. The secret key ( $K$ ) is obtained from the syndrome along with helper data ( $h$ ) as depicted in Fig. 3.

**Reproduction phase** In the reproduction phase, the noisy response along with helper data ( $h$ ) is given to a BCH decoder to obtain the secret key as shown in Fig. 4.

## 4 Simulation Results

The SETFF PUF is simulated using the Cadence tool. UMC 90 nm technology is used to design the PUF to generate a 128-bit response. 200 such instances are designed



**Fig. 5** SETFF PUF for IoT application

**Table 1** Uniformity and Uniqueness of SETFF PUF

Voltage (V)	Uniqueness (%)	Uniformity (%)
0.9 V	49.76	57.50
0.92 V	49.81	57
0.94	49.87	55
0.96	49.95	55
0.98	50.01	53.50
1	50.07	52.50

using Monte-Carlo. The responses are post-processed using MATLAB to obtain the PUF metrics. The analysis is carried out at a nominal voltage of 1 V and 27 °C. Table 1 shows the uniformity and uniqueness obtained at different voltage conditions. It is evident that they are very near to the ideal value (50%) and almost close to nominal voltage (1 V). At 0.9 V and 0.92 V, the uniformity value has slightly deviated from the nominal voltage. The reason can be the noise incurred in the design due to the bias from nominal voltage. The BCH error-correcting codes support reproducing the same responses.

Similarly, the uniformity and uniqueness were calculated at nominal voltage and different temperatures, as shown in Fig. 6. Figure 7 shows the reliability obtained for different voltages and temperatures. The voltage and the temperature were varied from 0.9 to 1 V and -40–100 °C. The average reliability was found to be 98.35% (ideal 100%). The worst-case reliability is at -40 °C with 96.48%. The worst-case BER is to the maximum of 3.52 which can be corrected by using the BCH error-correcting codes. Table 2 illustrates the reliability for different voltages and temperatures.

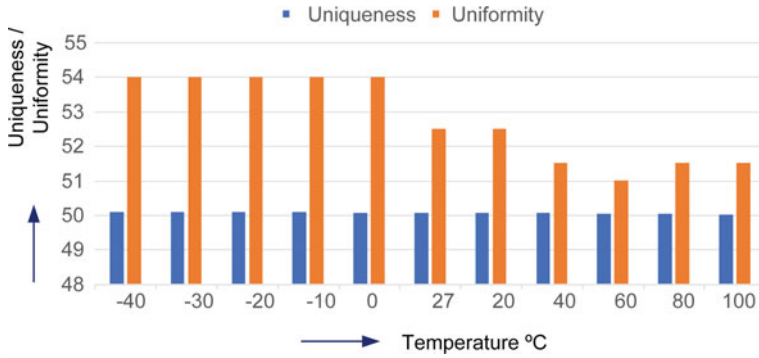


Fig. 6 Uniqueness and Uniformity of SETFF PUF at nominal voltage

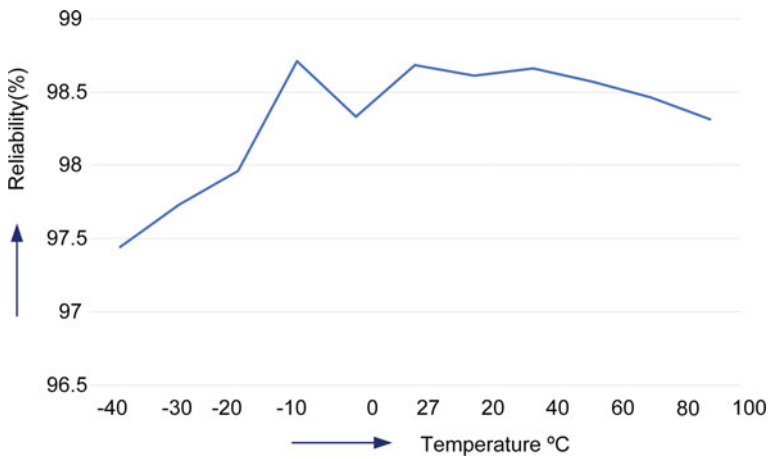


Fig. 7 Reliability of SETFF PUF

Table 2 Reliability of SETFF PUF

Temp °C / Volt (V)	0.9	0.92	0.94	0.96	0.98	1
-40	96.48	96.82	97.28	97.62	98.04	98.38
-30	96.71	97.07	97.54	97.93	98.38	98.74
-20	96.98	97.28	97.76	98.15	98.66	98.98
10	97.17	97.51	97.91	98.37	98.85	99.18
0	97.29	97.64	98.04	98.52	99.07	99.43
27	97.55	97.93	98.35	98.86	99.36	100
20	97.52	97.89	98.27	98.81	99.33	99.83
40	97.66	98	98.38	98.86	99.37	99.68
60	97.73	98.02	98.43	98.81	99.22	99.22
85	97.78	98.09	98.43	98.81	98.93	98.71
100	97.79	98.02	98.42	98.42	98.70	98.50

## 5 Conclusion

The SETFF PUF circuit was implemented, and the metrics like uniqueness, reliability, and uniformity were evaluated for different voltages and temperatures. From the above discussions, it is clear that the uniqueness for the SETFF PUF is about 50.07%, reliability is about 98.35%, and uniformity is about 52.50%. As the PUF metrics are close to the ideal value, the PUF responses of the PUF can be used to generate a secret key by using BCH error-correcting codes. SETFF PUF is best suited for resource-constrained IoT eco-system due to limited number of devices used and robust PUF metrics.

## References

1. Khan S, Shah AP, Chouhan SS, Gupta N, Pandey JG, Vishvakarma SK (2020) A symmetric D flip-flop based PUF with improved uniqueness. *Microelectron Reliab*, Elsevier
2. Khan S, Shah AP, Chouhan SS, Rani S, Gupta N, Pandey JG, Vishvakarma SK (2020) Utilizing manufacturing variations to design a tri-state flip-flop PUF for IoT security applications. *Analog Integr Circ Sig Process*, Springer Science
3. Sushma R, Murty NS (2018) Feedback oriented XORed flip-flop based Arbiter PUF. In: 2018 Third international conference on electrical, electronics, communication, computer technologies and optimization techniques (ICEECCOT), Dec 2018
4. Tanaka Y, Bian S, Hiromoto M, Sato T (1995) Coin flipping PUF: a novel PUF with improved resistance against machine learning attacks. In: *IEEE transactions on circuits and systems II: express briefs*, vol 65, issue no 5, pp 602–606, Mar 2018. van Leeuwen J (ed) *Computer science today. Recent trends and developments. Lecture Notes in Computer Science*, vol 1000. Springer-Verlag, Berlin Heidelberg New York
5. Wang D, Liu L, Wang B, Wei S (2018) Area-efficient delay-based PUF based on logic gates. In: 2018 10th International conference on communication software and networks, Oct 2018
6. Wang Q, Cui A, Qu G, Li H (2020) A new secure scan design with PUF-based key for authentication. In: 2020 IEEE 38th VLSI test symposium (VTS), June 2020
7. Zhang H, Jiang H, Eaker MR, Lezon KJ, Narasimham B, Mahatme NN, Massengill LW, Bhua BL (2018) Evaluation on flip-flop physical unclonable functions in a 14/16-nm Bulk FinFET technology. In: 2018 IEEE international reliability physics symposium (IRPS), May 2018
8. Mahalat MH, Mandal S, Mondal A, Sen B (2020) An efficient implementation of arbiter PUF on FPGA for IoT application. In: 2019 32nd IEEE international system-on-chip conference (SOCC), May 2020
9. Dinesh Kumar S, Thapliyal H (2020) Design of adiabatic logic-based energy-efficient and reliable PUF for IoT devices. *J Emerg Technol Comput Syst* 16(3), Article 34:18. <https://doi.org/10.1145/3390771>
10. Hemavathy S, Kanchana Bhaaskaran VS (2020) Design and analysis of secure quasi-adiabatic Tristate physical unclonable function. In: 2020 IEEE International symposium on smart electronic systems (iSES) (Formerly iNiS), pp 109–114. <https://doi.org/10.1109/iSES50453.2020.00034>
11. Hemavathy S, Bhaaskaran VK (2020) Design and analysis of secure quasi-adiabatic tristate physical unclonable function. In: 2020 IEEE Consumer electronics magazine. <https://doi.org/10.1109/MCE.2021.3117541>

# A Concise Survey on Solving Feature Selection Problems with Metaheuristic Algorithms



Rama Krishna Eluri and Nagaraju Devarakonda

## 1 Introduction

Machine learning made a better path for the management of complex tasks. To implement such complex tasks, we need to understand the type of features we need to consider. The concepts of feature extraction and selection will make a better implementation based on the concept and the problem statement. The best way to understand the feature selection process is to make the variables more reliable. The most important factors will give the most valuable accuracy. In most of the cases because of the higher dimension data, there is a chance to surpass the implementation of data mining concepts. To solve such cases, we need to consider feature selection. This pre-processing phase will be most helpful to monitor the higher dimension data and make them most useful for the prediction model implementation. There are several cases we have less accuracy with respect to the outliers. The outliers will damage the patterns in the variables and make the weak models. Because of this higher dimensional implementation, we need to focus on pre-processing which can help to make the models with at most accurately.

The implementation makes sense that there is a lot of difference in making the models more important. We have different types of optimizers which are most useful for implementation of reducing the higher dimensional data. The data is modified into smaller parts using the standard scaler or vectorization. After changing the data from different higher dimensions to the lower dimensions, then we can get to implement the machine learning models in an accurate manner. There are various methods developed to implement the dimensionality reduction. In those most important factors, we got a chain of implementations with metaheuristics. These metaheuristics will help the

---

R. K. Eluri (✉) · N. Devarakonda

School of Computer Science and Engineering, VIT-AP University, Amaravati, Andhra Pradesh, India

e-mail: [ramakrishna.20phd7014@vitap.ac.in](mailto:ramakrishna.20phd7014@vitap.ac.in)

developer to make the data more reliable and most prominent for the prediction model implementation. This article will work on the extensive review on the optimization mechanism to implement feature selection in a better way, and these feature selection procedures will help the researchers to understand more about the importance of data, and if there is any possibility that to remove the non-required fields which are causing effect to the modeling, we can get result of the importance of features in the modeling, and based on that result, we can remove a specific feature from the modeling if it is not required. The modeling concepts will be working with optimization techniques. Metaheuristics are playing a vital role in making the results more accurate. We are going to check the last few years' information on optimization and metaheuristics which are most important things in modeling. The metaheuristic models need to be divided into several parts based on their non-functional behavior. There are hundreds of metaheuristic algorithms, and to solve the binary class problem, we are taking the algorithms list of the same class type. In real-time scenario, the data will be large in flow and it will be complex task to handle such kind of bigger flow time to time. A specific dataset may contain larger values of columns (features) and the entries (rows). There is no rule to use all the features which are mentioned in the dataset for the prediction model implementation. There must be a scenario which will declare the features which are the most important. Some of the features must be avoided as they are not relevant to the problem statement, and they will degrade the accuracy of the modeling. The main aim of the feature reduction and selection procedure is to handle the large dataset with larger values and features and managing them to make a better accuracy model based on the problem statement. This problem-solving approach consists of three different patterns: feature reduction, selection, and constructing the features. The process of reduction and construction is to make a subset from the main set and make them to use for the trial and error method type of modeling, while the feature selection will be working on the selection of most required and prominent features. Related to machine learning, this is the most important and challenging task. From the set of  $n$  number of main sets features, we may have minimum of  $2^n$  subset features which are most important and useful. It will be tough task to evaluate each and every subset. The concept of subset is to bring a different combination of subsets to evaluate each subset in manual procedure. But if the data is larger than expected, the concept of subset will make you a clear false step to evaluate. Because, the subset count will be more and there must be a unique procedure to manage all the combination of the features. We have different search methods which are proposed from different methodologies and implementations. Greedy method is one of the search methods used for the feature selection and implementation. Exhaustive search and random search are also one of the best combinations to determine the feature selection procedure. The most important factor which is bothering the procedure of implementing the feature selection is cost. There will be more of computational cost involved in the feature selection if the dataset is too large to handle with the small scale. These metaheuristic algorithms are helping us to implement such kind of problem solving and kind of methodologies which can be most useful in near future. They are used for finding the best subset from the groups of sets. Since last few decades, the work on metaheuristics is being done to implement the concept

of optimization and feature selection. We need to process all the things related to the large data and make it most useful for the further research. These metaheuristic concepts will try to understand the importance of data in image classification, data mining, text mining, etc. and give us a useful pattern to recognize the best feature set.

There are several research works found on implementation of metaheuristics and the optimization mechanisms and most of the cases they work on the concept of genetic algorithms and different types of optimization methods like colony optimization and swarm analysis [1]. We also have the concept and research on feature selection in digital type of data that is multimedia. The multimedia information and data regarding the multimedia building are also the most important factors in recent times [2]. Even we can consider feature selection as one of the important factors in medical domain. The medical domain is having remarkable changes in data handling things. There are tons of data available for every second, and this kind of large flow of data must be used for the better research implementation [3]. Sharma and Kaur [4] explain about the algorithms which are inspired from the nature. Those are nature-based algorithms which are most successful in understanding the different patterns in the dataset and make them to use for the better cause. They can separate the binary information from other type of information in the dataset. In the case of prediction model implementation, we should not use the object type data for the modeling. They need to be converted into the form of numeric, then only the feature will be eligible for the modeling. The concept of optimization will work on better understanding on the data and make it available in all the cases [5] that used the metaheuristic algorithms for bioinformatics applications and DNA sequencing [6] for ensemble modeling. The concept of ensemble modeling is quite interesting because of implementing the model in recursive manner until it achieves the best accuracy in the modeling. We are trying to implement the better combination of the features. The weak features will be sent as the input to the next model. There is no pattern of using same type of model for all over the modeling. We can use different methods to design the final modeling. The concept of implementation of the modeling is the most important factor. The factor which is most important in the model is most fit and best fit among all the other features. These are the cases which can implement feature selection procedure for real-time industry-oriented concepts like Industry 4.0.

- The main strategies and purpose we are going to discuss in this article are mentioned below in the sequence order:
- This article is to give clear idea on problem-solving approaches in metaheuristic models and how they are impacting the other procedures in modeling.
- List of metaheuristic algorithms was given and discussed.
- How binary metaheuristic algorithm is being used to define the concept of feature selection.
- Vital factors for implementing metaheuristics.
- Challenges of developing metaheuristic models as well as the models which are not being used these metaheuristic models.
- Research gaps identified in the procedure of literature survey.



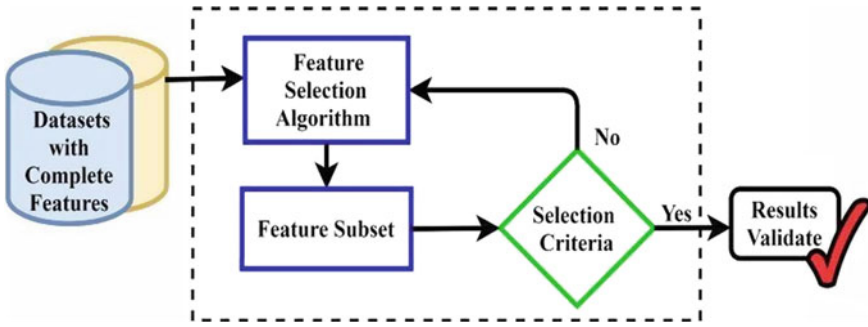


Fig. 1 Overall feature selection procedure [7]

In Fig. 1, we are going to discuss about the overall feature section architecture. In this architecture, we will be having set of features to be considered. First, we need to choose a dataset with complete features without deleting anything. We need to implement the feature selection algorithm using which are trying to identify the best features among those. After implementing the procedure, we need to take the subset of the features which are processed in the feature selection procedure, then we need to make it to use for the selection criteria. The selection criteria will check whether the subset is accurate and useful or not. If not, we need to move for the feature selection procedure implementation again. The procedure is recursive in this situation. If the result of subset is fine to be used, then we can directly conduct the modeling with those subsets. The results of the subsets in the form of accuracy will define the importance of these features. The implementation is the most important factor. We need to validate the subsets based in the accuracy we achieved with the dataset and the modeling procedure which we considered.

Figure 2 will try to make us to understand how the research and publication are going on and went well in last 10 years on metaheuristic models. This Fig. 2 will

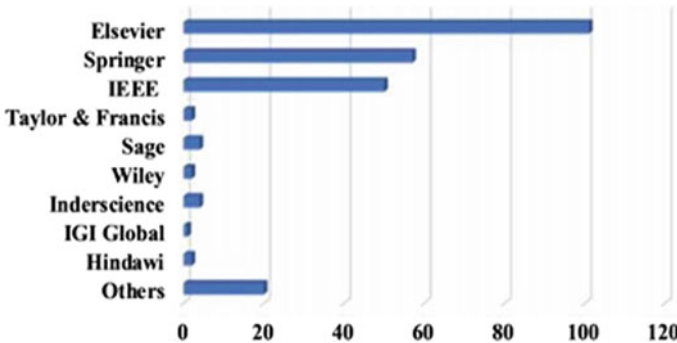


Fig. 2 Research articles published in different journals [7]

define the sample graphs of research work held on the same concept of metaheuristics, and we can follow the same kind of implementation procedure for the metaheuristic.

The next section will give a brief introduction to every possible aspect of the metaheuristics. We will be explaining about the methodologies available in building metaheuristics. The next part will be about literature survey on metaheuristics. In Sect. 4, we are doing to share data about challenges we face, and in next section, we are going to discuss about different case studies impacting the concept of metaheuristics and feature selection procedure. The final section will help to plot all the future scope related to this work.

## **2 Methodology and Background**

In this section, we are going to discuss about the procedure of implementing the feature selection, and the implementation will be helpful to implement different mathematical models. This section will help to make a note of the mathematical structures of the metaheuristic models and implementation.

### ***2.1 Feature Selection***

The procedure of feature selection will be included when there is a situation to handle the irrelevant, not organized, and not proper information in the dataset [8]. Feature selection is one of most important factors in machine learning to achieve better accuracy in real-time problems. There are various problems that can be solved with feature selection procedure, and all these factors play a vital role for the better accuracy of the modeling. There are different applications in medical imaging and processing [9]. The medical implementation in the concept of feature selection is a challenging task for the researchers, and not all the time, we will be getting the chance of implementing the better accuracy. There are other chances that can be considered for the improved accuracy result with the feature selection in their respective fields [10]. Text mining was discussed to make the users to understand the type of text, and insights of text are there. The situation may get into existence when there is a need to understand what the user is typing or trying to ask. Consider the example go search engine and how it is trying to give the result according to the human perspective. The result will be based on the search keywords. But, the procedure of gathering the results and making them to useful for the modeling is the most important and challenging task. The concept of text mining and sentimental analysis is the most important and useful [11]. Also, it describes about the work related to text mining. The sentiment analysis in natural language processing will help the users to make them to implement insights of the users search criteria. The criteria is to follow the tokenization. The tokenization will take the values in the form of keywords. The

keywords will try to help the researchers to analyze the requirement of the user when they search for a specific portion in the query.

For an instance, let us consider an example where we need to implement the feature selection concept. The feature selection is the most important to grab the features from the universal set of features. Let “S1” be the universal set of features; “F” represents the features, and “n” represents the feature number. The feature selection procedure for this kind of problems will be mentioned like below:

$$S1 = \{f(1), f(2), f(3) f(n)\}$$

- n* Number of features
- f* Features
- S1* Universal set

The objective is to select the best among the features which are being selected. The features need to be working with each other to make a best model. Then, we need to make a subset of features, and it denotes with *N* the subset of the sets which will be like mentioned below:

$$N = \{f(1), f(2), f(3) f(n)\}$$

Figure 3 will explain the working procedure of feature selection procedure. There are different components to be discussed when we are focusing on the feature selection architecture. The first thing is feature selection, and it is further divided into three different parts. The first one is filter methods. Using this filter method, we are trying to identify the most requested keyword from the search criteria. Then, we need to make the wrapper methods which hold all the information. And, the third

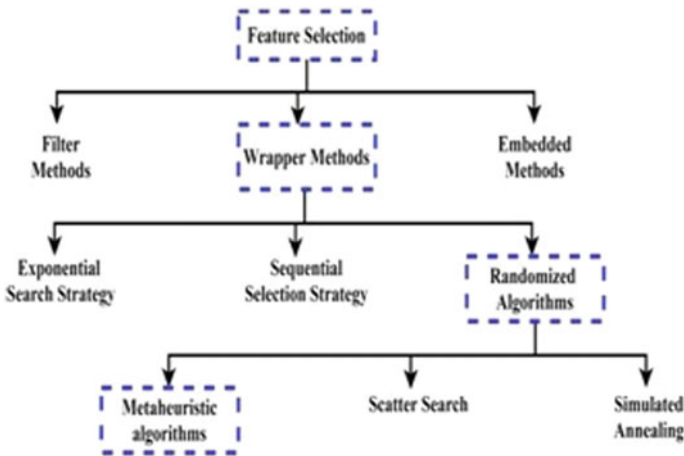


Fig. 3 Feature selection procedure [7]

method is embedded method in which is trying to make it possible for the more accuracy. From the wrapper method, we further enhance with exponential search strategy, sequential search strategy, and randomized algorithms. In this randomized algorithm, we are trying to implement the metaheuristic models, scatter search, and simulated annealing [12–20]. We cannot include all the detailed description of the concepts from Fig. 3. The detailed description of the concepts is available in [1].

## 2.2 *Metaheuristic Algorithms*

To avoid local optima, we can implement metaheuristic algorithms. These algorithms will be used for the implementation of optimization which can avoid the outliers in the data and make the dataset into most useful in the pre-processing stage. The pre-processing will help the researchers to make the dataset clean for subletting. The subletting will help in the work related to best accuracy model gathering and implementation. These are non-mathematical parts of logics which are accuracy, flexible, and simple [21]. There are different search methods like gradient decent models and stochastic models. But in the feature selection process using metaheuristic, we need not to take care of the gradient search methods, and all the operations will be done by the metaheuristic models. They are the concepts of reducing the error in the dataset and make that dataset most useful for the modeling. These are flexible to the mathematical model implementation. The model implementation will try to define the problem be like with single solution or multiple solutions. There is a chance of implementing metaheuristics with the concept of premature convergence. Then, we can make a better model with the data available. The data which is available for modeling must satisfy different rules for different results. Due to the stochastic knowledge and behavior, these models will remove all the barriers to build efficient models with local optima. The minimal values of the optima are better to make a better model, and all these models are based on the type of result we are acquiring. The results which are most useful for the modeling are the things which are based on the subset and the result of subset in the form of solution.

The metaheuristic features are further classified into two different types. First thing is based on single solution; other category is population solution that is multiple solutions. Let us discuss about those factors now.

### 2.2.1 **Single Solution**

This will be working on single solution occurred when the model was designed. Then, there will be a chance of not having regular check of the result in this procedure. The procedure includes checking the local optima and when there is a result will be used as the next input and iteration will continue. But, there will no regular checking of the improvement of the model result.

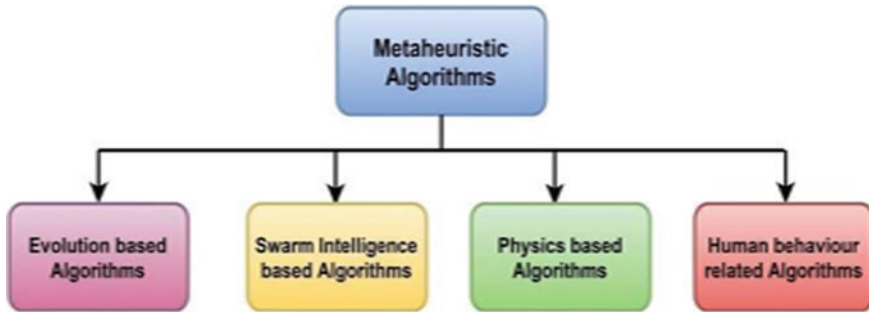


Fig. 4 Classification of metaheuristics [7]

### 2.2.2 Multiple Solutions

The multiple solution patterns are based on gathering different solutions from different populations. The different populations will help the user to make a better classification model. The prediction model can contain different results, and those results will be with respect to the iterations we are following while building the prediction model. Initially, it starts with the general population, and later, it turns into number of iterations and generations. For each and every concept, there are different iterations and generations, and all these can be solved using metaheuristics by understanding the result of the each and every subset of the result. Figure 4 will be explaining the categorization. In this categorization, we can feel so simple to implement the metaheuristics, and procedure will be simple and implementable. There are four different subparts in metaheuristic algorithms, and each and every algorithm is having different requirements. There are four different parts of optimizations in metaheuristics which are facing a most critical implementations and requirements. The results will be the most important and required.

Let us discuss about the things related to different types of metaheuristics in general. The implementation lies within the subset. The implementation makes a better chance for optimization. The optimization is required for the better implementation of the prediction models. Let us discuss about all the factors of the metaheuristics in detail now. The characteristics will help to understand the requirement in data handling in an effective manner.

## 2.3 Evolution Based

This will be a naturally generated solution or the subset. The randomly considered subset will try to analyze the insights, and we can get the result of the optimization based on the number of features being used. The evolution will be based on the total number of features to be considered in general from the set of large data, and the subset must be drawn randomly without any further implementation. The further

implementation will be based on the first set of solutions and the accuracy of those models. The accuracy will help to analyze whether the set we considered is correct or not, and if there is any chance of changing the model and the subsets, we can perform in the form of model selection and loss function calculation. In the loss function calculation, we can estimate the total loss we occurred in the process of generating the accurate result from the samples which we considered. The natural evolution will make the model more reliable, and most important thing is the randomly drawn subsets will not be much helpful in the future because of not having proper relation between them. This concept is mostly used in genetic algorithms and in which we need to discuss about [22] Darwin technique, [23] other important evaluation strategies, [24] generic programming, and [25] different search mechanisms.

### **Swarm-Based Modeling**

This swarm-based modeling is the concept of building the model with the most social behavior. The social behavior is different from the general behavior. There will be an intelligence agent who can understand the requirements of the users and trying to plot the difference between the general understanding and relations and the social responsibility and understanding. The situation will change time to time with respect to the user, and the users will try to judge the procedure based on behavior. There are different popular techniques which are being used over the time. These kinds of features can be useful for better implementation. Then, [26] discussed about the PSO, particle swarm optimization, [27] discusses about ant colony optimizations, [28] discusses about monkey optimization, [29] discusses honey bee optimization. All these strategies will help the researchers to work more on the real-time data. The real-time data can be most useful for the better implementation of the optimization mechanisms so that we can get a better accuracy from the data.

### **Physics Based**

In this kind of approach, we are checking the rules of physics. For an instance, we have a challenge of gathering information and plot it to a better prediction model using the concepts of physical entities, then we need to predict in this kind optimization mechanism. In this mechanism, we can discuss further about simulated annealing [30] and search criteria like harmony [31].

### **Human Behavior Algorithm**

In this human behavior, it is like an optimization mechanism where we can store the information based on the human genetic approaches. The DNS research can be a best example for the human behavior. We have lot more chances of understanding the human behavior with other general instances. We need to focus on the things like teaching learning algorithm also called as TLBO [32], and also, we are going to discuss about the other technique called league championship algorithm.

### 3 Metaheuristic Methods on FS

There are different concepts which are related to the optimization mechanisms. These optimization mechanisms will decrease the loss function, manage the data with the high-end quality, and maintain the better accuracy all over the prediction model implementation. In this category, we need to discuss about different other features we have in this metaheuristic methods. In this method, we need to look over the data in the form of the binary format. The binary format is easy way to classify any problem based on the requirements. As we discussed in previous section about four different types in metaheuristics, we are going to look over the methods in the table format. These tables will explain different algorithms which are most used in last 10 years. These algorithms achieved the highest accuracy in defining the category of the modeling. They are also defined in separate tables with respect to physics, human intelligence, and human behavior. Tables 1, 2, and 3 will define the articles and review on algorithms they are using [33–49].

#### 3.1 Evolution Algorithm

This kind of algorithm will try to understand the basic level of the problem statement and try to focus on the things like what are the different parameters we can consider to get an initial stage of the factors of subsets. The subsets which we are forming in this kind of algorithms will be having no proper ground truth. The ground truth of this kind of algorithms will be based on the output they achieved after the implementation. The implementation will help to make the algorithm more accurate without any latency and loss of logic. The specificity and sensitivity will be calculated and all the models with some cases of prediction accuracy. The accuracy will be defined in the rate of total number of positive instances it considered as the subset.

#### 3.2 Swarm Analysis

Swarm intelligent-based analysis is consisting of different sub-approaches where we can find all the aspects of the identifying the optimal result from the given set. The result of swarm analysis will be used on medical domain to gather information related to disease insights, and all the structural data formats can be used for the analysis of the swarm analysis. Some of the approaches will be mentioned below. All these approaches are the search criteria. The search criteria in intelligent analysis will be there, and we need to discuss about each and every important search criteria which are having most impact factor.

**Table 1** Swarm intelligence-based algorithms

Abbreviation	Algorithm	Year
BBA	Bumblebees algorithm	2009
PFA	Paddy field algorithm	2009
CS	Cuckoo search	2009
GSO	Group search optimizer	2009
CGS	Consultant guided search	2010
BA	Bat algorithm	2010
TCO	Termite colony optimization	2010
HS	Hunting search	2010
ES	Eagle strategy	2010
HSO	Hierarchical swarm optimization	2010
FA	Firefly algorithm	2010
FOA	Fruit fly optimization algorithm	2011
ECO	ECO-inspired evolutionary algorithm	2011
WSA	Weightless swarm algorithm	2011
FPA	Flower pollination algorithm	2012
BMO	Bird mating optimizer	2012
ACS	Artificial cooperative search algorithm	2012
KH	Krill herd algorithm	2012
FROGSIM	Japanese tree frogs calling algorithm	2012
opt Bees	The opt bees algorithm	2012
WSA	Wolf search algorithm	2012
TGSR	The great salmon run algorithm	2012
DE	Dolphin echolocation	2013
SSO	Swallow swarm optimization algorithm	2013
EVOA	Egyptian vulture optimization algorithm	2013
CSO	Chicken swarm optimization	2014
AMO	Animal migration optimization	2014
GWO	Gray wolf optimization	2014
SSO	Shark smell optimization	2014
ALO	Ant lion optimizer	2015
BSA	Bird swarm algorithm	2015
VCS	Virus colony search	2015
AAA	Artificial algae algorithm	2015
DA	Dragonfly algorithm	2015
DSOA	Dolphin swarm optimization algorithm	2016
CSA	Crow search algorithm	2016

(continued)



**Table 1** (continued)

Abbreviation	Algorithm	Year
WOA	Whale optimization algorithm	2016
MBF	Mouth brooding fish algorithm	2017
ABO	Artificial butterfly optimization	2017
SHO	Selfish herd optimizer	2017
GOA	Grasshopper optimization algorithm	2017
SSA	Salp swarm algorithm	2017
SHO	Spotted hyena optimizer	2017
EPO	Emperor penguin optimizer	2018
SSA	Squirrel search algorithm	2018
BOA	Butterfly optimization algorithm	2019
EPC	Emperor penguin colony	2019
HHO	Harris hawks optimizer	2019
GEO	Golden eagle optimizer	2021

### 1. Cuckoo Search

When the researchers tried to extract the bird emotions using machine learning and implementation of the sentimental analysis on the birds, then the project cuckoo was started. In this, they successfully extracted and recorded the voice of cuckoo bird. In this research, they implemented the binary class classification using the voice, and this can be a sort of unsupervised learning, and based on the data available, they started working on extracting the cosine waves from the voice of bird. Then, they implemented BCS to make the model more accurate and valuable.

### 2. BAT Model with SVM

This model will help to predict the information from the bat sounds and the actions. The actions define the V-shaped curve from the data and that can be implemented using SVM for classifying the best approach we can give for the bat prediction. It is a kind of real-world problem-solving approach. In this approach, we are trying to figure out the scenarios on which the data can be used to predict the position of bats. The binary classification implementation was done on this concept.

### 3. Flower Pollination

We need to develop a machine learning algorithm which can help to predict the pollination of the flowers based on the true events. We need to predict the result of the binary classification and try to identify the best flowers to implement pollination. The pollination procedure cannot be implemented on all the flowers, and we need to predict the scenarios where the flowers will be the inputs, and based on the binary variables of the flowers, we design a sigmoid function which can help us to predict the best flower which can subject to the pollination.

**Table 2** Physical-based algorithms

Abbreviation	Algorithm	Year
GSA	Gravitational search algorithm	2009
CSS	Charged system search	2010
GbSA	Galaxy-based search algorithm	2011
EMO	Electromagnetism optimization	2011
ACROA	Artificial chemical reaction optimization algorithm	2011
Spiral	Spiral optimization	2011
BH	Black hole algorithm	2012
WCA	Water cycle algorithm	2012
CSO	Curved space optimization	2012
RO	Ray optimization	2012
MBA	Mine blast algorithm	2013
GBMO	Gases Brownian motion optimization	2013
ACMO	Atmosphere clouds model	2013
KGMO	Kinetic gas molecule optimization	2014
CBO	Colliding bodies optimization	2014
WSA	Weighted superposition attraction	2015
LSA	Lightning search algorithm	2015
SCA	Sine cosine algorithm	2016
WEO	Water evaporation optimization	2016
MVO	Multi-verse optimizer	2016
LAPO	Lightning attachment procedure optimization	2017
ES	Electro-search algorithm	2017
TEO	Thermal exchange optimization	2017
F3EA	Find fix finish exploit analyze	2019

4. **KNN, NB, and SVM for FS**

Feature selection is the procedure where we can implement the best pre-processing procedure and we can find the best features for the prediction model implementation. We can use the traditional machine learning models to implement this concept, and there is a scenario that we can implement the concept of machine learning in the traditional search method. The traditional search method can help us to define the process of finding the best path for the prediction model implementation. The model that can help to predict the best feature from the binary features can make a best part of the prediction model implementation. The prediction model must have the robust nature without giving any issues with respect to the accuracy.

5. **Gray Wolf Algorithm**

Getting data from an animal is quite interesting, and we need to check for the alternatives when there is any situation in general life. We need to check for the things which are having most impact and that can be done based on the

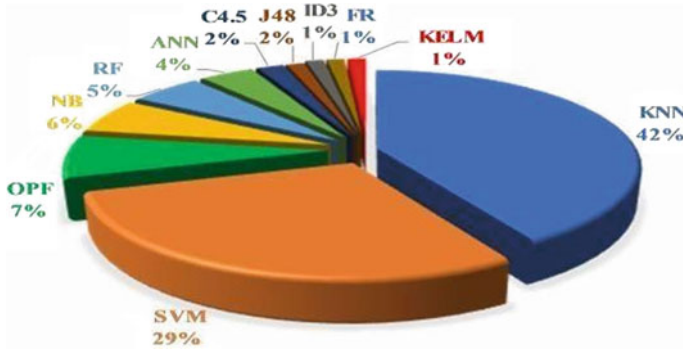
**Table 3** Human intelligence-based algorithms

Abbreviation	Algorithm	Year
LCA	League championship algorithm	2009
HIA	Human-inspired algorithm	2009
SEOA	Social emotional algorithm	2010
BSO	Brain storm optimization	2011
TLBO	Teaching learning-based optimization	2011
ASO	Anarchic society optimization	2012
SLC	Soccer league competition	2013
SBA	Social-based algorithm	2013
EMA	Exchange market algorithm	2014
GCO	Group counseling optimization	2014
JA	Jaya algorithm	2016
VPL	Volleyball premier league algorithm	2018
GSK	Gaining sharing knowledge-based algorithm	2019

search patterns. The search pattern wolf by nature is having more impact on the algorithm design. Researchers gathered the moves of the wolf and make them into the form of binary data. This data can be used for the prediction model design and implementation. This can be a helpful act for the researchers who are working the search patterns. K-NN algorithm is one of the best to form the clusters out of data, and we need to process that structure without any hesitation. The distance methodology can be most useful and that can be also useful in the form of search problems. Including all these kinds of factors, there are lot more optimization algorithms which are having the most impact on the search patterns and make the model more optimal. The list of the other kind of reputed algorithms is as follow:

6. **Ant Lion model**
7. **Krill herd model**
8. **Firefly model**
9. **Dragon fly**
10. **Whale model for optimization**
11. **Grasshopper optimization model**
12. **Salp swarm model**
13. **Penguin model for optimization**

All these are the different other kinds of algorithms which are most accurate in identifying the insights of the data. The most important factor in all these is to make the model more robust and cannot be collapsed. To make any perfect machine learning model, we need to check for the base models and then using all the optimizers is the second step. The base models for the classification are: SVM, K-NN, naïve Bayes, random forest, logistic regression, and decision trees. Including all these models, we also can use the performance evaluation-type model called XG-Boost algorithm.



**Fig. 5** Impact of each model

We have a chart that defines the list of most impactful models in the process of modeling. The process of modeling also includes the effective algorithms which are helping to understand the optimization mechanism and also the way they are helping to build the prediction models and implementation. There are different scenarios for understanding the reason behind using only the limited resources of modeling. The reason is, according to the data available, we need to modify the contents. Figure 5 will define the model procedure.

### 3.3 Evaluation Criteria

The modeling procedures need to be evaluated for the better performance, and we need to understand which methods will give the best accuracy and implementation state. We need to evaluate using the standard evaluation methods called confusion matrix. Using confusion matrix, we need to get the report regarding precision, recall, and *f1*-score. These are evaluation metrics available for better implementation of the prediction model. “Some general measure metrics are also used for checking the performance such as average fitness value (objective function value), worst and best fitness value, standard deviation of fitness values, and average number of selected features from the original datasets. These performance measures are used in [50, 51] to evaluate the performance”.

## 4 Future Work and Conclusion

The machine learning modeling is the best for implementing the life hacking models, and problem-solving approach in machine learning will help the real-world entities to solve the problems. We need to implement lot more things in the form of machine

learning. We need to make it possible to create data out of human imagination and interest. The interest human has on the small things cannot be helpful for creating a larger-than-life model. The purpose of data and the fastness of creating and solving the model are the most important thing. In every possible optimization mechanism, we considered we have SVM and random forest as the main characters. All these are the most important things. Since the dawn of the research in optimization, the same procedures are being used. That can be solved in the future if we do the research in a better manner. In this article, we discussed about the things which are most frequently used, and the wolf model is the most used and that is because of understanding the behavior of the wolf in a better manner.

## References

1. Xue B, Zhang M, Browne WN, Yao X (2016) A survey on evolutionary computation approaches to feature selection. *IEEE Trans Evol Comput* 20(4):606–626
2. Lee PY, Loh WP, Chin JF (2017) Feature selection in multimedia: the state-of-the-art review. *Image Vis Comput* 67:29–42
3. Remeseiro B, Bolon-Canedo V (2019) A review of feature selection methods in medical applications. *Comput Biol Med* 112:103375
4. Sharma M, Kaur P (2020) A comprehensive analysis of nature-inspired meta-heuristic techniques for feature selection problem. *Arch Comput Methods Eng* 1–25. <https://doi.org/10.1007/s11831-020-09412-6>
5. Asghar MZ, Khan A, Ahmad S, Kundi FM (2014) A review of feature extraction in sentiment analysis. *J Basic Appl Sci Res* 4(3):181–186
6. Saelys Y, Inza I, Larrañaga P (2007) A review of feature selection techniques in bioinformatics. *Bioinformatics* 23(19):2507–2517
7. Agrawal P, Abutarboush HF, Ganesh T, Mohamed AW (2021) Metaheuristic algorithms on feature selection: a survey of one decade of research (2009–2019). *IEE Access*
8. Liu H, Yu L (2005) Toward integrating feature selection algorithms for classification and clustering. *IEEE Trans Knowl Data Eng* 17(4):491–502
9. Ahmed S, Zhang M, Peng L (2013) Enhanced feature selection for biomarker discovery in LC-MS data using GP. In: *Proceedings of the IEEE congress on evolutionary computation*, pp 584–591
10. Aghdam MH, Ghasem-Aghaee N, Basiri ME (2009) Text feature selection using ant colony optimization. *Expert Syst Appl* 36(3):6843–6853
11. Ghosh A, Datta A, Ghosh S (2013) Self-adaptive differential evolution for feature selection in hyperspectral image data. *Appl Soft Comput* 13(4):1969–1977
12. Dash M, Liu H (1997) Feature selection for classification. *Intell Data Anal* 1(1–4):131–156
13. Guyon I, Elisseeff A (2003) An introduction to variable and feature selection. *J Mach Learn Res* 3:1157–1182
14. Liu H, Motoda H, Setiono R, Zhao Z (2010) Feature selection: an ever evolving frontier in data mining. In: *Feature selection in data mining*, Hyderabad, India, pp 4–13
15. Hoque N, Bhattacharyya DK, Kalita JK (2014) MIFS-ND: a mutual information-based feature selection method. *Expert Syst Appl* 41(14):6371–6385
16. Xu Z, King I, Lyu MR-T, Jin R (2010) Discriminative semi-supervised feature selection via manifold regularization. *IEEE Trans Neural Netw* 21(7):1033–1047
17. Tang J, Alelyani S, Liu H (2014) Feature selection for classification: a review. In: *Data classification: algorithms and applications*, p 37

18. Jović A, Brkić K, Bogunović N (2015) A review of feature selection methods with applications. In: Proceedings of the 38th international convention on information and communication technology, electronics and microelectronics (MIPRO), pp 1200–1205
19. Bebis G, Miller R (2004) Object detection using feature subset selection. *Pattern Recognit* 37(11):2165–2176
20. Jain AK, Duin RPW, Mao J (2000) Statistical pattern recognition: a review. *IEEE Trans Pattern Anal Mach Intell* 22(1):4–37
21. Mirjalili S, Mirjalili SM, Lewis A (2014) Grey wolf optimizer. *Adv Eng Softw* 69:46–61
22. Holland JH (1992) Genetic algorithms. *Sci Amer* 267(1):66–73
23. Rechenberg I (1978) Evolutionsstrategien. In: *Simulationsmethoden der Medizin und Biologie*. Springer, Berlin, Germany, pp 83–114
24. Holland JH (1992) *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence*. MIT Press, Cambridge, MA, USA
25. Glover F (1986) Future paths for integer programming and links to artificial intelligence. *Comput Oper Res* 13(5):533–549
26. Kennedy J, Eberhart R (1995) Particle swarm optimization. In: Proceedings of the IEEE international conference on neural networks (ICNN), vol 4, pp 1942–1948
27. Dorigo M, Maniezzo V, Coloni A (1996) Ant system: optimization by a colony of cooperating agents. *IEEE Trans Syst, Man, Cybern B, Cybern* 26(1):29–41
28. Karaboga D (2005) An idea based on honey bee swarm for numerical optimization. Department of Computer Engineering, Erciyes University, Kayseri, Turkey, Technical Report 06, 2005
29. Zhao R-Q, Tang W-S (2008) Monkey algorithm for global numerical optimization. *J Uncertain Syst* 2(3):165–176
30. Kirkpatrick S, Gelatt CD, Vecchi MP (1983) Optimization by simulated annealing. *Science* 220(4598):671–680
31. Woo Geem Z, Hoon Kim J, Loganathan GV (2001) A new heuristic optimization algorithm: harmony search. *Simulation* 76(2):60–68
32. Rao RV, Savsani VJ, Vakharia DP (2012) Teaching–learning-based optimization: an optimization method for continuous non-linear large scale problems. *Inf Sci* 183(1):1–15
33. Bolón-Canedo V, Alonso-Betanzos A (2019) Ensembles for feature selection: a review and future trends. *Inf Fusion* 52:1–12
34. Liu H, Motoda H (1998) *Feature extraction, construction selection: a data mining perspective*, vol 453. Springer, USA
35. Olorunda O, Engelbrecht AP (2008) Measuring exploration/exploitation in particle swarms using swarm diversity. In: Proceedings of the IEEE congress on evolutionary computation (IEEE world congress on computational intelligence), pp 1128–1134
36. Lin L, Gen M (2009) Auto-tuning strategy for evolutionary algorithms: balancing between exploration and exploitation. *Soft Comput* 13(2):157–168
37. Mohamed AW, Hadi AA, Mohamed AK (2020) Gaining-sharing knowledge based algorithm for solving optimization problems: a novel nature-inspired algorithm. *Int J Mach Learn Cybern* 11:1501–1529
38. Storn R, Price K (1997) Differential evolution—a simple and efficient heuristic for global optimization over continuous spaces. *J Global Optim* 11(4):341–359
39. Kashan AH (2009) League championship algorithm: a new algorithm for numerical function optimization. In: Proceedings of the international conference on soft computing and pattern recognition, pp 43–48
40. Civicioglu P (2012) Transforming geocentric Cartesian coordinates to geodetic coordinates by using differential search algorithm. *Comput Geosci* 46:229–247
41. Civicioglu P (2013) Backtracking search optimization algorithm for numerical optimization problems. *Appl Math Comput* 219(15):8121–8144
42. Comellas F, Martinez-Navarro J (2009) Bumblebees: a multiagent combinatorial optimization algorithm inspired by social insect behavior. In: Proceedings of the 1st ACM/SIGEVO summit on genetic and evolutionary computation, pp 811–814

43. Premaratne U, Samarabandu J, Sidhu T (2009) A new biologically inspired optimization algorithm. In: Proceedings of the international conference on industrial and information systems (ICIIS), pp 279–284
44. Yang X-S, Deb S (2009) Cuckoo search via Lévy flights. In: Proceedings of the world congress on nature and biologically inspired computing (NaBIC), pp 210–214
45. He S, Wu QH, Saunders JR (2009) Group search optimizer: an optimization algorithm inspired by animal searching behavior. *IEEE Trans Evol Comput* 13(5):973–990
46. Iordache S (2010) Consultant-guided search: a new Metaheuristic for combinatorial optimization problems. In: Proceedings of the 12th annual conference on genetic and evolutionary computation, pp 225–232
47. Yang X-S (2010) A new metaheuristic bat-inspired algorithm. In: Nature inspired cooperative strategies for optimization. Springer, Berlin, Germany, pp 65–74
48. Hedayatzadeh R, AkhavanSalmassi F, Keshtgari M, Akbari R, Ziarati K (2010) Termite colony optimization: a novel approach for optimizing continuous problems. In: Proceedings of the 18th Iranian conference on electrical engineering, pp 553–558
49. Oftadeh R, Mahjoob MJ, Shariatpanahi M (2010) A novel metaheuristic optimization algorithm inspired by group hunting of animals: hunting search. *Comput Math Appl* 60(7):2087–2098
50. Salimi H (2015) Stochastic fractal search: a powerful metaheuristic algorithm. *Knowl-Based Syst* 75:1–18
51. Dhivyaprabha TT, Subashini P, Krishnaveni M (2018) Synergistic fibroblast optimization: a novel nature-inspired computing algorithm. *Front Inf Technol Electron Eng* 19(7):815–833

# A SUMO Simulation Study on VANET-Based Adaptive Traffic Light Control System



Malose Mathiane, Chunling Tu, Pius Adewale Owola,  
and Mukatshung Claude Nawej

## 1 Introduction

Traffic congestion is a condition that happens on road networks when vehicles go slower than normal due to the greater physical usage of vehicles (traffic) on the road at that moment. Traffic congestion, often known as gridlock, can occur because of blocked roads, substandard roads, road accidents, a lack of a proper vehicle control traffic lighting system, or inappropriate driving by road users, among other things. Because of the sluggishness of traffic and the increased number of cars queued, this will make the ride longer. The road is one such overstretched infrastructure, a condition that has led to a rise in traffic. Even though traffic signals have been used reliably to regulate commuter's movement, management of traffic has remained a topic of concern in numerous real-world urban locations around the world. A traffic control light is a signaling system that commands the flow of cars [1], by traffic light control (TLC) units at road intersections. Although these traffic light settings may change few times per day, depending on the time of day, they are still considered to be fixed pre-time settings. In this case, traffic management is far from optimal, since the traffic density in real time is not considered. To minimize the time, delay due to the traffic lights, a system that regulates traffic lights operations on intersections based on the density of traffic flow is needed. There are several developed approaches for traffic optimization systems, including fixed period lighting systems, green wave lighting systems, real-time traffic signal optimization systems,

---

M. Mathiane · C. Tu (✉) · P. A. Owola · M. C. Nawej  
Computer Systems Engineering, Tshwane University of Technology, Pretoria, South Africa  
e-mail: [duc@tut.ac.za](mailto:duc@tut.ac.za)

P. A. Owola  
e-mail: [owolawipa@tut.ac.za](mailto:owolawipa@tut.ac.za)

M. C. Nawej  
e-mail: [nawejmc@tut.ac.za](mailto:nawejmc@tut.ac.za)



and real-time intersection traffic optimization systems [2]. Vehicular ad hoc networks (VANET) have recently become more common and widely deployed around the world on all roads. Most modern cars are fitted with wireless modules that provide contact points and communication control points for vehicles to connect with each other [3, 4]. According to literature [5], the most prominent wireless communication research subject is to improve inter-vehicle and roadside communication. VANET helps road vehicles to report traffic congestion, abrupt stops, and other dangerous road conditions to other vehicles [6]. IEEE 802.11p is a wireless connectivity update to the IEEE 802.11 standard that has been approved by the IEEE and was used in wireless access in vehicle environments (WAVE) [7]. Navigation systems enable vehicles to know their geographical location.

Intelligent transportation systems (ITS) are proposing the WAVE standards to describe an architecture that collectively enables vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) wireless communication [7]. For traditional traffic light control systems, even if there are no vehicles on one side, the traffic signals will still illuminate for a certain period. Vehicles must wait for the period to be completed, which is a loss of time. Adaptive traffic control systems' main purpose is to control traffic flow at an intersection based on real-time density, by changing the traffic lights in time [8], because a system that restricts traffic based on traffic flow reduces time waste [9].

Due to safety considerations, such a research area needs intensive simulation study before being tested on the route. Based on a simulation environment setup using SUMO [10], this research paper proposes a VANET-based adaptive traffic signal control system. As critical components of VANET, the road topology, vehicle movement, and the networking are simulated. According to the density of traffic flow, the timing of the signal lights is optimized, to minimize traffic congestions and other related problems.

The rest of this paper is documented using the following sections: related work which will present the review of the relevant literature, modeling of the adaptive traffic control system, SUMO simulation model and experimental design, and experimental results which focus on data analyses and discussion thereafter followed by conclusion and future work.

## 2 Related Work

To assess the density of road vehicles, several forms of traffic control systems have been implemented. A system operating on issues related to traffic, such as traffic jams, a fair latency period of vehicle stoppage, or forcible passage was demonstrated [12]. Based on the density program code and lot of cars, an adaptive control signal system dynamically modified the waiting time (red signal) and was compared to the standard traffic light control [13]. The emphasis was given to the density-based traffic system and smart streetlight systems [11]. A smart traffic control device using image

processing as a method for density calculation was suggested [12]. A novel adaptable centralized crossroad for smart traffic systems was developed, relying on smart VANETS including meaning vehicle task and controller function [13]. An intelligent system was proposed using VANET to detect traffic congestion in real time. A newly adaptable centralized intersection control was established that depends on VANETs for ITS, in which the vehicle task and the controller task were included in the intersection control. For city traffic flow optimization, a smart intersect light system was proposed [2], after testing in a real-world context, which was revealed to achieve the most realistic and accurate results. SUMO was extended by a rerouting technique that dynamically modifies the cars' route during simulation time to eliminate delays caused by random road incidents [9], by combining traffic control interface (TraCI) and its Python API with SUMO. A clever solution to traffic congestion was proposed, by changing the itineraries of vehicles delayed in traffic and rerouting them to an alternate route.

### 3 VANET Model for Adaptive Traffic Control System

In a VANET, vehicles communicate with each other and infrastructure across a 1km range utilizing the short-range radio signal DSRC (5.9 GHz). This is an ad hoc connection in which each linked node can roam around. The roadside units (RSU) connect to other network devices and act as routers between the vehicles in the vicinity. Each vehicle is equipped with an on-board unit (OBU), which communicates with RSU via DSRC radios.

#### 3.1 Modeling

Signaling equipment, such as traffic lights, traffic signals, stoplights, or robots, is placed at road intersections, pedestrian crossings, and other areas to control traffic flows. Figure 1 shows how an adaptive traffic light operates at the intersection.

SUMO cars are generated at random at a set interval of time, and an algorithm can be used to manage traffic lights, measure density, and prioritize lanes with more vehicles. The adaptive traffic light algorithm is depicted below:

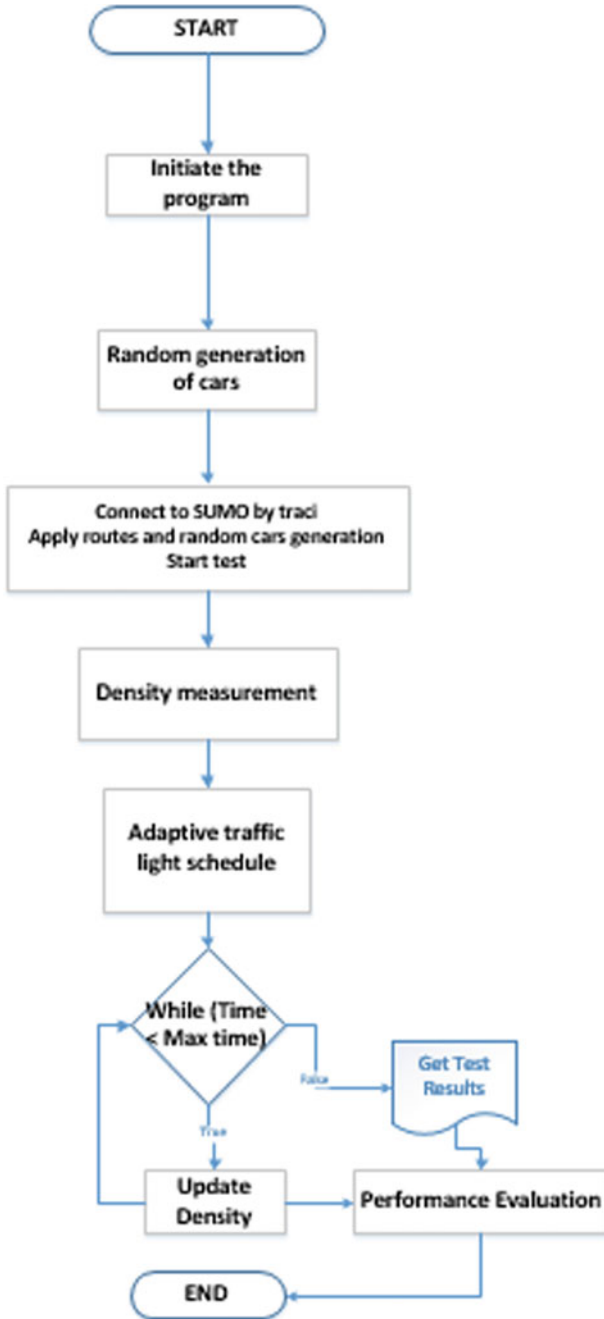


Fig. 1 Adaptive traffic signal control system's flow diagram

```

Preconditioning
<lane data, vehicle density, and signal data Compute the traffic signal time>
Simulation condition
<Execute the traffic signal time>
a traffic lights
    {
        begin
            Calculate the current lane data. Begin looping.
            Calculate the vehicle density for each lane.
        end.
    }
Priority of Signals ()
    {
        With the present traffic signal, calculate the final vehicle density.
        Assign traffic signal timing based on vehicle density.
    }

```

The number of messaged vehicles on the road is used to determine the density of traffic. The result of the traffic density measurement is subsequently utilized to evaluate traffic light timing. In Eq. 1, the traffic density is calculated (1).

$$V(t) = \frac{r + 1}{L_L L_n} \tag{1}$$

where  $r$  is the number of messaged vehicles,  $L_L$  represents the lane length, and  $L_n$  denotes the number of lanes.

### 3.2 Performance Index

The lost time  $T_L$  is the amount of time that no vehicle passes an intersection when the traffic light is displaying a green (go) signal. The average time loss  $\bar{T}_L$  on each vehicle at the intersection can be calculated as Eq. (2).

$$\bar{T}_L = \frac{T_L}{N} \tag{2}$$

where  $N$  is the number of vehicles on all sides of the intersection.

Waiting time (standby time)  $T_w$  refers to the length of time that vehicles take at the intersection waiting for the green light. The average waiting time  $T-w$  is calculated by dividing the time that vehicles were inactive in the traffic network by the number of vehicles as shown by Eq. (3).

$$\bar{T}_w = \frac{T_w}{N} \tag{3}$$

where  $N$  is the number of vehicles on all sides of the intersection.

### 4 SUMO Simulation Model and Experiments Design

Using TraCI, SUMO provides the features for traffic control, traffic light control, and lane detection. Figure 2 shows the interaction between SUMO and TraCI.

TraCI is a TCP-based client/server application that allows you to operate a traffic simulation. SUMO acts as the server, and the external script (the “controller”) acts as the client. TraCI provides commands to manage the traffic using Python when running the simulation in SUMO as a server and controlling the signal based on the current simulation state.

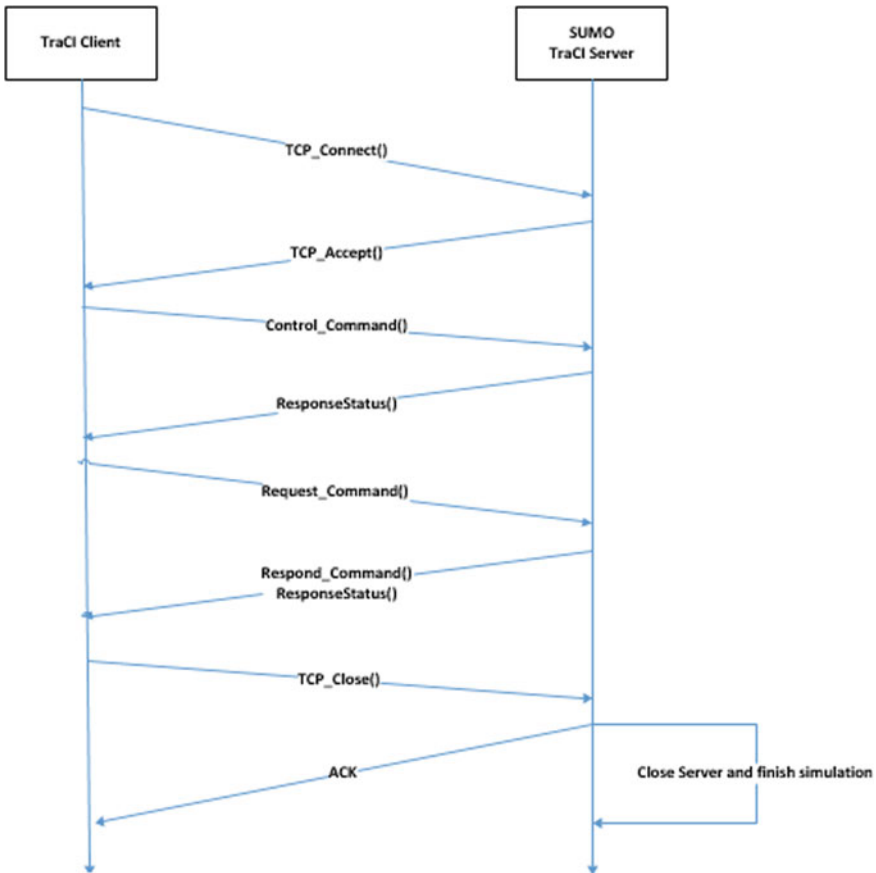


Fig. 2 Integrating SUMO and TraCI for traffic control

## 4.1 Simulation Steps

Traffic intersection is designed using SUMO to create traffic crossroad, and TraCI is used to control traffic based on traffic load and prioritize lanes at the intersection. Below are the steps on how to set up simulation in SUMO:

- The input files are supplied to SUMO to prepare the simulation network.
- Two files containing geometric network information will be converted to SUMO nodes and links, and “nod.xml” and “edg.xml” are the two files, respectively.
- The file that contains the vehicle types, the vehicles, and the corresponding roads details for every car is called as “rou.xml”, containing the “con.xml” script included to specify permitted network traffic movements.
- The “det.xml” file defines the detector position. The files of cars, the edges, and links have been put together in the simulator using the program NETCONVERT to create the geometric network.
- Combined with the route file and detector file, the network file is used to develop a “sumo.cfg” configuration file and fed to simulation graphical user interface application to generate simulation.
- The “output.xml” file stores the simulated data from SUMO in real time and can be used for result analyses and the plotting of the graphs using the Python “Matplotlib” library.
- SUMO begins with TraCI, where SUMO acts as a server and TraCI acts as a client. The two exchanges information via TCP port.
- Plexe APIs expands the TraCI methods provided by regular veins to add platooning functionality.
- Anaconda, a Python and R programming language distribution, makes scientific computing package management and deployment easier.

## 4.2 Vehicles and Intersections

The simulation is loaded using Anaconda for Python. A traditional traffic signal control is built using SUMO as shown in Fig. 3.

A vehicle is loaded in the simulation environment at the beginning of each lane every six seconds with a probability of 0.3. Simulation is executed for 1 h, and Fig. 4 shows the simulation result presented by the SUMO graphical interface.

The SUMO GUI, as seen in Fig. 4, clearly illustrates the parameters being recorded in real time while the simulation is running and may be used for data analyses. On the SUMO interface, right click and then select “show parameters” to view the current relevant vehicle-related information parameters.

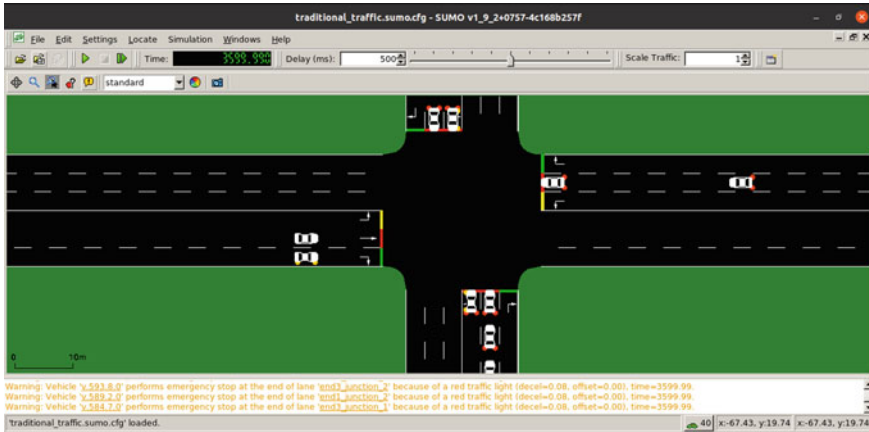


Fig. 3 SUMO traffic simulation program working environment for a traditional traffic light control

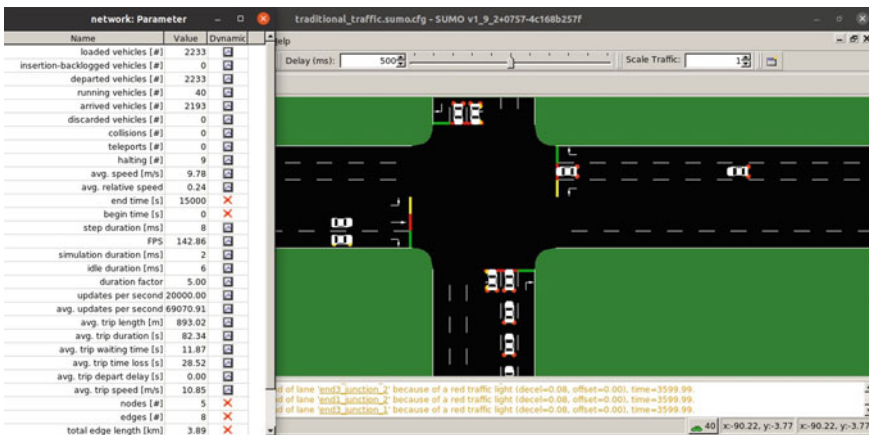


Fig. 4 SUMO traffic simulation program results representation using the graphical interface

### 4.3 Adaptive Traffic Light Control

The proposed algorithm is intended to make the most efficient use of the available resources for traffic management and time allocation. The first consideration is that each vehicle must cross the crossing, regardless of the number of vehicles present. As a result, traffic will always take a minimum amount of time to go forward. The method tries to schedule vehicles at a junction by assessing the number of vehicles on each road and dispersing this traffic using multiple parameters to dramatically reduce individual waiting times. An adaptive traffic signal is built using SUMO as shown in Fig. 5.

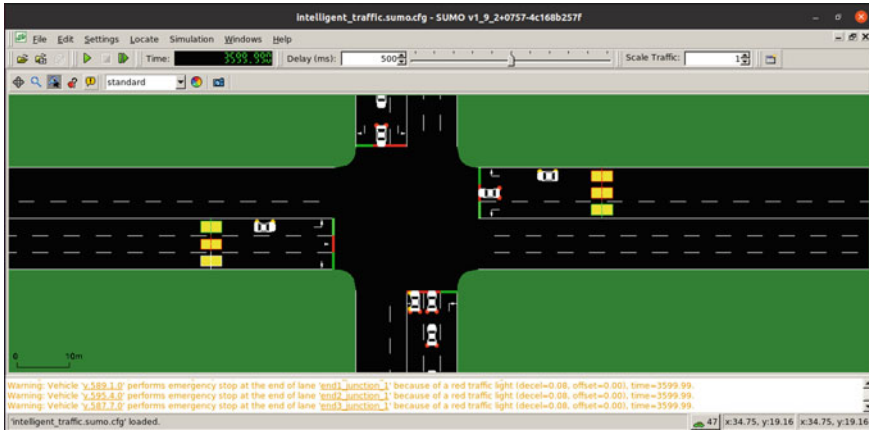


Fig. 5 SUMO traffic simulation program working environment for an adaptive traffic light control

Vehicle is loaded in the simulation at the beginning of each lane every six seconds with a probability of 0.3. The simulation was executed for 1 h, and Fig. 6 shows the simulation result presented by the SUMO graphical interface.

Figure 6 shows the SUMO GUI, which shows the parameters being recorded in real time while the simulation is ongoing and can be utilized for data analysis. To access the current relevant vehicle-related information parameters, right click on the SUMO interface and select “show parameters.”

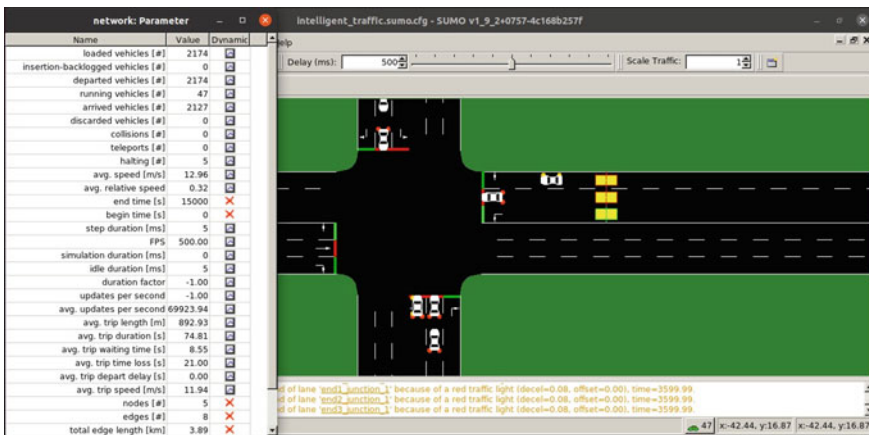


Fig. 6 SUMO traffic simulation program results representation using the graphical interface



**Table 1** Simulation input parameters

Parameters	Crossroad scenario
Channel type	Wireless
Vehicle length	4 M
Inter-vehicle distance	6
Number of lanes	12
Traffic type	TCP
MAC protocol	802.11p
Simulation time	1 h
Platoon steps	600
Platoon probability	0.3
Platoon size	1
Data bit rate	0.5 MB/s
Vehicle speed	16.6 m/s
Maximum acceleration	2.6

#### ***4.4 Parameters Used During Simulation***

During the simulation of the traffic light, the parameters listed below are utilized as input to the simulation environment, and Table 1 illustrates each parameter:

#### ***4.5 Experimental Results***

The suggested adaptive traffic light control system and the fixed-time traffic light system were separately tested in the SUMO mimicking a real-time environment with same densities and number of vehicles. Average waiting times and average time loss were compared in the experiments.

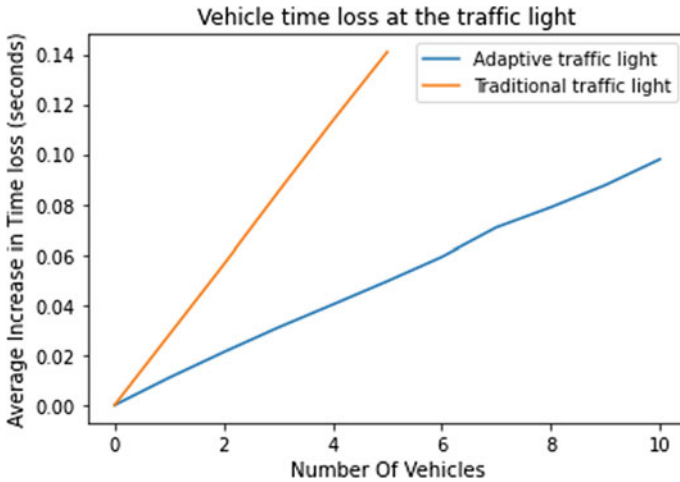
#### ***4.6 Average Time Loss***

The overall average time loss for traditional traffic light is 33.31 s after using the formula to determine the average time loss, and the average time loss for an adaptive traffic light is 21.07 s after using the formula to compute the average time loss. The average time for every 200 cars per sample was derived from the number of cars in total loaded throughout the simulation extracted using the “output.xml” to better demonstrate the performance of the traffic light as given in Table 2.

**Table 2** Average loss time  $\bar{T}_L$  for the traditional and the proposed adaptive traffic light control systems

Period	Traditional	Adaptive
	0.028087	0.010931
	0.056114	0.021191
	0.084977	0.031010
1 h	0.113412	0.040196
	0.140900	0.049484

200 vehicles as sample, six samples are shown



**Fig. 7** Average lost time at an intersection for a traditional and adaptive traffic light

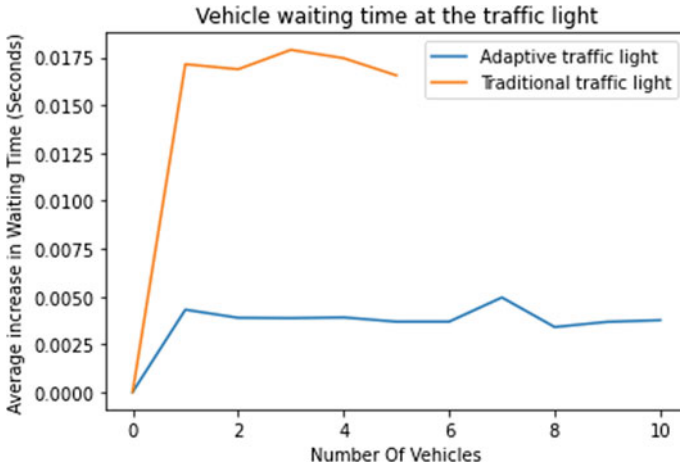
Figure 7 shows a comparison of traditional and adaptive traffic light control systems, with the adaptive traffic light having a lower number of cars losing time waiting at the traffic intersection.

### 4.7 Average Waiting Time

The overall average waiting time for traditional traffic light is 3.15 s after using the formula to determine the average waiting loss, and the average time loss for an adaptive traffic light is 0.61 s after using the formula to compute the average waiting time. The average waiting time for every 200 cars per sample was derived from the number of cars in total loaded throughout the simulation extracted using the “output.xml” file to better demonstrate the performance of the traffic light as given in Table 3.

**Table 3** Average waiting  $\bar{T}_w$  for the traditional and the proposed adaptive traffic light control systems

Period	Traditional	Adaptive
	0.000000	0.000000
	0.016552	0.003679
	0.017140	0.004307
1 h	0.016874	0.003882
	0.017889	0.003864
	0.017453	0.003901



**Fig. 8** Average waiting time at an intersection for a traditional and adaptive traffic light

Figure 8 shows a comparison of standard and adaptive traffic light control systems, with the adaptive traffic light having a lower number of cars losing time waiting at the traffic intersection.

## 5 Conclusion

This paper aimed to demonstrate a VANET-based adaptive traffic light control system simulation study using SUMO. The traditional and adaptive traffic light control systems were simulated and compared. Combined with SUMO, TraCI and its Python API were used, including “plexapi” and Anaconda for Python. The vehicle density at intersections was used to propose a traffic light control system. The simulation results compared the traditional and adaptive traffic light regulation ideas. The vehicle average waiting time and average time loss were obtained.

## 6 Future Work

Future study could concentrate on optimizing vehicle entry into the system and establishing a more realistic test environment, as well as testing artificially intelligent traffic planning techniques on the same system. Future projects could include employing artificial intelligence technological techniques to prioritize emergency cars at crossings. In addition, an emergency vehicle may be able to take a shorter route to its destination.

## References

1. Ahamed A, Vakilzadian H (2020) Impact of direction parameter in performance of modified AODV in VANET. *J Sens Actuator Netw* 9:40–40
2. Akhter M, Ahsan N, Quaderi SJS, Forhad MA, Sumit SH, Rahman MR (2020) A sumo-based simulation framework for intelligent traffic management system. *J Traffic Logistics Eng* 8
3. Celik Y, Karadeniz AT (2018) Urban traffic optimization with real time intelligence intersection traffic light system. *Int J Intell Syst Appl Eng* 6:214–219
4. Devi RB, Reddy DK, Sravani E, Srujan G, Shankar S, Chakrabartty S (2017) Density based traffic signal system using Arduino Uno. In: *International conference on inventive computing and informatics (ICICI)* pp 426–429
5. Eze EC, Zhang SJ, Liu EJ, Eze JC (2016) Advances in vehicular ad-hoc networks (VANETs): challenges and roadmap for future development. *Int J Autom Comput* 13:1–18
6. Ghazal B, Elkhatab K, Chahine K, Kherfan M (2016) Smart traffic light control system. In: *Third international conference on electrical, electronics, computer engineering and their applications (EECEA)*, pp 140–145
7. Hasan S, Elhadeif M (2019) Performance evaluation of citywide intersections traffic control algorithm in VANETs-based. In: *International conference on internet of vehicles*, pp 37–46
8. Khumara MAD, Fauziyyah L, Kristalina P (2018) Estimation of urban traffic state using simulation of urban mobility (SUMO) to optimize intelligent transport system in smart city. In: *International electronics symposium on engineering technology and applications*, pp 163–169
9. Kumari A, Patil D, Maurya V, Phalle S, Hule S, Poonia RC (2018) A performance evaluation of routing protocols for vehicular ad hoc networks with swarm intelligence. *Int J Syst Assur Eng Manage* 9:830–835
10. Sable T, Parate N, Nadkar D, Shinde S (2020) Density and time-based traffic control system using video processing. *ITM Web Conf* 3028
11. Saeed I, Elhadeif M (2018) A distributed in VANETs-based intersection traffic control algorithm. In: *Advanced multimedia and ubiquitous engineering*, pp 343–351
12. Sakiz F, Sen S (2017) A survey of attacks and detection mechanisms on intelligent transportation systems: VANETs and IoV. *Ad Hoc Netw* 61:33–50
13. Yan G, Rawat DB (2017) Vehicle-to-vehicle connectivity analysis for vehicular ad-hoc networks. *Ad Hoc Netw* 58:25–35

# FINNger: To Ease Math Learning for Children using Hand Gestures by Applying Artificial Intelligence



Pallavi Malavath, Nagaraju Devarakonda, and Zdzislaw Polkowski

## 1 Introduction

Artificial intelligence is considered as simulation of intelligence of humans which is processed by machines, called computer systems. The expert systems like natural language processing-NLP, speech recognition, and machine vision focuses on hand gestures in many applications. The main objective of artificial intelligence is to make computers perform intellectual tasks like decision-making, perception, problem solving, and analyzing human communication so that it can translate in any language. Kids have astonishing capability to use latest smart devices like smart phones, tablets, etc. It is a known fact that children are facing difficulty in learning mathematical-related subjects at pre-school. That is why an easy-to-follow method would be to make usage of this new technology while developing/teaching our children. Our focus is extensively on alphabetization so that children can overcome from the difficulties in learning various subjects especially mathematics. In this work, we proposed a basis for instinctual application that provides a lot of ease while using technological applications and enhances the children knowledge and improves capability to learn new concepts in a low age (Fig. 1).

---

P. Malavath (✉) · N. Devarakonda  
School of Computer Science and Engineering (Scope), VIT-AP, Amaravati, India  
e-mail: [pallavimalavath.20phd7126@vitap.ac.in](mailto:pallavimalavath.20phd7126@vitap.ac.in)

Z. Polkowski  
Technical Sciences, Jan Wyzykowski University, Polkowice, Poland  
e-mail: [z.polkowski@ujw.pl](mailto:z.polkowski@ujw.pl)

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2022  
T. Sengodan et al. (eds.), *Advances in Electrical and Computer Technologies*,  
Lecture Notes in Electrical Engineering 881,  
[https://doi.org/10.1007/978-981-19-1111-8\\_20](https://doi.org/10.1007/978-981-19-1111-8_20)

239



**Fig. 1** Example for hand recognition

## 2 Literature Survey

Naqvi et al. [1] proposed a hybrid technique which computes convex hull based on Grahams scan algorithm and quick hull algorithm. Castro et al. [2] used Jarvis March and Quick Hull algorithm in order to get convex hull for set of points which can be maintained as small data structure which is in a compact manner called  $k^2$ -tree. Alshamrani et al. [3] proposed a technique for preprocessing based on Graham scan algorithm for computing a convex hull for a set of points in 2D space by enhancing Graham scan algorithm in addition to the filtering techniques for a fastest convex hull computing. Cinque and Di Maggio [4] proposed an structured parallel algorithm for designing a convex hull which is realized by the BSP technique with complexity of  $O(nh/p)$ . Panvar et al. [5] presented a real-time technique for hand gesture recognition by detecting different shapes, orientation, fingers status, centroid, thumb related to folded fingers and raised fingers, and corresponding location in a given image. Tan et al. [6] introduce a convex hull algorithm based on Bresenham algorithm combined with Jarvis March in which the convex hull algorithm can be applied directly to the pixels in the binary image instead of finding the edge detection, thereby reducing the allocation of resources. Alzubaidi [7] proposed a method for finding the minimum bounding of a 2D convex polygon by using Jarvis March and Graham scan methods. Bermudez et al. [8] investigate the usage of different battery technologies and influence of the media access layer (MAC) and physical (PHY) layers by using Jarvis March method which inspects the effect of LoRa's MAC and physical layers. Gurav and Kadbe [9] proposed a technique which uses different hand gestures which controls the robo for offices and household and used AdaBoost algorithm.

## 3 Proposed Application

In this work, we present a technique which builds a desktop application where children could easily use their camera or any image capturing device, and they just need to show their hands to the screen and computer or any device which can capture image can be able to identify how many fingers that kid raised and can be able to compute

**Fig. 2** Example of 2 raised fingers of kid's hand



basic arithmetic operations with it. Preferably in future work, this technique should be able to work in smart phones, and also, this would be a benefit to non-developed countries. In this work, we proposed two different techniques:

- Contour computing and finding its convex hull by using Jarvis March algorithm.
- Implementing a convolutional neural network (CNN) which is trained on more number of images so that the trained model should generalize and can differentiate colors, hand sizes, different backgrounds, and different illumination effects (Fig. 2).

### **3.1 Our Methodology**

See Fig. 3.

## **4 Datasets**

For developing the application, we have taken two datasets:

- To find the convex hull of the contours and to test the detection of fingertips.
- To train the neural network.

### **A Koryakin/fingers**

This dataset is developed by Pavel Koryakin, and it is available on <https://kaggle.com> with the public domain license. It contains 21,600 black and white  $128 \times 128$  images belongs to both right and left hands, from 0 to 5 raised fingers. These 21,600 images are split 18,000 and 3600 images. 18,000 images are used

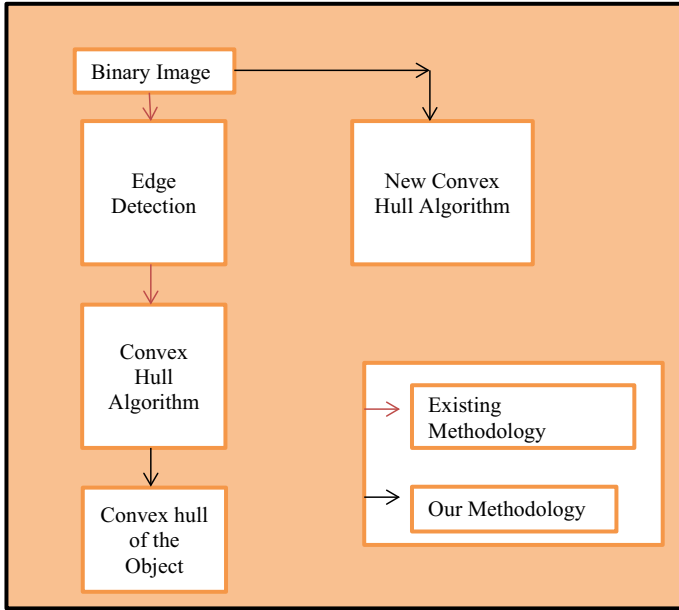


Fig. 3 Proposed methodology

to train the dataset, while 3600 is used for validation set/test set. Some of the example images can be seen in Fig. 4.

**B Custom dataset**

We observe that the mentioned dataset is not closely related to reality because images have a particular illumination and specific poses. So, in proposed system, we created a custom dataset in which images will have different illuminations and real-life backdrop setups. We created a custom dataset with 3200. Out of 3200 images, 2850 has been used for training dataset, and 360 images are used for validation test/test set. We have developed a custom application which can be able to generate entire dataset in 15 min. This makes task easier. Our proposed application asks for number of images that we wish to generate, and it asks for



Fig. 4 Available images on [10] dataset



**Fig. 5** Example images of custom dataset



which hand (right/left) and quantity of numbers we are planning to take photos, and finally, it will start taking photos using Webcam. Some samples of pictures can be seen in Fig. 5. These images can also have the presence of body of the particular person, not only hand, very closely related to the real-life situations.

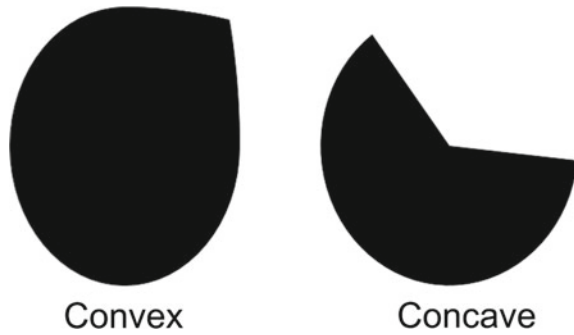
## 5 A Implementation

### 5.1 Convex Hull

Convex hull can be divided into two parts. First one is convex, and a convex object is an object with no interior angles greater than  $180^\circ$ . A shape which is not convex is called concave or non-convex. Second one is hull, and it is a shape of convex hull with tight fitting boundary around the shape. Figure 6 shows the objects of convex and concave structure.

For constructing convex hull, the problem is about developing, constructing, articulating, and including a given set of points in a specific plane by polygon capsule known as *convex polygon*. It is a very simple polygon without self-intersection, and no edges go beyond the polygon. A boundary of a convex set can be formed by *convex curve*.

**Fig. 6** Convex object and concave object



## 5.2 Implementation of Convex Hull

To find the convex hull for a given specific set of points, we need some algorithms, and those are called *gift wrapping algorithms*. In our proposed work, we have used Jarvis March algorithm. This algorithm comes under gift wrapping algorithms. Jarvis March algorithm has the complexity  $O(nh)$ , where “h” is number of points in convex hull, and “n” is number of input points. To apply this technique for images, we need to follow some steps:

- (a) We need to binarize the input image.
- (b) Find contours and convex hull.
- (c) Detecting the fingertips.

Jarvis March algorithm determines the points related to the convex hull in the original order. It works fast when the key points are less. Let  $X_0$  be labeled as the leftmost point and  $X_1$  is first point in counterclockwise in point of view from  $X_0$ . Now,  $X_2$  will be the first point in counterclockwise in  $X_1$  point of view and so on.

### 5.2.1 Jarvis March Algorithm

Finding  $X_{i+1}$  will take linear time. Here, while loop will be executed  $n$  times. More precisely, while loop will be executed  $h$  times, here  $h$  = number of convex hull vertices. Jarvis March time complexity is  $O(nh)$ . The best case will be  $h = 3$ . The worst case is  $h = n$ , when the points are, for example, arranged on a circumference of a circle (Figs. 7 and 8).

## 5.3 Binary Mask Preparation

Binarization can be done in three steps:

- (a) Conversion of input image on to gray scale.

- 1) Initialize  $p$  as left most point
- 2) Do following while we do not come back to the first or left most point
  - a) The next point  $q$  is the point such that the triplet  $p, q, r$  is counterclockwise for any point  $r$ .
  - b)  $next[p] = q$  (store  $q$  as next of  $p$  in the output convex hull).
  - c)  $p = q$  (set  $p$  as  $q$  for next iteration)

Fig. 7 Jarvis March for algorithm

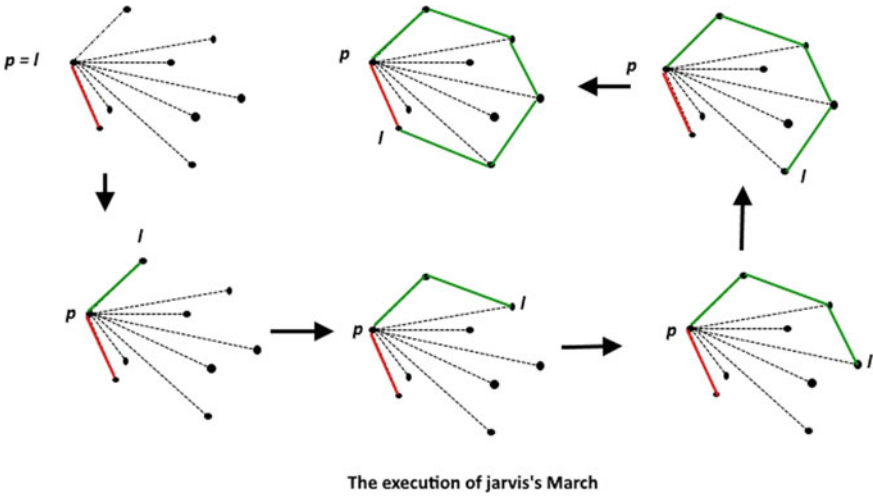
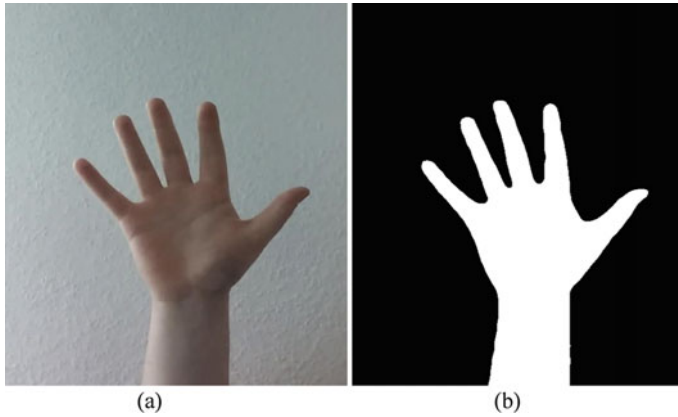


Fig. 8 Detailed execution of Jarvis March algorithm

- (b) Blur the image to remove noise.
- (c) Threshold to binarize the image.

We need to create the binary mask of hand to find the contour of a hand and then convert the input frames in BGR format to hue, lightness, saturation (HLS) color space. It will give the color information of a given image. We need to take care of proper hue value of the skin, and later, we can adjust the lightness and saturation. In this work, the hue range is  $0^{\circ}$ – $30^{\circ}$ , and saturation range is around 6–65% by using a color picker. To overcome the noises like small gaps and single pixels, we will improve the mask of the hand, thereby smoothing the mask by blurring and keeping a threshold value to get a binary mask (Fig. 9).



**Fig. 9** a Frame, b mask

#### 5.4 Find Contours and Finding the Convex Hull of the Hand

We find the contour of the hand by using *findContour* function related to OpenCV. By finding the contours, we will get the boundary points of hand. *findContour* function finds the connections and gives contour as a list. *ConvexHull* function is used to find the convex hull of the contours. Finally, to visualize convex hulls, we use OpenCV's *drawContours* function.

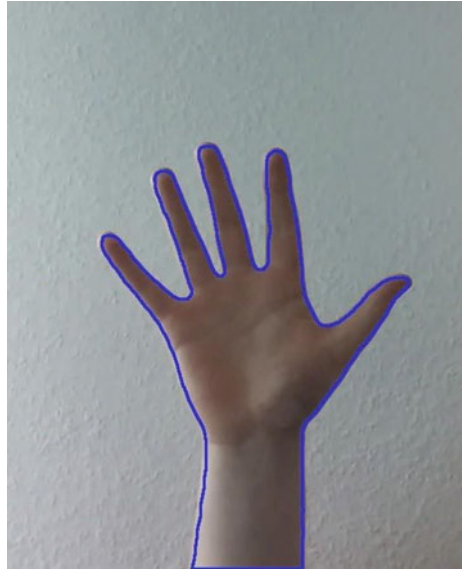
After detecting the hand contour, an algorithm called Jarvis March is used to determine the convex hull and detection of fingertips. For this, we need to find the convex hull and convexity-defect regions (deviation of objects from the hull) of the contour of the hand. We will allot each point in the neighborhood to the different cluster and later select the optimal central key point of every cluster. The edge points are indicated by red circles as in Fig. 11.

#### 5.5 Detection of Fingertips

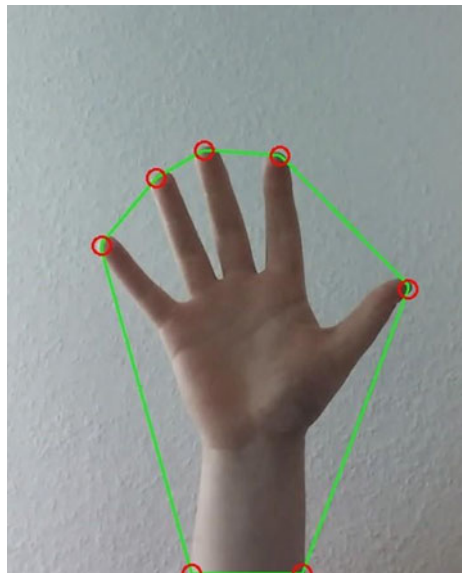
Now, after computing the convexity defects of a hand contour, it gives defect regions of the contour of the hand which are labeled as starting point  $P_0$ , end point  $P_1$ , defect point  $P_2$ . Defect points are indicated by circles in green color and are detected in the "valley" position between the two points of convex hull related to hand contour.

When implementing the common approach, a problem is encountered, and the **alpha angle** between the fingers is very less to be considering as fingertips. If we continue with the common approach problem will encounter if single finger is raised by children, it cannot be detected. To overcome this problem, we came up with another idea by considering **beta** angle. It works well. The defect points areas are returned to the array of vectors. The vector entry indicates the index point in a hand contour.

**Fig. 10** Contour of the hand



**Fig. 11** Convex hull of a hand



The entry 0 indicates starting point, the entry 1 indicates ending point, and the entry 2 indicates the defect points. We allocate every hull point by its 2 neighboring defect points and then its discards the points which do not have 2 neighbors, and they should not consider those points which are located at a fingertip every time. After getting hull points, we can find angle **beta** by 2 corresponding vectors  $pt \rightarrow d_1$  &  $pt \rightarrow d_2$

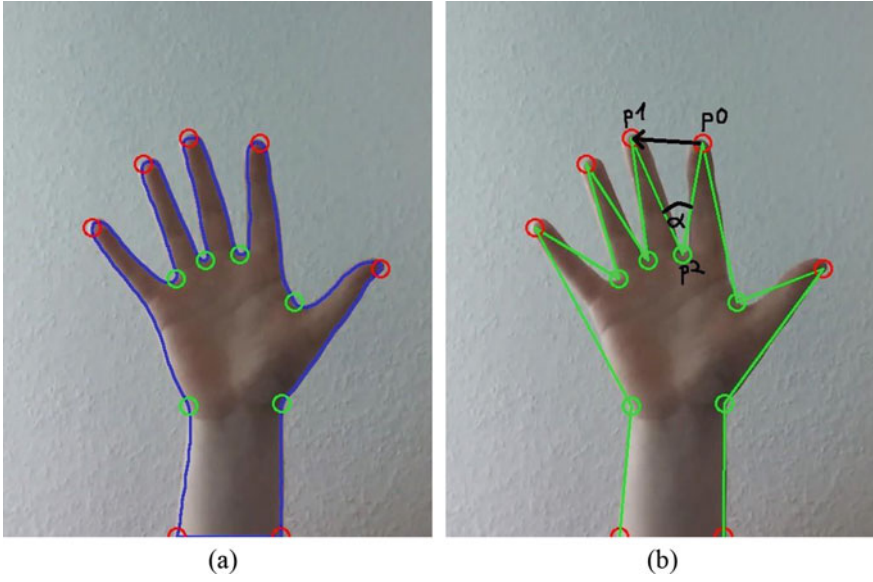
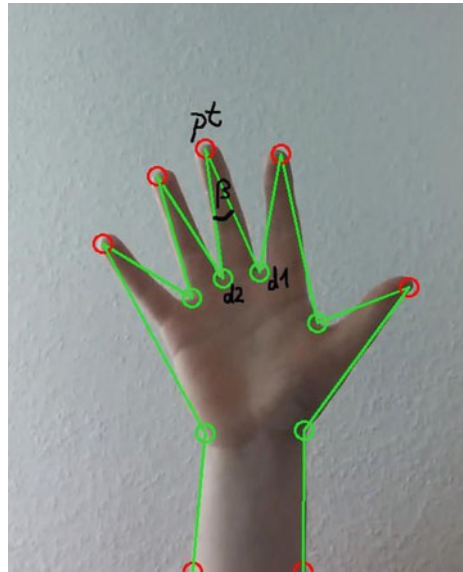


Fig. 12 a Defect points, b common approach

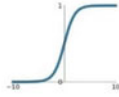
Fig. 13 Proposed approach



## Activation Functions

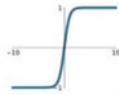
**Sigmoid**

$$\sigma(x) = \frac{1}{1+e^{-x}}$$



**tanh**

$$\tanh(x)$$



**ReLU**

$$\max(0, x)$$



**Leaky ReLU**

$$\max(0.1x, x)$$



**Maxout**

$$\max(w_1^T x + b_1, w_2^T x + b_2)$$

**ELU**

$$\begin{cases} x & x \geq 0 \\ \alpha(e^x - 1) & x < 0 \end{cases}$$

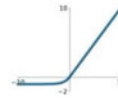


Fig. 14 Possible activation functions [4]

for every vertex while applying law of cosines. Finally, depends on the sharpness of the **beta** angle, we can determine how many fingers are raised.

### A Convolutional Neural network

CNN is related to deep learning algorithm which takes an input image and allocates weights to various objects, and it can be able to differentiate from one another. It is the major part of neural networks. This model allows us to extricate higher representations from the input image. A survey and rough overview about CNN with explanation can be seen in [11].

- **Activation**

Activation function is a part of artificial neural network which helps to learn various patterns in the data. In various parts of the network, a neural network has many activation functions, and it has high impact on neural network learning (Fig. 14).

- **Batch Normalization Layer**

These layers were instigated on [12] to circumvent internal co-variant shift and are defined as a change in the distribution of activation functions during training. When the data pass from one layer to another layer, it creates numerical instability due to variations in the range. This layer re-scales and re-centers the layer values in sustainable range.

- **Max Pooling Layer**

It is a pooling technique which selects max element from the feature map of the input image which is covered by a filter. The outcome of the max pooling layer is a feature map with outstanding features of the feature map. Max pooling layer decreases the shape/size of the input image. Max pooling explanation is shown in Fig. 15. This pooling layer has kernel size  $k \in \mathbb{R}^p$  which is recapitulate through the input image and reduces the input image dimensionality. Assume, a matrix  $M \in \mathbb{R}^{p \times q}$ , if we apply  $(i, j)$  kernel-sized pooling layer, then a new matrix will come  $N \in \mathbb{R}^{[p/i] \times [q/j]}$  (Fig. 16).

- **Dropout Layer**

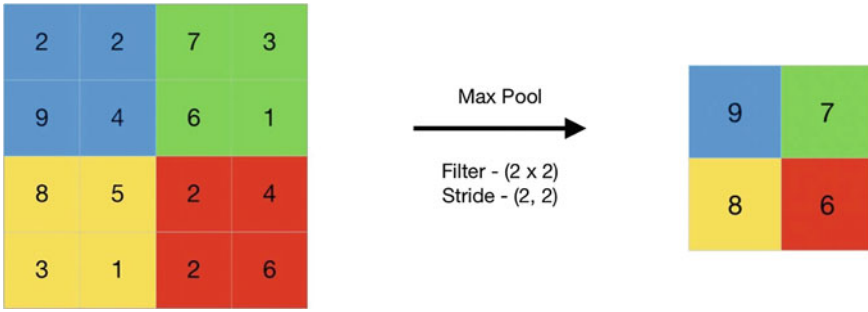


Fig. 15 Example of a 2D max pooling application using a  $2 \times 2$  kernel [13]

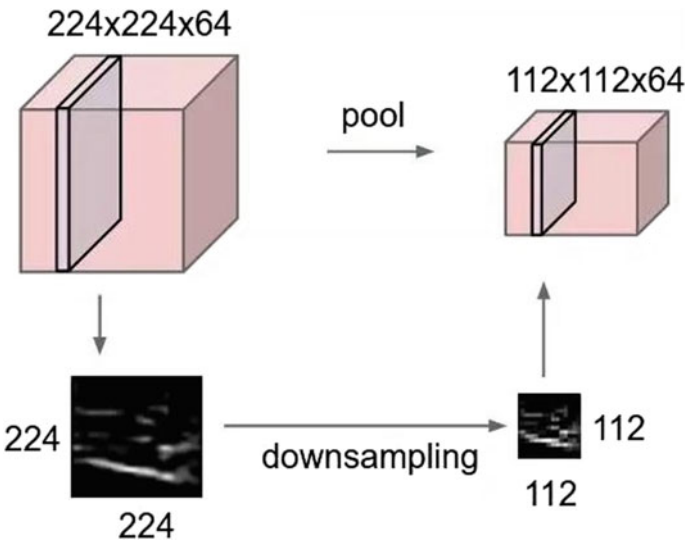


Fig. 16 Example of a 2D max pooling down sampling

As mentioned in [14], to overcome the problem of over-fitting, these dropout layers are used. It randomly drops the previous layer outputs, i.e.; it drops every feature with some probability  $p$  using Bernoulli distribution samples. These dropout layers can be used in any layer.

### B The Proposed Model FINNger

The proposed model FINNger was implemented in Python by using PyTorch [15, 13]. We have trained several models and tried to detect quantity of raised fingers (numbers) and find the difference between left/right hand on the input image. The dropout layers and batch normalization layers do not alter the output shape and also input shape. In this, we have used (4,



4) kernel size for every step of convolutional layer. It consists of several building blocks having an input image with dimensions (3, 96, 96):

1. **C<sup>1</sup> Block:** Convolutional block-1 consists of:
  - I **C<sup>1</sup>c<sup>1</sup>:** 2D (2 dimensional) convolutional layer with an outcome matrix (64, 96, 96), with rectified linear function (ReLU) activation function.
  - II **C<sup>1</sup>b<sup>1</sup>:** Batch normalization layer.
  - III **C<sup>1</sup>c<sup>2</sup>:** 2D (2 dimensional) convolutional layer with an outcome matrix (64, 96, 96), with rectified linear function (ReLU) activation function.
  - IV **C<sup>1</sup>b<sup>2</sup>:** Batch normalization layer.
  - V **C<sup>1</sup><sub>max-p</sub>:** 2 dimensional max pooling layer having  $2 \times 2$  kernel, with an outcome matrix (64, 48, 48).
  - VI **C<sup>2</sup>d<sub>0.2</sub>:** Dropout layer having  $p = 0.2$  (randomly setting), 20% of input values are set to 0.
2. **C<sup>2</sup> Block:** Convolutional block-2 consists of:
  - I **C<sup>2</sup>c<sup>1</sup>:** 2D (2 dimensional) convolutional layer with an outcome matrix (128, 48, 48), with rectified linear function (ReLU) activation function.
  - II **C<sup>2</sup>b<sup>1</sup>:** Batch normalization layer
  - III **C<sup>2</sup>c<sup>2</sup>:** 2D (2 dimensional) convolutional layer with an outcome matrix (128, 48, 48), with rectified linear function (ReLU) activation function.
  - IV **C<sup>2</sup>b<sup>2</sup>:** Batch normalization layer
  - V **C<sup>2</sup><sub>max-p</sub>:** 2 dimensional max pooling layer having  $2 \times 2$  kernel, with an outcome matrix (128, 24, 24).
  - VI **C<sup>2</sup>d<sub>0.3</sub>:** Dropout layer having  $p = 0.2$  (randomly setting), 30% of input values are set to 0.
3. **C<sup>3</sup> Block:** Convolutional block-3 consists of:
  - I **C<sup>3</sup>c<sup>1</sup>:** 2D (2 dimensional) convolutional layer with an outcome matrix (128, 24, 24), with rectified linear function (ReLU) activation function.
  - II **C<sup>3</sup>b<sup>1</sup>:** Batch normalization layer
  - III **C<sup>3</sup>c<sup>2</sup>:** 2D (2 dimensional) convolutional layer with an outcome matrix (128, 48, 48), with rectified linear function (ReLU) activation function.
  - IV **C<sup>3</sup>b<sup>2</sup>:** Batch normalization layer
  - V **C<sup>2</sup><sub>3max-p</sub>:** 2 dimensional max pooling layer having  $2 \times 2$  kernel, with an outcome matrix (128, 12, 12).
  - VI **C<sup>3</sup>d<sub>0.4</sub>:** Dropout layer having  $p = 0.4$  (randomly setting), 40% of input values are set to 0.

4. FC<sup>1</sup>: Fully connected layer, it is responsible for the outcome of classification, which gives output—(6) shaped vector. Finally, we apply “log softmax” function to the outcome which is shown the following equation:

$$\text{LogSoftmax}(i) = \log(\exp x_i) / \sum_j \exp x_j \quad (1)$$

## 6 Results

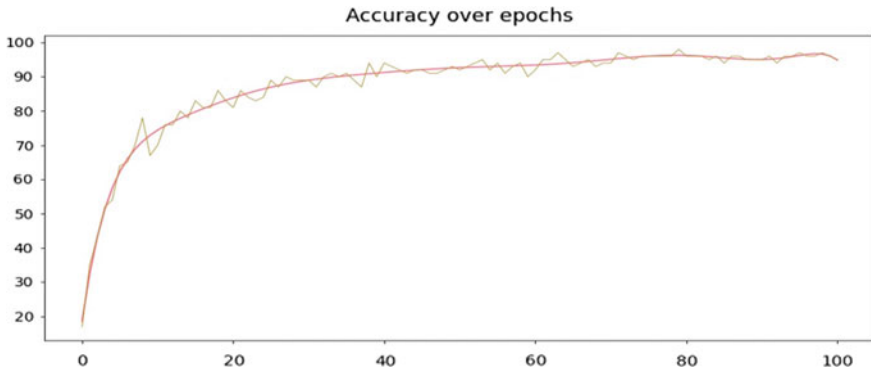
### 6.1 *Finding the Convex Hull of the Contours and Testing the Detection of Fingertips*

We have evaluated all the input images related to koryakinp/fingers dataset to examine the performance of convex hull of hand contours. We achieved good results with high accuracy about 85% using the test dataset images and 83.9% using training dataset images.

### 6.2 *FINNger*

We have taken [12]’s dataset, and we faced some real-time issues like lighting and background issues, even though we have achieved 93% accuracy on validation set after 10 epochs. Later, we tried our custom dataset along with [12] dataset, and it ended up with minor issues. The main problem is with background and illumination effects in real time. Then, we have changed our approach by using only custom dataset instead of mixing both. Because of small dataset, we implemented data augmentation to overcome the problem. In our approach, data augmentation was very simple, and it just creates horizontal flip to the input image and random translations. As we did not use rotation augmentation, if we tilt hand, the accuracy decreases radically. By using Adam [15] optimizer, we have trained our application for about 100 epochs. Choosing 0:0003 for learning and 0:0001 for weight decay as hyperparameters, we trained our model. These parameters were tested first, and they worked very well with good results. With the validation dataset, we attain 98% high accuracy which is shown in Fig. 17 with circa 75 epochs and oscillating to-and-fro to 95%.

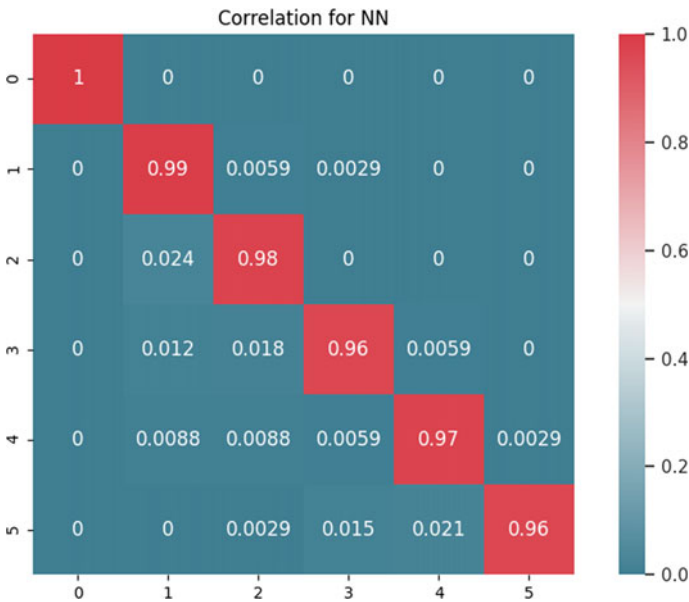
By representing correlation matrix which is shown in Fig. 14, we have high correlation, and when mislaid the value, and likely to get the results wrong because we



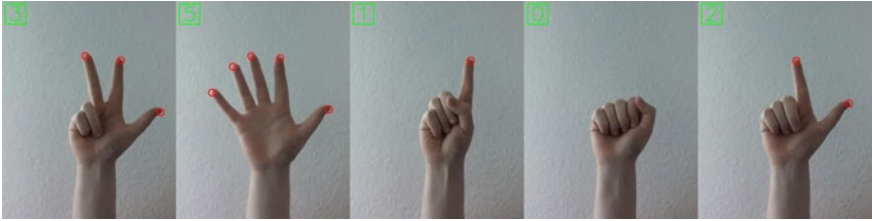
**Fig. 17** Proposed model with accuracy from the 100 epochs, the brown line shows accuracy, and orange line shows smoothed polynomial

are having few numbers because our application can classify the images with 6 categories [0, 1, 2, 3, because we are having few numbers because our application can classify the images with 6 categories [0, 1, 2, 3, 4, 5] (Fig. 18).

We got expected results with children’s hands also. On the left top, there is a value that our proposed neural network presently expects it is being shown to it. Below, calculation will be displayed. To perform calculations, press enter on the terminal when we want to get the result. For debugging purpose, the output values by the



**Fig. 18** Expected values (lines) with correlation and the output (columns)



**Fig. 19** Detecting how many fingers are raised

**Fig. 20** UI from our application



network are displayed on the bottom side. An example is shown in Fig. 20, in which it exactly detects number 2. Based on the output from the network, it is very confident that the detected number is correct and it is 2 (Fig. 19).

## 7 Future Work

To get better results,

- By using the dataset, [12] create a new synthetic dataset by adding custom backgrounds, different illumination, translation generating, and rotation.
- To make the application more accessible, create mobile application.

## 8 Conclusion

With our proposed application, we achieved good results by using convex hull technique with stable position and lighting settings. It worked for homogeneous images in the dataset achieved good results with kids hands but with minimal issues like illumination effects and rotation. By implementing the combination of above mentioned techniques in “future work,” we can get better outcomes and can be used in real life.

We focus broadly on alphabetization so that children can overcome from the difficulties in learning various subjects especially mathematics. In this work, we proposed a basis for instinctual application that provides a lot of ease while using technological applications and enhances the children knowledge, thereby implementing Jarvis March algorithm in artificial intelligence for the finger tip detection which can improve the capability to learn new concepts and also implementing a novel CNN to achieve a new technique so named FINNger.

## References

1. Naqvi SZZ, Raza M, Haider W (2020) A new approach to compute convex hull. *Innovative Syst Des Eng* 2(3). [www.iiste.org](http://www.iiste.org) ISSN 2222-1727 (Paper) ISSN 2222-2871 (Online)
2. Castro JF, Romero M, Gutiérrez G, Caniupán M (2020) Efficient computation of the convex hull on sets of points stored in a k-tree compact data structure. *Knowl Inf Syst* 62:4091–4111
3. Alshamrani R, Alshehri F, Kurdi H (2020) A preprocessing technique for fast convex hull computation. In: *The 11th international conference on ambient systems, networks and technologies (ANT)*, Warsaw, Poland
4. Cinque L, Di Maggio C (2002) A BSP realisation of Jarvis’s algorithm. In: *Proceedings 10th international conference on image analysis and processing*. IEEE, Venice, Italy. <https://doi.org/10.1109/ICIAP.1999.797603>, ISBN:0-7695-0040-4
5. Panvar M (2012) Hand gesture recognition based on shape parameters. In: *International conference on computing, communication and applications*. IEEE, Dindigul, India. INSPEC Accession Number: 12655227, <https://doi.org/10.1109/ICCCA.2012.6179213>
6. Tan GY, Woun BS, Wong YP (2016) A hybrid convex hull algorithm for fingertips detection. *Indian J Sci Technol* 9(28). <https://doi.org/10.17485/ijst/2016/v9i28/97821>
7. Alzubaidi AMN (2014) Minimum bounding circle of 2D convex hull. *Int J Sci Res (IJSR)* 3(9):364–367
8. Bermudez E, Sadok DFH (2021) Application of a computational geometry algorithm to improve the energy consumption in a LoRaWAN network. *Int J Digit Soc (IJDS)* 12(1)
9. Gurav RM, Kadbe P (2015) Real time finger tracking and contour detection for gesture recognition using OpenCV. In: *International conference on industrial instrumentation and control (ICIC)*. College of Engineering Pune, India. <https://doi.org/10.1109/IIC.2015.7150886>
10. Paszke A et al (2019) PyTorch: an imperative style, high-performance deep learning library. In: *NEURIPS2019*. Available at <http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf>
11. Li Z et al (2020) A survey of convolutional neural networks: analysis, applications, and prospects. *CoRR*. Available at <https://arxiv.org/abs/2004.02806>
12. Koryakin P (2019) koryakinp/fingers Kaggle’s Dataset, Kaggle. Available at <https://www.kaggle.com/koryakinp/fingers>
13. Science Wiki (2018) Max-pooling/pooling—computer science Wiki. Available at <https://computer-science-wiki.org/index.php?title=Max-pooling/Pooling&oldid=7839>

14. PyTorch Website Documentation. Available at <https://pytorch.org/>
15. Silver L (2019) Smartphone ownership is growing rapidly around the world, but not always equally. Pew Research Center. Available at <https://www.pewresearch.org/global/wp-content/uploads/sites/2/2019/02/Pew-Research-CenterGlobal-Technology-Use-20182019-02-05.pdf>

# Control of Rice Blast Pathogens Using Back Propagation Fuzzy Neural Network



S. Anita Shanthi, G. Sathiyapriya, and L. Darwin Christdhas Henry

## 1 Introduction

Cheng et al. [1] analyzed the characteristics and mathematical concepts of BPFNN and suggested several methods for improving BP algorithms. Ishibuchi et al. [2] proposed an architecture of MLBPFNN for classifying problems based on fuzzy vectors. In the proposed NN a fuzzy vector was mapped to a fuzzy number and the activation function extended to an input output relation. Scheibel et al. [3] described an algorithm which combined the centers and variance of Gaussian nodes to improve the response of a RBF neural network. Wen Jin Wei et al. [4] proposed a new method of BP algorithm and a new model of controllable FFNN. Lee [5] applied FNN to the synthesis of fuzzy automata. Fuzzy numbers were introduced by Zadeh and distances between fuzzy numbers by Dijkman et al. [6]. Ishibuchi et al. [7] developed FNN architecture whose weights were taken as trapezoidal fuzzy numbers. Requena [8] developed DPI of fuzzy numbers based on trained ANN. Pathinathan et al. [9] defined pentagonal fuzzy number and some basic operations on pentagonal fuzzy numbers. Virgin Raj et al. [10] found out the quality of fertilizer used to get an expected output, using triangular and pentagonal fuzzy number BPNN. Anita Shanthi et al. [11] applied trapezoidal fuzzy number in RBFNN and developed an algorithm based on this concept.

Nattapatphon Kongcharoen et al. [14] studied the fungicidal activities of various fungicides against rice blast. Miah et al. [15] evaluated the efficiency of different

---

S. Anita Shanthi (✉) · G. Sathiyapriya

Department of Mathematics, Annamalai University, Annamalainagar, Chidambaram, Tamilnadu 608002, India

e-mail: [shanthi.anita@yahoo.com](mailto:shanthi.anita@yahoo.com)

L. Darwin Christdhas Henry

Department of Plant Pathology, Annamalai University, Annamalainagar, Chidambaram, Tamilnadu 608002, India

fungicides in controlling blast disease. Based on these concepts the notion of BPFNN is developed and applied to minimize the disease incidence and enhance the productivity of rice crops.

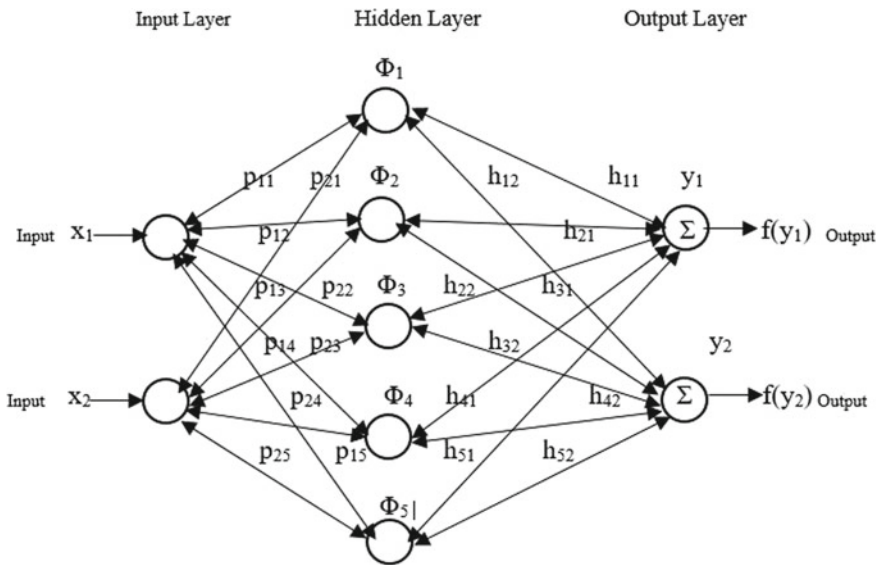
## 2 Back Propagation Neural Network (BPNN)

Back propagation is a short form for “Backward propagation of errors”. It is a standard method of training ANN. In this network, first feed forward the inputs  $x_i$  and the input layer weights  $p_{ij}$  to find the Gaussian activation function for the hidden layer function. Then the sum of the product of Gaussian function  $\Phi_{ij}$  and the output layer weights  $h_{ij}$  are used to determine the output function and it is activated using sigmoid activation function. If the calculated output  $f(y)$  is not equal to the assumed target value, then the error with respect to all the weights in the network are calculated. The weights are updated and then back propagated to get the expected output as shown in Fig. 1.

**Definition 2.1** The Gaussian activation function is used in the hidden layer.

$$\Phi_{ij} = \exp\left(\frac{-(p_{ij} - m_i)^2}{\sigma_i^2}\right)$$

where  $p_{ij}$ 's are the weights of the input layer,  $m_i$  and  $\sigma_i$  are mean and variance of the Gaussian activation function.



**Fig. 1** Backpropagation fuzzy neural network



**Definition 2.2** The outputs of BPF neural network is the sum of the product of Gaussian activation function and the output layer weights.

$$y_i = \sum_{j=1}^m h_{ij} \Phi_{ij} \quad i = 1, 2, \dots, 5 \quad j = 1, 2.$$

The sigmoid activation function is used in the output layer,

$$f(y_i) = \frac{1}{1 + \exp(-y_i)}$$

**Definition 2.3** For output layer, the error is defined using error formula,

$$e_{O_i} = (T_i - f(y_i)) * f(y_i) * (1 - f(y_i))$$

where  $T_i$ 's are assumed target values and  $f(y_i)$ 's are the calculated output. The weight correction formula for output layer is,

$$h_{ij}^1 = h_{ij} - \delta e_{O_i} f(y_i)$$

where  $\delta$  is the assumed learning rate.

**Definition 2.4** For input layer, the error

$$e_{ij} = \Phi_{ij} * (1 - \Phi_{ij}) * h_{ij}$$

where  $\Phi_{ij}$  is the Gaussian activation function. The weight correction formula for input layer is,

$$p_{ij}^1 = p_{ij} - \delta e_{ij} \Phi_{ij}$$

The expected target values are fixed. Inputs are feed forwarded. The weights of the input layer are taken to be any fuzzy number. Here the input layer weights are taken as pentagonal fuzzy numbers. Gaussian activation function is used in the hidden layer. The outputs of BPF neural network is the sum of the product of Gaussian activation function and the output layer weights. Sigmoid activation function is used to calculate the output values. If the output value is not equal to the target assumed the error is found. The input and output weights are then updated using suitable weight correction formula until the desired output is reached.

### 3 Algorithm

Step 1: Initialize the expected target output  $T_1, T_2$  and take inputs  $x_1 = 0, x_2 = 1$ .

Step 2: Construct the pentagonal fuzzy numbers  $P_1, P_2$  which are the weights of the input layer and  $H_1, H_2$  the weights of the output layer.

Step 3: Apply the Gaussian activation function to the network.

Step 4: Determine the output function using Definition 2.2.

Step 5: Apply the sigmoid activation function to the output layer. If the calculated output is not equal to the expected output, then go to the next step.

Step 6: Find the error using error formula and update the weights.

Step 7: Repeat the Steps from Step 3 to Step 5. Stop the process when the calculated output is equal to the expected output. Otherwise repeat the process.

### 4 Application

Rice blast is one of the most destructive diseases of rice. Blast symptoms can occur on leaves, nodes, neck and on the panicles. Leaf spots are typically elliptical (eye shaped), with brown to red-brown margins and gray-white centers Fig. 2. The most serious damage occurs when the fungus attacks nodes at the base of earhead Fig. 3. The stems often break at the diseased node. This stage of the disease is referred to as “rotten neck”. The rotten neck prevents the flow of water and nutrients to the grains which stops developing. The grains of infected plants may become completely chaffy Fig. 4, depending on the stage of head developed and the infection (Fig. 5).

**Fig. 2** Leaf spot



**Fig. 3** Fungus attacked at earheads



**Fig. 4** Rotten neck



**Fig. 5** *Pyricularia oryzae*

### 4.1 Example

Let  $x_1, x_2$  denote two rice fields affected by rice blast pathogens.  $P_1, P_2$  are two pentagonal fuzzy numbers denoting five natural products viz., sheep urine, goat urine, hen litter, goat dung and cow dung which are liquefied and sprayed to control the pathogens affecting the two rice fields, at different concentrations. The effectiveness of the application of fungicides are recorded in the hidden layer.  $H_1$  and  $H_2$  are two pentagonal fuzzy numbers denoting five different types of fungicides viz., Azoxystrobin, Carbendazim, Flutriafol, Mancozeb and Thiophanate methyl sprayed to the two rice fields, at different concentrations. The effectiveness of the application of natural products and fungicides sprayed against rice blast disease is found. If the result obtained does not reach the expected output then proceed to the next step. The error is calculated and the quantity of fungicides and natural products are updated till the expected output (reduced incidence of pathogens) is obtained.

**Step 1:** Assume the expected target value  $T_1 = 0.1$  and  $T_2 = 0.15$  and inputs  $x_1 = 0$  and  $x_2 = 1$ .

**Step 2:** Construct the pentagonal fuzzy number which are input layer weights,

$$P_1 = (0.1, 0.15, 0.23, 0.29, 0.35)$$

$$P_2 = (0.06, 0.13, 0.28, 0.33, 0.4)$$

Output layer weights,

$$H_1 = (0.12, 0.2, 0.2, 0.31, 0.5)^T$$

$$H_2 = (0.11, 0.18, 0.25, 0.29, 0.48)^T$$

**Step 3:** Apply the Gaussian activation function to the input layer weights.

$$m_1 = \frac{1}{5}[1.12] = 0.224.$$

$$\sigma_1 = [(0 - 0.224) + (1 - 0.224)] = 0.552 \quad \sigma_1^2 = 3047.$$

$$\Phi_{11} = \exp\left(\frac{-(p_{11} - m_1)^2}{\sigma_1^2}\right) = \exp\left(\frac{-(0.1 - 0.224)^2}{0.3047}\right) = 0.9508.$$

Similarly,

$$\Phi_{12} = 0.9822, \quad \Phi_{13} = 0.9998, \quad \Phi_{14} = 0.9858, \quad \Phi_{15} = 0.9492.$$

$$m_2 = \frac{1}{5}[1.2] = 0.24.$$

$$\sigma_2 = [(0 - 0.24) + (1 - 0.24)] = 0.52 \quad \sigma_2^2 = 2704.$$

$$\Phi_{21} = \exp\left(\frac{-(p_{21} - m_2)^2}{\sigma_2^2}\right) = \exp\left(\frac{-(0.11 - 0.24)^2}{0.2704}\right) = 0.8871.$$

Similarly,

$$\Phi_{22} = 0.9562, \Phi_{23} = 0.9941, \Phi_{24} = 0.9705, \Phi_{25} = 0.9097.$$

**Step 4:** Determine the output function

$$y_i = \sum_{i=1}^m h_{ij} \Phi_{ij}$$

$$y_1 = h_{11} \Phi_{11} + h_{21} \Phi_{12} + h_{31} \Phi_{13} + h_{41} \Phi_{14} + h_{51} \Phi_{15} = 1.2907.$$

$$y_2 = h_{12} \Phi_{21} + h_{22} \Phi_{22} + h_{32} \Phi_{23} + h_{42} \Phi_{24} + h_{52} \Phi_{25} = 1.2363.$$

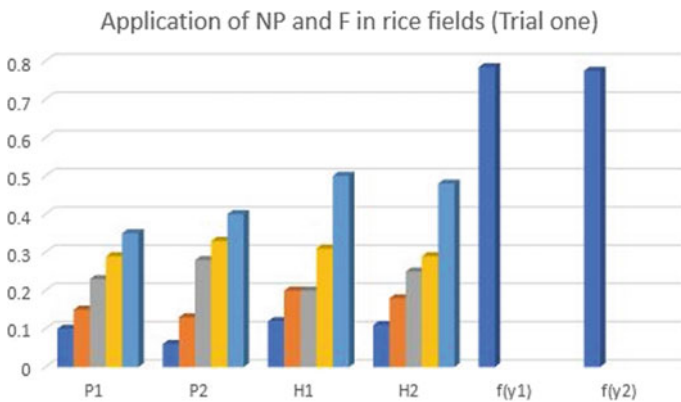
**Step 5:** Apply the sigmoid activation function to the output layer.

$$f(y_i) = \frac{1}{1 + \exp(-y_i)}, \quad f(y_1) = \frac{1}{1 + \exp(-y_1)} = 0.7843, \quad f(y_2) = 0.7749.$$

If the output is not equal to the target value, then go to the next step (Fig. 6).

**Step 6:** To find error,

$$E = \frac{1}{2}(T_1 - f(y_1))^2 + \frac{1}{2}(T_2 - f(y_2))^2 = 0.8588.$$



**Fig. 6** Trial one

Choose the learning rate

$$\delta = 0.5.$$

For output unit, error

$$e_{O_1} = -0.1158, \quad e_{O_2} = -0.109.$$

$$\delta_1 = \delta e_{O_1} = -0.05789$$

and

$$\delta_2 = \delta e_{O_2} = -0.0545$$

$$\delta_1 f(y_1) = -0.0454$$

$$\delta_2 f(y_2) = -0.0422$$

Update the output layer weights,

$$h_{11}^1 = h_{11}(\text{odd}) - \delta e_{O_1} f(y_1) = 0.0746.$$

Similarly,

$$h_{21}^1 = 0.1546, \quad h_{31}^1 = 0.1546, \quad h_{41}^1 = 0.2646, \quad h_{51}^1 = 0.4546.$$

$$h_{12}^1 = h_{12}(\text{odd}) - \delta e_{O_2} f(y_2) = 0.0678.$$

Similarly,

$$h_{22}^1 = 0.1378, \quad h_{32}^1 = 0.2078, \quad h_{42}^1 = 0.2478, \quad h_{52}^1 = 0.4378.$$

For input unit,

Error of  $P_1$

$$e_{11} = 0.0056, \quad e_{12} = 0.0035, \quad e_{13} = 0, \quad e_{14} = 0.0043, \quad e_{15} = 0.0241.$$

Error of  $P_2$

$$e_{21} = 0.0110, e_{22} = 0.0075, e_{23} = 0.0015, e_{24} = 0.0083, e_{25} = 0.0394.$$

Update the input layer weights,

$$p_{11}^1 = p_{11}(\text{odd}) - \delta e_{11} \Phi_{11} = 0.1027.$$

Similarly,

$$p_{12}^1 = 0.1517, p_{13}^1 = 0.2300, p_{14}^1 = 0.2921, p_{15}^1 = 0.3614.$$

$$p_{21}^1 = 0.0648, p_{22}^1 = 0.1336, p_{23}^1 = 0.2807, p_{24}^1 = 0.3340, p_{25}^1 = 0.4179.$$

Using the updated weights, repeat the process from Step 3 to Step 5.

Apply the Gaussian activation function,

$$m_1 = 0.2276, \sigma_1^2 = 0.2968.$$

$$\Phi_{11} = 0.9488, \Phi_{12} = 0.9808, \Phi_{13} = 0.9999, \Phi_{14} = 0.9861, \Phi_{15} = 0.9414.$$

$$\Phi_{21} = 0.8801, \Phi_{22} = 0.9519, \Phi_{23} = 0.9954, \Phi_{24} = 0.9705, \Phi_{25} = 0.8918.$$

Determine the output function  $y_1 = 1.0659$  and  $y_2 = 1.0285$ .

Apply the sigmoid function,  $f(y_1) = 0.7438$  and  $f(y_2) = 0.7366$ .

Using the updated weight, the output is 0.7438 and 0.7366 which is not equal to the expected output. Repeating the process (using MATLAB coding)  $n = 149$  times, using the updated weight, the outputs are  $y_1 = 0.1$  and  $y_2 = 0.15$ , which is the expected output (Figs. 7 and 8).

## 5 Conclusion

Two rice fields are considered. Five natural products taken as input layer weights at different concentrations are sprayed to control the rice blast pathogens affecting the two fields. Later five fungicides at different concentrations are sprayed to both of the fields. The effectiveness of the application of NP and F at the end of first trial is 0.7843 and 0.7749 which is the level of incidence of the disease. This level of incidence is not assumed target value. Hence back propagation is carried out in which the error is calculated and the quantity of natural products and fungicides to be applied are revised. The level of incidence of rice blast pathogens in the two fields after the application of revised weights are 0.7438 and 0.7366 which is still not a minimum. Hence this process is repeated 149 times using MATLAB coding until a

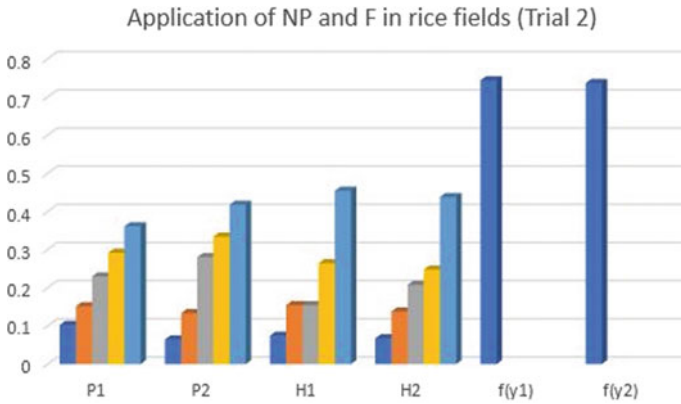


Fig. 7 Trial two

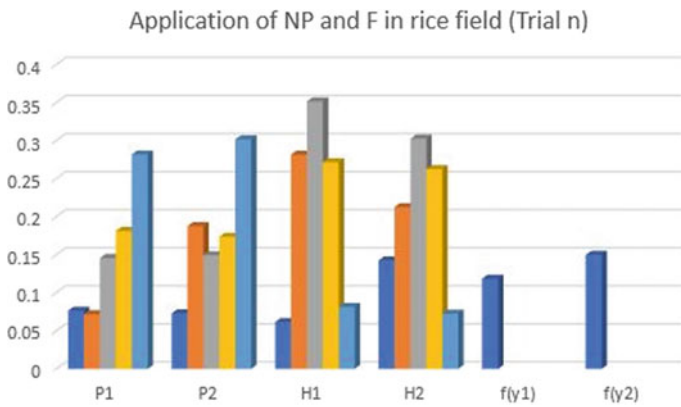


Fig. 8 Trial n

minimum level of incidence of pathogens 0.1 and 0.15 in each field is reached. It is also found that after a certain stage spraying of more natural products and less of fungicides helps in minimizing the disease rate and hence enabling the healthy growth of plants.

### References

1. Li J, Cheng J-H, Shi J-Y, Huang F (2012) Brief introduction of back propagation neural network algorithm and its improvement. *Adv Comput Sci Inform Eng* 2:553–558
2. Ishibuchi H, Fujioka R, Tanaka H (1992) An architecture of neural network for input vectors fuzzy numbers. In: *Proceedings of fuzzy IEEE' 92*, pp 1293–1300
3. Scheibel F, Steele NC, Low R (1999) Centre and variance selection for Gaussian radial basis



- function artificial neural networks. Artificial neural nets and genetic algorithms. In: Proceedings of the international conference in Portoroz, Slovenia, pp 141–147
4. Jin-Wei W, Jia-Li Z, Si-wei L, Zhen H (2000) The improvements of BP neural network learning algorithm. In: Proceedings of international conference on signals processing, vol 3, pp 1647–1649
  5. Lee SC Fuzzy neural network, mathematical biosciences. 23:151–177
  6. Dijkman JG, Van Haeringen H, de Lange SJ (1983) Fuzzy numbers. *J Math Anal Appl* 92:301–341
  7. Ishibuchi H, Morioka K, Tanaka H (1994) A fuzzy neural network with trapezoidal fuzzy weights. In: Proceedings of 1994 IEEE, 3rd international fuzzy systems conference, pp 228–233
  8. Requena B, Blanco A, Delgado M, Verdegay JL (1995) A decision personal index of fuzzy numbers based on neural network. *Fuzzy Sets Syst* 73:185–199
  9. Pathinathan T, Ponnivalavan K (2014) Pentagonal fuzzy numbers. *Int J Comput Algorithm Integr Intell Res (IIR)* 3:1003–1005
  10. Virgin Raj A, Nandhini V (2018) Application of pentagonal and triangular fuzzy numbers in neural network [BP algorithm]. *Int J Curr Adv Res* 7:168–172
  11. Anita Shanthi S, Sathiyapriya G (2020) Application of trapezoidal fuzzy number in radial basis function of neural network. *Adv Math Sci J* 9:1851–1858
  12. Sivanandam SN, Sumathi S, Deepa SN (2007) Introduction to fuzzy logic using MATLAB. Springer, Berlin Heidelberg
  13. Zhou K (1992) K: neural network models and MATLAB simulation program design. Tsinghua University Press, Beijing
  14. Kongcharoen N, Kaewsalong N, Dathoup T (2020) Efficacy of fungicides in controlling rice blast and dirty particle diseases in Thailand. *Sci Rep*
  15. Miah G, Rafil Y, Ismail R, Sahebil M, Hashemi FSG, Yusuff O, Usman MG (2017) Blast diseases intimidation towards rice cultivation: a review of pathogen and strategies to control. *J Animal Plant Sci* 27(4):1058–1066

# Comparison of Machine Learning Algorithms for Bearing Failures Classification and Prediction



Yasser N. Aldeoes, Prasad Ghockle, and Shilpa Y. Sondkar

## 1 Introduction

The maintenance cost of rotary machine can be reduced significantly if health status of roller bearings can be diagnosed at early stage. The health status of roller bearings can be monitored by observing two parameters: (a) vibration and (b) speed of rotation. The vibration and speed of rotation can be observed through IoT sensors deployed for rotary machine. The data obtained from these sensors can be analyzed by employing machine learning methods which can classify the bearing failures in four common classes, namely: (a) bearing health conditions (HC), (b) inner race fault (IF), outer race fault (OF), and (d) ball bearing fault (BF) [1]. Vibration signals monitoring and analysis are used to predict and diagnose bearings. Vibration analysis provides early information about faults for future monitoring purposes [2]. Almost 40% of the bearing failure is because of continuous stress which causes wear and tear that begins with a small pit or spalls and produces vibration. So by analyzing these vibration signals, we can detect bearing fault in the early stages [3]. Furthermore, the condition monitoring and fault diagnosis one of the main difficulties in maintenance of wind turbine blade bearings because the rotation speeds of blade bearings are very slow, and the fault signals are very weak [4]. Artificial neural networks (ANNs) [5], support vector machines (SVMs) [6], K-nearest neighbor (KNN), random forest (RF) [7]. All these algorithms are applied in the same field. Machine learning (ML)

---

Y. N. Aldeoes (✉) · P. Ghockle  
Department of Computer Engineering, Vishwakarma University, Pune, India  
e-mail: [yasser.aldeoes-043@vupune.ac.in](mailto:yasser.aldeoes-043@vupune.ac.in)

P. Ghockle  
e-mail: [prasad.gokhale@vupune.ac.in](mailto:prasad.gokhale@vupune.ac.in)

S. Y. Sondkar  
Department of Instrumentation Engineering, Vishwakarma Institute of Technology, Pune, India  
e-mail: [shilpa.sondkar@vit.edu](mailto:shilpa.sondkar@vit.edu)

algorithms are used for various purposes like data mining, image processing, predictive analytics [8], etc. Based on what has been mentioned above, the purpose and objective of this current journal paper are to use a machine learning algorithm classification learner apps for comparison, or more specifically, to compare and predict the bearing capacity while decreasing and increasing the speed and knowing which of the models gives the highest error accuracy. To this end, we first collected the bearing datasets in four cases including healthy bearing and bearing with inner race fault and bearing with outer race fault and a bearing with ball fault from the experimental site. Meantime, receiving operating characteristic (ROC) curves are used to estimate the classification comparison accuracy of the models. The rest of the paper is organized as follows: Sect. 2 presents the literature review discussing bearing failures of different machine learning algorithms, Sect. 3 presents the methodology used including bearing diagnostics and data description, Sect. 4 presents the comparison results of the classifier machine learning algorithms, and Sect. 5 presents the conclusion.

## 2 Literature Review

In an environment of manufactures and industry no machine without vibration, the vibrations signal indicates there is a failure of part of the machine or the machine itself, procedure and classify the type of the failure are identified by vibration analysis. Vibration analysis is the most common method for determining the type and severity of any faults in components of machines (like bearings and gears), as well as any maintenance options related to the machine [9]. However, the detection of early fault and classification of the health condition of bearings remains as complicated topics across various industries. Documented classification methods require correct selection to implement. So, Yanfei et al. presented an innovative diagnosis model using the complementary ensemble empirical mode decomposition (CEEMD) with kernel support vector machines (kernel SVMs) to evaluate the health condition of bearings in terms of defect severity [10, 11]. Te Han et al. describe the compare some machine learning algorithms like random forest, artificial neural networks, and support vector machine for the intelligent diagnosis of rotating machinery, and the results indicate that random forest outperforms in terms of recognition accuracy, stability, and robustness to features on the comparative classifiers [12]. Present a comprehensive review of AI algorithms including the following methods:  $k$ -nearest neighbor, Naive Bayes, support vector machine, artificial neural network, and deep learning in rotating machinery fault diagnosis. Then, AI algorithms in industrial applications [13] proposed new method 1D-CNN-based VAF algorithm to bearing fault diagnosis classify for different bearing faults based on the vibroacoustic signals collected by an accelerometer and a microphone, and the experiment results show higher accuracies of the proposed method in fault classification for different SNRs in bearing fault diagnosis [14]. Proposed deep neural networks to realize rotating speed fluctuation problems and classification and obtain a higher accuracy than other methods by the

advantages of deep learning, an intelligent fault diagnosis [15]. The work implementation artificial neural network model for rotor oscillations on non-linear bearing supports for the multistage centrifugal compressor and comparison of the bearing stiffness use of ANN and physical experiment accuracy for practical purposes [16]. ML can identify the location and size of the bearing failure using different classification approaches such as ANN, SVM, and logistic regression for the condition monitoring of the cylindrical roller bearing [17]. The researchers are describes and comparison the fault classification for vibration signals measured in roller bearings and gearboxes by SVM, the results show that both algorithm's accuracies are as high as 99.3% for the gearbox dataset and 100% for the roller bearings [18]. Proposed method using support vector machine algorithm for early detection and classification of bearing faults and clarify the training SVM using time and frequency domain to bearing fault detection and classification [19]. The author refers to a review of the various most commonly used and perfect machine learning algorithms. Highlight the advantages and disadvantages of machine learning algorithms from an application perspective to assist in decision-making and how to implement specific application requirements and choose the appropriate learning algorithm for them [20]. Singh et al. use support vector machine to classify the bearing faults of induction motor such as ball, cage, and outer race faults showed the accuracy of 100% for a ball and 85.7% for cage and 100 for outer race [21, 22]. Develop approach based on convolutional neural networks, and a random forest classifier for bearing fault detection from the data itself that got results accuracy of 93.61% and the latter an accuracy of 87.25% [23]. Proposed a new strategy-based support vector machines (SVMs) to reduce noise effect in bearing fault diagnosis systems. In this paper, the experiments indicate to employed the bearing toolbox (BEAT) to reduce the error rates [24].

### 3 Methodology

The goal of this paper is to compare different type of machine learning algorithms to identify which is the most suitable algorithm to classify the bearing failure into four types: (a) bearing health conditions (HC), (b) inner race fault (IF), outer race fault (OF), and (d) ball bearing fault (BF). We have considered dataset of "bearing vibration data under time-varying rotational speed conditions" contributed by Huan Huang, Natalie Baddour from Department of Mechanical Engineering University of Ottawa [25]. Authors (Huan Huang, Natalie Baddour) had acquired the data by deploying *NIUSB – 6212BNC* data acquisition boards with accelerometer measuring vibration data and encoder measuring rotation speed data.

### 3.1 Bearing Diagnostics

Figure 1 shows the front and side view of a typical roller bearing which consists of four components: (1) inner race, (2) balls, (3) cage, and (4) outer race. The part with failure generates vibrations in a particular frequency band which can be used to classify the bearing failure into four types. The four probable bearing failing frequencies are (1) ball pass frequency outer race (BPFO) generated when rolling elements roll across a defect in the outer ring, (2) ball pass frequency inner race (BPFI) frequency generated when rolling elements roll across a defect in the inner ring, (3) fundamental train frequency (FTF) of the cage, and (4) ball spin frequency (BSF) circular frequency of each rolling element as it spins. They have the following definitions:

Ball pass frequency, inner race:

$$BPFI = \frac{nf_r}{2} \left\{ 1 - \frac{d}{D} \cos \alpha \right\} \tag{1}$$

Ball pass frequency, outer race:

$$BPFO = \frac{nf_r}{2} \left\{ 1 + \frac{d}{D} \cos \alpha \right\} \tag{2}$$

Fundamental train frequency (cage speed):

$$FTF = \frac{f_r}{2} \left\{ 1 - \frac{d}{D} \cos \alpha \right\} \tag{3}$$

Ball (roller) spin frequency:

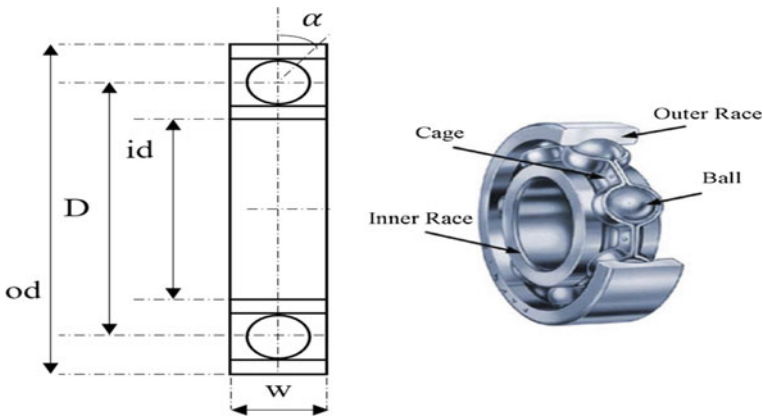


Fig. 1 Bearing faults types in front and side view

$$BSF = \frac{Df_r}{d} \left\{ 1 - \left( \frac{d}{D} \cos \alpha \right)^2 \right\} \tag{4}$$

$D$  is pitch diameter,  $\alpha$  is contact angle between the ball and the cage,  $d$  ball diameter,  $f_r$  rotating speed of bearing (Hz),  $n$  number of rolling. Vibration signal analysis is the used for fault diagnosis in rotating machines are bearings, gears, etc. The failures can happen in any element of the rolling bearing in above components. Almost bearing failures are recognized as cracks in the outer race, inner race with a rough surface's, and corrosion in balls [26]. Analysis of vibration data is often time and frequency-dependent. These analyzes are useful for detecting and determining trends in system vibration levels for early diagnosis of bearing faults which helps in understanding and monitoring them before the catastrophic errors of the bearing occur, so the previously mentioned equations contribute and help in discovering and diagnosing the bearing failures [27]. It is feasible to avoid malfunctions and save routine maintenance expenses by employing sensors to determine when equipment verification is required. Control of industrial equipment is carried out remotely and in real-time thanks to embedded sensors connected to the Internet. In this situation, for detecting the expected faults, the location of vibration sensors placed closer to the bearing is very important based on the rotational speed of the machines. The machine learning algorithms can be applied for automatic detection and classification of faults.

### 3.2 Dataset

Four different rolling bearing datasets are used in this paper to comparison the five suggested method's (decision trees, SVM, KNN, ensemble, ANN) accuracy and advantage. Table 1 contains the details on the four datasets. This is an open experiment bearing dataset given by University of Ottawa, Ottawa, Ontario, Canada [28]. The motor running at variable speed, on experiment two bearings right side was a healthy bearing, and the left side was unhealthy. The accelerometer sensor fixed on left side bearing to collect vibration signals and to measure rotating speed used tachometer to validate the results. In the dataset, there are four fault conditions, i.e., bearing health conditions (HC), faulty (inner race fault (IF)), faulty (outer race fault (OF)),

**Table 1** Data descriptions

Dataset	Fault conditions	Speed	Number of samples
1	HC	Increase from 15.4 to 24.8 Hz, then decrease to 19.1 Hz	300
2	IF	Increase from 14.8 to 21.7 Hz, then decrease to 13.6 Hz	300
3	OF	Increasing from 13.3 to 26.3 Hz	300
4	BF	Increasing from 14.0 to 24.5 Hz	100

and bearing with ball fault (BF). Moreover, each condition includes 2,000,000 data points, which are organized into 300 samples at last. The data are all sampled at 200 kHz, and the sampling duration is 10 s.

## 4 Results

In our experiments, the classification learner application in MATLAB R2021a used to compare five types of machine learning to determine the best method for bearing failures. Both vibration data and speed data, the signals are sampled at 200 kHz, and the sampling time duration is 10 s. A total sampling of a healthy bearing and bearing with inner race fault and bearing with outer race fault used is 300 samples while the total sampling for a bearing with ball fault is 100 samples.

### 4.1 Comparison Results of the Classifier Machine Learning Algorithms

The classification performance was compared by machine learning algorithms without using the PCA to determine the best method. From the results, we can find there is five widely algorithms used: decision trees, SVM, KNN, ensemble, ANN. The results of comparing the classification using 19 machine learning algorithms are shown in show in Tables 2, 3, 4, and 5.

Firstly, in Table 1 that classifier for healthy bearing, we can see the highest classification level of accuracy is 85.3% achieved by fine K-nearest neighbor (KNN) (weighted KNN) with time of 1.4 s followed by fine Gaussian SVM with 81.3% with time of 1.2 s, compared with decision tree and ensemble and neural Network, and they have low level of accuracy and higher time.

Table 3 shows the comparison of the classification results for a bearing with inner race fault, all algorithms maintained good performance in accuracy, the highest accuracy is 93.2% (attained by fine Gaussian SVM, fine KNN) with low time followed by ensemble (boosted trees) accuracy is 92.2%. The lowest accuracy of SVM is achieved by linear SVM is 64% while the lowest accuracy for KNN is 68%. We note from the table that the results of SVM and KNN obtained nearly the same results, and also decision trees and neural network obtained nearly the same results the highest is 83.1% attained by (fine-medium tree, medium – wide NN).

In Table 4, we can noticed from the best results obtained for a bearing with outer race fault by applying classification learner apps, the best classifier is K-nearest neighbor KNN, the highest accuracy is 92% achieved by weighted KNN with time 1.5 s followed by SVM (linear SVM), and neural network (trilayered NN) accuracy is 89.3% with time 2 s.

**Table 2** Comparison results of classification learner apps for a healthy bearing

Classifiers	Classifier type	Accuracy (%) (validation)	Cost validation	Prediction speed (Obj/s)	Training time
Decision trees	Fine tree	73.3	20	6600	10.094
	Medium tree	70.7	22	5300	1.9452
	Coarse tree	72.0	21	3600	1.2995
SVM	Linear SVM	64.0	27	7000	1.8682
	Quadratic SVM	72.0	21	7500	3.6004
	Cubic SVM	69.3	23	5300	23.806
	Fine Gaussian SVM	81.3	14	10,000	1.2653
KNN	Fine KNN	78.7	16	4000	1.2838
	Medium KNN	69.3	23	3700	1.5221
	Coarse KNN	61.3	29	3400	1.463
	Weighted KNN	85.3	11	2100	1.4098
Ensemble	Boosted trees	70.7	22	910	3.088
	Bagged trees	76.0	18	760	3.4689
	RUS boosted trees	72.0	21	1000	3.0238
Neural network	Narrow NN	70.7	NA	7200	7.0323
	Medium NN	65.3	NA	6200	8.2529
	Wide NN	73.3	NA	3000	9.3554
	Bilayered NN	74.7	NA	3300	12.828
	Trilayered NN	66.7	NA	6200	10.956

In Table 5, we can see the comparison of classification results for a bearing with ball fault that the highest classification accuracy is 100% which was created by SVM achieved by linear SVM. It is the best result flowed by decision trees and neural network, and the accuracy is 96%. Generally, all algorithms showed the nearest performance, and all accuracy is obtained more than 90%.

From the above tables noticed the average of all values for these performance measures, as well as prediction speed and training time differ from one model to the next during the training process; some models provide very high accuracy with short training time, while others provide more than 95% final accuracy, indicating that the model has a good overall performance. To understand the characteristics of the predictions models and classifying faults results for the training dataset for all bearings status, the faults of all bearing conditions are plotted by using prediction-validation in a two-dimension for each vibration signal and rotating speed shown in Fig. 2.

After getting the results from the machine learning algorithms for bearing faults condition, bearing health conditions, faulty (inner race fault), faulty (outer race fault),



**Table 3** Comparison results of classification learner apps for a bearing with inner race fault

Classifiers	Classifier type	Accuracy (%) (validation)	Cost validation	Prediction speed (obj/s)	Training time
Decision trees	Fine tree	83.1	13	8200	8.0012
	Medium tree	83.1	13	7700	2.0577
	Coarse tree	76.6	18	6200	1.5977
SVM	Linear SVM	64.9	27	9200	1.3422
	Quadratic SVM	78.4	16	6800	1.7239
	Cubic SVM	79.7	16	8300	6.5732
	Fine Gaussian SVM	93.2	7	6800	1.094
KNN	Fine KNN	93.2	8	4800	1.2231
	Medium KNN	68.8	24	2600	2.4637
	Cubic KNN	68.8	24	3100	1.3513
	Weighted KNN	88.3	9	4400	1.406
Ensemble	Boosted trees	92.2	6	1300	2.7451
	Bagged trees	85.7	11	950	3.8373
	RUS boosted trees	84.4	12	1300	2.5437
Neural network	Narrow NN	72.7	NA	11,000	5.8344
	Medium NN	83.1	NA	10,000	6.1773
	Wide NN	83.1	NA	6400	9.3554
	Bilayered NN	70.1	NA	12,000	6.2873
	Trilayered NN	80.1	NA	12,000	5.605

and bearing with ball fault are compared on the bar graph to display a perfect comparison. The bar graph for accuracy is shown in Fig. 3. In the bearing health condition, decision trees (D trees) obtained an accuracy of 0.73, SVM has an accuracy of 0.81, KNN has an accuracy of 0.85, ensemble has an accuracy of 0.76, and neural network (NN) has an accuracy of 0.74. In the inner race fault section, decision trees obtained an accuracy of 0.83, SVM has an accuracy of 0.93, KNN has an accuracy of 0.93, ensemble has an accuracy of 0.92, and neural network has an accuracy of 0.83. In the outer race fault section, decision trees obtained an accuracy of 0.84, SVM has an accuracy of 0.89, KNN has an accuracy of 0.92, ensemble has an accuracy of 0.84, and neural network has an accuracy of 0.89. In the ball fault condition, decision trees obtained an accuracy of 0.96, SVM has an accuracy of 1 (means 100%), KNN has an accuracy of 0.96, ensemble has an accuracy of 0.96, and neural network has an accuracy of 0.96. (Fig. 3).

**Table 4** Comparison results of classification learner apps for a bearing with outer race fault

Classifiers	Classifier type	Accuracy (%) (validation)	Cost validation	Prediction speed (obj/s)	Training time
Decision trees	Fine tree	70.7	22	1400	12.698
	Medium tree	73.3	20	7900	0.6764
	Coarse tree	84.0	12	9700	0.4918
SVM	Linear SVM	89.3	8	6700	2.0485
	Quadratic SVM	76.0	18	10,000	0.7162
	Cubic SVM	84.0	12	4800	6.1923
	Fine Gaussian SVM	85.3	11	8500	1.067
KNN	Fine KNN	86.7	10	5100	2.5392
	Medium KNN	84.0	12	4200	1.2921
	Cubic KNN	84.0	12	2800	2.5875
	Weighted KNN	92.0	6	4800	1.5359
Ensemble	Boosted trees	76.0	18	1000	9.0954
	Bagged trees	84.0	12	1000	3.6146
	RUS boosted trees	82.7	13	210	3.8974
Neural network	Narrow NN	85.3	NA	8300	9.2167
	Medium NN	82.7	NA	6700	6.6261
	Wide NN	86.7	NA	6300	11.871
	Bilayered NN	88.0	NA	9200	7.6733
	Trilayered NN	89.3	NA	6800	9.5868

## 4.2 Confusion Matrices

The confusion matrix for the testing set is also used to illustrate the classification model’s comparison performance of bearing faults. In the meanwhile, the confusion matrix also called the error matrix allows us to calculate the indices with precision, also used to further demonstrate the model’s performance, and to understand how the presently selected classifier performed in each class and helps to identify the areas where the classifier has performed poorly.

Figure 4a shows the confusion matrix for healthy bearing applied by KNN classifier precision is 92.3%. (b) shows the confusion matrix of inner race fault by SVM classifier is 94.6%. (c) shows the confusion matrix of outer race fault by KNN classifier is 90%. (d) shows the confusion matrix for a bearing with ball fault is 100% by applying ensemble classifier indicates the best results for the classification model has a very good performance.

**Table 5** Comparison results of classification learner apps for a ball fault

Classifiers	Classifier type	Accuracy (%) (validation)	Cost validation	Prediction speed (obj/s)	Training time
Decision trees	Fine tree	96.0	1	2300	9.7217
	Medium tree	96.0	1	2600	2.1439
	Coarse tree	96.0	1	2800	1.7757
SVM	Linear SVM	100.0	0	2200	2.098
	Quadratic SVM	96.0	1	2300	1.6661
	Cubic SVM	96.0	1	2700	1.6749
	Fine Gaussian SVM	80.0	5	2200	1.3982
KNN	Fine KNN	92.0	2	1300	1.218
	Medium KNN	80.0	5	1300	1.3622
	Cubic KNN	80.0	5	1100	1.5274
	Weighted KNN	96.0	1	1300	1.5247
Ensemble	Boosted trees	48.0	13	2100	1.5642
	Bagged trees	96.0	1	300	3.4039
	RUS boosted trees	48.0	13	2000	2.9424
Neural network	Narrow NN	96.0	NA	3000	1.7564
	Medium NN	96.0	NA	2400	2.5248
	Wide NN	96.0	NA	2500	2.2081
	Bilayered NN	96.0	NA	1100	1.7643
	Trilayered NN	92.0	NA	2000	2.769

### 4.3 Receiver Operating Characteristics (ROC)

Finally, the receiver operating characteristics (ROC) curve is frequently utilized to compare different classifiers to provide a clearer understanding of the model classification performances. It is easy to find the best classify performance, the area under the curve (AUC) is calculated, and a bigger area denotes a better classifier. In fact, the curve represents the relationship between the FP rate (x axis) and the TP rate (y axis). The top-left portion of the picture corresponds to the model's categorization capacity, which has a high TP rate and a low FP rate [18]. In Fig. 5, it shows the ROC curves for different models, (a) the ROC curve for healthy condition bearing shows the KNN best classifier and highest AUC = 0.91, (b) the ROC curve for a bearing with inner race fault showed the AUC = 0.93, (c) ROC curve for a bearing with outer race fault noticed the value of AUC obtained the AUC = 0.88, (d) we see through receiver operating characteristics (ROC) curve noticed for ball fault the SVM classifier superiority over ensemble and KNN, because they climb more quickly toward top left.

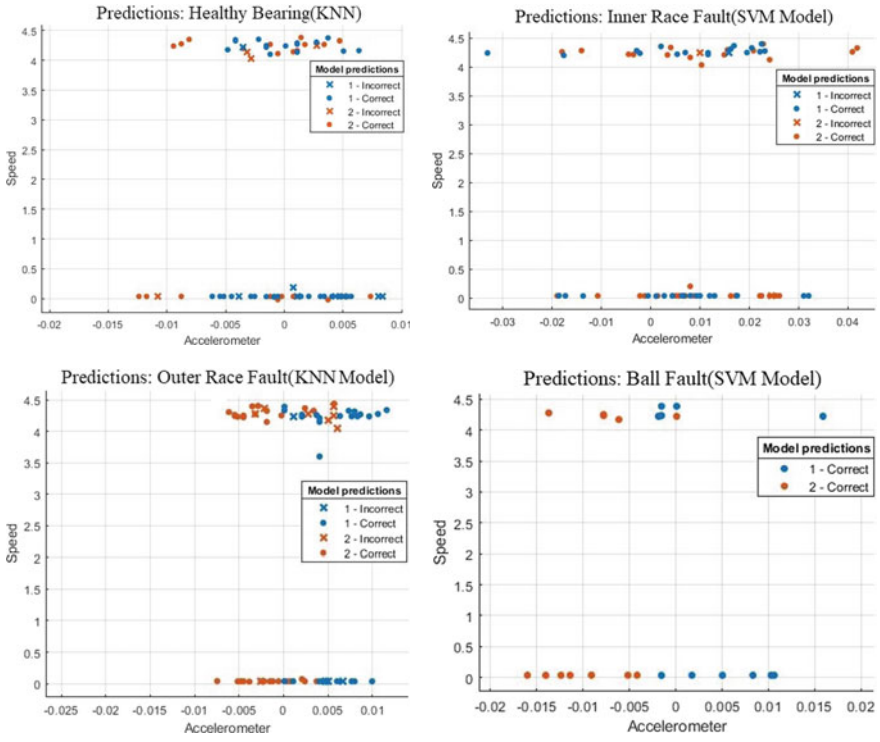


Fig. 2 Predictions of models

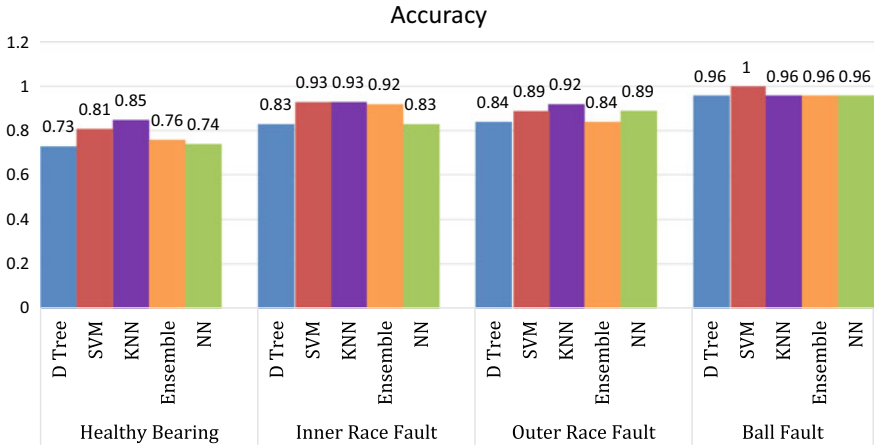


Fig. 3 Classifiers versus accuracy showing variations of bearing conditions

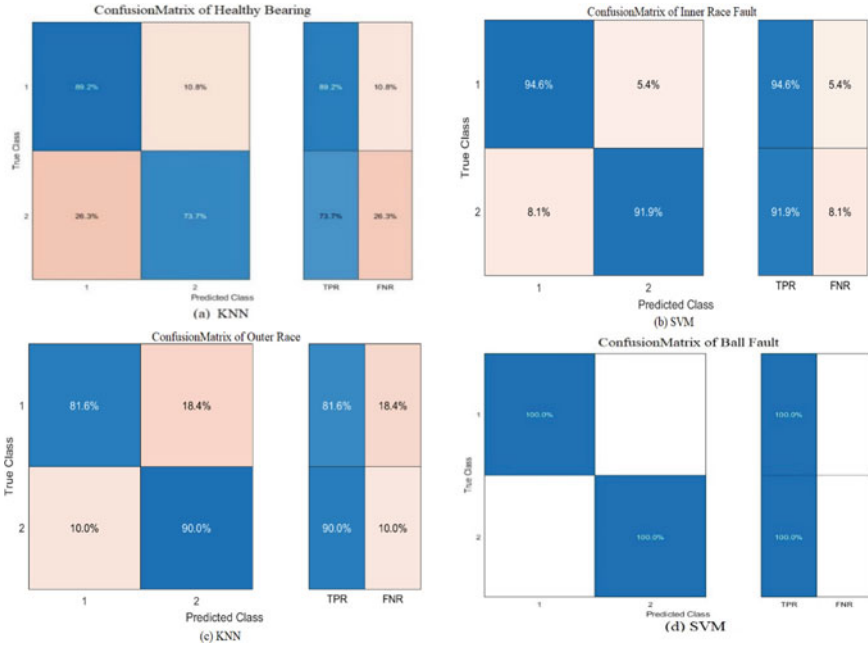


Fig. 4 Confusion matrices for different algorithms

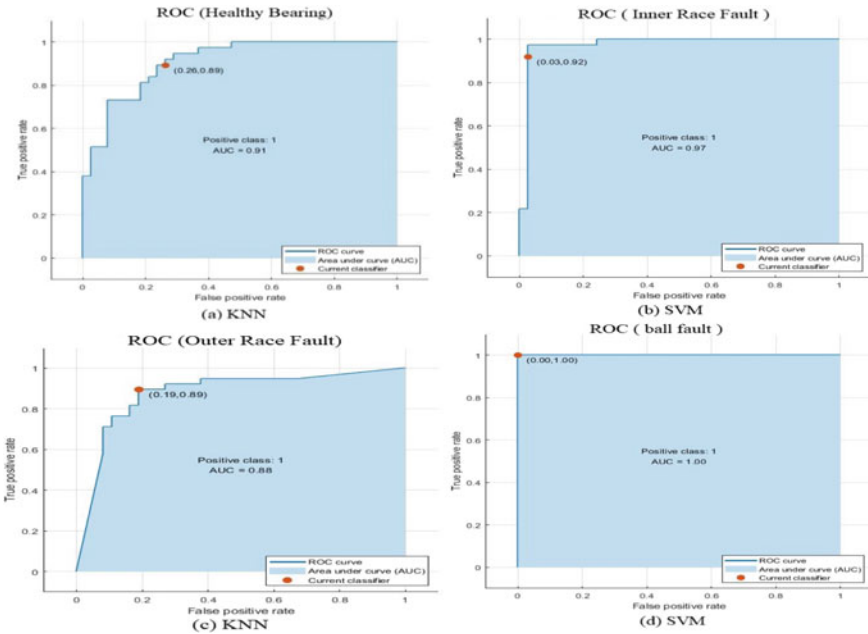


Fig. 5 ROC curves for different algorithms

SVM got the highest AUC = 1.00 with true position rate TPR = 1 and false position rate FPR = 0, while ensemble got AUC = 1 with true position rate TPR = 1 and false position rate FPR = 0.08333 and KNN got AUC = 0.83 with true position rate TPR = 0.91667 and false position rate FPR = 0.25 that is main the SVM method exactly classify the bearing with ball fault condition.

## 5 Conclusion

The current paper is comparison a bearing faults classification model based on machine learning algorithms. The paper proposed the decision trees, support vector machine, K-nearest neighbor, ensemble, neural network algorithms which describes every condition of bearing failures by giving the estimate tables. When machine learning algorithms are executed on a bearing dataset, the most important parameter is the accuracy. Hence, the result shows that K-nearest neighbor classifier with bearing healthy gave the highest accuracy of 85.3%, for SVM it is 81.3%, while the decision trees and ensemble and neural network classifiers obtained nearly the same results between 64 and 76%. The second bearing condition is a bearing with inner race fault the results (identified) showed best both KNN and SVM classifier have highest classification accuracy rates 93.2%. The third condition of bearing with outer race fault the best result achieved by KNN classifier accuracy is 92%, flowed by SVM and neural network classifiers are 89.3%. The fourth condition of bearing status a ball fault the SVM classifier attains the highest accuracy of 100%, and it may be because of small training datasets are taken. While others classifiers have been performed with great accuracy and precision rate more than 95%. Therefore, the comparison results for the bearing conditions show that SVM is the best 100% in ball fault diagnostic for accurate validation in the training but not so good in the testing while KNN shows the 96% higher efficiency in training and testing which proves that ranking feature method helps in getting the good results in training and testing. For another bearing conditions identified by the KNN classifier, the highest performance in the training validation and test compared with SVM, ensemble, neural network, and decision trees.

## References

1. Li X, Jiang H, Wang R, Niu M (2021) Rolling bearing fault diagnosis using optimal ensemble deep transfer network. *Knowl-Based Syst* 213
2. Geramitchioski T, Trajcevski L (2011) Prediction of rolling bearings defect in motor-fan using vibration signal analysis 3(3):211–216
3. Kulkarni S, Bewoor A (2016) Vibration based condition assessment of ball bearing with distributed defects. *J Meas Eng* 4(2):87–94
4. Liu Z, Wang X, Zhang L (2020) Fault diagnosis of industrial wind turbine blade bearing using acoustic emission analysis. *IEEE Trans Instrum Meas* 69(9):6630–6639

5. Abiodun OI, Jantan A, Omolara AE, Dada KV, Mohamed NAE, Arshad H (2018) State-of-the-art in artificial neural network applications: a survey. *Heliyon* 4(11)
6. Nalepa J, Kawulok M (2019) Selecting training sets for support vector machines: a review. *Artif Intell Rev* 52(2):857–900
7. Shah K, Patel H, Sanghvi D, Shah M (2020) A comparative analysis of logistic regression, random forest and KNN models for the text classification. *Augment Hum Res* 5(1)
8. Sarker IH (2021) Machine learning: algorithms, real-world applications and research directions. *SN Comput Sci* 2(3):1–21
9. Kumar S, Goyal D, Dang RK, Dhama SS, Pabla BS (2018) Condition based maintenance of bearings and gears for fault detection-a review. *Mater Today Proc* 5(2):6128–6137
10. Lu Y, Xie R, Liang SY (2020) CEEMD-assisted kernel support vector machines for bearing diagnosis. *Int J Adv Manuf Technol* 106(7–8):3063–3070
11. Goyal D, Choudhary A, Pabla BS, Dhama SS (2020) Support vector machines based non-contact fault diagnosis system for bearings. *J Intell Manuf* 31(5):1275–1289
12. Han TJ, Zhao D, Wang Q, Lei Yin, K (2017) Comparison of random forest , artificial neural networks and support vector machine for intelligent diagnosis of rotating machinery. *Trans Inst Measur Contro*
13. Liu R, Yang B, Zio E, Chen X (2018) Artificial intelligence for fault diagnosis of rotating machinery: a review. *Mech Syst Sig Process* 108:33–47
14. Wang X, Mao D, Li X (2021) Bearing fault diagnosis based on vibro-acoustic data fusion and 1D-CNN network. *Meas J Int Meas Confed* 173(September)
15. Han B, Ji S, Wang J, Bao H, Jiang X (2021) An intelligent diagnosis framework for roller bearing fault under speed fluctuation condition. *Neurocomputing* 420:171–180
16. Pavlenko I, Simonovskiy V, Ivanov V, Zajac J, Pitel J (2019) Application of artificial neural network for identification of bearing stiffness characteristics in rotor dynamics analysis. Springer International Publishing vol 1
17. Parmar U, Pandya DH (2021) Comparison of the supervised machine learning techniques using WPT for the fault diagnosis of cylindrical roller bearing 13(2):50–56
18. Medina R, Macancela JC, Lucero P, Cabrera D, Sánchez RV, Cerrada M (2020) Gear and bearing fault classification under different load and speed by using Poincaré plot features and SVM. *J Intell Manuf*
19. Senanayaka JSL, Kandukuri ST, Van Khang H, Robbersmyr KG (2017) Early detection and classification of bearing faults using support vector machine algorithm. In: *Proceeding of 2017 IEEE Work Electric Machines and Drives Control Diagnosis, WEMDCD 2017*, pp 250–255
20. Ray S (2019) A quick review of machine learning algorithms. In: *Proceedings of the International Conference on Machine Learning, Big Data, Cloud and Parallel Computing Trends, Perspectives Prospects COMITCON 2019*, pp 35–39
21. Singh M, Shaik AG (2019) Faulty bearing detection, classification and location in a three-phase induction motor based on stockwell transform and support vector machine. *Meas J Int Meas Confed* 131:524–533
22. Glowacz A (2019) Fault diagnosis of single-phase induction motor based on acoustic signals. *Mech Syst Sig Proc* 117:65–80
23. Janssens O et al (2016) Convolutional neural network based fault detection for rotating machinery. *J Sound Vib* 377:331–345
24. Batista L, Badri B, Sabourin R, Thomas M (2013) A classifier fusion system for bearing fault diagnosis. *Expert Syst Appl* 40(17):6788–6797
25. Huang H, Baddour N (2018) Bearing vibration data collected under time-varying rotational speed conditions. *Data Br* 21(November):1745–1749
26. Kankar PK, Sharma SC, Harsha SP (2011) Fault diagnosis of ball bearings using machine learning methods. *Expert Syst Appl* 38(3):1876–1886
27. Wei Y, Li Y, Xu M, Huang W (2019) A review of early fault diagnosis approaches and their applications in rotating machinery, pp 1–26

# A Parallelized Algorithm for Finding 3-D Mesh Illumination



Varun M. Khachane, V. Harshitha, Parthivi Khullar, S. V. Rishikesh, C. H. Mani Kumar, and K. V. Vineetha

## 1 Introduction

Many applications such as computer graphics, radar cross-section analysis and reconstruction of 3-D geometric models of objects use triangular meshes extensively [1–4]. Triangular meshes are the surface construction of an object using only triangles, which are connected to each other by common edges or corners [4]. Due to its simplicity, flexible structure, good boundary adaptability and efficiency, triangular meshes have advantages over other polygon meshes [5, 6]. A mesh of an object can consist of tens of thousands of faces (triangles) and performing operations on the mesh means operating on every face. One such operation is finding the portion of the mesh which is illuminated by an electromagnetic point source. This operation can be used while creating games, radar simulation and in reconstruction of 3-D models with respect to light source [6, 7].

When an algorithm is built for determining illumination of 3-D triangular mesh, it should satisfy all the boundary conditions. It is important that the algorithm should produce proper results even if the mesh contains convexity or concavity. Also, the algorithm should be efficient in terms of both computational speed and memory requirements. Using techniques such as multithreading and multiprocessing enables every face (triangle) of the mesh to be processed in parallel [8–10]. Machine learning techniques can be used for determining the illumination of the meshes. However, these techniques increase the computational time [11–14].

This paper illustrates a method of obtaining the illuminated and non-illuminated portion of the mesh. First the information of every face (triangle) is stored in a data structure and the surface normal of those faces is found. An electromagnetic plane

---

V. M. Khachane · V. Harshitha · P. Khullar · S. V. Rishikesh · C. H. Mani Kumar · K. V. Vineetha (✉)

Computer Science and Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Bengaluru, India  
e-mail: [jain\\_vineetha@blr.amrita.edu](mailto:jain_vineetha@blr.amrita.edu)



wave source is specified, which is a point source with 3-D co-ordinates. The angle between the surface normal of the face and the ray from the wave source is calculated, which incidents on the centroid of the face (triangle). The proposed method processes every face independently, and the order in which the faces are processed does not affect the output, so it is convenient to parallelize the process. The list containing the information of all the faces is divided into equal parts and assign each part to different processes. These processes run in parallel resulting in decreasing the execution time.

## 2 Methodology

To perform any operation on the triangular mesh every face of the mesh needs to be operated. The data of every face (triangle) in a mesh, such as the co-ordinates of vertices, is stored in mesh files. Mesh files have different file formats (.obj, .stl, .ply, 0.3ds) and every file format has its own way of storing the data. Python file functions are used to extract the data. In this proposed technique .obj file format is used and in this format every line is an independent set of data and is differentiated by using abbreviations, for example 'v' is used for vertices and f' is used for faces in the beginning of every line. The vertices are stored in the  $x, y, z$  format and the faces (triangles) have the respective index of the vertices and it is stored in a counter-clockwise fashion.

The following three steps are performed to obtain the illuminated portion of the mesh:

- i. Extracting the vertices and faces from the mesh file.
- ii. Finding the surface normal of the face.
- iii. Finding the angle between the incident ray and surface normal.

Python file functions are used to extract the vertices and face data from the .obj file format and store them in separate lists. Using the vertices of the faces (triangles) the side vectors of the face,  $\bar{V}$  and  $\bar{W}$  are found.

$$\bar{V} = \bar{a} - \bar{b} \quad (1)$$

$$\bar{W} = \bar{b} - \bar{c} \quad (2)$$

where  $\bar{V}$  and  $\bar{W}$  are the side vectors and  $\bar{a}, \bar{b}, \bar{c}$  are the vertices of the face (triangle) arranged in a counter-clockwise fashion.

Normal of the triangle is a line which is perpendicular to the plane in which the triangle lies. To get the normal vector  $\bar{N}$  the cross product of the two side vectors  $\bar{V}$  and  $\bar{W}$  are found.

$$N_x = V_y W_z - V_z W_y \quad (3)$$

$$N_y = V_z W_x - V_x W_z \quad (4)$$

$$N_z = V_x W_y - V_y W_x \quad (5)$$

Here  $\overline{V} = V_x \hat{x} + V_y \hat{y} + V_z \hat{z}$  and  $\overline{W} = W_x \hat{x} + W_y \hat{y} + W_z \hat{z}$  are the side vectors of the face (triangle) and  $\overline{N} = N_x \hat{x} + N_y \hat{y} + N_z \hat{z}$  is the normal to the face of the mesh.

The normal will always face outwards as the vertices of the face (triangle) are stored in counter-clockwise fashion.

For any object to be visible the light rays from the source should reflect from the surface of the object into the eyes and for this to happen the angle between the incident ray and the reflected ray should be less than  $180^\circ$ . Using the same concept here, for the mesh object to be illuminated by the electromagnetic plane wave source, the angle between the incident ray and the normal of the face (triangle) should be less than or equal to  $90^\circ$ .

In the algorithm, the incident ray is considered as the vector between the centroid of the face (triangle) and the wave source. If the angle made by the normal and the incident ray is less than  $90^\circ$  that face falls in the visible region. This operation is performed with all the faces to obtain the illuminated portion of the mesh. To calculate the angle between the vectors, the dot product of the two vectors,  $\overline{N}$  and  $\overline{I}$  are found.

$$\overline{N} \cdot \overline{I} = N_x I_x + N_y I_y + N_z I_z \quad (6)$$

Here  $\overline{N} = N_x \hat{x} + N_y \hat{y} + N_z \hat{z}$  is the normal vector of the face and  $\overline{I} = I_x \hat{x} + I_y \hat{y} + I_z \hat{z}$  is the incident ray vector.

When there are tens of thousands of faces (triangles) in a mesh, iteratively processing all the faces one by one is computationally costly. To decrease the computation time, parallel processing techniques are used. The faces are stored in a list data structure. In sequential execution, the entire list is passed to one process, whereas in parallel processing the list is divided into equal parts and one part of the list is passed to each thread/process. Now the code runs in parallel, and each thread/process has to operate on smaller number of faces. The output from all threads/processes are stored in two common lists. One list contains the data of all the faces in the illuminated surface while the other list contains data of all the faces in the non-illuminated surface.

### 3 Results

To test the feasibility and accuracy of the proposed technique, the algorithm is tested with various input meshes with complex shapes which include both concave and

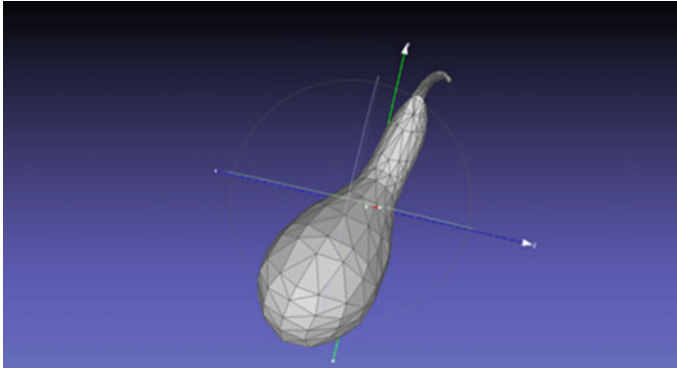
convex surfaces. The algorithm successfully finds illuminated and non-illuminated portions of the mesh.

Figure 1 shows the triangular mesh of an object which has both convex and concave surfaces. This mesh is used as our input mesh because of its complex structure and because it tests all the boundary conditions.

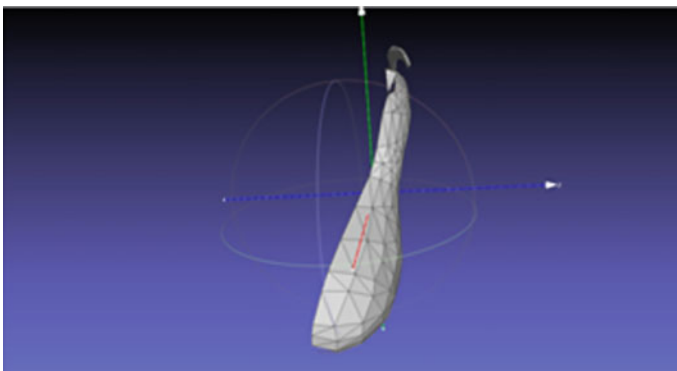
Figure 2. is the illuminated portion of the mesh. The wave source is placed at the  $(0, 0, 10)$  co-ordinates and the algorithm operates on the object mesh to find the illuminated portion. The result is obtained as shown in Fig. 2.

Figure 3 illustrates the non-illuminated part of the object mesh. The wave source is placed at the  $(0, 0, 10)$  co-ordinates and the algorithm operates on the object mesh to find the non-illuminated portion. The result is obtained as shown in Fig. 3.

After successfully getting the output for object in Fig. 1, the algorithm is operated on various meshes to get runtimes of different optimization techniques. The runtimes of sequential, multithreading, and multiprocessing execution are calculated and compared.



**Fig. 1** Triangular mesh of a complex object



**Fig. 2** Illuminated mesh of gourd object

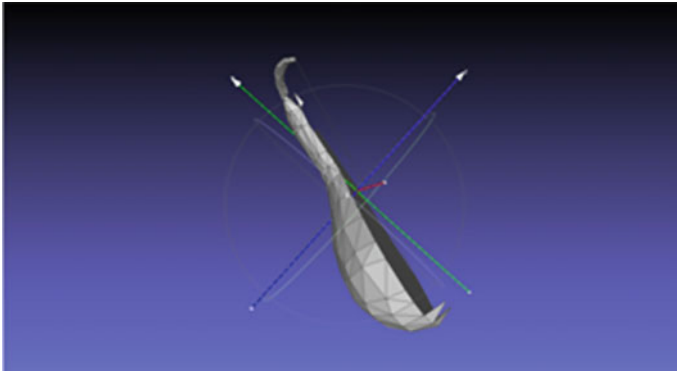


Fig. 3 Non-illuminated mesh of gourd object

Figure 4 represents Time versus Number of threads for meshes with different number of faces that are executed using multithreading. It is observed that multithreading does not help reduce the execution time of the algorithm. With every new thread there is an overhead time. As threads run concurrently, they only increase the execution time. As the number of threads increase, it is observed that the run time of the algorithm also increases, and this is not a feasible option. Hence, a better solution is required to reduce the execution time of the algorithm. Therefore, multiprocessing libraries are chosen to analyze if they can help reduce the execution time of the algorithm.

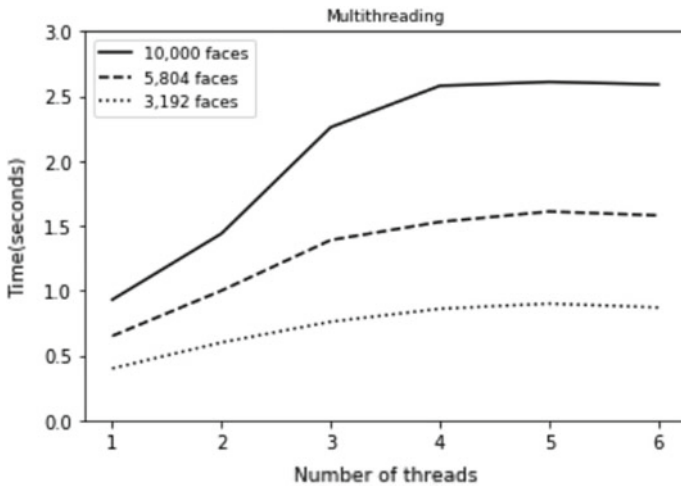


Fig. 4 Execution time of different meshes (multithreading line graph)

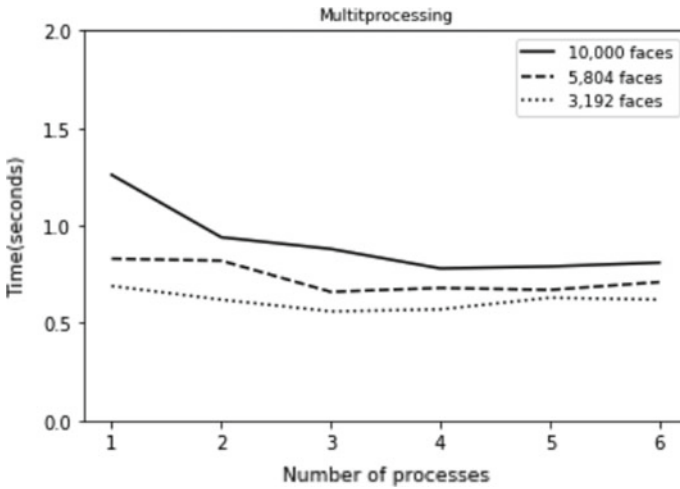


Fig. 5 Execution time of different meshes (multiprocessing line graph)

Figure 5 represents Time versus Number of processes for meshes with different number of faces that are executed using multiprocessing. It is observed that in multiprocessing, with an increase in the number of processes, the execution time decreases. There is an overhead with every new process invoked but unlike the threads, processes run in parallel, so the execution time reduces and compensates for the overhead.

To check the efficiency of the algorithms, a comparison of three different meshes and the execution time of these different optimization techniques has been performed.

Figure 6 represents Time versus Number of threads/processes for an object mesh with 3192 faces (triangles), executed with different optimization techniques.

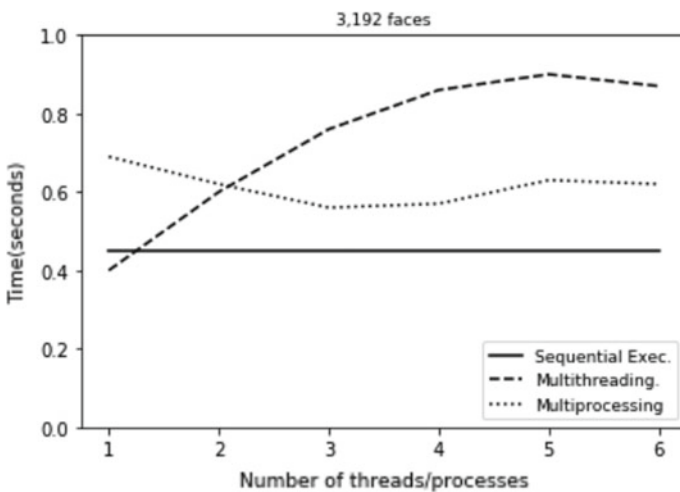


Fig. 6 Execution time comparison with different execution techniques (3192 faces)

Figure 7 represents Time versus Number of threads/processes for an object mesh with 5804 faces (triangles), executed with different optimization techniques.

Figure 8 represents Time versus Number of threads/processes for an object mesh with 10,000 faces (triangles), executed with different optimization techniques.

Figure 9 represents the Face versus Time graph of sequential execution and multiprocessing. Three different meshes with an increasing number of faces (triangles) were executed and compared. In multiprocessing the instance of the code with a

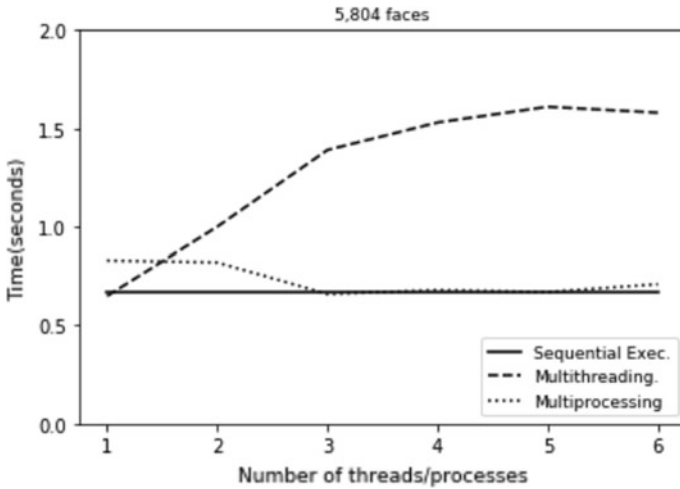


Fig. 7 Execution time comparison with different execution techniques (5804 faces)

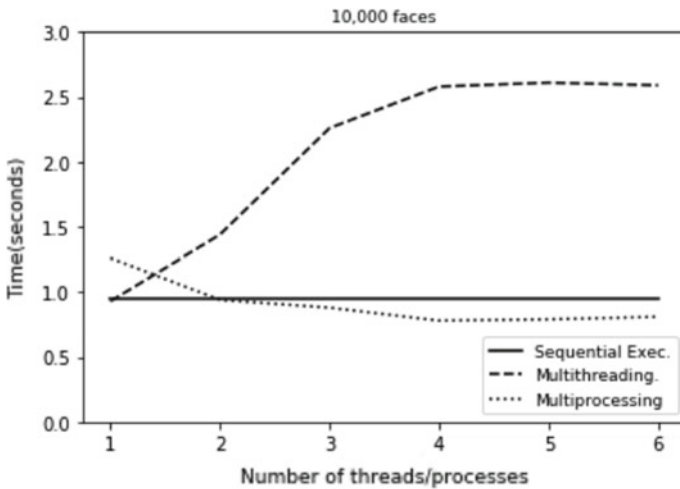
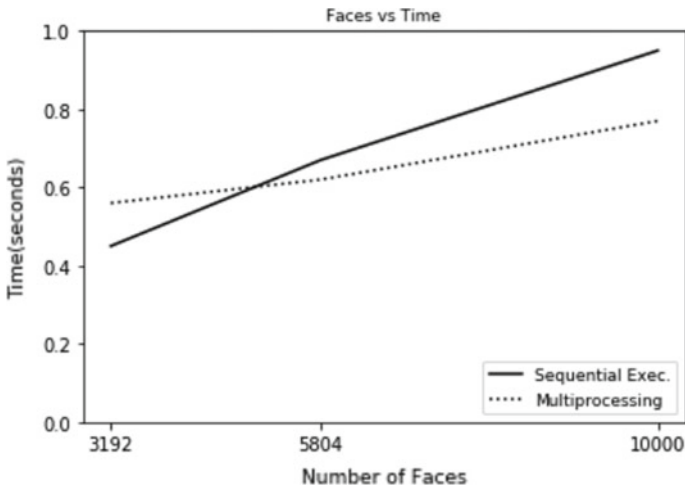


Fig. 8 Execution time comparison with different execution techniques (10,000 faces)



**Fig. 9** Execution time comparison of different meshes using sequential execution and multiprocessing

different number of processes is executed. Hence, the instance with least execution time is considered.

From the above graphs it is observed that when the faces are less in number, multiprocessing takes more execution time and sequential takes less execution time. As the number of faces increases, the execution time between the sequential and multiprocessing execution reduces. This is because multithreading and multiprocessing executions have an extra overhead time of invoking a new thread or process, respectively. So, with every new thread or process invocation, extra overhead is added into the execution time. As the number of faces increases, the reduced execution time of multiprocessing compensates for the overhead time and helps in reducing the execution time.

The execution of the algorithm for finding the illuminated portion of a 3-D mesh with multiprocessing gives a faster way to process the object mesh and get the illuminated portion.

## 4 Conclusion

In multithreading and multiprocessing there is a function invocation that is assigned to a thread or process, respectively. Both threads and processes when invoked add an overhead to the execution time, thereby increasing the total execution time. When the number of faces (triangles) in a mesh is less, the overhead only increases the execution time, which makes the sequential execution of code the better option. When a mesh has a large number of faces, sequential execution takes a lot of time to

process all the faces. So, by dividing the list containing all the faces and executing them parallelly, not only does the execution time reduce but also compensates for the overhead time of the process invocation.

In this paper, a technique to find the portion of a 3-D mesh that is illuminated and not illuminated by the electromagnetic plane wave source is successfully achieved. It is feasible to run the code in serial execution when the object mesh consists of fewer faces, and parallel execution of the code when the object mesh consists of many faces.

## References

1. Sun J, Jiang Y, Jiang J, Bai X (2016) Triangular mesh construction based on point cloud matrix and edge feature extraction. In: 2016 IEEE information technology, networking, electronic and automation control conference, pp 275–279
2. Liu G, Huo J, Xu L, Shi X, Yang X, Chen Z (2019) RCS analysis of UAV groups. In: 2019 international conference on microwave and millimeter wave technology (ICMMT), pp 1–2
3. Aiswarya G, Valsaraj N, Vaishak M, Nair TU, Karthik V, Nair JJ (2017) Content-based 3D image retrieval using point cloud library a novel approach for the retrieval of 3D images. In: 2017 international conference on communication and signal processing (ICCSP), pp 0817–0820
4. Peng Y, Gao C (2009) Influencing factors on vertex normal accuracy for triangular meshes. In: 2009 international conference on intelligent human-machine systems and cybernetics, pp 347–350
5. Sakr NM, Youssef BAB, Hassan YF, Atta EH (2018) An effective method for hole filling in 3D triangular meshes. In: 2018 IEEE international symposium on signal processing and information technology (ISSPIT), pp 1–7
6. Ng K, Low Z (2014) Simplification of 3D triangular mesh for level of detail computation. In: 2014 11th international conference on computer graphics, imaging and visualization, pp 11–16
7. Zheng Z (2012) Simplification algorithm for general meshes. In: 2012 fifth international joint conference on computational sciences and optimization, pp 53–56
8. Fernando E, Murad DF, Wijanarko BD (2018) Classification and advantages parallel computing in process computation: a systematic literature review. In: 2018 international conference on computing, engineering, and design (ICCED), pp 143–147
9. Liu J, Wu Y, Marsaglia J (2012) Making learning parallel processing interesting. In: 2012 IEEE 26th international parallel and distributed processing symposium workshops and PhD forum, pp 1307–1310
10. Guntha R, Hariharan B, Rangan PV (2016) Multi-agent framework for automatic deployment and state restoration in iterative software development process. In: 2016 international conference on advances in computing, communications and informatics (ICACCI), pp 1833–1838
11. Purushothaman A, Vineetha KV, Kurup DG (2018) Fall detection system using Artificial Neural Network. In: 2018 second international conference on inventive communication and computational technologies (ICICCT). IEEE, pp 1146–1149
12. Dev A, Jain V (2021) Identifying phished website using multilayer perceptron. In: Tripathy A, Sarkar M, Sahoo J, Li KC, Chinara S (eds) Advances in distributed computing and machine learning. Lecture notes in networks and systems, vol 127. Springer, Singapore
13. Vineetha KV, Kurup DG (2017, September) Direct demodulator for amplitude modulated signals using artificial neural network. In: international symposium on intelligent systems technologies and applications. Springer, Cham, pp 204–211



14. Bharadwaja V, Ananmy R, Nikhil S, Vineetha KV, Shah J, Kurup DG (2018) Implementation of artificial neural network on raspberry pi for signal processing applications. In: 2018 international conference on advances in computing, communications and informatics (ICACCI), pp 1488–1491

# An Useful Survey on Supervised Machine Learning Algorithms: Comparisons and Classifications



B. Abhishek and Amit Kumar Tyagi

## 1 Introduction

These days, ML is one of the most popular fields in computer science, with a wide range of applications. It is the technique of discovering meaningful patterns in data automatically. Machine learning technologies are designed to enable algorithms to accustom [19]. Machine learning has emerged to become one of the cornerstones as a result of advancements in information technology, a very fundamental, if often overlooked, aspect of our lives. With an ever generation of information, smart data analysis is expected to become ever more common as a critical component of technological advancement.

Machine learning (ML) has a variety of applications, the most important data mining is one of them. When conducting research or seeking to draw connections between various characteristics, people are more susceptible to making errors [9, 16].

Machine learning and Data mining are considered as pairs that may be used to extract a variety of insights using the right learning methods. Because to the advancement of smart and nano-technology, data mining and machine learning have advanced significantly, piquing interest in finding underlying patterns to derive value. Whilst combining machine learning, information theory, statistics and computers has resulted in a strong mathematical basis and a set of immensely effective tools. The algorithms are categorised in to a typology which depends on projected result of the algorithm. Supervised learning is used to create the function that maps inputs to expected output. Machine learning algorithms have become increasingly sophisticated as a result of unprecedented data generation. This has necessitated the

---

B. Abhishek (✉) · A. K. Tyagi  
School of Computer Science and Engineering, Vellore Institute of Technology, Chennai, India  
e-mail: [abhishek.b2020@vitstudent.ac.in](mailto:abhishek.b2020@vitstudent.ac.in)

A. K. Tyagi  
e-mail: [amitkrtyagi025@gmail.com](mailto:amitkrtyagi025@gmail.com)

use of a variety of supervised and unsupervised machine methodologies for learning. The purpose of classification tasks is to persuade the technology to understand how to use a categorization system that we have devised [21], supervised learning used frequently.

The accessibility buried within Big Data is perfectly suited for machine learning. Because it is based on data and evolves at a machine level, ML ensures extracting significance from large and diverse data sources by removing less reliance on individual tracks. Machine learning is well adapted to the challenges of dealing with several data sources, a wide range of variables, and vast amounts of data, and ML thrives on expanding data sets. Extra data fed into a machine learning framework, the better it can be taught and the better the insights it can produce. ML is brilliant at finding and displaying hidden patterns in data because it is free of the constraints of individual level reasoning and study [15].

The classification problem is a common supervised learning challenge in which the learner must look at several input–output samples that transforms learning process, divide a vector into one of several classes. The process of learning a series of rules from instances in order to build a classifier that can generalise to new situations is known as learning by deduction (examples in a training set). Figure 1 shows a way for applying supervised machine learning to a true problem.

This study concentrated on distinguishing machine learning algorithms and determining efficient approach with the highest precision and accuracy. In addition to testing the efficacy, on big and small datasets to classify them accurately and to provide an insight into the development of supervised ML models.

Remainder of work is divided onto the following sections: Sect. 2 comprises a literature review that discusses the categorization of various supervised learning algorithms; the procedure is presented in part 3, the findings are shown in Sect. 4, and the inference and future research recommendations are presented in Sect. 5.

## 2 Literature Survey

### 2.1 Classification of Supervised Learning Algorithms

Given below are the types of SML algorithms which deals with classification, according to [21] are detailed below. As per the assertion, linear classifier achieves its goal on producing a judgment on the basis of a linear value features' combinations.

- (1) Linear Classifiers: Using linear (hyperplane) decision boundaries, linear classification models divide input vectors into classes [6]. In ML, purpose of classification is to organise things with similar feature values are clubbed together into groups [23]. As it is the quickest classifier and where speed is an issue, then linear classifier is used [21]. The convergence rate amongst data set variables and on the other hand, it is influenced by the margin. The margin is a parameter

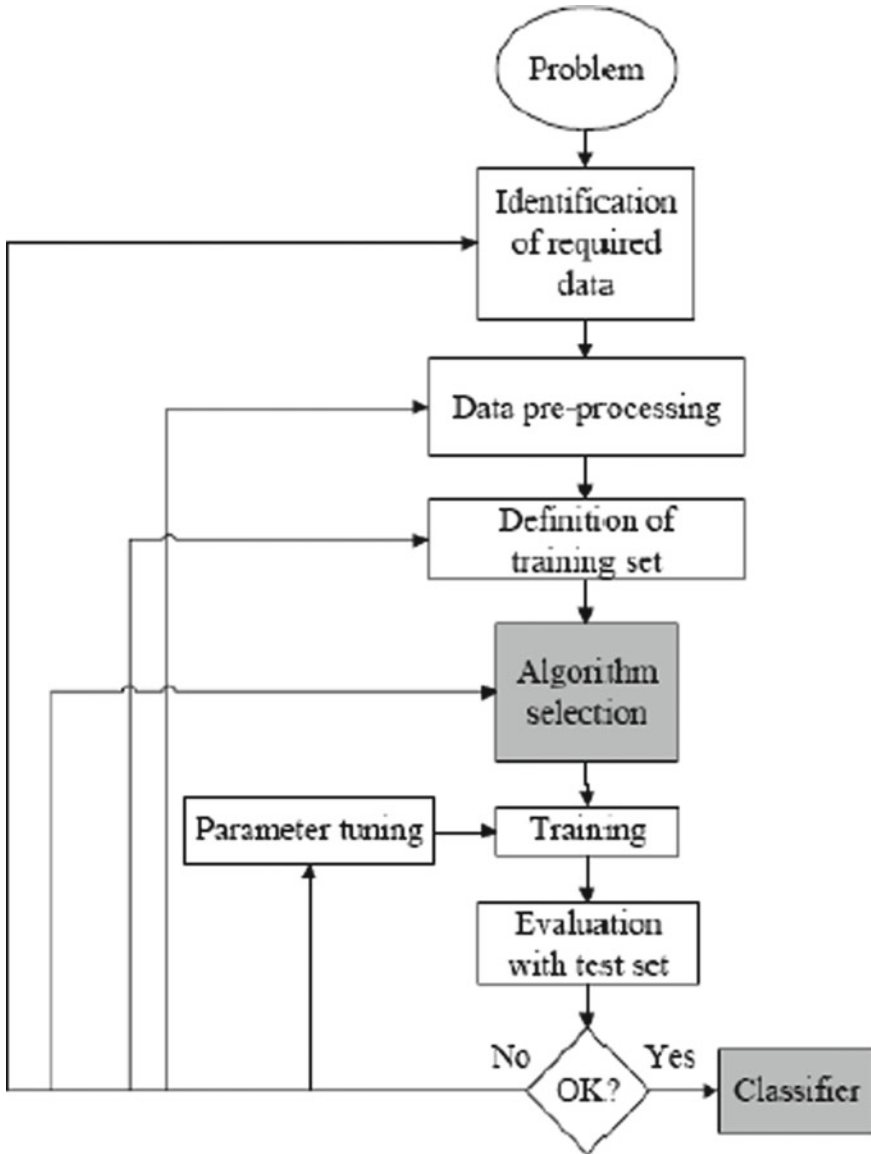


Fig. 1 Supervised machine learning mechanisms

that determines how easily a dataset can be separated linearly, and to solve a classification problem in a simple manner [18].

- (2) **Logistic regression:** A supervised learning classification approach for predicting the probability of a target variable. Location of the class boundary, as well as probabilities of class change with distance from the boundary are

commonly mentioned in a specialised approach which is called as logistic regression. Probabilistic claims set logistic regression apart from other classification algorithms. It generates more exact, detailed forecasts and may be customised in a number of ways; nevertheless, such exact predictions may turn out to be erroneous. Ordinary Least Squares (OLS) and logistic regression are two methods for predicting outcomes. The outcome of logistic regression prediction, on the other hand, is a binary result [13]. In logistic regression, the linear interpolation approach is applied [11].

- (3) **Naive Bayesian (NB) Networks:** NB is one of the basic networks and is a combination of directed acyclic graphs with only parent node and having many childrens' (representing identified nodes), using the premise that child nodes are self-contained with reference to their parents [7]. As a result, calculating probability distributions is the basis of the Naive Bayes model (independence) [14]. Bayes classifiers are generally outperformed by other, more sophisticated algorithms for learning (such as ANNs).
- (4) **Multi-layer Perceptron:** Rather of dealing with a non-convex, unconstrained minimization problem as is the case with traditional neural network training [21], the network's weights are calculated using a quadratic programming problem with linear restriction. Perceptron is [17] used in other algorithms also. It is mostly used to learn from a set of training examples by continuously executing it over them until it discovers an always correct prediction vector. The labels on the test set are then predicted using this prediction rule [9].
- (5) **Support Vector Machines (SVMs):** SVMs have recently been popular in supervised machine learning approaches [24]. Multi-layer perceptron neural networks and SVM models have many commonalities. The idea of "margin" defined as distance between two classes on either side of a hyperplane, underpins SVMs. It has been demonstrated that decreasing the predicted generalisation error by widening the margin and determining the separation between the separating hyperplane and the occurrences on either side [9].
- (6) **K-means:** For handling the well-known clustering problem, K-means is considered to be the easiest unsupervised learning algorithms, according to [2, 22]. The approach uses a simple and uncomplicated method for classifying a given sample based on preset no. of clusters (assuming k clusters). When labelled data is not available, the K-Means technique is used [1]. A generic method for transforming poor estimates into highly precise prediction rules.
- (7) **Decision Trees (DT):** It arranges instances on the basis of their feature values. Each node of the tree represents a feature in an instance to be classified, and each branch indicates a value that the node can adopt. This is a prediction model used in decision tree learning that converts observations about an object into judgements about its goal value. It is utilised in machine learning and data mining. Classification trees or regression trees are two terms for similar tree models [20]. After DT have been pruned using a validation set, post-pruning procedures are often utilised in decision tree classifiers to evaluate their performance. Any node in the sorted training instances can be eliminated and given the most common class [9].

- (8) **Neural Networks (NN):** Bishop [2] Several NN that can do multiple tasks such as regression and/or classification in one go, despite the fact that most networks can only perform one at a time. This network is a single output variable in the overwhelming majority of cases. The input and activation functions of an artificial neural network, as well as the network architecture and the weight of each input link, are all critical components (ANN). The ANN's behaviour is governed by the present values of the weights since the first two features are constant. The net's weights are then considerably adjusted to get the net's output values closer to the anticipated output values. Algorithms may be used to teach a network in a variety of ways [12].
- (9) **Bayesian Network:** A graphical depiction of a collection of variables' probability relationships (features). The most well-known statistical learning approaches are Bayesian networks [9]. This has a drawback of not being suited for datasets with many features [4]. Prior understanding of a Bayesian network's topology, also known as domain knowledge, which could form/node in the following manner:
1. Designating as a root, i.e. one with no children.
  2. Defining to be a leaf, i.e. one without children.
  3. Designating as the direct cause or effect.
  4. Declare that it is not directly connected to one another.
  5. Determining the independence of two nodes based on a set of criteria.
  6. Declare that a node in the ordering happens before another node by supplying partial node ordering.
  7. Creating a comprehensive node sorting strategy.

## 2.2 Features of Machine Learning (ML) Algorithms

ML approaches that are supervised can be used in a variety of fields. Authors [18, 25] include a number of ML application-oriented studies.

When dealing with multi-dimensions and continuous data, neural networks and SVMs perform substantially good. Whilst dealing discrete/categorical characteristics, logic-based systems, in other hand, tends to project in a superior manner. To achieve optimal prediction accuracy, neural network models and SVMs require a sample in larger size, whereas NB may just require a small dataset.

The fact that KNN is particularly sensitive to irrelevant information is well-known; this property can be explained by the algorithm's design. Furthermore, non-essential elements may render neural network training useless, if not impossible. When it comes to problems that need diagonal partitioning, most decision tree methods fall short. The instance space is divided orthogonally to one variable's axis and parallel to the other axes. As a result, after partitioning, all of the regions that arise are hyperrectangles.

Naive Bayes (NB) utilises very minimal storage capacity throughout the RAM required to retain the prior and conditional probabilities in both training stage as well

as in classification stage which is absolute minimum. The fundamental KNN method requires a lot of storage during the training phase, and the execution space is at least as big as the training region. Final classifier, on the other hand, is generally a very condensed description of the data, execution space is frequently significantly lower than training space for all non-lazy learners. In addition, whereas rule algorithms cannot be utilised as incremental learners, Naive Bayes and the KNN can. Missing values are not taken into account whilst calculating probabilities, as a result, in Naive Bayes, they have no impact on the final choice. Neural networks and KNN, on the other hand, function only with full records.

Finally, the operational profiles of Decision Trees and NB vary; whilst one is quite precise, whilst the other is not, and vice versa. The operational profiles of SVM and ANN are likewise comparable. Over all datasets, no single learning system can consistently outperform others. Varied data sets of different sorts of variables and quantities of occurrences influence the type of method which performs exceptionally good. According to the no-free-lunch hypothesis, on all data sets, no single learning system will outshine another [10]. The comparison of several learning algorithms is shown in Table 1.

### 3 Methodology

The National Institute of Diabetes and Digestive and Kidney Diseases provided the relevant data for the study. This data was chosen for its correctness, as well as the fact that it has been anonymised (de-identified), ensuring confidentiality. There are eight attributes, plus one class, for a total of nine. The following are the numeric values for all attributes:

1. How many times you have been pregnant
2. In an oral glucose tolerance test, plasma glucose concentration after 2 h
3. Blood pressure in the diastole (mm Hg)
4. The thickness of the skin folds on the triceps (mm)
5. Insulin serum ( $\mu$  U/ml) after 2 h
6. BMI (weight in kilogrammes divided by height in metres)
7. Pedigree function in diabetes
8. Your age (years)
9. Variable in the class.

Table 2 displays the overall number of cases included in this study, with 500 of them testing positive for diabetes and 268 testing negative.

WEKA 3.7.13 was used to perform a comparative examination of various supervised machine learning methods. The values 1 s in the class distribution (class variable) were changed to YES, indicating that they had been tested for DIABETES POSITIVELY, and the values 0 s to NO, indicating that they had been tested for DIABETES NEGATIVELY. Random Forest, Decision Table, neural

**Table 1** Learning algorithm contrasts (\*\*\*\*—best performance, \*—worst performance) [9]

	Decision Trees	Neural networks	Naïve Bayes	kNN	SVM	Rule-learners
Accuracy in general	**	***	*	**	****	**
Speed of learning with respect to number of attributes and the number of instances	***	*	****	****	*	**
Speed of classification	****	****	*	****	****	
Tolerance to missing values	***	*	****	*	**	**
Tolerance to irrelevant attributes	***	*	**	**	****	**
Tolerance to redundant attributes	**	**	*	**	***	**
Tolerance to highly interdependent attributes (e.g. parity problems)	**	***	*	*	***	**
Dealing with discrete/binary/continuous attributes	****	*** (not discrete)	*** (not continuous)	*** (not directly discrete)	** (not discrete)	*** (not directly continuous)
Tolerance to noise	**	**	***	*	**	*
Dealing with danger of overfitting	**	*	***	***	**	**
Attempts for incremental learning	**	***	****	****	**	*
Explanation ability/transparency of knowledge/classification	****	*	****	**	*	****
Model parameters handling	***	*	****	***	*	***

**Table 2** Class distribution: (1—“tested positive for diabetes”) and (0—“tested negative for diabetes”)

Class	Value number of instances	Converted value (attribute)
0	500	NO
1	268	YES

networks (Perceptron), SVM, Naive Bayes, JRip and Decision Tree were the seven classification methods utilised in this study (J48).

This study was able to anticipate precision and accuracy assured in terms of distinct machine learning algorithms by setting parameters by using 2 alternative sets of no. of occurrences.



## 4 Results

Various machine learning methods were classified and compared using WEKA. The results are shown in Table 3 with nine qualities and factors taken into account.

The TIME it takes to develop the model is referred to as TIME. The Mean Absolute Error (MAE) is a metric that indicates how accurate a prognosis or prediction is in terms of the actual outcome. The kappa statistic is a statistic for comparing observed and expected accuracy (Random Chance): YES indicates that you have been diagnosed with diabetes. NO indicates that a diabetes test was negative. Further, Table 4 displays the classification results using six attributes and a contrasting the various machine learning techniques and settings examined.

The TIME it takes to develop the model is referred to as TIME. The MAE is a metric that actually gives a picture that how close a forecast or prediction is to the actual result.

Kappa statistic is a measure which actually compares the observed accuracy to that of anticipated accuracy (Random Chance). YES indicates that the patient has been diagnosed with diabetes. NO indicates that a diabetes test was negative.

### 4.1 Discussions

Table 3 shows a comparison of findings of 768 samples and of 9 features. It may be seen that if compared to MAE, algorithms are having a higher kappa statistic. In addition, occurrences that have been correctly classified outnumber those that have been mistakenly classified. This indicates that predictive analysis is more reliable with larger data sets. As demonstrated in Table 3, SVM and NB require a large sample size to attain maximum prediction accuracy, whereas DT and Decision Table have the least precision.

Table 4 gives a comparative results for 384 instances and 6 attributes. When compared to MAE, neural networks, JRip, and J48 all have lower kappa statistics, which do not give precision and accuracy. It shows that JRip, J48, and neural networks perform well with smaller datasets have a far lower percentage of correctly categorised occurrences than wrongly classified ones. SVM and RF, on the other hand, demonstrate high accuracy and precision with fewer data sets. In comparison to JRip and Decision Tree, Decision Table took longer to build the model. As a result, saving time does not imply accuracy. If kappa statistic is smaller than Mean Absolute Error (MAE), algorithm will not show any precision and accuracy. As a result, algorithm with such features cannot be applied for that data set since precision and accuracy will be lacking.

Table 6 compares precision for larger and data sets that are smaller, with SVM indicating algorithm with the best prediction. Table 5 also demonstrates that SVM is the most precise algorithm. Data sets that are smaller. Tables 7 and 8 compare the percentage of successfully identified and erroneously categorised data given the

**Table 3** Shows a comparison of different categorization algorithms using a larger data set and more attributes

Algorithm	Time (s)	Correctly classified (%)	Incorrectly classified (%)	Test mode	Attributes	No of instances	Kappa statistic	MAE	Precision of YES	Precision of NO	Classification
Decision Table	0.23	72.3958	27.6042	10-fold crossvalidation	9	768	0.3752	0.341	0.619	0.771	Rules
Random Forest	0.55	74.7396	25.2604	10-fold crossvalidation	9	768	0.4313	0.3105	0.653	0.791	Trees
Naive Bayes	0.03	76.3021	23.6979	10-fold crossvalidation	9	768	0.4664	0.2841	0.678	0.802	Bayes
SVM	0.09	77.3438	22.6563	10-fold crossvalidation	9	768	0.4682	0.2266	0.740	0.785	Functions
Neural networks	0.81	75.1302	24.8698	10-fold crossvalidation	9	768	0.4445	0.2938	0.653	0.799	Functions
JRip	0.19	74.4792	25.5208	10-fold crossvalidation	9	768	0.4171	0.3461	0.659	0.780	Rules
Decision Tree (J48)	0.14	73.8281	26.1719	10-fold crossvalidation	9	768	0.4164	0.3158	0.632	0.790	Tree

**Table 4** Shows a comparison of different categorization algorithms using a smaller data set and fewer attributes

Algorithm	Time	Correctly classified %	Incorrectly classified %	Test mode	Attributes	No of instances	Kappa statistic	MAE	Precision of YES	Precision of NO	Classification
Decision Table	0.09	67.9688	32.0313	10-fold crossvalidation	6	384	0.3748	0.3101	0.581	0.734	Rules
Random Forest	0.42	71.875	28.125	10-fold crossvalidation	6	384	0.3917	0.3438	0.639	0.761	Trees
Naïve Bayes	0.01	70.5729	29.4271	10-fold crossvalidation	6	364	0.352	0.3297	0.633	0.739	Bayes
SVM	0.04	72.9167	27.0833	10-fold crossvalidation	6	384	0.3837	0.2708	0.711	0.735	Functions
Neural networks (Perceptron)	0.17	<b>59</b>	<b>41</b>	10-fold crossvalidation	6	384	<b>0.1156</b>	<b>0.4035</b>	0.444	0.672	Functions
JRip	0.01	<b>64</b>	<b>36</b>	10-fold crossvalidation	6	384	<b>0.2278</b>	<b>0.4179</b>	0.514	0.714	Rules
Decision Tree (J48)	0.03	<b>64%</b>	<b>36</b>	10-fold crossvalidation	6	384	<b>0.1822</b>	<b>0.4165</b>	0.519	0.685	Tree

**Table 5** Using smaller and larger data sets, illustrate how different algorithms order the precision of positive and negative diabetes

Smaller dataset 384		
Algorithm	Precision of YES (positive diabetes)	Precision of NO (negative diabetes)
SVM	0.711	0.735
Random Forest	0.639	0.761
Naïve Bayes	0.633	0.739
Decision Table	0.581	0.734
Decision Tree (J48)	0.519	0.685
JRip	0.514	0.714
Neural networks (Perceptron)	0.444	0.672

time it took to build the model for smaller and larger datasets. Table 7 reveals that Naive Bayes and JRip are the algorithms with the fastest time to build, however JRip has a lower percentage of correctly classified samples, indicating the time it took to create as a model is not synonymous with accuracy. Similarly, at a time of 0.04 s, SVM has the best level of accuracy. When compared to Table 8, the third correctly categorised algorithm was neural networks (perceptron). This indicates that a neural network works better with a large dataset than it does with a small one. In addition, the findings show that Decision Table struggles with huge datasets.

Table 9 depicts mean and standard deviation for qualities studied, with concentration of Plasma glucose (attribute 2) having the highest mean and Diabetes pedigree function (attribute 7) having the lowest mean, showing a significant effect on a small data set.

In the last, in [26–41] authors have recommended authors to read these research efforts, to know more information about the role of AI, Computer Vision, or Machine

**Table 6** Using smaller and larger data sets, illustrate how different algorithms order the precision of positive and negative diabetes

Large data set 768		
Algorithm	Precision of YES (positive diabetes)	Precision of NO (negative diabetes)
SVM	0.74	0.785
Naïve Bayes	0.678	0.802
JRip	0.659	0.78
Random Forest	0.653	0.791
Neural networks (Perceptron)	0.653	0.799
Decision Tree (J48)	0.632	0.79
Decision Table	0.619	0.771

**Table 7** Show rankings, i.e. correctly classified and erroneously classified data sets based on the time it took to develop these model using various techniques for smaller and larger data sets, respectively

Smaller dataset 384			
Algorithm	Time (s)	Correctly classified (%)	Incorrectly classified (%)
SVM	0.04	72.92	27.08
Random Forest	0.42	71.88	28.13
Naïve Bayes	0.01	70.57	29.43
Decision Table	0.09	67.97	32.03
JRip	0.01	<b>64</b>	<b>36</b>
Decision Tree (J48)	0.03	<b>64</b>	<b>36</b>
Neural networks (Perceptron)	0.17	<b>59</b>	<b>41</b>

**Table 8** Show rankings, i.e. correctly classified and erroneously classified data sets based on the time it took to develop the model using various techniques for smaller and larger data sets, respectively

Large data set 768			
Algorithm	Time (s)	Correctly classified (%)	Incorrectly classified (%)
SVM	0.09	77.34	22.66
Naïve Bayes	0.03	76.30	23.70
Neural networks (Perceptron)	0.81	75.13	24.87
Random Forest	0.55	74.74	25.26
JRip	0.19	74.48	25.52
Decision Tree (J48)	0.14	73.83	26.17
Decision Table	0.23	72.40	27.60

**Table 9** A descriptive analysis of different dataset attributes

Attribute number	Mean	Standard deviation
1	3.8	3.4
2	120.9	32.0
3	69.1	19.4
4	20.5	16.0
5	79.8	115.2
6	32.0	7.9
7	0.5	0.3
8	33.2	11.8

learning techniques respect to these sensitive areas/useful applications like health-care, agriculture, etc. We hope that readers/researchers will find suitable problem for themselves from these research work.

## 5 Conclusion and Future Scope

The parameters for ML classification must be fine-tuned thoroughly, and the data collection must have a large number of occurrences. It takes time to build the algorithm's model, but it also requires accuracy and precise classification. As a result, the best learning approach for one set of data could not work for another set of data with logically different features. In machine learning classification, the most essential question is not whether one learning algorithm is better than another, but rather under what conditions one technique may surpass another on a specific application job. In this direction, meta-learning is moving, with researchers attempting to discover functions that link datasets and algorithm performance [12]. Meta-learning does this by representing the features of learning issues with a collection of traits called as meta-qualities, with the objective of discovering links between these traits and the effectiveness of learning algorithms. The number of examples, the proportion of categorical characteristics, the fraction of missing data, and the entropy of class are all used to construct learning tasks.

Brazdil et al. [3] given a dataset with a long list of data and statistical measurements. After gaining a deeper grasp of each method's strengths and limitations, the option of combining to solve an issue, two or more algorithms should be investigated. The idea is to balance off the flaws of one method with the advantages of another. SVM, NB, and RF machine learning algorithms provide great precision and accuracy regardless of the number of characteristics and data examples. Machine learning algorithms must be precise, accurate, and have a low error rate in order to be used for supervised predictive machine learning.

Large data quantities should be processed in a distributed processing environment, according to this research. This will result in a higher level of correlation between the variables, resulting in a better model output.

## References

1. Alex S, Vishwanathan SVN (2008) Introduction to machine learning. Published by the press syndicate of the University of Cambridge, Cambridge, United Kingdom. Cambridge University Press 2008. ISBN 0-521-82583-0
2. Amit Kumar Tyagi, Poonam Chahal, "Artificial Intelligence and Machine Learning Algorithms", Book: Challenges and Applications for Implementing Machine Learning in Computer Vision, IGI Global, 2020.DOI: 10.4018/978-1-7998-0182-5.ch008
3. Bishop CM (1995) Neural networks for pattern recognition. Clarendon Press, Oxford, England; Oxford University Press, Inc. New York, NY, USA ©1995 ISBN 0198538642

4. Brazdil P, Soares C, da Costa J (2003) Ranking learning algorithms: using IBL and meta-learning on accuracy and time results. *Mach Learn* 50(3):251–277. Copyright ©Kluwer Academic Publishers. Manufactured in The Netherlands. <https://doi.org/10.1023/A:1021713901879>
5. Cheng J, Greiner R, Kelly J, Bell D, Liu W (2002) Learning Bayesian networks from data: an information-theory based approach. *Artif Intell* 137:43–90. Copyright © 2002. Published by Elsevier Science B.V. All rights reserved
6. Domingos P, Pazzani M (1997) On the optimality of the simple Bayesian classifier under zero-one loss. *Mach Learn* 29:103–130. Copyright © 1997 Kluwer Academic Publishers. Manufactured in The Netherlands
7. Elder J (n.d) Introduction to machine learning and pattern recognition. Available at LASSONDE University EECS Department York
8. Good IJ (1951) Probability and the weighing of evidence. *Philosophy* 26(97):163–164. Published by Charles Griffin and Company, London 1950. Copyright © The Royal Institute of Philosophy. <https://doi.org/10.1017/S0031819100026863>
9. Hastie T, Tibshirani R, Friedman JH (2001) The elements of statistical learning. In: Data mining, inference, and prediction. Springer, New York
10. Hormozi H, Hormozi E, Nohooji HR (2012) The classification of the applicable machine learning methods in robot manipulators. *Int J Mach Learn Comput (IJMLC)* 2(5):2012. <https://doi.org/10.7763/IJMLC.2012.V2.189pp.560-563>
11. Kanungo T, Mount DM (2002) A local search approximation algorithm for k-means clustering. In: Proceedings of the eighteenth annual symposium on Computational geometry. ACM Press, Barcelona, Spain
12. Kotsiantis SB (2007) Supervised machine learning: a review of classification techniques. *Informatica* 31:249–268. Retrieved from IJS
13. Lemnaru C (2012) Strategies for dealing with real world classification problems. Unpublished PhD thesis, Faculty of Computer Science and Automation, Universitatea Technica, Din Cluj-Napoca
14. Logistic Regression, pp 223–237. Available at <https://www.stat.cmu.edu/~cshalizi/uADA/12/lectures/ch12.pdf>
15. Neocleous C, Schizas C (2002) Artificial neural network learning: a comparative review. In: Vlahavas IP, Spyropoulos CD (eds) Methods and applications of artificial intelligence. Hellenic conference on artificial intelligence SETN 2002. Lecture notes in computer science, vol 2308. Springer, Berlin, Heidelberg, pp 300–313. [https://doi.org/10.1007/3-540-46014-4\\_27](https://doi.org/10.1007/3-540-46014-4_27)
16. Newsom I (2015) Data analysis II: logistic regression
17. Nilsson NJ (1965) Learning machines. McGraw-Hill, New York. Published in: *J IEEE Trans Inform Theor* 12(3):407–407. <https://doi.org/10.1109/TIT.1966.1053912>
18. Pradeep KR, Naveen NC (2017) A collective study of machine learning (ML) algorithms with big data analytics (BDA) for healthcare analytics (HcA). *Int J Comput Trends Technol (IJCTT)* 47(3):149–155. ISSN 2231-2803, <https://doi.org/10.14445/22312803/IJCTT-V47P121>. Available from IJCTT website <http://www.ijcttjournal.org/2017/Volume47/number-3/IJCTT-V47P121.pdf>
19. Setiono R, Loew WK (2000), FERNN: an algorithm for fast extraction of rules from neural networks. *Appl Intell*
20. Shalev-Shwartz S, Ben-David S (2014) Understanding machine learning from theory to algorithms
21. Taiwo OA (2010) Types of machine learning algorithms. In: Zhang Y (ed) New advances in machine learning. InTech, University of Portsmouth United Kingdom, pp 3–31. ISBN 978-953-307-03406
22. Tyagi, Amit Kumar and G, Rekha, Machine Learning with Big Data (March 20, 2019). Proceedings of International Conference on Sustainable Computing in Science, Technology and Management (SUSCOM), Amity University Rajasthan, Jaipur - India, February 26-28, 2019, Available at SSRN: <https://ssrn.com/abstract=3356269> or <http://dx.doi.org/10.2139/ssrn.3356269>

23. Timothy Jason Shepard PJ (1998) Decision fusion using a multi-linear classifier. In: Proceedings of the international conference on multisource-multisensor information fusion
24. Vapnik VN (1995) The nature of statistical learning theory, 2nd ed. Springer, pp 1–20. Retrieved from website <https://www.andrew.cmu.edu/user/kk3n/simplicity/vapnik2000.pdf>
25. Witten IH, Frank E (2005) Data mining: practical machine learning tools and techniques, 2nd ed. Morgan Kaufmann Publishers, San Francisco, CA, U.S.A. © 2005 Elsevier Inc. Retrieved. ISBN 0-12-088407-0



# Image Encryption and Decryption Using Chaotic Bimodal Quadratic Map



H. Soumya Babu , K. Gopakumar, and N. Vijayakumar

## 1 Introduction

In recent years, the transmission of digital images through the Internet has been increasing. The use of traditional ciphers like DES, RSA, IDEA, and AES in real-time digital image encryption is difficult and requires high computational power due to the special characteristics of digital images like strong correlation between adjacent pixels, data redundancy etc. Chaotic system with properties like sensitivity to initial conditions, ergodicity, and randomness find an immense application in such cryptographic applications. The chaos-based techniques are extensively studied in recent years, since their properties lead to the potential cryptography. The chaotic encryption systems are preferred compared to other systems, to achieve a combination of high security and speed and also ease of implementation.

## 2 Literature Review

Chaotic systems are deterministic nonlinear systems that are highly sensitive to initial conditions. It, in turn, shows a pseudo-random nature. The high sensitivity to initial conditions makes the output of the system totally different for even a small change in the input. An important characteristic that makes chaotic systems ideal for encryption

---

H. Soumya Babu (✉)  
LBS Centre for Science and Technology, Kerala, India  
e-mail: [soumiyam@gmail.com](mailto:soumiyam@gmail.com)

K. Gopakumar  
A P J Abdul Kalam Technological University, Kerala, India

N. Vijayakumar  
Department of ECE, GEC BartonHill, Kerala, India

is its pseudo-random characteristic that seems to be random from the attackers point of view, whereas definable for a decoder from the view of cipher system so as to make the decryption simpler. A few of the efforts undertaken in this area are summarized briefly. In 1989, a chaotic image algorithm was first proposed by Matthews [1]. L. Zhang et al., in 2005, proposed an image encryption approach based on chaotic maps [2]. Later, in 2006, a new image encryption method was proposed by Pareek and Patidar which used two logistic maps in the chaotic regime along with a 80-bits external key [3]. This method uses initial conditions, derived from the external secret key by providing weightage to its bits corresponding to their position in the key. A new image encryption scheme was then presented by Gao and Chen in 2006, which shuffles the pixel positions of the image by the use of an image shuffling matrix and uses the state combination of two chaotic systems to confuse the relationship between the plain and cipher image [4]. Furthermore, R.Raja Kumar et al., in 2011, proposed an encryption algorithm that used combination of Logistic and Henon maps to expand the parameters [5]. An improved image encryption algorithm based on two chaotic logistic maps along with a 80-bit secret key to derive the initial condition was presented by Purwar et al. [6]. A variety of chaotic image encryption algorithms were developed in the later years [7–10]. The recent literature review shows that chaos-based image encryption is found to be a better choice.

Logistic map [11], tent map [12], and Baker chaotic map [13] are some of the chaotic systems of low-dimension that formed the basis of the initially developed image encryption schemes. Later, the researchers used combinations and deformations of these basic chaotic maps along with some mathematical manipulations to design image encryption schemes, like the combination of 1-D tent map and logistic map [14], baker map, and logistic map [15]. The combination of maps to develop such encryption systems increased the security of the algorithms and had potential cryptographic applications.

### 3 Background

In this paper, a combination of a well-known one-dimensional and two-dimensional chaotic maps, namely the logistic map and the difference map, respectively, are used for image encryption. The sequences generated from the iterated maps, operating in the chaotic regime are used to generate the key to develop the image encryption scheme.

#### 3.1 *The Logistic Map*

The logistic map forms one of the simplest one-dimensional chaotic maps, represented by a simple dynamical equation, with complex chaotic behavior. The logistic map is mathematically defined by the following equation:

$$y_{n+1} = ky_n(1 - y_n) \quad (1)$$

where  $k$  is the bifurcating parameter in the range  $(0, 4]$ ,  $y_{n+1}$  is the next chaotic output sequence, and  $y_n$  is the present input [16, 17].

### 3.2 The Difference Map

The difference map is a bimodal quadratic map developed as an amplification of the difference between the one-dimensional logistic map and tent map. Based on the selection of the bifurcating parameter, the map exhibits a variety of behaviors [18]. Difference map,  $F_D(y, \alpha)$  is defined as the difference between the logistic map,  $F_L(y, 4)$  and tent map  $F_T(y, 2)$  multiplied by the parameter  $\alpha \in [0, 4]$ .

$$F_D(y, \alpha) = \alpha[F_L(y, 4) - F_T(y, 2)] \quad (2)$$

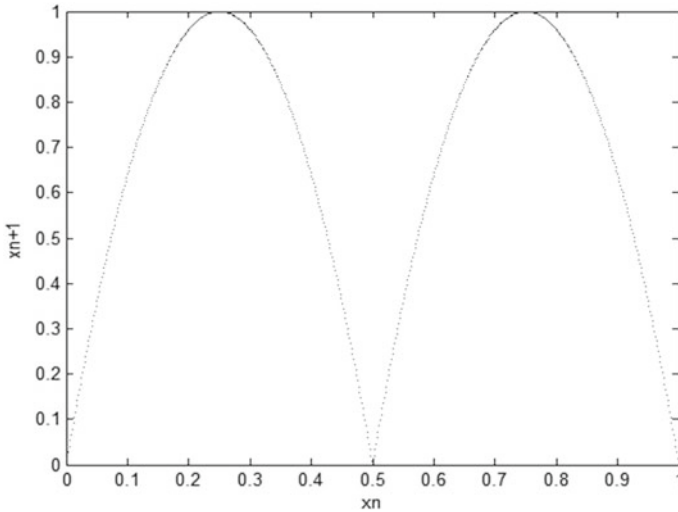
$$F_D(y, \alpha) = \begin{cases} 2\alpha y(1 - 2y) & \text{for } y < 0.5 \\ 2\alpha(y - 1)(1 - 2y) & \text{for } y \geq 0.5, \end{cases} \quad (3)$$

Depending on the value of the bifurcation parameter  $\alpha$ , the difference map behaves either as a bimodal map or a unimodal map. For some specific value of the initial condition  $y_0 \in [0, 1]$ ,  $F_D(y, \alpha)$  behaves as unimodal map after the first iteration for  $\alpha = 2$ . For specific value of bifurcating parameter  $\alpha \in [0, 2]$ , the map resembles logistic map, though it oscillates in the interval  $[0, 0.5]$ , whereas for  $\alpha \in [2, 4]$ , it behaves as a bimodal map, and it oscillates in the interval  $[0, 1]$ . Figure 1 shows the difference map plotted for  $\alpha = 4$  using MATLAB.

## 4 Implementation of Proposed Image Encryption Scheme

The image encryption scheme proposed in this paper uses two chaotic maps, the 1-D logistic map and the 2-D difference map to generate the key needed for the encryption and decryption process. The proposed encryption scheme involves two main steps: (1) The formation of an encryption key and (2) the encryption of the image using the aforementioned key.

For a grayscale image, for each pixel, there can be many shades of gray. For an 8-bit grayscale image,  $2^8 = 256$  shades of gray can be represented in each pixel. Thus, larger the number of bits used to represent the image, larger than the number of shades of gray represented in each pixel, thus representing a better detailed image. First, a highly random sequence is generated using the logistic map equation represented by Eq. (1), with bifurcating parameter,  $a = 3.99$ . The parameter is chosen to ensure that



**Fig. 1** Difference map for bifurcating parameter  $\alpha = 4$

the map is in a chaotic regime, and hence, a chaotic sequence is generated from it. Since each number in a sequence is used to encrypt each pixel of the image, the length of the sequence of the logistic map function is equal to the product of the number of rows and columns of the image. Each number in the sequence is converted into an 8-bit binary by multiplying each number in the sequence with 255 and rounding off to the nearest integer. The sequence thus obtained will be able to represent the 255 shades of gray and is further converted to 8-bit binary numbers to form the first part of the key, 'K1.' A similar key 'K2' is obtained from the 2-D chaotic difference map for a particular value of the initial condition. The final key 'K' for the encryption is obtained by the bitwise XOR operation of the sequences K1 and K2. Each pixel contains a single shade of gray out of the 256 shades of gray. Therefore, each pixel in the original image will be converted to a number between 0 and 255, and finally convert each number to its 8-bit binary form, say P. The sequence 'P' is then bitwise XORed with the generated key 'K' to get the encrypted sequence P', from which the encrypted image can be reconstructed with further mathematical operations. The original image can be further decrypted from the encrypted image by XORing it with the same key 'K' used for encryption. Due to the fact that the chaotic maps are highly sensitive to initial conditions, even a bit change in the original key can result in an entirely different result and hence apt for secure communication applications.

## 5 Results and Security Analysis

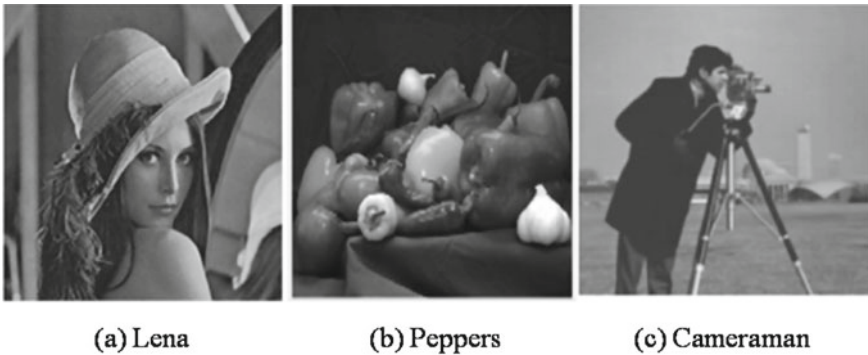
In order to analyze the performance of the proposed algorithm for image encryption, a set of three images were chosen, the image sizes of the input plain images were normalized to  $256 \times 256$  for convenience. Figure 2 shows the three plain images used for the simulation.

### 5.1 Statistical Analysis

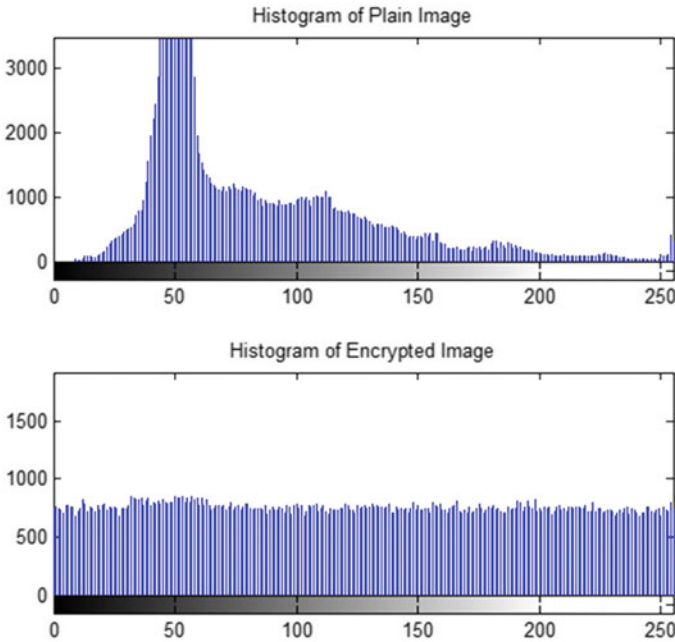
In this section, the proposed method is analyzed using various tests to determine the encryption quality and also to find its stability via statistical attacks. The test includes histogram analysis, maximum deviation analysis, irregular deviation analysis, and calculation of NPCR, UACI, and MAE [20].

#### 5.1.1 Histogram Analysis

The histogram of an image is a graphical representation that shows the pixel count in an image at the different intensity values. Every pixel of the 8-bit grayscale image has  $[0, \dots, 255]$  value range, for each grayscale input image with 256 gray levels. Figure 3 represents the histogram of the original image and the encrypted image. Uniform histogram of the encrypted image indicates the reduced possibility of statistical attacks.



**Fig. 2** Plain images used for encryption  $\alpha = 4$



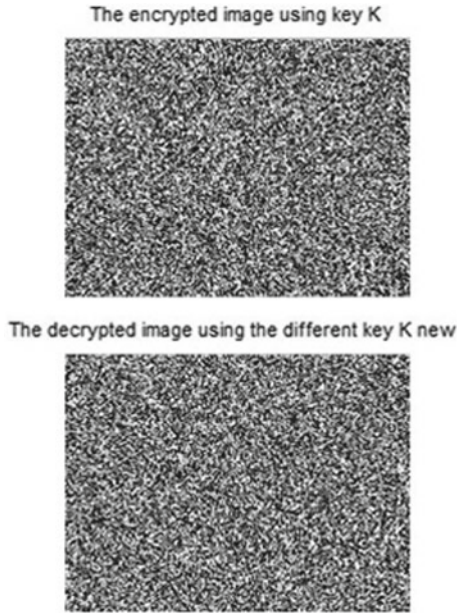
**Fig. 3** Histogram of the plain and encrypted image

### 5.1.2 Key Sensitivity Analysis

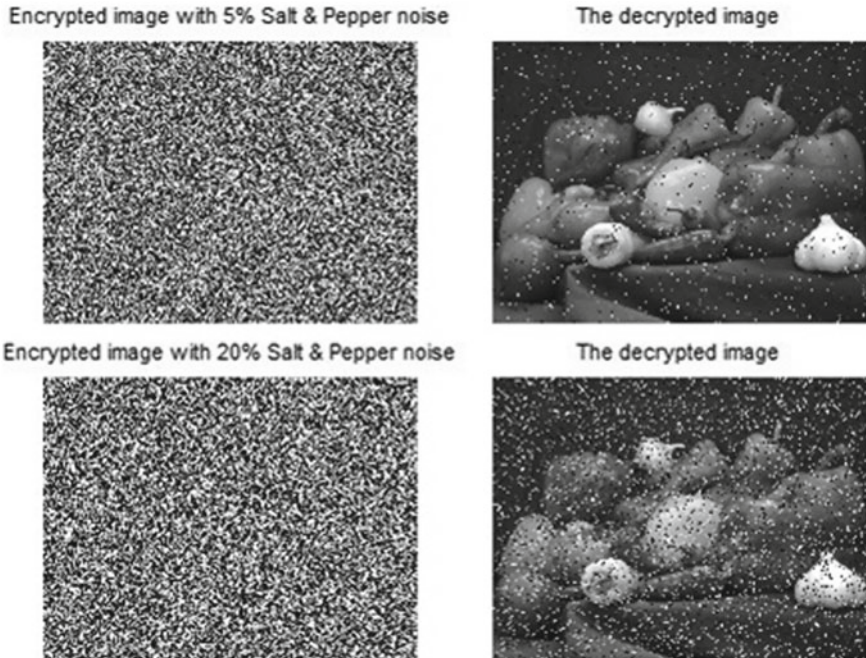
A secure image encryption process needs to be highly sensitive toward even small changes in the key, resulting in a completely different result, even for a single bit change in the main key. In this paper, the use of a key with a very small change in the main key to decrypt the image resulted in an absolutely different decrypted image. Figure 4 shows the input image encrypted with the original key 'K' and the result of decrypting the same image with a new key ' $K_{new}$ ', obtained from the original key with just one bit change.

### 5.1.3 Addition of Salt and Pepper Noise

In this analysis, the original plain image is first transmuted to its encrypted form using the original key, and then, salt and pepper noise is added to the image with different noise rates. It was observed that the image decrypted from the encrypted image, added with even 20% salt and pepper noise, was comparable to the decrypted image obtained from the original encrypted image without any noise. Figure 5 shows the encrypted image added with salt and pepper noise at different rates of 5% and 20% and the corresponding decrypted images.



**Fig. 4** The encrypted image using key K and decrypted image using a different key Knew



**Fig. 5** Encrypted image added with salt and pepper noise at different rates of 5 corresponding decrypted images

**Table 1** PSNR amount applying different rates of salt and pepper noise

PSNR amount	Rate of salt and pepper noise
71.31	1
64.21	15
61.21	10
58.18	20

#### 5.1.4 Peak Signal-to-Noise Ratio (PSNR)

PSNR is a parameter that measures the amount of equality of the two given images. It is mostly used as a measure of the quality of the image, as the ratio between the maximum power of a signal and the power of the corrupting noise. Here, it is measured between the original image and the encrypted image added with different rates of noise.

$$\text{PSNR} = 10 \times \log \frac{255^2}{\sum_{i=1}^m \sum_{j=1}^n \left[ \frac{|P(i,j) - C(i,j)|^2}{(m \times n)} \right]} \quad (4)$$

where  $P(i, j)$  and  $C(i, j)$  represent the plain image and the encrypted image, respectively. Table 1 shows the amount of PSNR obtained by adding salt and pepper noise to the pepper image at noise rates of 1, 5, 10, and 20%.

Table 1 shows the amount of PSNR obtained by adding salt and pepper noise to the pepper image at noise rates of 1, 5, 10, and 20%.

#### 5.1.5 Correlation Coefficient Analysis

Smaller values of the correlation coefficient indicate the success of the encryption process. A unity value for the correlation coefficient indicates that the encrypted image is the same as that of the original image and that the mentioned encryption process failed to hide the details of the original image. Similarly, a zero value for the correlation coefficient indicates that the original image and its encrypted form are totally different. Thus, smaller value of the correlation coefficient indicates the success of the encryption process. The following equation can be used to measure the correlation coefficient [20].

$$\text{CC} = \frac{\text{cov}(m, n)}{\sigma_m \sigma_n} \quad (5)$$



where

$$\text{cov}(m, n) = \sum_{i=1}^N (m_i - E(m))(n_i - E(n)) \quad (6)$$

$$\sigma_x = \sum_{i=1}^N (m_i - E(m))^2)^{1/2} \quad (7)$$

$$\sigma_y = \sum_{i=1}^N (n_i - E(n))^2)^{1/2} \quad (8)$$

$$E(m) = \frac{1}{N} \sum_{i=1}^N m_i \quad (9)$$

and  $m, n$  are the plain and cipher image grayscale pixel values.

## 5.2 Maximum Deviation Analysis

The maximum deviation analysis is carried out to determine the quality of the proposed algorithm for encryption [19]. In this analysis, the maximum deviation between the plain image and the cipher image is calculated. Initially, plain image and cipher image histograms are measured, then the absolute difference between the histograms are measured, and then, maximum deviation is obtained using the Eq. (10).

The maximum deviation (MD) is calculated as

$$\text{MD} = \frac{h_0 + h_{255}}{2} + \sum_{i=1}^{254} h_i \quad (10)$$

Table 2 summarizes the maximum deviation calculated for three encrypted test images, and higher values of maximum deviation indicate that there is a large deviation between the original plain image and the encrypted cipher image.

## 5.3 Irregular Deviation Analysis

This analysis is based on the amount of deviation caused on the encrypted image by the encryption process [20]. The irregular deviation can be calculated using equation (11).

**Table 2** Results of CC MD and ID FACTORS

Image	Correlation Coefficient (CC)	Maximum Deviation (MD)	Irregular Deviation (ID)
Lena	0.0088	170646	268918
Peppers	0.0011	163392	193478
Cameraman	0.0023	185446	228918

$$ID = \sum_{i=0}^{255} |H(i) - \frac{1}{256} \sum_{i=0}^{255} h_i| \tag{11}$$

where the values of  $H$  and  $D$  are defined as  $H = \text{histogram}(D)$  and  $D = |P - C|$ , where  $P$  and  $C$  represents the original and encrypted image, respectively, and  $h_i$  represents the amplitude of the absolute difference histogram at the value  $i$ . A low value of ID indicates a better encryption algorithm. Table 2 shows the values of CC, MD, and ID factors calculated for the three input images.

### 5.4 Differential Analysis

The resistance to differential attacks in image encryption techniques is usually measured by three tests, namely number of pixels change rate (NPCR), unified average changing intensity (UACI), and mean absolute error (MAE) [20].

#### 5.4.1 NPCR

It represents the number of pixels change rate of the encrypted image while one pixel of the plain image is changed. NPCR is calculated using the following equations

$$NPCR = \frac{\sum_{i=0}^m \sum_{j=0}^n d(i, j)}{m \times n} \times 100\% \tag{12}$$

where

$$d(i, j) = \begin{cases} 0, & C1(i, j) = C2(i, j) \\ 1, & C1(i, j) \neq C2(i, j) \end{cases} \tag{13}$$

$C1(i, j)$  and  $C2(i, j)$  represent the matrices of the encrypted images obtained by encrypting two plain images with just a single-bit difference, whereas  $m$  and  $n$  represent the height and width by pixels of the encrypted images.

**Table 3** Results of NPCR and UACI tests

Image	Lena	Peppers	Camerman
NPCR %	99.36	99.32	98.65
UACI %	30.23	29.41	30.34
MAE	82	79	85

### 5.4.2 UACI

It represents the average intensity of the differences between the plain image  $P(i, j)$  and encrypted image  $C(i, j)$ .

$$UACI = \frac{1}{m \times n} \frac{\sum_{i=1}^m \sum_{j=1}^n |P(i, j) - C(i, j)|}{255} \times 100\% \quad (14)$$

### 5.4.3 MAE

MAE represents the mean absolute error. There must be a substantial difference shown by the encrypted image when compared with the corresponding plain image. MAE is a key technique to measure this difference in a secure encryption system. A better image security is ensured with larger values of MAE.

The results of NPCR and UACI of the three encrypted images are calculated, and the results are shown in Tables 3. As shown in the table, the NPCR values are close to the ideal value, which indicates that the proposed scheme for encryption is highly sensitive to even minor changes in the original image. The UACI test results obtained for the various images are also found to be close to the ideal results, indicating that a single pixel change can result in a large rate of influence, thus indicating the strength of image encryption algorithms/ciphers with respect to differential attacks.

## 6 Conclusion

In this paper, an image encryption and decryption scheme based on chaotic maps, the 1-D logistic map and 2-D difference map have been proposed and simulated using MATLAB. The test images were initially encrypted with a secret key  $K$  and then decrypted, using exactly the same key  $K$  and a slightly different key  $K_{new}$ . Since the keys were generated using the logistic map and difference map operating in the chaotic regime, even a small variation in the key will affect the process of decryption thus being cryptographically stable to difference attacks. The analyses includes determining the correlation coefficient, maximum deviation, irregular deviation, key sensitivity, and PSNR that validates the performance of the proposed encryption system. The uniform distribution of the histogram of the encrypted image indicates a reduction in the possibility of static attacks.

## References

1. Matthews R (1989) On the derivation of a “chaotic” encryption algorithm. *Cryptologia* 13(1):29–42
2. Zhang L, Liao X, Wang X (2005) An image encryption approach based on chaotic maps. *Chaos Fractals* 24(3):759–765
3. Pareek N, Patidar V, Sud K (2006) mage encryption using chaotic logistic map. *Image Vis Comput* 24(9):926–934
4. Gao T, Chen Z (2008) Image encryption based on a new total shuffling algorithm. *Solitons Fractals* 38:213–220
5. Raja KR, Sampath A, Indumathi P (2011) Enhancement and analysis of chaotic image encryption algorithms. *Comput Sci Inf Technol* 1(7)
6. Purwar R, Priyanka (2013) An improved image encryption scheme using chaotic logistic maps. *Comput Sci Inf Technol* 2(3):220–227
7. Assad SEI, Farajallah M (2016) A new chaos-based image encryption system. *Signal Process* 41:144–157
8. Pak C, Huang L (2017) A new color image encryption using combination of the 1d chaotic map. *Signal Process* 138:129–137
9. Cao C, Sun T, Liu W (2018) A novel bit-level image encryption algorithm based on 2d-licm hyperchaotic map. *Signal Process* 143:122–133
10. Moafimadani SS, Chen Y, Tang C (2019) A new algorithm for medical color images encryption using chaotic systems. *Entropy* 21(6)
11. Akhshani A, Akhavan A, Lim C, Hassan Z (2012) An image encryption scheme based on quantum logistic map. *Commun Nonlinear Sci Numerical Simul* 17(12):4653–4661
12. Li C, Luo G, Qin K, Li C (2017) An image encryption scheme based on chaotic tent map. *Nonlinear Dyn* 87(01)
13. Mao Y, Chen G, Lian S (2004) A novel fast image encryption scheme based on 3d chaotic baker maps. *Int J Bifurcation Chaos* 14(10):3613–3624
14. Ramasamy P, Ranganathan V, Kadry S, Damasevi R, Bla Zauskas T (2019) An image encryption scheme based on block scrambling, modified zigzag transformation and key generation using enhanced logistic–tent map. *Entropy* 21(7)
15. Luo Y, Yu J, Lai W, Liu L (2019) A novel chaotic image encryption algorithm based on improved baker map and logistic map. *Multimedia Tools Appl* 1–21
16. Hamdi M (2015) A very efficient pseudo-random number generator based on chaotic maps and s-box tables. *Int J Comput Control Quantum Inf Eng* 09(2):481–485
17. May R (1976) Simple mathematical models with very complicated dynamics. *Nature* 26:457
18. Garc’a-Mart’nez M, Campos-Canton I, Campos-Cant E, Celikovsk’ y S (2013) Simple mathematical models with very complicated dynamics, vol 74, no 3, pp 457–830
19. Nien HH, Changchien SK, Wu SY, Huang CK (2008) A new pixel-chaotic-shuffle method for image encryption. In: 10th International conference on control, automation, robotics and vision, pp 883–887
20. Fishawy NFE, Zaid O (2007) Quality of encryption measurement of bitmap images with rc6, mrc6, and rijndael block cipher algorithms. *Int J Netw Secur* 5:241–251

# A Novel March XR Algorithm, Design, and Test Architecture for Memories



I. G. Matri, Aishwaraya, N. Shreya, Saroja V. Siddamal,  
and Suneeta V. Budihal

## 1 Introduction

Memories are the important module in every embedded system. Hence, it becomes important to check if the read and write operations in the memory are happening as expected. Sometimes, due to manufacturing defects or due to unexpected changes in the hardware, certain cells or address locations of the memory become faulty because of which memory produces an unexpected results. Therefore, there is a necessity to check if the memory is infected with the faults. Hence, an inbuilt testing circuitry known as built-in self-test (BIST) is designed which checks for faults as per the patterns or algorithms incorporated in it. The BIST employed for memory is known as MBIST and is interfaced with the memory that needs to be tested.

In order to guide the BIST to move to the different states of algorithm, an FSM is designed which generates the states of an algorithm as required. Rather than implementing multiple algorithms, which can increase the testing time or consume larger area, it is appreciated to implement one algorithm which covers most of the faults, reducing testing time and area. Hence, the authors propose a novel March XR algorithm which results in larger fault coverage.

---

I. G. Matri · Aishwaraya · N. Shreya · S. V. Siddamal (✉) · S. V. Budihal  
KLE Technological University, Hubballi, India  
e-mail: [sarojavs@kletech.ac.in](mailto:sarojavs@kletech.ac.in)

S. V. Budihal  
e-mail: [suneeta\\_vb@kletech.ac.in](mailto:suneeta_vb@kletech.ac.in)

## 2 Literature Survey

The authors in [1] discuss the implementation of BIST circuit for multiple RAMs in a SoC; as the present day, SoCs have more than one memory. This is done by testing multiple RAMs in parallel using a BIST circuit. When the BIST enters the test mode, the controller sends the signal to the sequencer which generates the signal for the test pattern generator to generate the test patterns based on the March C algorithm. The paper also discusses the usage of configurable CLFSR for generating the addresses in required sequence, which contributes to the area reduction. Paper [2] deals with reducing the switching activity when testing SRAM. The built architecture is tested in VHDL. The authors in [3] discuss the usage of BIST circuits in the automotive field. As functional safety and reliability are the two main important aspects of the automotive industry, there is a need for constant testing of electronic controlled units (ECUs). This paper uses the flexible BIST called Star Memory System of Synopsys for carrying out the memory test in power-on mode and mission mode. The paper [4] discusses the BIST architecture to implement three different March algorithms, namely March C, March Y, and March A. The paper also brings out the comparison between March C, March Y, and March A algorithms. March A has more states and hence occupies more area, consumes more time and power. March Y has least number of states and hence consumes least area, least time and power, while March C has greater fault coverage compared to March Y, but it consumes more power, time, and area. In [5], the authors discuss the implementation of BIST for multi-ported memory using microcode architecture with two variants, namely synchronous p-MBIST and asynchronous p-MBIST. Both synchronous and asynchronous p-MBISTs consist of a microcode instruction storage (MIS) unit, comparator, and test collar. The paper also discusses memory repair using redundancy array logic. In [6], built-in self-test is designed for RAM using VHDL. Different blocks used for BIST are implemented separately to check correctness and then later combined together for final design. RAM is circuit under test; LFSR is used to implement test pattern generator, and comparator is used for test response analyzer. BIST controller unit is also designed. The design is implemented on FPGA Spartan-6 board. In paper [7], the authors have worked on BIST methods used for fault. Methods used in papers for each category are mentioned. Types of faults covered, type of memory, and algorithm used in different papers are also listed. The method proposed in [8] takes advantage of MBIST architecture and algorithm to estimate power consumption of SRAM. Realistic power consumed is estimated by mimicking the normal behavior of logic on memory. Therefore, the goal is to reduce the normal mode and BIST mode memory activity differences. The proposed method is evaluated using commercial modem chip and algorithm using Python and C++ languages. In paper [9], a new algorithm which is the improvement of March C+ algorithm is proposed. The algorithm is used for testing of low-voltage SRAM. Fault models such as read destructive coupling fault, write destructive coupling fault, and write disturb fault which are not identified by March C+ algorithm are aimed to be detected in new March LV algorithm. Layout of SRAM with BIST is shown, and under SMIC 40 nm LL CMOS

process, the low-voltage 8 T SRAM is tested. The work in [10] refers to the circuit's ability to test itself. The examination can be taken either online or offline. A test may be run simultaneously or non-concurrently with regular operation in an online environment. Stuck-at faults, transition faults, coupling faults, neighborhood pattern-sensitive faults, dynamic faults, and so on can all be found in memory. The Verilog software tool is used to build 256X8 (2 KB) SRAM memory for fault detection. This built memory will be used to evaluate a circuit. The analysis testing is carried out during run time. The authors in [11] have discussed about semiconductor memories. Semiconductor memories are becoming increasingly important as technology advances. A new test architecture is proposed in this paper that uses a response analyzer and checker to detect a fault on a chip, as well as a modified MARCH C algorithm is proposed. The authors in [12] discuss the hardware implementation scan test architecture. March tests, in which fixed sequences of March elements are used, are the most widely used memory test algorithms. These operations are carried out in either decreasing or increasing order of the memory address. BIST controller, memory under test (MUT), checker, and response analyzer make up this architecture.

The rest of the paper is organized as follows. Section 3 discusses various memory faults and existing algorithms. In Sect. 4, a novel March XR algorithm is explained. Section 5 discusses the results and finally the conclusion.

### 3 Memory Faults and Algorithms

Different types of faults can be found in memory such as single cell faults and multi-cell faults. Single cell faults are as follows: stuck at faults (SAFs), transition faults (TFs), write destructive faults (WDFs), read destructive faults (RDFs): incorrect read faults (RDFs), deceptive read destructive faults (DRDFs). Some of two cell faults are as follows:

coupling faults (CFs), inversion coupling faults (CFin), idempotent coupling faults (CFid), static coupling faults (CFst), write destructive coupling faults (CFwd), read destructive coupling faults (CFrd), incorrect read coupling faults (CFir), and deceptive read destructive coupling faults (CFdrd).

#### Algorithms for testing

A numbers of algorithms are available to test memories. The algorithms used to test memory are March algorithms, RAW algorithms, Zero-one algorithm (MSCAN), galloping pattern (GALPAT), walking pattern (WALPAT), sliding, butterfly algorithm, moving inversion (MOVI) algorithm, surround disturb (SD) algorithm, and many more. When the values from memory are accessed, the March algorithm follows March up and March down pattern. The size and the word length are determined by these algorithms. Table 1 shows various algorithms to test memory.

**Table 1** Algorithms to test memory

Algorithm	Description
MATS	{ $\Downarrow$ (w0); $\Downarrow$ (r0, w1); $\Downarrow$ (r1)}
MATS+	{ $\Downarrow$ (w0); $\Uparrow$ (r0, w1); $\Downarrow$ (r1, w0)}
MATS++	{ $\Downarrow$ (w0); $\Uparrow$ (r0, w1); $\Downarrow$ (r1, w0, r0)}
MARCH X	{ $\Downarrow$ (w0); $\Uparrow$ (r0, w1); $\Downarrow$ (r1, w0); m (r0)}
MARCH RAW1	{ $\Downarrow$ (w0); $\Downarrow$ (w0, r0); $\Downarrow$ (r0); $\Downarrow$ (w1, r1); $\Downarrow$ (r1); $\Downarrow$ (w1, r1); $\Downarrow$ (r1); $\Downarrow$ (w0, r0); $\Downarrow$ (r0)}
MATS	{ $\Downarrow$ (w0); $\Downarrow$ (r0, w1); $\Downarrow$ (r1)}

### 4 Proposed March XR Algorithm

Since the algorithm that is incorporated in the BIST plays an important role in detecting the faults, it becomes important to select the algorithm which can detect maximum faults. Rather than implementing multiple algorithms, which can increase the testing time or consume larger area, it is appreciated to implement one algorithm which covers most of the faults reducing testing time and area. Hence, the authors propose a novel MARCH XR algorithm as shown in Table 2 which results in larger fault coverage.

The finite state machine for the March XR algorithm is as shown in Fig. 1. The algorithm is tested on a two-port memory RAM. For every port, at the corresponding state, flags are set. For the FSM shown, the architecture is built as shown in Fig. 2.

BIST controller is designed as an FSM where in each state different read and write operations are carried out according to the proposed algorithm MARCH XR on memory. Different faults have been created in memory to test the controller. The architecture is FSM-based BIST controller. The modules present in the architecture are FSM, data generator, address generator, control generator, comparator, memory wrapper, and the memory to be tested which is instantiated in the controller.

**Table 2** Proposed March XR algorithm

Algorithm	Description
Proposed March XR	$\Downarrow$ W0 $\Downarrow$ (R0 W1) $\Uparrow$ (R1 W0) $\Downarrow$ R0 $\Downarrow$ W0 $\Downarrow$ (W0 R0) $\Downarrow$ R0 $\Downarrow$ W1 $\Downarrow$ (W1 R1) $\Downarrow$ R1

where

$\Downarrow$  W0  $\Downarrow$  (R0 W1)  $\Uparrow$  (R1 W0)  $\Downarrow$  R0  $\Uparrow$  W0  $\Downarrow$  (W0 R0)  $\Uparrow$  W1  $\Downarrow$  (W1 R1)

W0—Write 0 W1—Write 1 R0—Read 0 R1—Read 0

$\Downarrow$ —Increasing or decreasing address value

$\Downarrow$ —Decreasing address value

$\Uparrow$ —Increasing address value



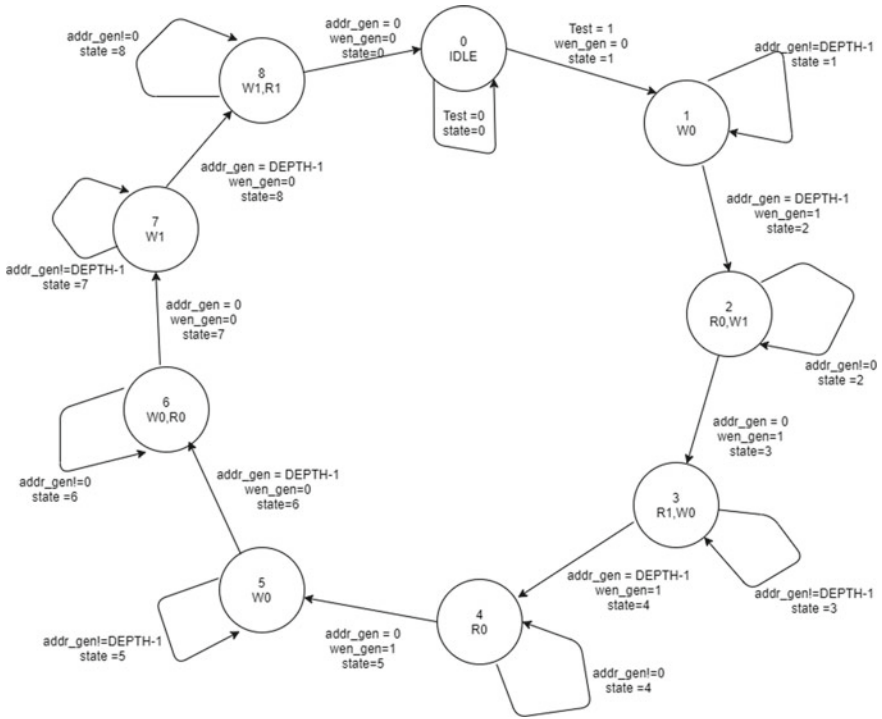


Fig. 1 FSM for proposed March XR algorithm

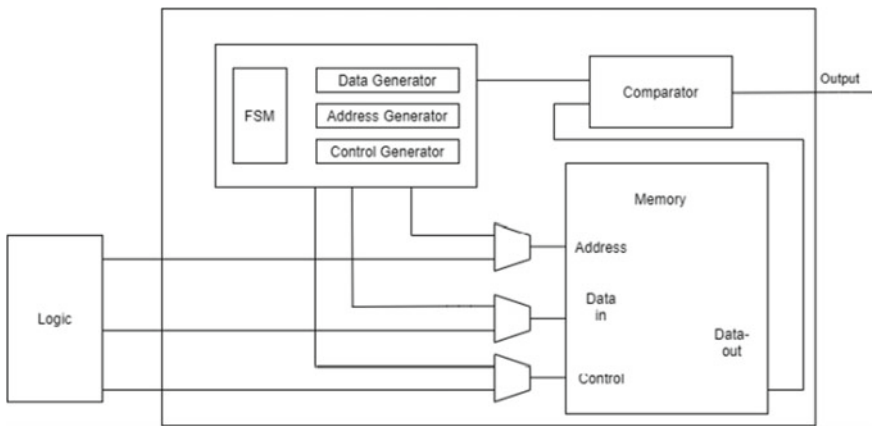


Fig. 2 Proposed architecture for March XR algorithm

## 5 Results and Discussion

In this section, the proposed March XR algorithm is evaluated for various test coverages. To test various faults at different states, flags are set—f2, g2, f3, g3, etc., at various states at port A and B. To test, various faults are created of the required type. Table 3 shows various faults detected at each state. Proposed March XR algorithm covers all the faults like destructive faults, coupling faults, transition faults, and stuck at faults as shown in Table 4. The proposed algorithm is simulated and synthesized using Xilins ISE 14.2 and implemented on Spartan-6 FPGA. A dual-port SRAM of size 1024\*8 is considered for testing. Various faults are introduced at different memory location. Flags are set based on the faults detected. Figure 3a–c shows the results of memory tested for faults. Figure 3a shows the BIST pass output. Both the ports of memory are tested.

Figure 3b shows the read coupling fault occurring at address 75. As shown in the figure along with in correct coupling fault at the same address other faults like stuck at 0, coupling faults also occur.

At address 127, static coupling faults are tested as shown in Fig. 3c. Other faults like transition, read destructive, and other coupling faults have also occurred at this address. Proposed XR algorithm tests most of the faults in the memory. A comparative study is made with respect to the existing algorithms. As shown in Table 4 MATS algorithm only covers faults like AF, RDF, and IRF; using this algorithm, it is not possible to cover other dynamic faults mentioned in table. Similarly, MATS+, MATS++ algorithms cover few faults. The proposed novel March XR algorithm detects all mentioned faults.

## 6 Conclusion

The authors have presented a novel March XR algorithm. Dual-port SRAM of size 1024\*8 is used as test memory. With FSM-based BIST architecture and March XR algorithm, most of the faults like destructive faults, transition faults, stuck at faults, and coupling fault are detected. These faults are created different memory locations and later tested. A single March XR covers most of the faults; hence, the need to design different algorithms is eliminated. Implemented BIST controller is faster in testing and consumes less area.

**Table 3** Faults covered in each state

Faults	f2,g2	f3,g3	f4,g4	f6,g6	f7,g7	f9,g9	f 10,g10
Stuck at 0		✓				✓	✓
Stuck at 1	✓		✓	✓	✓		
Write destructive fault 0				✓	✓		
Write destructive fault 1						✓	✓
Incorrect read fault 0	✓		✓	✓	✓		
Incorrect read fault 1		✓				✓	✓
Read destructive fault 0	✓		✓	✓	✓		
Read destructive fault 1		✓				✓	✓
Deceptive read destructive fault 0					✓		
Deceptive read destructive fault 1							✓
Idempotent coupling fault 0-10		✓					
Idempotent coupling fault 0-11	✓						
Idempotent coupling fault 1-00		✓					
Idempotent coupling fault 1-01			✓				
Inversion coupling fault 0-1			✓	✓			
Incorrect read coupling fault	✓		✓	✓	✓		
Read destructive coupling fault	✓		✓	✓	✓		
Deceptive RD coupling fault					✓		
Static coupling fault 00		✓					
Static coupling fault 01	✓		✓		✓		
Static coupling fault 10		✓					✓
Static coupling fault 11	✓						
Transition coupling fault 010		✓					
Transition coupling fault 011		✓					
Transition coupling fault 100			✓				
Transition coupling fault 101			✓				
Write destructive coupling fault				✓	✓		
Dynamic Incorrect read fault 0	✓		✓	✓			
Dynamic Incorrect read fault 1		✓				✓	
Dynamic Read destructive fault 0	✓		✓	✓	✓		
Dynamic Read destructive fault 1		✓				✓	✓
Dynamic Deceptive RD fault 0					✓		
Dynamic Deceptive RD fault 1							✓
Dynamic IR coupling fault 0	✓		✓	✓			
Dynamic RD coupling fault 0	✓		✓	✓	✓		
Dynamic DRD coupling fault 0					✓		

```
C:\Users\Maitri\Desktop\new\new mem>iverilog bist algc.v mem.v
VCD info: dumpfile example1.vcd opened for output.
BIST PASS - MARCH XR algorithm - Port A
BIST PASS - MARCH XR algorithm - Port B
```

**a**

```
the fault has occurred at 75 address
All the cells have fault
Incorrect read fault 0
Read destructive fault 0
Stuck at 1 fault
Incorrect read coupling fault 00
Read destructive coupling fault 00
Read destructive coupling dynamic fault 00
Dynamic read destructive fault 0

BIST FAIL - MARCH XR algorithm - Port A
```

**b**

```
the fault has occurred at 127 address
the fault has occurred at 27 cell
Idempotent coupling fault 010
Incorrect read coupling fault 10
Static coupling fault 00
Transition coupling fault 010
Transition coupling fault 011
Read destructive coupling fault 10

BIST FAIL - MARCH XR algorithm - Port A

the fault has occurred at 255 address
the fault has occurred at 27 cell
```

**c**

**Fig. 3** a Memory tested for no fault. b Incorrect read coupling fault. c Static coupling faults

**Table 4** Comparison of traditional March algorithm with proposed March XR

Algorithm	AF	CF	Dynamic	RDF	WDF	DRDF	IRF	Transition
MATS	✓	×	×	✓	×	×	✓	Few
MATS+	✓	×	×	✓	×	×	✓	Few
MATS++	✓	Few	Few	✓	×	×	✓	✓
MARCH X	✓	Few	Few	✓	×	×	✓	✓
MARCH C	✓	Few	Few	✓	×	×	✓	✓
Proposed March XR	✓	✓	✓	✓	✓	✓	✓	✓

## References

1. John PK, Rony Antony P (2017) BIST architecture for multiple RAMs in SoC. *Procedia Comput Sci* 115:159–165. ISSN 1877-0509, <https://doi.org/10.1016/j.procs.2017.09.121>
2. Wu Y, Ivanov A (2006) Low power SoC memory BIST. In: 2006 21st IEEE international symposium on defect and fault tolerance in VLSI systems, Arlington, VA, USA, pp 197–205. <https://doi.org/10.1109/DFT.2006.39>
3. Sargsyan D (2017) ISO 26262 compliant memory BIST architecture. In: 2017 computer science and information technologies (CSIT), Yerevan, Armenia, pp 78–82. <https://doi.org/10.1109/CSITechnol.2017.8312145>
4. Singh A, Kumar GM, Aasti A (2020) Controller architecture for memory BIST algorithms. In: 2020 IEEE international students’ conference on electrical, electronics and computer science (SCEECS), Bhopal, India, pp 1–5. <https://doi.org/10.1109/SCEECS48394.2020.43>
5. Patil SR, Musle DB (2017) Implementation of BIST technology for fault detection and repair of the multi-ported memory using FPGA. In: 2017 international conference of electronics, communication and aerospace technology (ICECA), Coimbatore, India, pp 43–47. <https://doi.org/10.1109/ICECA.2017.8212849>
6. Tewary T, Dey S, Roy S (2020) Realization of built-in self test (BIST) enabled memory (RAM) using VHDL and implementation in Spartan6 FPGA board. In: 2020 IEEE VLSI device circuit and system (VLSI DCS), Kolkata, India, pp 322–326. <https://doi.org/10.1109/VLSIDCS47293.2020.9179879>
7. Pundir KS, Sharma OP (2017) Fault tolerant reconfigurable hardware design using BIST on SRAM: a review. In: 2017 international conference on intelligent computing and control (I2C2), Coimbatore, India, pp 1–16. <https://doi.org/10.1109/I2C2.2017.8321907>
8. Zhao Y, Chen L (2019) Power estimation of embedded SRAMs using BIST algorithms. In: 2019 IEEE international conference on electron devices and solid-state circuits (EDSSC), Xi’an, China, pp 1–2. <https://doi.org/10.1109/EDSSC.2019.8754261>
9. Cai Z, Wang Y, Liu S, Lv K, Wang Z (2019) A novel BIST algorithm for low-voltage SRAM. In: 2019 IEEE international test conference in Asia (ITC-Asia), Tokyo, Japan, pp 133–138. <https://doi.org/10.1109/ITC-Asia.2019.00036>
10. Tak-Barekar S, Mali M, Khade A (2020) Sensitizing defects in static random access memory using built-in self-test. *Solid State Technol* 63(6). Publication Year: 2020
11. Nekara H, Siddamal SV (2020) Design and implementation of novel scan architecture for test power reduction. *Procedia Comput Sci* 171:2556–2562
12. Vannal NS, Siddamal SV, Bidaralli SV, Bhille MS (2015) Design and testing of combinational logic circuits using built in self test scheme for FPGAs. In: 2015 fifth international conference on communication systems and network technologies (CSNT), 4–6 April 2015, pp 903–907. <https://doi.org/10.1109/CSNT.2015.151>

# Analysis of Heterogeneous Queueing Model with Unreliable Server and Working Vacation



M. Seenivasan, H. Manikandan, and K. S. Subasri

## 1 Prologue

The queueing theory applied until now all most the various real life environments like bank, ticket counters, manufacturing system, communication network system, and computer system, industrial and so on. Many researchers have investigated the analytical aspect of heterogeneous servers or single server with intermittently available server queueing models throughout the different techniques. After completion in hand service, interruption occurs in service period at random length of time is called intermittently available service. The queueing model with an intermittently available server with single channel bulk service was analyzed by Chaudhry [1]. Ammar Sherif [2] investigated queue with implement behavior in transient analysis of a two-heterogeneous servers. Multi-server queueing with single WV was studied by Lin and Ke [3]. Two heterogeneous server queue with intermittently server was analyzed by Seenivasan et al. [4]. Levy and Yechiali [5] explained multiserver queueing system with server's vacations. A study on multi-server queueing simulation was studied by Shanmugasundaram and Punith [6].

Recently many authors investigated catastrophic customers in queueing model. If the catastrophe occurs, all arrivals are destroyed or leave the system immediately. A queueing network with catastrophes and product from solution was analyzed by Chao [7]. Vinodhini and Vindhya [8] Explained about transient solution of multiserver model with catastrophes and impatient customers. M/M/C queueing system with

---

M. Seenivasan (✉)

Mathematics Wings-DDE, Annamalai University, Annamalainagar, Tamilnadu 608 002, India

e-mail: [emseeni@yahoo.com](mailto:emseeni@yahoo.com)

H. Manikandan

Department of Chemistry, Annamalai University, Annamalainagar, Tamilnadu 608 002, India

K. S. Subasri

Department of Mathematics, Annamalai University, Annamalainagar, Tamilnadu 608 002, India

occasionally suffers disastrous failure and all customer lost and the system restarts immediately. Markovian queue having two heterogeneous servers with catastrophes was studied by Indra and Vijay Rajan [9]. In many real life situation, the server can be work at different rate during the vacation period is called queue with working vacation. This queueing model is using in manufacturing and service sectors. As well as queueing models where the server may subject to breakdowns and sent to repaired [10]. Was analyzes single server M/G/1 queueing system subject to breakdown -some reliability and queueing problems. A brief and useful explanation about queueing system with more vacations was studied by Doshi [11]. Queueing Analysis, A Foundation of performance Analysis of vacation and priority systems was explained by Takagi [12]. M/M/1 queue with working vacation (M/M/1/MW) was analyzed by Servi and Finn [13]. Working vacations queueing models with multiple types of server breakdowns was explained by Jain and Jain [14]. Vinod [15] Was analyzed about the service facility may have one or more identical servers that are subject to breakdowns and repairs in unreliable queueing systems. Aissani and Artalejo [16] Was explained on the single server retrial queue subject to breakdown. Seenivasan and Indumathi [17] Was studied a retrial queueing model with unreliable server. Transient and steady-state analysis of an M/M/1 queue with balking, catastrophes and unreliable server was explained by Tarabia [18]. Non markovian retrial queue subject to disasters and server failures was investigated by Wang and Li [19]. Alfa [20] Discussed in vacation models in discrete time. Boudali and Economou [21] was examined optimal and equilibrium balking strategies in the single server Markovian queue with catastrophes.

The rest of paper has been arranged as follows. In Sect. 2 describe the construction of model. In Sect. 3 is given the matrix geometric structure and solution. Also we derived the stability condition of the model and some formulas for evaluation are given in sub section of 3. Numerical study has been done and various evaluation of result has been calculated in Sect. 4. Finally, in Sect. 5 given the conclusion for this model.

## 2 Construction of the Model

The following assumption describe proposed for this model:

- Arriving customer of the system according to Poisson processes with rate  $\lambda$ .
- There are two heterogeneous servers providing the service with different rates. Customer comes one by one receives the service based on first in first out queue discipline. Both servers follow exponential distribution with service rates  $\mu_1$  for server1 and  $\mu_2$  for server2 (i.e.,  $\mu_1 \neq \mu_2$ ).
- Server2 goes to implements some important or odd jobs while queue length is greater than zero. But before going to implements these odd jobs the server2 first complete in his hand service. There are no customers in the queue, or equal to may take vacations follow the exponential distribution with rate  $\theta$ . If server2 saw the

arriving customers in his vacation period, then he serve at the low rate or server2 immediately interrupt his vacation go to regular working state and serve with the normal service rate.

- Intermittently available time of server2 follows exponential distribution with rate  $\nu$ .
- Server2 may be fail or subject to breakdowns at random follows exponential distribution with three different rates depending the states of server2. If the breakdown occurs server2 is serving in regular working state, then the service with rate  $\alpha_1$ . If the breakdown occurs server2 is serving in intermittently available state, then the service with rate  $\alpha_2$ . If the breakdown occurs server2 is serving in working vacation state, then the service rate with  $\alpha_0$ .
- Once the system is breakdown, then it enters the repaired immediately. The repair time follows exponential distribution with rates  $\beta_0$  for breakdown occurs only on server2 is serve in working state,  $\beta_1$  for breakdown occurs only on server2 is serve in working vacation state and  $\beta_2$  for break down occurs only on server2 is in intermittently state respectively.
- Catastrophe occurs follows exponential distribution with rate  $\xi$ .
- All processes are mutually independent and identical to each other.

In below figure shows that structure of transition diagram of this model (Fig. 1).

In the system, let us consider  $K(t)$  and  $n(t)$  be the status of server2 and the number of customers at time  $t$  respectively, described as follows:

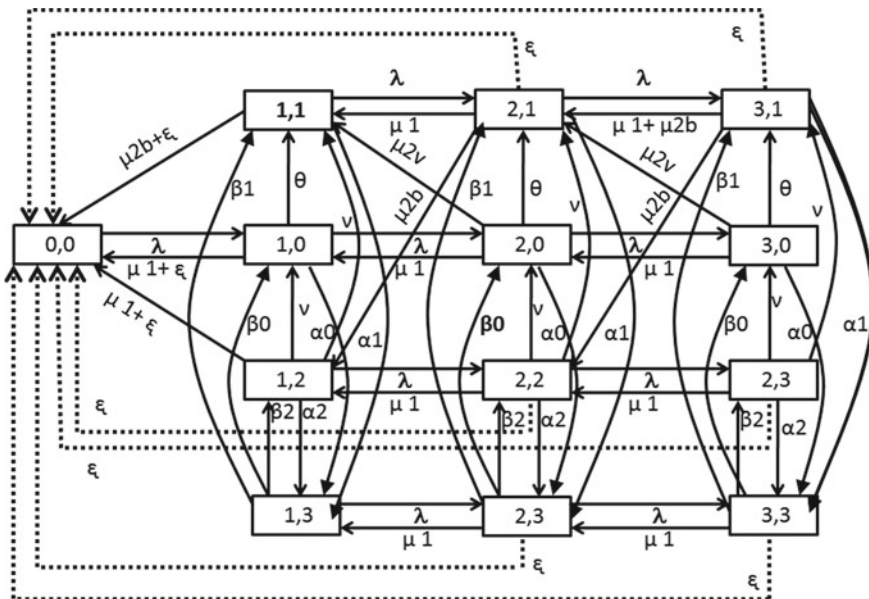


Fig. 1 Transition diagram of the model



$$K(t) = \begin{cases} 0, & \text{when server2 is in working vacation} \\ 1, & \text{when server2 is in regular working state} \\ 2, & \text{when server2 is in intermittently available} \\ 3, & \text{when sever2 is in breaking down} \end{cases}$$

$K(t)$  and  $n(t)$  be the system state. Let the Markov process (MP) be  $\{(n(t), K(t)), t \geq 0\}$  with state space arranged in lexicographical order as given below:

$$\Omega = \{(0, 0)\} \cup \{(i, j); i \geq 1, j \equiv 0, 1, 2, 3\}$$

$Q$  Matrix (called infinitesimal generator Matrix) is defined as

$$Q = \begin{pmatrix} B_{00} & B_{01} & \dots & \dots & \dots & \dots & \dots & \dots \\ B_{10} & B_{11} & A_0 & \dots & \dots & \dots & \dots & \dots \\ A_3 & A_1 & B_1 & A_0 & \dots & \dots & \dots & \dots \\ A_3 & \dots & A_2 & B_2 & A_0 & \dots & \dots & \dots \\ A_3 & \dots & \dots & A_2 & B_2 & A_0 & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \end{pmatrix}$$

where,

$$B_{00} = (-\lambda); B_{01} = (\lambda \ 0 \ 0 \ 0);$$

$$A_3 = \begin{pmatrix} \xi \\ \xi \\ \xi \\ \xi \end{pmatrix}, B_{10} = \begin{pmatrix} \mu_1 + \xi \\ \mu_{2b} + \xi \\ \mu_1 + \xi \\ \mu_1 + \xi \end{pmatrix}; A_0 = \begin{pmatrix} \lambda \ 0 \ 0 \ 0 \\ 0 \ \lambda \ 0 \ 0 \\ 0 \ 0 \ \lambda \ 0 \\ 0 \ 0 \ 0 \ \lambda \end{pmatrix}$$

$$A_1 = \begin{pmatrix} \mu_1 & \mu_{2v} & 0 & 0 \\ 0 & \mu_1 & \mu_{2b} & 0 \\ 0 & 0 & \mu_1 & 0 \\ 0 & 0 & 0 & \mu_1 \end{pmatrix}; A_2 = \begin{pmatrix} \mu_1 & \mu_{2v} & 0 & 0 \\ 0 & \mu_1 + \mu_{2b} & \mu_{2b} & 0 \\ 0 & 0 & \mu_1 & 0 \\ 0 & 0 & 0 & \mu_1 \end{pmatrix};$$

$$B_{11} = \begin{pmatrix} -(\lambda + \mu_1 + \xi + \theta + \alpha_o) & \theta & 0 & \alpha_o \\ 0 & -(\lambda + \mu_{2b} + \xi + \alpha_1) & 0 & \alpha_1 \\ v & v & -(\lambda + \mu_1 + 2v + \xi + \alpha_2) & \alpha_2 \\ \beta_0 & \beta_1 & \beta_2 & -(\lambda + \mu_1 + \beta_0 + \beta_1 + \beta_2 + \xi) \end{pmatrix};$$

$$\begin{aligned}
 B_1 &= \begin{pmatrix} -(\lambda + \mu_1 + \mu_2 v + \xi + \theta + \alpha_o) & \theta & 0 & \alpha_o \\ 0 & -(\lambda + \mu_1 + \mu_2 b + \xi + \alpha_1) & 0 & \alpha_1 \\ v & v & -(\lambda + \mu_1 + 2v + \xi + \alpha_2) & \alpha_2 \\ \beta_0 & \beta_1 & \beta_2 & -(\lambda + \mu_1 + \beta_0 + \beta_1 + \beta_2 + \xi) \end{pmatrix} \\
 B_2 &= \begin{pmatrix} -(\lambda + \mu_1 + \xi + \theta + \alpha_o) & \theta & 0 & \alpha_o \\ 0 & -(\lambda + \mu_1 + 2\mu_2 b + \xi + \alpha_1) & 0 & \alpha_1 \\ v & v & -(\lambda + \mu_1 + 2v + \xi + \alpha_2) & \alpha_2 \\ \beta_0 & \beta_1 & \beta_2 & -(\lambda + \mu_1 + \beta_0 + \beta_1 + \beta_2 + \xi) \end{pmatrix}
 \end{aligned}$$

### 3 Matrix Geometric Solution

Here  $\Pi$  is given by  $\Pi = (\Pi_0, \Pi_1, \Pi_2, \dots)$  is the static probability matrix such that  $\Pi Q = 0$  and  $\Pi e = 1$ . where  $\Pi_0 = (\pi_{00}, \Pi_1 = (\pi_{10}, \pi_{11}, \pi_{12}, \pi_{13}), \Pi_2 = (\pi_{20}, \pi_{21}, \pi_{22}, \pi_{23})$  and  $\Pi_i = (\pi_{i0}, \pi_{i1}, \pi_{i2}, \pi_{i3})$  for  $i \geq 1$ .

By satisfying steady state equations, sub vectors  $\Pi_i$ 's are given.

$$\Pi_0 B_{00} + \Pi_1 B_{10} + \Pi_2 (1 - R)^{-1} A_3 = 0 \tag{1}$$

$$\Pi_0 B_{01} + \Pi_1 B_{11} + \Pi_2 A_1 = 0 \tag{2}$$

$$\Pi_0 A_0 + \Pi_2 (B_1 + R A_2) = 0 \tag{3}$$

$$\Pi_i A_0 + \Pi_{i+1} B_2 + \Pi_{i+2} A_2 = 0 \tag{4}$$

And

$$\Pi_i = \Pi_2 R^{i-2} \text{ for } i \geq 3 \tag{5}$$

Let us denote rate matrix be  $R$ . Using matrix equation [22] it is irreducible non-negative solution.

$$R^2 A_2 + R B_2 + A_0 = 0 \tag{6}$$

The normalizing equation is given by

$$\Pi_0 e + \Pi_1 e + \Pi_2 (1 - R)^{-1} e = 1 \tag{7}$$

Here 'e' is unit matrix.

From Eq. (6)

$$R = -[A_0 + R^2 A_2] B_2^{-1} \tag{8}$$

$R$  matrix is found by iterative procedure with  $R(0) = 0$

$$R_{n+1} = -[A_0 + R_n^2 A_2] B_2^{-1} \tag{9}$$

Using the above equation to solve  $R$  matrix and check all values of  $R$  will monotonically increasing. And  $R$  is convergent to  $-B_2^{-1}$  and  $[A_0 + R_n^2 A_2]$  is positive.

By considering matrix geometric procedure the steady state is obtained.  $\Pi$  partitioned as  $\Pi = (\pi_0, \pi_1, \pi_2, \pi_3)$  is a stationary probability vector of the (reducible) matrix generator from  $B = A_2 + B_2 + A_0$ . Then  $B$  can be written as

$$B = \begin{pmatrix} -S_1 & \mu_{2v} + \theta & 0 & \alpha_0 \\ 0 & -S_2 & \mu_{2b} & \alpha_1 \\ v & v & -S_3 & \alpha_2 \\ \beta_0 & \beta_1 & \beta_2 & -S_4 \end{pmatrix}$$

where,

$$\begin{aligned} S_1 &= \mu_2 + \xi + \theta + \alpha_0, \\ S_2 &= \mu_{2b} + \alpha_1 + \xi, \\ S_3 &= 2v + \alpha_2 + \xi, \\ S_4 &= \beta_0 + \beta_1 + \beta_2 + \xi \end{aligned}$$

In MP  $B$  matrix is minimal generator. Also unique equations is fulfilled by  $\Pi$

$$\Pi B = 0 \quad \text{and} \quad \Pi e = 1 \tag{10}$$

We have,

$$\Pi_0 = [1 + \omega_0 + \omega_1 + \omega_2]^{-1} \tag{11}$$

$$\Pi_1 = \omega_0 \Pi_0 \tag{12}$$

$$\Pi_2 = \omega_1 \Pi_0 \tag{13}$$

$$\Pi_3 = \omega_2 \Pi_0 \tag{14}$$

where,

$$\begin{aligned} \omega_0 &= \frac{(S_1 + \mu_{2v} + \theta)(\beta_2 \alpha_2 - S_3 S_4) + (\mu_{2b} \alpha_2 + S_3 \alpha_0)(\beta_0 - \beta_1)}{S_2(\beta_2 \alpha_2 - S_3 S_4) - (\mu_{2b} \alpha_2 + S_3 \alpha_1)(\beta_0 - \beta_1)} \\ \omega_1 &= \frac{S_1}{v} - \frac{S_1 + \mu_{2v} + \theta}{\beta_0 - \beta_1} \end{aligned}$$

$$\omega_2 = \frac{S_2[(S_1 + \mu_{2v} + \theta)(\beta_2\alpha_2 - S_3S_4) + (\mu_{2b}\alpha_2 + S_3\alpha_0)(\beta_0 - \beta_1)]\beta_0}{(\beta_0 - \beta_1)[S_2(\beta_2\alpha_2 - S_3S_4) - (\mu_{2b}\alpha_2 + S_3\alpha_1)(\beta_0 - \beta_1)]\nu}$$

$$= \frac{S_1 + \mu_{2v} + \theta}{\beta_0 - \beta_1}$$

$$- \frac{S_2[(S_1 + \mu_{2v} + \theta)(\beta_2\alpha_2 - S_3S_4) + (\mu_{2b}\alpha_2 + S_3\alpha_0)(\beta_0 - \beta_1)]}{(\beta_0 - \beta_1)[S_2(\beta_2\alpha_2 - S_3S_4) - (\mu_{2b}\alpha_2 + S_3\alpha_1)(\beta_0 - \beta_1)]}$$

Now applying the stability condition  $\Pi A_0 e < \Pi A_2 e$  for QBD process in  $Q$  matrix. We get

$$\lambda[1 + (\omega_0 + \omega_1 + \omega_2)(1 + \Pi_0)] < \mu_1[1 + (\omega_0 + \omega_1 + \omega_2)(1 + \Pi_0)] + \omega_2 \Pi_0 \mu_2 \tag{15}$$

### 3.1 Evaluation of the Results

The evaluation of result is used to find the probabilities of server states expressed below:

- Prob. of system is empty ( $P_{emp}$ ) =  $\Pi_0$ .
- Prob. of server2 is in intermittently available state ( $P_{2int}$ ) =  $\sum_{j=1}^{\infty} \pi_{j2} + \Pi_0$ .
- Prob. of server2 is in working state ( $P_{2working}$ ) =  $1 - P_{2int}$ .
- Prob. of server2 is in Breakdown state ( $P_{2bd}$ ) =  $\sum_{j=1}^{\infty} \pi_{j3}$ .
- Prob. of server2 is in working vacation state ( $P_{2WV}$ ) =  $\sum_{j=1}^{\infty} \pi_{j0} + \Pi_0$ .
- Expected no. of customer in system (MNS)  $\Pi = \Pi_1 e + \Pi_2 (I - R)^{-1} e + \Pi_2 (I - R)^{-2} e$ .

## 4 Numerical Study

In this section, we have illustrated numerical results by the described method of this paper, to gain and understand the execution of parameter  $\lambda$  on the specific probabilistic description characteristics of the system.

We find the probability vectors in varying values of  $\lambda$  and other values are fixed, which are in Tables 1, 2, 3, 4 and 5.

**Example 1** Let us consider  $\lambda = 0.01, \mu_1 = 0.5, \mu_{2v} = 0.4, \mu_{2b} = 0.6, \nu = 0.04, \alpha_0 = 0.3, \alpha_1 = 0.2, \alpha_2 = 0.4, \beta_0 = 0.2, \beta_1 = 0.8, \beta_2 = 0.6, \xi = 0.02$  and  $\theta = 0.7$  and the

Values of  $R$  matrix is

**Table 1** Probability vectors for  $\lambda = 0.01$

$\Pi_i$	$\pi_{i0}$	$\pi_{i1}$	$\pi_{i2}$	$\pi_{i3}$	Total
$\Pi_0$	0.96336				0.96336
$\Pi_1$	0.01426	0.01233	0.00006	0.00002	0.02667
$\Pi_2$	0.00102	0.00007	0.00001	0.00004	0.00114
$\Pi_3$	0.00006	0.00004	0.00000	0.00002	0.00012
$\Pi_4$	0.00000	0.00000	0.00000	0.00000	0.00000
$\Pi_5$	0.00000	0.00000	0.00000	0.00000	0.00000
Total					<b>0.99129</b>

**Table 2** Probability vectors for  $\lambda = 0.05$

$\Pi_i$	$\pi_{i0}$	$\pi_{i1}$	$\pi_{i2}$	$\pi_{i3}$	Total
$\Pi_0$	0.92323				0.92323
$\Pi_1$	0.02033	0.02429	0.01076	0.01135	0.06673
$\Pi_2$	0.00089	0.00145	0.00099	0.00071	0.00404
$\Pi_3$	0.00002	0.00006	0.00007	0.00004	0.00019
$\Pi_4$	0.00001	0.00003	0.00005	0.00002	0.00006
$\Pi_5$	0.00000	0.00000	0.00000	0.00000	0.00000
Total					<b>0.99425</b>

**Table 3** Probability vectors for  $\lambda = 0.1$

$\Pi_i$	$\pi_{i0}$	$\pi_{i1}$	$\pi_{i2}$	$\pi_{i3}$	Total
$\Pi_0$	0.84588				0.84588
$\Pi_1$	0.00307	0.07004	0.00054	0.07534	0.14622
$\Pi_2$	0.00014	0.00379	0.00040	0.00035	0.00469
$\Pi_3$	0.00001	0.00022	0.00007	0.00005	0.00035
$\Pi_4$	0.00000	0.00001	0.00001	0.00000	0.00002
$\Pi_5$	0.00000	0.00000	0.00000	0.00000	0.00000
Total					<b>0.99716</b>

$$R = \begin{pmatrix} 0.00531 & 0.00242 & 0.00068 & 0.00108 \\ 0.00007 & 0.00548 & 0.00037 & 0.00059 \\ 0.00048 & 0.00140 & 0.01138 & 0.00234 \\ 0.00066 & 0.00269 & 0.00348 & 0.00569 \end{pmatrix}$$

Assignee matrix  $R$  in Eqs. (1), (2) and (3) the sub vectors  $\pi_{i0}, \pi_{i1}, \pi_{i2}$  and  $\pi_{i3}$  are obtained. Using the normalization condition  $\Pi_0e + \Pi_1e + \Pi_2[I - R]^{-1}e = 1$  for the numerical rates from above, the vector  $\Pi_2$  is given by  $\Pi_2 = (0.00102, 0.00007, 0.00001, 0.00004)$ . The remaining vectors  $\Pi_i$ 's are obtained from the relation  $\Pi_i =$

**Table 4** Probability vectors for  $\lambda = 0.2$

$\Pi_i$	$\pi_{i0}$	$\pi_{i1}$	$\pi_{i2}$	$\pi_{i3}$	Total
$\Pi_0$	0.72312				0.72312
$\Pi_1$	0.00464	0.19718	0.00329	0.03263	0.23774
$\Pi_2$	0.00050	0.02955	0.00212	0.00011	0.03228
$\Pi_3$	0.00011	0.00325	0.00092	0.00049	0.00477
$\Pi_4$	0.00003	0.00041	0.00029	0.00019	0.00092
$\Pi_5$	0.00001	0.00006	0.00008	0.00003	0.00018
$\Pi_6$	0.00000	0.00001	0.00002	0.00001	0.00004
$\Pi_7$	0.00000	0.00000	0.00001	0.00000	0.00001
$\Pi_8$	0.00000	0.00000	0.00000	0.00000	0.00000
Total					<b>0.99906</b>

**Table 5** Probability vectors for  $\lambda = 0.5$

$\Pi_i$	$\pi_{i0}$	$\pi_{i1}$	$\pi_{i2}$	$\pi_{i3}$	Total
$\Pi_0$	0.42701				0.42701
$\Pi_1$	0.14340	0.17684	0.04599	0.00573	0.37196
$\Pi_2$	0.03360	0.06099	0.00037	0.01092	0.10588
$\Pi_3$	0.00809	0.02132	0.00845	0.00662	0.04448
$\Pi_4$	0.00226	0.00809	0.00732	0.00365	0.02132
$\Pi_5$	0.00079	0.00348	0.00497	0.00203	0.01127
$\Pi_6$	0.00035	0.00169	0.00311	0.00116	0.00631
$\Pi_7$	0.00018	0.00090	0.00188	0.00067	0.00363
$\Pi_8$	0.00010	0.00051	0.00113	0.00040	0.00214
$\Pi_9$	0.00006	0.00029	0.00067	0.00023	0.00125
$\Pi_{10}$	0.00003	0.00017	0.00040	0.00014	0.00074
$\Pi_{11}$	0.00002	0.00010	0.00024	0.00008	0.00042
$\Pi_{12}$	0.00001	0.00006	0.00014	0.00005	0.00026
$\Pi_{13}$	0.00001	0.00004	0.00008	0.00003	0.00016
$\Pi_{14}$	0.00000	0.00002	0.00005	0.00002	0.00009
$\Pi_{15}$	0.00000	0.00001	0.00003	0.00001	0.00005
$\Pi_{16}$	0.00000	0.00001	0.00002	0.00001	0.00004
$\Pi_{17}$	0.00000	0.00000	0.00001	0.00000	0.00001
$\Pi_{18}$	0.00000	0.00000	0.00001	0.00000	0.00001
$\Pi_{19}$	0.00000	0.00000	0.00000	0.00000	0.00000
Total					<b>0.99937</b>

$\Pi_2 R^{i-2}$ ,  $i = 3, 4, 5, \dots$  and shown in Table 1. Columns 2, 3, 4 and 5 are constitute the four values of  $\Pi_i$ ,  $i = 0, 1, 2, 3, \dots$  and the total values of four components given in last column. Finally the total probability is verified as  $0.99129 \approx 1$ .

**Example 2** Let us consider  $\lambda = 0.05$ ,  $\mu_1 = 0.5$ ,  $\mu_{2v} = 0.4$ ,  $\mu_{2b} = 0.6$ ,  $\nu = 0.04$ ,  $\alpha_0 = 0.3$ ,  $\alpha_1 = 0.2$ ,  $\alpha_2 = 0.4$ ,  $\beta_0 = 0.2$ ,  $\beta_1 = 0.8$ ,  $\beta_2 = 0.6$ ,  $\xi = 0.02$  and  $\theta = 0.7$  and the values of  $R$  matrix is

$$R = \begin{pmatrix} 0.002619 & 0.01211 & 0.00378 & 0.00552 \\ 0.00036 & 0.02754 & 0.00233 & 0.00306 \\ 0.00237 & 0.00695 & 0.05628 & 0.01157 \\ 0.00326 & 0.01337 & 0.01731 & 0.02815 \end{pmatrix}$$

Assignee matrix  $R$  in Eqs. (1),(2) and (3) sub vectors  $\pi_{i0}$ ,  $\pi_{i1}$ ,  $\pi_{i2}$  and  $\pi_{i3}$  are obtained. Using the normalization condition  $\Pi_0 e + \Pi_1 e + \Pi_2 [I - R]^{-1} e = 1$  for the numerical rate from above, the vector  $\Pi_2$  is given by  $\Pi_2 = (0.00089, 0.00145, 0.00099, 0.00071)$ . The rest of  $\Pi_i$ 's are found from  $\Pi_i = \Pi_2 R^{i-2}$ ,  $i = 3, 4, 5, \dots$  are shown in Table 2. Remaining columns are constitute four values of  $\Pi_i$ ,  $i = 0, 1, 2, 3, \dots$  and the total values of four components given in last column. Finally the total probability is verified as  $0.99425 \approx 1$ .

**Example 3** Let us consider  $\lambda = 0.1$ ,  $\mu_1 = 0.5$ ,  $\mu_{2v} = 0.4$ ,  $\mu_{2b} = 0.6$ ,  $\nu = 0.04$ ,  $\alpha_0 = 0.3$ ,  $\alpha_1 = 0.2$ ,  $\alpha_2 = 0.4$ ,  $\beta_0 = 0.2$ ,  $\beta_1 = 0.8$ ,  $\beta_2 = 0.6$ ,  $\xi = 0.02$  and  $\theta = 0.7$  and the values of  $R$  matrix is

$$R = \begin{pmatrix} 0.05145 & 0.02456 & 0.00853 & 0.01104 \\ 0.00075 & 0.05409 & 0.00562 & 0.00617 \\ 0.00469 & 0.01480 & 0.01115 & 0.02288 \\ 0.00641 & 0.02708 & 0.03506 & 0.05559 \end{pmatrix}$$

Assignee matrix  $R$  in Eqs. (1), (2) and (3) sub vectors  $\pi_{i0}$ ,  $\pi_{i1}$ ,  $\pi_{i2}$  and  $\pi_{i3}$  are obtained. Using the normalization condition  $\Pi_0 e + \Pi_1 e + \Pi_2 [I - R]^{-1} e = 1$  for the numerical rates from above, the vector  $\Pi_2$  is given by  $\Pi_2 = (0.00014, 0.00379, 0.00040, 0.00035)$ . The remaining vectors  $\Pi_i$ 's are obtained from the relation  $\Pi_i = \Pi_2 R^{i-2}$ ,  $i = 3, 4, 5$  are shown in Table 3. Remaining columns are constitute four values of  $\Pi_i$ ,  $i = 0, 1, 2, 3, \dots$  and the total values of four components given in last column. Finally the total probability is verified as  $0.99716 \approx 1$ .

**Example 4** Let us consider  $\lambda = 0.2$ ,  $\mu_1 = 0.5$ ,  $\mu_{2v} = 0.4$ ,  $\mu_{2b} = 0.6$ ,  $\nu = 0.04$ ,  $\alpha_0 = 0.3$ ,  $\alpha_1 = 0.2$ ,  $\alpha_2 = 0.4$ ,  $\beta_0 = 0.2$ ,  $\beta_1 = 0.8$ ,  $\beta_2 = 0.6$ ,  $\xi = 0.02$  and  $\theta = 0.7$  and the values of  $R$  matrix is

$$R = \begin{pmatrix} 0.09795 & 0.04894 & 0.02035 & 0.02168 \\ 0.00154 & 0.10658 & 0.01507 & 0.01272 \\ 0.00863 & 0.03089 & 0.21626 & 0.04428 \\ 0.01169 & 0.05413 & 0.07133 & 0.10811 \end{pmatrix}$$

Assignee matrix  $R$  in Eqs. (1), (2) and (3) sub vectors  $\pi_{i0}, \pi_{i1}, \pi_{i2}$  and  $\pi_{i3}$  are obtained. Using the normalization condition  $\Pi_0e + \Pi_1e + \Pi_2[I - R]^{-1}e = 1$  for the numerical rates from above, the vector  $\Pi_2$  is given by  $\Pi_2 = (0.00050, 0.02955, 0.00212, 0.00011)$ . The remaining vectors  $\Pi_i$ 's are obtained from the relation  $\Pi_i = \Pi_2R^{i-2}, i = 3, 4, 5$  and shown in Table 4. Remaining columns are constitute four values of  $\Pi_i, i = 0, 1, 2, 3, \dots$  and the total values of four components given in last column. Finally the total probability is verified as  $0.99906 \approx 1$ .

**Example 5** Let us consider  $\lambda = 0.5, \mu_1 = 0.5, \mu_{2v} = 0.4, \mu_{2b} = 0.6, v = 0.04, \alpha_0 = 0.3, \alpha_1 = 0.2, \alpha_2 = 0.4, \beta_0 = 0.2, \beta_1 = 0.8, \beta_2 = 0.6, \xi = 0.02$  and  $\theta = 0.7$  and the values of  $R$  matrix is

$$R = \begin{pmatrix} 0.22332 & 0.12610 & 0.07484 & 0.05371 \\ 0.00472 & 0.25528 & 0.06136 & 0.03386 \\ 0.02003 & 0.08662 & 0.49588 & 0.09951 \\ 0.02712 & 0.013578 & 0.18360 & 0.24878 \end{pmatrix}$$

Assignee matrix  $R$  in Eqs. (1), (2) and (3) sub vectors  $\Pi_0, \Pi_1, \Pi_2$  and  $\Pi_3$  are obtained. Using the normalization condition  $\Pi_0e + \Pi_1e + \Pi_2[I - R]^{-1}e = 1$  for the numerical rates from above, the vector  $\Pi_2$  is given by  $\Pi_2 = (0.03360, 0.06099, 0.00037, 0.01092)$ . The remaining vectors  $\Pi_i$ 's are obtained from the relation  $\Pi_i = \Pi_2R^{i-2}, i = 3, 4, 5$  and shown in Table 5. Remaining columns are constitute four values of  $\Pi_i, i = 0, 1, 2, 3, \dots$  and the total values of four components given in last column. Finally the total probability is verified as  $0.99937 \approx 1$ .

### 4.1 Evaluation of the Results

From the probability vectors, then find some results of evaluation such as prob. of system is empty, prob. of server2 is in regular working, prob. of server2 is in WV, prob. of server2 is in intermittently available, prob. of server2 is in breakdown and MNS. Varying the parameter  $\lambda = 0.01$  to  $0.5$ , the performance measures shown in Table 6.

**Table 6** Effect of  $\lambda$  on performance measures

$\lambda$	0.01	0.05	0.1	0.2	0.5
$P_{emp}$	0.95336	0.92323	0.84588	0.72312	0.42701
$P_{2int}$	0.96343	0.93510	0.84690	0.72985	0.50160
$P_{2working}$	0.03550	0.0649	0.15310	0.27015	0.49840
$P_{2wv}$	0.97870	0.94448	0.84911	0.75358	0.61591
$P_{2bd}$	0.96344	0.93535	0.92162	0.75658	0.45511
MNS	0.02898	0.07544	0.15689	0.31827	0.91885



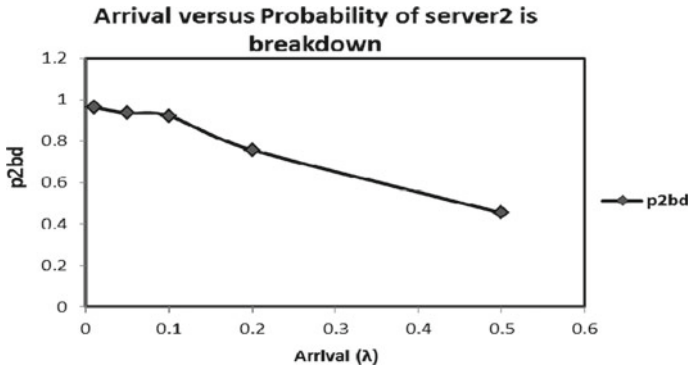


Fig. 2 Arrival versus probability of system is empty

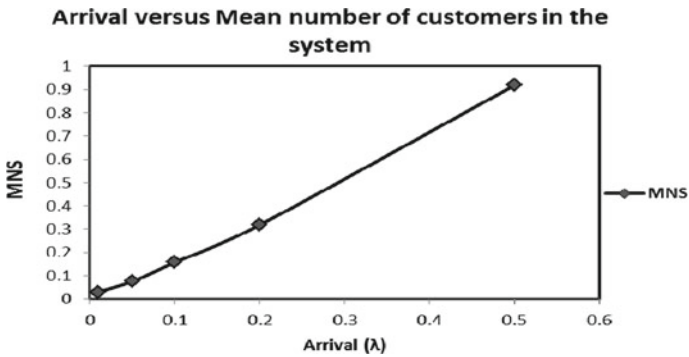


Fig. 3 Arrival versus probability server2 is intermittently available

If arrival rate increasing with parameter  $\lambda$ , then the both  $p_{emp}$  and  $p_{2int}$  decreasing correspondingly, as shown in Figs. 2 and 3. Then  $p_{2working}$  to be increased as exhibit in Fig. 4. Then the probability of server2 is working vacation and breaking downs are decreased as appear in Figs. 5 and 6. And also if arrival rate increasing, then MNS increasing correspondingly as exhibit in Fig. 7.

### 5 Conclusion

In this paper, we studied M/M/2 heterogeneous servers, server1 always available and server2 is intermittently available server and unreliable server. In this model server2 is in working vacation or to vacation interruption. Steady state probability vectors are found using by matrix geometric procedure. Finally analyzed the effect of performance measures with different parameters.

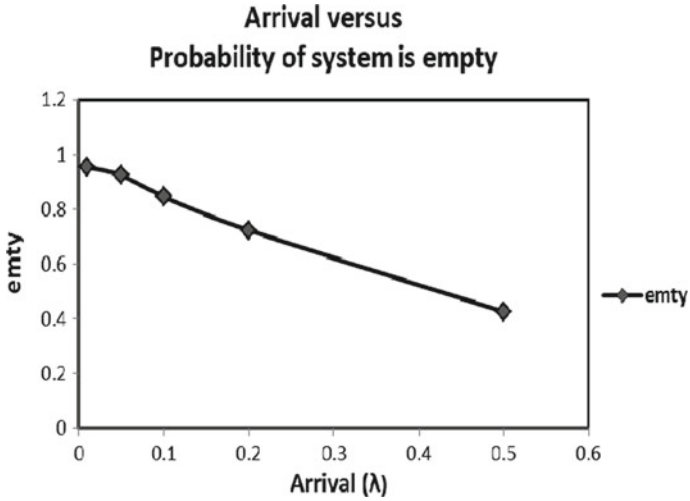


Fig. 4 Arrival versus probability of server2 is working

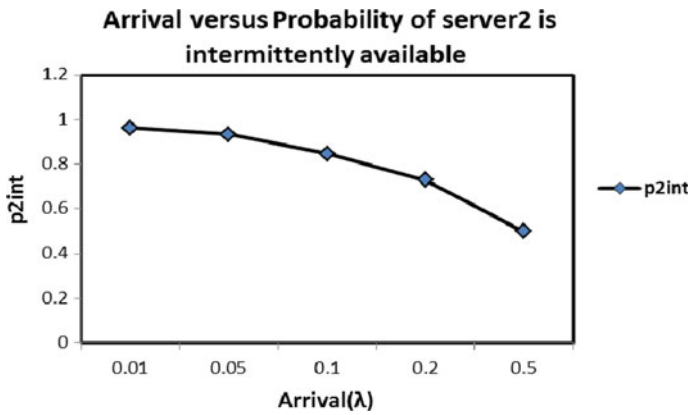


Fig. 5 Arrival versus probability of server2 is working vacation

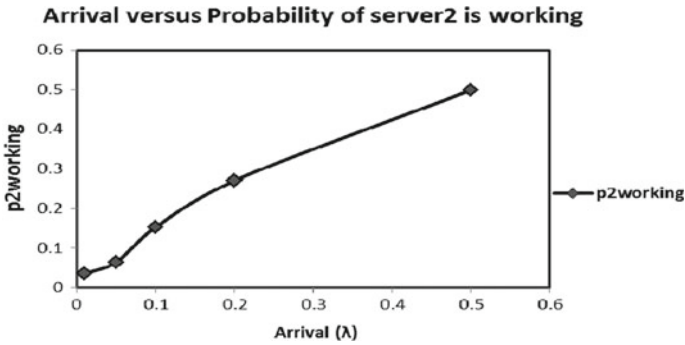


Fig. 6 Arrival versus probability of server2 is breakdown

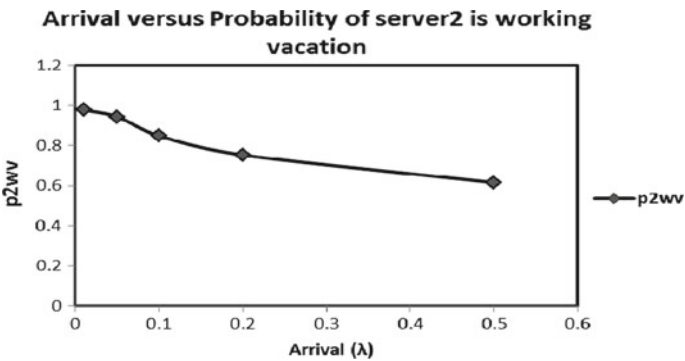


Fig. 7 Arrival versus mean number of customers in the system

## References

1. Chaudhry ML (1972) Single channel constant capacity bulk service queueing process with an intermittently available server. *Infor* 10(3):284–291
2. Ammar Sherif I (2014) Transient analysis of a two-heterogeneous servers queue with impatient behavior. *J Egyptian Mathem Soc* 22:90–95
3. Lin C, Ke J (2005) Multiserver queue with single working vacation. *Appl Math* 33:201–209
4. Seenivasan M, Indumathi M, Chakravarthy VJ (2020) Performance analysis of two heterogeneous server queueing model with intermittently obtainable server using matrix geometric method. *J Phys* 1724
5. Levy Y, Yechiali U (1976) An M/M/s queue with server’s vacation. 14:153–163
6. Shanmugasundaram S, Punith S (2014) A study on multiserver queueing simulation. *J Sci Res* 3(7):1519–1521
7. Chao X (1995) A queueing network model with catastrophes and product from solution. *Operation Res Lett* 18:75–79
8. Vinodhini GAF, Vindhya V (2014) Transient solution of a multi-server queue with catastrophes and impatient customers when system is down. *Appl Mathem Sci* 8(92):4584–4592
9. Indra, Rajan V (2017) Queueing analysis of Markovian queue having two heterogeneous servers with catastrophes using matrix geometric technique. *Int J Statist Syst* 12(2):205–212 ISSN., 0973–2675

10. Tang Y (1997) Single server M/G/1 queueing system subject to breakdown some reliability and queueing problems. *Microelectron Reliab* 37:315–321
11. Doshi BT (1986) Queueing system with vacations—a survey. *Queueing Syst* 1:29–66
12. Takagi H (1991) *Queueing analysis, a foundation of performance analysis of vacation and priority systems, part 1. vol 1.* Elsevier Science Publishers
13. Servi LD, Finn SG (2002) M/M/1 queue with working vacation (M/M/1/MW). *Perf Eval* 50:41–52
14. Jain M, Jain A (2010) Working vacations queueing models with multiple types of server breakdowns. *Appl Math Model* 34(1):1–30
15. Vinod B (1985) Unreliable queueing systems. *Comput Oper Res* 12:322–340
16. Aissani A, Artalejo JR (1998) On the single server retrial queue subject to breakdown. *Queueing Sys* 30:309–321
17. Seenivasan M, Indumathi M (2018) A retrial queueing model with unreliable server in K policy. *Adv Algebra Anal-Trends Mathem* 361–372
18. Tarabia AMK (2011) Transient and steady-state analysis of an M/M/1 queue with balking, catastrophes, server failures and repairs. *J Ind Manag Optim* 7:811823
19. Wang J, Li J (2008) Transient analysis of an M/G/1 retrial queue subject to disasters and server failures. *Eur J Oper Res* 189:1118–1132
20. Alfa A (2003) Vacation models in discrete time. *Queueing Syst* 44:5–30
21. Boudali O, Economou A (2012) Optimal and equilibrium balking strategies in the single server Markovian queue with catastrophes. *Eur J Oper Res* 218(3):708715
22. Neuts MF (1981) *Matrix-geometric solution in stochastic models, Johns Hopkins series in the mathematical series., vol 2.* Johns Hopkins University Press, Baltimore, USA

# Performance Analysis of Single Server Low-priority Queue Based on Electronic Transmitter



R. Ramesh and M. Seenivasan

## 1 Introduction

A transmitter is a small discrete electronic apparatus. A transceiver contains the transmitter and a receiver. The transmitters are frequently contracted as “XMTR” or “TX” by technological archives. On over a length, the transmitter intends the radio communication of data. The transmitter collects the inputs for to transmit by the electronic sign structures, namely a voice gesture, a visual gesture and binary codes gestures from an analog computer, and so on.... The transmitters did the modulation process with the help of transmission frequent waves (transform gesture). All of these gestures have been increased to transform on various assorted combinations. More modulation forms are also used. Here, the signals (data) are providing (arrival) to the transmitter by the form of a queue in the Poisson process, and it transmits (service) the data by exponential distribution (Fig. 1).

For the time being, we confront huge priority-based issues in the queueing ambient which are transmission of data, hospitals, banks, medical shops, petrol stations, dialing the phone calls, and so on.... Underscore the significance of time administration is the uttermost concentrate of all investigators. The queueing models are taking very emphatic role in this present situation. The definitions, theorems [3, 5, 9], and queueing models are hugely momentous for our analysis determination. On this perpetual environment, frequently, we pertain the fuzzy logic and applications [10, 22]. Once in a while, we touch the priority conjecture [15] in the queueing situations.

---

R. Ramesh

Department of Mathematics, Arignar Anna Government Arts College, Musiri, Tamil Nadu, India

M. Seenivasan (✉)

Mathematics Wings—DDE, Annamalai University, Annamalainagar, Chidambaram, Tamil Nadu, India

e-mail: [emseeni@yahoo.com](mailto:emseeni@yahoo.com)

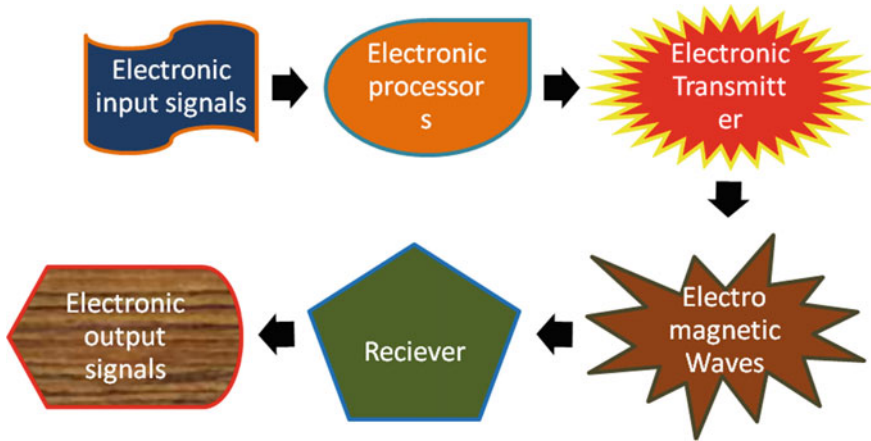


Fig. 1 Process of electronic transmitter

In some tasks, the priorities are accepted immediately, and in some other tasks, it takes more time. These performances could be computed with fuzzy syllogism [7, 13].

Altogether, the customers of class 1 (with priority) and customers of class 2 (without priority) are serviced by same server. But the cost measures are entirely different for these customers [19]. In this everlasting situation, the preemptive priority-based customers are basically intended as the customers of special class 1. This was the additional privilege for the above-mentioned customers. But the non-preemptive priority-based customers [14, 16] are always normal in consideration. Howbeit, practically, these customers of class 1 are finer than class 2 customers.

Multifarious authors have so far employed multihued ranking techniques [1, 2, 4, 6, 12] for finding the fuzzy queues performances. Robust ranking technique [11] is weighed as a well-known ranking technique. But the distance-based ranking methods are very fashionable [16–18, 20, 21].

The centroid of incenters [8] ranking method of working is utilized for estimating the electronic transmitter’s performances in the shape of queuing models with fuzzy parameters. This strategy is vastly contented for appraising the concrete classical utilities for these queueing models.

## 2 Preambles

### 2.1 Fuzzy Set

The fuzzy set  $\tilde{A} = \{(x, \phi_{\tilde{A}}(x)); x \in U\}$  can be figured by the membership function  $\phi_{\tilde{A}}$  tends the members of universal set  $U$  to unit segment  $[0, 1]$ . (i.e.) The mapping

$\phi_{\tilde{A}} : U \rightarrow [0, 1]$   $[0, 1]$  is named as the degree of membership function of  $\tilde{A}$  and  $\phi_{\tilde{A}}(x)$  is termed as the membership utility for  $x \in U$  in the fuzzy set  $\tilde{A}$ .

### 2.2 Triangular Fuzzy Number

The triangular fuzzy number  $\tilde{A}(x)$  is represented by  $\tilde{A}(a_1, a_2, a_3; 1)$  with the membership function

$$\phi_{\tilde{A}}(x) = \left\{ \begin{array}{ll} \frac{x-a_1}{a_2-a_1}, & a_1 \leq x \leq a_2 \\ 1, & x = a_2 \\ \frac{x-a_3}{a_2-a_3}, & a_2 \leq x \leq a_3 \\ 0_x & \text{or else} \end{array} \right\}$$

### 2.3 Trapezoidal Fuzzy Number

The trapezoidal fuzzy number  $\tilde{A}(x)$  is represented by  $\tilde{A}(a_1, a_2, a_3, a_4; 1)$  and its membership function

$$\phi_{\tilde{A}}(x) = \left\{ \begin{array}{ll} \frac{x-a_1}{a_2-a_1}, & a_1 \leq x \leq a_2 \\ 1, & a_2 \leq x \leq a_3 \\ \frac{x-a_4}{a_3-a_4}, & a_3 \leq x \leq a_4 \\ 0 & \text{or else} \end{array} \right\}$$

## 3 Non-preemptive Priority Fuzzy Queues

Let us consider the  $K$ -th priority signals (data) appeared in a single passage queue with Poisson stream and their rate in fuzzy parameter be  $\tilde{\lambda}_k$ , ( $k = 1, 2, 3, \dots, r$ ). The service is done by the exponential distribution with rate in fuzzy parameter  $\tilde{\mu}$ . This service is followed by FCFS discipline with their respective priorities. Appraise the dispersal of service for this priority is exponential mean rate  $1/\tilde{\mu}_k$  units. Without break of priorities, the service should be begun and finished with in a new entry is allowed. All of we know that  $\rho_k = \frac{\lambda_k}{\mu_k}$  ( $1 \leq k < r$ ),  $\sigma_k = \sum_{i=1}^k \rho_i$  ( $\sigma_0 \approx 0, \sigma_r \approx \rho$ ). At  $\sigma_r = \rho < 1$ , the system will be stationary. Let  $\phi_{\tilde{\lambda}}, \phi_{\tilde{\mu}}$  indicate the membership functions for  $\tilde{\lambda}, \tilde{\mu}$ . Then, we own the following fuzzy sets:  $\tilde{\lambda}_i = \{(s, \phi_{\tilde{\lambda}}(s)) | s \in S$  and  $\tilde{\mu} = \{(t, \phi_{\tilde{\mu}}(t)) | t \in T$ , where  $S$  and  $T$  are crisp rates of arrival and service individually.

Consider  $F(S, T)$  be the system habit of attentiveness and  $s$  and  $t$  are fuzzy numerals. Now, let us consider the performance expedients of priority queues (non-preemptive). By using steady-state contexts  $\rho_k = \frac{\lambda_k}{\mu_k} < 1$ , and all other elemental queuing theory formulas.

The mean number of units in the queue  $L_q = \sum_{i=1}^r L_q^{(i)} = \sum_{i=1}^r \frac{\lambda_i \sum_{k=1}^r \frac{\rho_k}{\mu_k}}{(1-\sigma_{i-1})(1-\sigma_i)}$

The time spent by an unit in queue  $W_q = \sum_{i=1}^r \frac{\lambda_i w_q^{(i)}}{\lambda}$ , where  $w_q^{(i)} = \frac{\sum_{k=1}^r \frac{\rho_k}{\mu_k}}{(1-\sigma_{i-1})(1-\sigma_i)}$

Contemplate a system of four priority queues (i.e., the arrival rates  $\tilde{\lambda}_i, i = 1, 2, 3, 4$ ) with single server.

By this conception, the FM/FM/1/L queue will be minimized as M/M/1 model along with uniform rates of service.

(i.e.)  $\tilde{\mu}_1 = \tilde{\mu}_2 = \tilde{\mu}_3 = \tilde{\mu}_4 = \tilde{\mu}$ ,

We knew that

$$\tilde{\rho}_1 = \frac{\tilde{\lambda}_1}{\tilde{\mu}_1}, \tilde{\rho}_2 = \frac{\tilde{\lambda}_2}{\tilde{\mu}_2}, \tilde{\rho}_3 = \frac{\tilde{\lambda}_3}{\tilde{\mu}_3} \text{ and } \tilde{\rho}_4 = \frac{\tilde{\lambda}_4}{\tilde{\mu}_4}$$

Since  $\tilde{\rho} = \tilde{\rho}_1 + \tilde{\rho}_2 + \tilde{\rho}_3 + \tilde{\rho}_4$ ,

$$\tilde{\lambda} = \tilde{\lambda}_1 + \tilde{\lambda}_2 + \tilde{\lambda}_3 + \tilde{\lambda}_4; \tilde{\rho} = \frac{\tilde{\lambda}_1 + \tilde{\lambda}_2 + \tilde{\lambda}_3 + \tilde{\lambda}_4}{\tilde{\mu}}$$

Also

$$w_q^{(i)} = \frac{\frac{\tilde{\rho}_1 + \tilde{\rho}_2 + \tilde{\rho}_3 + \tilde{\rho}_4}{\tilde{\mu}}}{(1 - \sigma_{i-1})(1 - \sigma_i)} \text{ and } L_q^{(i)} = \frac{(\tilde{\rho}_1 + \tilde{\rho}_2 + \tilde{\rho}_3 + \tilde{\rho}_4) \tilde{\rho}_i}{(1 - \sigma_{i-1})(1 - \sigma_i)}$$

The average number of signals (data) in the system

$$L_s = \frac{L_q}{\rho} = \frac{\sum_{i=1}^v L_q^{(i)}}{\sum_{i=1}^v \rho_i}$$

The waiting time of each data in the system

$$W_s = \frac{L_s}{\lambda} = \frac{\sum_{i=1}^v L_s^{(i)}}{\sum_{i=1}^v \lambda_i}$$

From which we deduce that

$$w_q^{(1)} = \frac{\tilde{\rho}}{\tilde{\mu} - \tilde{\lambda}_1}; w_q^{(2)} = \frac{\tilde{\rho}}{(1 - \tilde{\rho}_1 - \tilde{\rho}_2)(\tilde{\mu} - \tilde{\lambda}_1)};$$

$$w_q^{(3)} = \frac{\tilde{\rho}}{(1 - \tilde{\rho}_1 - \tilde{\rho}_2 - \tilde{\rho}_3)(\tilde{\mu} - \tilde{\lambda}_1 - \tilde{\lambda}_2)}$$



$$\begin{aligned}
 w_q^{(4)} &= \frac{\tilde{\rho}}{(1 - \tilde{\rho}_1 - \tilde{\rho}_2 - \tilde{\rho}_3)(\tilde{\mu} - \tilde{\lambda})}; \\
 L_q^{(1)} &= \frac{\tilde{\rho} \tilde{\rho}_1}{(1 - \tilde{\rho}_1)}; L_q^{(2)} = \frac{\tilde{\rho} \tilde{\rho}_2}{(1 - \tilde{\rho}_1 - \tilde{\rho}_2)(1 - \tilde{\rho}_1)}; \\
 L_q^{(3)} &= \frac{\tilde{\rho} \tilde{\rho}_3}{(1 - \tilde{\rho}_1 - \tilde{\rho}_2 - \tilde{\rho}_3)(1 - \tilde{\rho}_1 - \tilde{\rho}_2)}; \\
 L_q^{(4)} &= \frac{\tilde{\rho} \tilde{\rho}_4}{(1 - \tilde{\rho}_1 - \tilde{\rho}_2 - \tilde{\rho}_3)(1 - \tilde{\rho})}; \\
 L_s^{(1)} &= \frac{L_g^{(1)}}{\tilde{\rho}_1}; L_s^{(2)} = \frac{L_g^{(2)}}{\tilde{\rho}_2}; L_s^{(3)} = \frac{L_g^{(3)}}{\tilde{\rho}_3}; L_s^{(4)} = \frac{L_g^{(4)}}{\tilde{\rho}_4}; \\
 W_s^{(1)} &= \frac{L_g^{(1)}}{\lambda_1}; W_s^{(2)} = \frac{L_g^{(2)}}{\lambda_2}; W_s^{(3)} = \frac{L_g^{(3)}}{\lambda_3}; W_s^{(4)} = \frac{L_g^{(4)}}{\lambda_4}
 \end{aligned}$$

### 4 Centroid of Incenters Ranking Method

On numerous optimization sums, the “fuzzy numerals sort technique” takes a huge contribution in the world. In the past periods, more inventors were used centroid grounded routine to sort the fuzzy numerals. Because the centroid is an equipoise tip (gravity point) for every planar shape. Here, we suggest a new incenter-based ranking approach for attaining the performances of priority queues (non-preemptive) under fuzzy environment. Consider the trapezoid PABS, and this subdivided toward ΔPAC, ΔCBS, and ΔPCS. We get  $I_1, I_2,$  and  $I_3$  (incenters) from ΔPAC, ΔCBS, and ΔPCS. We made a Δ ( $I_1 I_2 I_3$ ) with these non-collinear incenters. Then, we can find the centroid G from Δ ( $I_1 I_2 I_3$ ) which is contemplated as the reference point for defining the ranking function. G will be equidistant to  $I_1, I_2,$  and  $I_3$  (Fig. 2).

Account a generalized trapezoidal fuzzy number  $\tilde{A} = (a_1, a_2, a_3, a_4; w)$ .

The incenters of the triangles PAC, CBS, and PCS are

$$\begin{aligned}
 I_1(x_{i1}, y_{i1}) &= \left( \frac{a_1\alpha_1 + a_2\beta_1 + \frac{a_2+a_3}{2}\gamma_1}{\alpha_1 + \beta_1 + \gamma_1}, \frac{w(\beta_1 + \gamma_1)}{\alpha_1 + \beta_1 + \gamma_1} \right); \\
 I_2(x_{i2}, y_{i2}) &= \left( \frac{a_1\alpha_2 + \frac{a_2+a_3}{2}\beta_2 + a_4\gamma_2}{\alpha_2 + \beta_2 + \gamma_2}, \frac{w\beta_2}{\alpha_2 + \beta_2 + \gamma_2} \right) \text{ and} \\
 I_3(x_{i3}, y_{i3}) &= \left( \frac{\frac{a_2+a_3}{2}\alpha_3 + a_3\beta_3 + a_4\gamma_3}{\alpha_3 + \beta_3 + \gamma_3}, \frac{w\beta_2}{\alpha_3 + \beta_3 + \gamma_3} \right)
 \end{aligned}$$

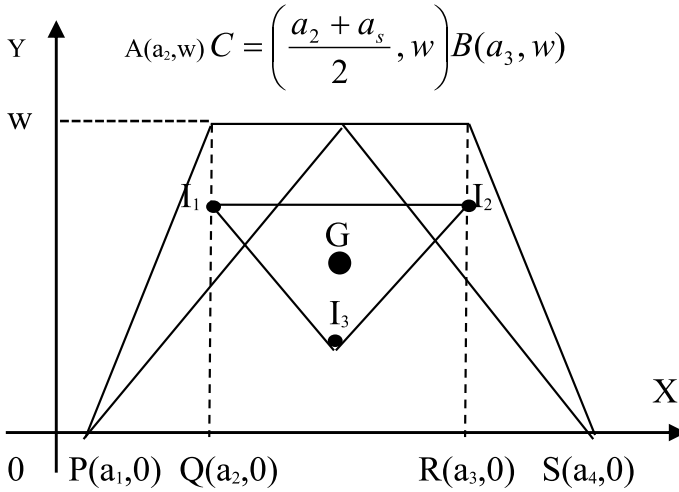


Fig. 2 Centroid of incenters

respectively.  
Where

$$\alpha_1 = (a_3 - a_2);$$

$$\beta_1 = \frac{1}{2}\sqrt{(a_2 + a_3 - 2a_1)^2 + 4W^2};$$

$$\gamma_1 = \sqrt{(a_2 - 2a_1)^2 + W^2};$$

$$\alpha_2 = \frac{1}{2}\sqrt{(2a_4 - a_2 - a_3)^2 + 4W^2};$$

$$\beta_2 = (a_4 - a_1);$$

$$\gamma_2 = \frac{1}{2}\sqrt{(a_2 + a_3 - 2a_1)^2 + 4W^2};$$

$$\alpha_3 = \sqrt{(a_4 - a_3)^2 + W^2};$$

$$\beta_3 = \frac{1}{2}\sqrt{(2a_4 - a_2 - a_3)^2 + 4W^2};$$

$$\gamma_3 = (a_3 - a_2)$$

The centroid  $G_{\tilde{A}}(\tilde{x}_0, \tilde{y}_0)$  of  $\Delta (I_1 I_2 I_3)$  with incenters  $I_1, I_2,$  and  $I_3$  of the generalized trapezoidal fuzzy number.

$$\tilde{A} = (a_1, a_2, a_3, a_4; w) \text{ as } G_{\tilde{A}}(\tilde{x}_0, \tilde{y}_0) = \left( \frac{x_{i1}+x_{i2}+x_{i3}}{3}, \frac{y_{i1}+y_{i2}+y_{i3}}{3} \right)$$

The ranking function for the generalized trapezoidal fuzzy number.

$\tilde{A} = (a_1, a_2, a_3, a_4; w)$  which is mapping every fuzzy number to real numerals is termed by  $R(\tilde{A}) = \tilde{x}_0 * \tilde{y}_0$ .

As a special case, for a generalized triangular fuzzy number  $\tilde{A} = (a_1, a_2, a_4; w)$ . (i.e.,  $a_2 = a_3$ ), the centroid  $G_A$ , with incenters  $I_1, I_2,$  and  $I_3$  as

$$G_{\tilde{A}}(\tilde{x}_0, \tilde{y}_0) = \left( \frac{x_{i1}+x_{i2}+x_{i3}}{3}, \frac{y_{i1}+y_{i2}+y_{i3}}{3} \right).$$

The ranking function of the generalized triangular fuzzy number.

$\tilde{A} = (a_1, a_2, a_4; w)$  which maps the set of all fuzzy numbers to a set of real numbers is defined as  $R(\tilde{A}) = \tilde{x}_0 * \tilde{y}_0$ .

### 5 Real-life Application

In this pandemic situation of COVID-19, disease cases news take place in KODAI FM 100.5 radio station, where some of the news are awaiting to transmit. At this state, each news is under exigency to transmit, and the radio station allows (non-preemptive priority only) the COVID-19 disease cases news only. But some of the news agents want to transmit their data with some low priorities, especially for festival business advertisements. In the present critical moment, we are calculating the mean waiting duration, moderate queue length along four priorities. Besides, we are analyzing how the above quantities have prolonged by those priorities.

#### 5.1 For Trapezoidal Fuzzy Number

Hold the data providing rates for 1st, 2nd, 3rd, and 4th priority entities are

$\tilde{\lambda}_1 = [1, 2, 4, 5:1], \tilde{\lambda}_2 = [2, 3, 5, 6:1], \tilde{\lambda}_3 = [3, 4, 6, 7:1], \tilde{\lambda}_4 = [4, 5, 7, 8:1]$  and clone transmission rate  $\tilde{\mu} = [22, 23, 25, 26:1]$  all per hour at once.

The membership function for  $[1, 2, 4, 5:1]$  can be

$$\varphi_{\tilde{\lambda}}^-(x) = \begin{cases} \frac{(x-1)}{(2-1)}, & 1 \leq x \leq 2 \\ 1, & 2 \leq x \leq 4 \\ \frac{(x-5)}{(4-5)}, & 4 \leq x \leq 5 \\ 0, & \text{or else} \end{cases}$$

Alike, all the remaining trapezoidal fuzzy numbers are followed by the similar form. At present, the ranking indices of  $\tilde{\lambda}$  and  $\tilde{\mu}$  by our proposed approach are

$$\begin{aligned}
 R(\tilde{\lambda}_1) &= R(1, 2, 4, 5:1) = \left(\frac{1.89 + 3 + 5.35}{3}\right) \times \left(\frac{0.64 + 0.47 + 0.64}{3}\right) = 1.97 \\
 R(\tilde{\lambda}_2) &= R(2, 3, 5, 6:1) = \left(\frac{2.89 + 4 + 5.1}{3}\right) \times \left(\frac{0.64 + 0.47 + 0.64}{3}\right) = 2.006 \\
 R(\tilde{\lambda}_3) &= R(3, 4, 6, 7:1) = \left(\frac{3.89 + 5 + 6.1}{3}\right) \times \left(\frac{0.64 + 0.47 + 0.64}{3}\right) = 2.89 \\
 R(\tilde{\lambda}_4) &= R(4, 5, 7, 8:1) = \left(\frac{4.89 + 6 + 7.1}{3}\right) \times \left(\frac{0.64 + 0.47 + 0.64}{3}\right) = 3.47 \\
 R(\tilde{\mu}) &= R(22, 23, 25, 26:1) = \left(\frac{22.89 + 24 + 25.1}{3}\right) \times \left(\frac{0.64 + 0.47 + 0.64}{3}\right) = 1.97
 \end{aligned}$$

and

$$\begin{aligned}
 R(\tilde{\lambda}) &= 10.26, R(\tilde{\rho}_1) = 0.141, R(\tilde{\rho}_2) = 0.144, \\
 R(\tilde{\rho}_3) &= 0.2, R(\tilde{\rho}_4) = 0.24, R(\tilde{\rho}) = 0.74
 \end{aligned}$$

Now,

Mean waiting duration of 1st priority in the queue =  $w_q^{(1)} = \frac{\tilde{\rho}}{\tilde{\mu} - \tilde{\lambda}_1} = 0.061$

Mean waiting duration of 2nd priority in the queue =  $w_q^{(2)} = \frac{\tilde{\rho}}{(1 - \tilde{\rho}_1 - \tilde{\rho}_2)(\tilde{\mu} - \tilde{\lambda}_1)} = 0.086$

Mean waiting duration of 3rd priority in the queue =  $w_q^{(3)} = \frac{\tilde{\rho}}{(1 - \tilde{\rho}_1 - \tilde{\rho}_2 - \tilde{\rho}_3)(\tilde{\mu} - \tilde{\lambda}_1 - \tilde{\lambda}_2)} = 0.144$

Mean waiting duration of 4th priority in the queue =  $w_q^{(4)} = \frac{\tilde{\rho}}{(1 - \tilde{\rho}_1 - \tilde{\rho}_2 - \tilde{\rho}_3)(\tilde{\mu} - \tilde{\lambda})} = 0.393$

The mean number of data in the queue about 1st priority =  $L_q^{(1)} = \frac{\tilde{\rho} \tilde{\rho}_1}{(1 - \tilde{\rho}_1)} = 0.121$

The mean number of data in the queue about 2nd priority =

$$L_q^{(2)} = \frac{\tilde{\rho} \tilde{\rho}_2}{(1 - \tilde{\rho}_1 - \tilde{\rho}_2)(1 - \tilde{\rho}_1)} = 0.173$$

The mean number of data in the queue about 3rd priority =

$$L_q^{(3)} = \frac{\tilde{\rho} \tilde{\rho}_3}{(1 - \tilde{\rho}_1 - \tilde{\rho}_2 - \tilde{\rho}_3)(1 - \tilde{\rho}_1 - \tilde{\rho}_2)} = 0.402$$

The mean number of data in the queue about 4th priority =

$$L_q^{(4)} = \frac{\tilde{\rho} \tilde{\rho}_4}{(1 - \tilde{\rho}_1 - \tilde{\rho}_2 - \tilde{\rho}_3)(1 - \tilde{\rho})} = 1.326$$

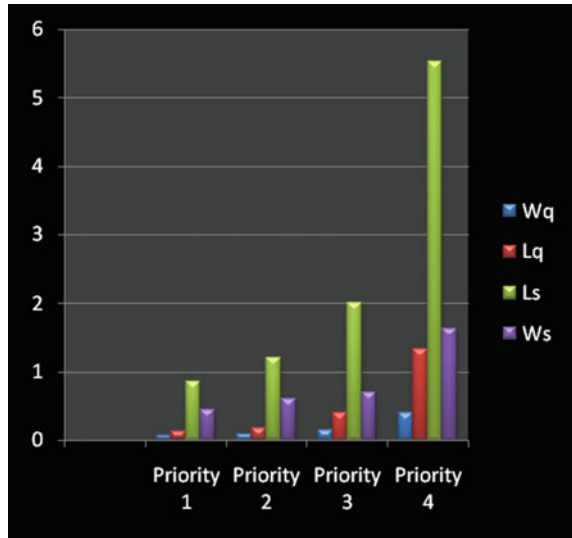
The mean number of data in the system about 1st priority =  $L_s^{(1)} = \frac{L_q^{(1)}}{\tilde{\rho}_1} = 0.858$

The mean number of data in the system about 2nd priority =  $L_s^{(2)} = \frac{L_q^{(2)}}{\tilde{\rho}_2} = 1.2$

**Table 1** Performance measures of  $W_q, L_q, L_s,$  and  $W_s$  (units)

Priorities	$W_q$ (units)	$L_q$ (units)	$L_s$ (units)	$W_s$ (units)
1	0.061	0.121	0.858	0.44
2	0.086	0.173	1.2	0.6
3	0.144	0.402	2.01	0.7
4	0.393	1.326	5.53	1.62

**Fig. 3** Performance measures of  $W_q, L_q, L_s,$  and  $W_s$  (units)



The mean number of data in the system about 3rd priority =  $L_s^{(3)} = \frac{L_g^{(3)}}{\rho_3} = 2.01$

The mean number of data in the system about 4th priority =  $L_s^{(4)} = \frac{L_g^{(4)}}{\rho_4} = 5.53$

Mean waiting duration of 1st priority in the system =  $W_s^{(1)} = \frac{L_g^{(1)}}{\lambda_1} = 0.44$

Mean waiting duration of 2nd priority in the system =  $W_s^{(2)} = \frac{L_g^{(2)}}{\lambda_2} = 0.6$

Mean waiting duration of 3rd priority in the system =  $W_s^{(3)} = \frac{L_g^{(3)}}{\lambda_3} = 0.7$

Mean waiting duration of 4th priority in the system =  $W_s^{(4)} = \frac{L_g^{(4)}}{\lambda_4} = 1.62$  (Table 1 and Fig. 3).

### 5.2 For Triangular Fuzzy Number

Conceive the data providing rates for 1st, 2nd, 3rd, and 4th priority entities are

$\tilde{\lambda}_1 = [1, 4, 5 : 1], \tilde{\lambda}_2 = [2, 5, 6 : 1], \tilde{\lambda}_3 = [3, 6, 7 : 1], \tilde{\lambda}_4 = [4, 7, 8 : 1]$  and clone transmission rate  $\tilde{\mu} = [22, 25, 26:1]$  all per hour at once.

The membership function for [1, 4, 5:1] can be

$$\varphi_{\lambda}^{-}(x) = \begin{cases} \frac{(x-1)}{(4-1)}, & 1 \leq x \leq 4 \\ 1, & x = 4 \\ \frac{(x-5)}{(4-5)} & 4 \leq x \leq 5 \\ 0, & \text{or else} \end{cases}$$

Furthermore, all the remaining trapezoidal fuzzy numbers are followed by the similar form. At present, the ranking indices of  $\tilde{\lambda}$  and  $\tilde{\mu}$  by our proposed approach are

$$\begin{aligned} R(\tilde{\lambda}_1) &= R(1, 2, 5 : 1) = \left(\frac{2 + 2.12 + 2}{3}\right) \times \left(\frac{1 + 0.47 + 1}{3}\right) = 1.67 \\ R(\tilde{\lambda}_2) &= R(2, 3, 6 : 1) = \left(\frac{3 + 3.12 + 3}{3}\right) \times \left(\frac{1 + 0.47 + 1}{3}\right) = 2.49 \\ R(\tilde{\lambda}_3) &= R(3, 4, 7 : 1) = \left(\frac{4 + 4.12 + 4}{3}\right) \times \left(\frac{1 + 0.47 + 1}{3}\right) = 3.31 \\ R(\tilde{\lambda}_4) &= R(4, 5, 8 : 1) = \left(\frac{5 + 5.12 + 5}{3}\right) \times \left(\frac{1 + 0.47 + 1}{3}\right) = 4.13 \\ R(\tilde{\mu}) &= R(22, 23, 26 : 1) = \left(\frac{23 + 23.12 + 23}{3}\right) \times \left(\frac{1 + 0.47 + 1}{3}\right) = 18.89 \end{aligned}$$

and

$$\begin{aligned} R(\tilde{\lambda}) &= 11.42, R(\tilde{\rho}_1) = 0.088, R(\tilde{\rho}_2) = 0.131, R(\tilde{\rho}_3) = 0.175, \\ R(\tilde{\rho}_4) &= 0.209, R(\tilde{\rho}) = 0.604 \end{aligned}$$

From basic queuing theory formulas,

Mean waiting duration of 1st priority in the queue =  $w_q^{(1)} = \frac{\tilde{\rho}}{\tilde{\mu} - \tilde{\lambda}_1} = 0.035$   
 Mean waiting duration of 2nd priority in the queue =

$$w_q^{(2)} = \frac{\tilde{\rho}}{(1 - \tilde{\rho}_1 - \tilde{\rho}_2)(\tilde{\mu} - \tilde{\lambda}_1)} = 0.044$$

Mean waiting duration of 3rd priority in the queue =

$$w_q^{(3)} = \frac{\tilde{\rho}}{(1 - \tilde{\rho}_1 - \tilde{\rho}_2 - \tilde{\rho}_3)(\tilde{\mu} - \tilde{\lambda}_1 - \tilde{\lambda}_2)} = 0.068$$

Mean waiting duration of 4th priority in the queue =

$$w_q^{(4)} = \frac{\tilde{\rho}}{(1 - \tilde{\rho}_1 - \tilde{\rho}_2 - \tilde{\rho}_3)(\tilde{\mu} - \lambda)} = 0.80$$

The mean number of data in the queue about 1st priority =  $L_q^{(1)} = \frac{\tilde{\rho} \tilde{\rho}_1}{(1 - \tilde{\rho}_1)} = 0.058$   
 The mean number of data in the queue about 2nd priority =

$$L_q^{(2)} = \frac{\tilde{\rho} \tilde{\rho}_2}{(1 - \tilde{\rho}_1 - \tilde{\rho}_2)(1 - \tilde{\rho}_1)} = 0.110$$

The mean number of data in the queue about 3rd priority =

$$L_q^{(3)} = \frac{\tilde{\rho} \tilde{\rho}_3}{(1 - \tilde{\rho}_1 - \tilde{\rho}_2 - \tilde{\rho}_3)(1 - \tilde{\rho}_1 - \tilde{\rho}_2)} = 0.223$$

The mean number of data in the queue about 4th priority =

$$L_q^{(4)} = \frac{\tilde{\rho} \tilde{\rho}_4}{(1 - \tilde{\rho}_1 - \tilde{\rho}_2 - \tilde{\rho}_3)(1 - \tilde{\rho})} = 0.527$$

The mean number of data in the system about 1st priority =  $L_s^{(1)} = \frac{L_g^{(1)}}{\tilde{\rho}_1} = 0.61$

The mean number of data in the system about 2nd priority =  $L_s^{(2)} = \frac{L_g^{(2)}}{\tilde{\rho}_2} = 0.84$

The mean number of data in the system about 3rd priority =  $L_s^{(3)} = \frac{L_g^{(3)}}{\tilde{\rho}_3} = 1.27$

The mean number of data in the system about 4th priority =  $L_s^{(4)} = \frac{L_g^{(4)}}{\tilde{\rho}_4} = 2.52$

Mean waiting duration of 1st priority in the system =  $W_s^{(1)} = \frac{L_g^{(1)}}{\lambda_1} = 0.31$

Mean waiting duration of 2nd priority in the system =  $W_s^{(2)} = \frac{L_g^{(2)}}{\lambda_2} = 0.34$

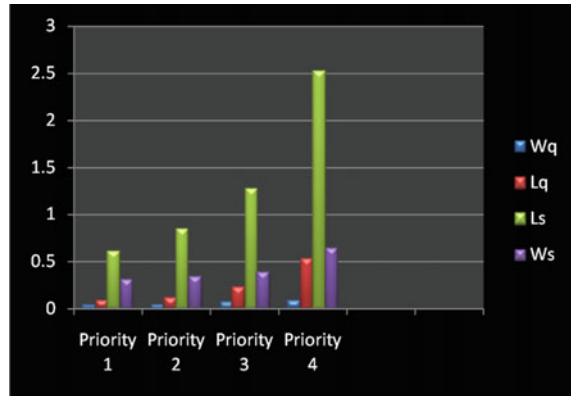
Mean waiting duration of 3rd priority in the system =  $W_s^{(3)} = \frac{L_g^{(3)}}{\lambda_3} = 0.38$

Mean waiting duration of 4th priority in the system =  $W_s^{(4)} = \frac{L_g^{(4)}}{\lambda_4} = 10.64$  (Table 2 and Fig. 4).

**Table 2** Performance measures of  $W_q, L_q, L_s,$  and  $W_s$  (units)

Priorities	$W_q$ (units)	$L_q$ (units)	$L_s$ (units)	$W_s$ (units)
1	0.035	0.088	0.61	0.31
2	0.044	0.110	0.84	0.34
3	0.068	0.223	1.27	0.38
4	0.080	0.527	2.52	0.64

**Fig. 4** Performance measures of  $W_q$ ,  $L_q$ ,  $L_s$ , and  $W_s$  (units)



## 6 Outcome

This article has diagnosed with afresh blueprint for expedients of performance of electronic transmitter in the formation of single server queuing model with fuzzy parameters by “fuzzy numerals sort technique” which are simple and more accurate. This method is very elemental in calculation not only gives the agreeable results to distinct problems, but also furnishes more legitimate results. In order to strengthen this proposed tactic, a real case study will be tropical in offing investigation of electronic transmitter.

## References

- Allaviranloo T, Jahantigh MA, Hajjighasemi S (2013) A new distance measure and ranking method for generalized trapezoidal fuzzy numbers. *Math Probl Eng*
- Azman FN, Abdullah L (2012) Ranking fuzzy numbers by centroid method. *Malays J Fundam Appl Sci* 3:117–121
- Bose S (2008) An introduction to queueing systems. Kluvar Academic/Plenum Publishers, New York
- Chu TC, Tsao CT (2002) Ranking fuzzy numbers with an area between the centroid point and original point. *Comput Math Appl* 43(1–2):111–117
- Cooper R (1990) Introduction to queueing theory, 3rd edn. CEE Press, Washington
- Dat LQ, Yu VF, Chou SY (2012) An improved ranking method for fuzzy numbers based on the centroid index. *Int J Fuzzy Syst* 14:413–419
- Devaraj J, Jayalakshmi D (2012) A fuzzy approach to non-preemptive priority queues. *Int J Math Arch* 2704–2712
- Ganesh AH, Shobana AH, Ramesh R (2021) Identification of critical path for the analysis of bituminous road transport network using integrated FAHP–FTOPSIS method. *Mater Today: Proc* 37:193–206
- Janos S (2016) Basic queueing theory. Globe Edit Publishers, Omniscriptum GMBH, Germany
- Klir GJ, Yuvan B (2005) Fuzzy sets and fuzzy logic theory and applications. Prentice Hall of India



11. Palpandi B, Geetharamani G (2013) Computing performance measures of fuzzy non-preemptive priority queues using robust ranking technique. *Appl Math Sci* 7:5095–5102
12. Parandin N, Araghi MAF (2008) Ranking of fuzzy numbers by distance method. *J Appl Math* 19:47–55
13. Pardo MJ, de la Fuente D (2007) Optimizing a priority discipline queueing model using fuzzy set theory. *Comput Math Appl* 54:267–281
14. Pekoz EA (2002) Optimal policies for multi-server non-preemptive priority queues. *Queueing Syst* 42(1):91–101
15. Ramesh R, Kumaraghuru S (2014) Priority disciplined queueing models with fuzzy parameters. *J Math Comput Sci* 4:594–602
16. Ramesh R, Kumaraghuru S (2017) Analysis of performance in four non-preemptive priority fuzzy queues by centroid of centroids ranking method. *Int J Comput Tech* 4:12–20
17. Phani Bushan Rao P, Ravi Shankar N (2011) Ranking fuzzy numbers with a distance method using circumcenter of centroids and an index of modality. *Adv Fuzzy Syst*, 2011
18. Rao PPB, Shankar NR (2012) Ranking generalized fuzzy numbers using area, mode, spreads and weights. *Int J Appl Sci Eng* 10:41–57
19. Ritha W, Lilly R (2010) Fuzzy queues with priority discipline. *Appl Math Sci* 4(12):575–582
20. Shankar NR, Sarathi BP, Babu SS (2013) Fuzzy critical path method based on a new approach of ranking fuzzy numbers using centroid of centroids. *Int J Fuzzy Syst Appl* 2:16–31
21. Wang YJ, Lee HS (2008) The revised method of ranking fuzzy numbers with an area between the centroid and original points. *Comput Math Appl* 55(9):2033–2042
22. Zadeh LA (1978) Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets Syst* 1:3–28

# Improved HELBP in Harsh Lightning Variations for Face Recognition



Shekhar Karanwal

## 1 Introduction

The motivation of face recognition (FR) develops because of its need in security, surveillance, biometric authentication, and telecommunication. A good FR is that which attains remarkable results in unfavorable challenges. The unfavorable challenges encompass lightning, emotion, occlusion, and pose. Literature suggests that local descriptors achieve finer results than the global ones in these conditions. Recently, Ahuja et al. [2] developed the FR in lightning and pose changes based on LBP and ELM. LBP is utilized for the extraction of features from distinct facial parts, which are merged for the developing of entire LBP size. Further, ELM is deployed for matching. Koley et al. [3] invented FR (heterogeneous and homogeneous) for lightning and noise. Specifically, the novel descriptor Gammadion binary pattern (GBP) is launched for generation of GBPSC images. The integration of GBPSC with CNN defeats the results of several ones. Siddiqui et al. [4] in their proposed work utilize LBP for extracting the features, and then, IFLDA is brought into play for compaction. This method gives the excellent outcomes in harsh lightning changes. Bindu et al. [5] presented the FR in which Co-HOG features are integrated with gradient magnitude features. Co-HOG is the HOG variant in which the orientations gradients pairs are utilized for creation of the block, as opposed to HOG which utilizes the alone gradient. Kumar et al. [6] proposed the ensemble-based FR for distinct challenges. Precisely, the D-LGS is introduced which utilizes the bilinear interpolation for enhancing the pixel density for descriptive image formation from the image (input). Further, genetic algorithm (GA) is used for the reduction.

Tong et al. [7] introduced the FR (for different challenges) by merging the multi-mirror symmetry with LBP, called as MMLBP. The following advantages are offered

---

S. Karanwal (✉)

CSE Department, Graphic Era University (Deemed), Dehradun, India

e-mail: [shekhar.karanwal@gmail.com](mailto:shekhar.karanwal@gmail.com)

by MMLBP: In heterogeneous illumination conditions, the illumination is adaptively compensated, and it generates features which are much nearer to those in controlled surroundings. Bhattacharya et al. [8] give the LFP for FR. LFP encodes the texture information directionally of the input image in compact way; as the result, more robust code is attained than many other methods. Every micro-pattern structure is computed by Van der Waal's force among the pixel pairs. Dahmouni et al. [9] develop its FR in numerous challenging environments. First, a novel descriptor is presented called as EVBP. The objective of EVBP is to merge LBP and the proposed model on virtual electric field. In proposed model, the neighbor of every pixel is considered as VE charges grid which is balanced electrostatically. Then, LBP is employed for representation of EVBP pixels.

In [1], Nguyen et al. invented horizontal elliptical LBP (HELBP) for FR. HELBP is advancement of LBP in which there is usage of elliptical horizontal pixels for comparison, rather than circular neighborhood pixels. In this work, it has been analyzed that in extreme lightning changes, the HELBP histogram performance (extracted globally) is not up to mark. In contrast to histogram feature, the map feature yields huge enhancement in accuracy. This novel feature is named as HELBP map. Besides this, two more descriptors are also evaluated for comparison, i.e., LBP [10] and MBP [11]. PCA [12] is deployed to all for feature reduction, and then, SVMs [13] are used for matching. The outcome of HELBP map is much better than others which are histogram based. Two utilized datasets are YB [14] and EYB [14]. Rest content is given as: LBP, HELBP, and MBP are derived in Sect. 2; invented descriptor is reported in Sect. 3; results are disposed in Sect. 4 with conclusion in Sect. 5.

## 2 Description of LBP, HELBP, and MBP

### 2.1 Local Binary Pattern (LBP)

LBP [10] descriptor is invented originally for textures. It has been owned efficiently in other applications too. In LBP, the neighbors gray value in  $3 \times 3$  patch is availed for comparison with the center gray value. The 1 is set to those positions whose gray value is bigger or equal to center gray value else 0 is set. There is emergence of 8-bit pattern after thresholding, which is transformed to LBP code by granting weights. After adopting the equal methodology for all locations (positions), there is the formation of LBP image. The LBP image forms the feature dimension of 256. Figure 1 gives the complete LBP example, and Fig. 2 shows map image with histogram. LBP with P neighbors at radius R is given in Eq. 1.  $VR,p$  designates the location-wise gray pixels of neighbors, and  $VC$  designates center gray pixel location.

$$\text{LBP}_{P,R}(x_c) = \sum_{p=0}^{P-1} k(V_{R,p} - V_C)2^p \quad (1)$$

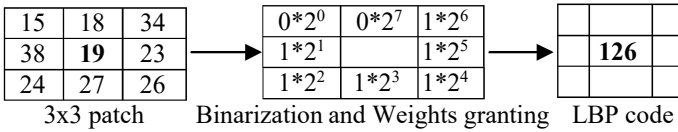


Fig. 1 LBP example

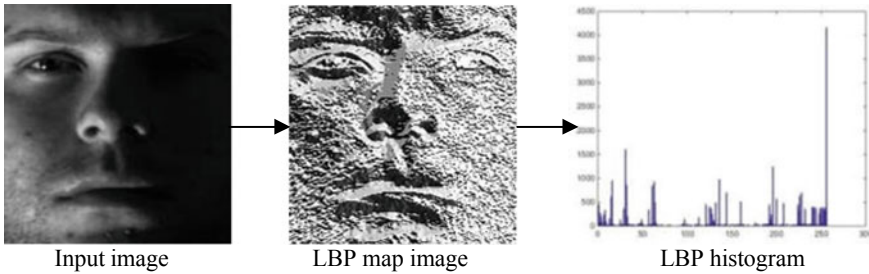


Fig. 2 LBP map image with histogram

$$k(y) = \begin{pmatrix} 1 & y \geq 0 \\ 0 & y < 0 \end{pmatrix}$$

### 2.2 Horizontal Elliptical LBP (HELBP)

Nguyen et al. [1] launched HELBP for FR. In HELBP, the horizontal neighbors gray value in  $3 \times 5$  patch is available for comparison with the center gray value. The 1 is set to those positions whose gray value is bigger or equal to center gray value else 0 is set. There is emergence of 8-bit pattern after thresholding, which is transformed to HELBP code by granting weights. After adopting equal methodology for all locations, there is formation of HELBP image. The HELBP image forms the feature dimension of 256. Figure 3 gives complete HELBP example, and Fig. 4 shows map image with histogram. HELBP with P neighbors at radius R1 and R2 is given in Eq. 2.  $V_{R_1, R_2, p}$  designates location-wise gray pixel of neighbors, and VC specifies center pixel location.

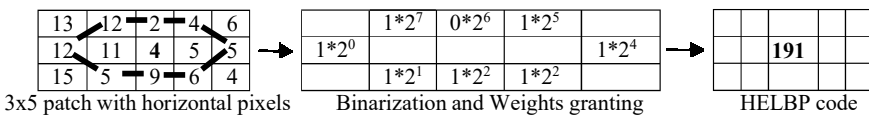


Fig. 3 HELBP example

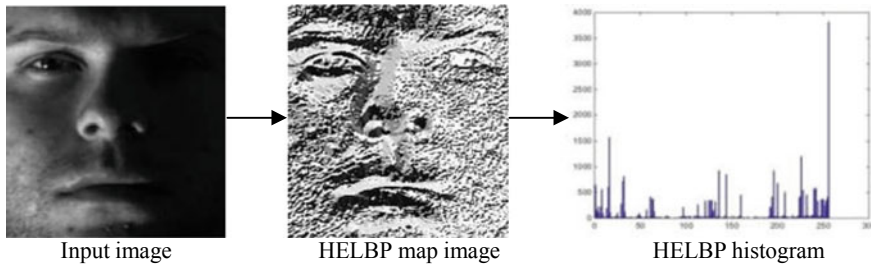


Fig. 4 HELBP map image with histogram

$$\text{HELBP}_{P,R_1,R_2}(x_c) = \sum_{p=0}^{P-1} k(V_{R_1,R_2,p} - V_C)2^p \tag{2}$$

$$k(y) = \begin{pmatrix} 1 & y \geq 0 \\ 0 & y < 0 \end{pmatrix}$$

### 2.3 Median Binary Pattern (MBP)

MBP [11] is developed for textures. In MBP, all pixels gray value in  $3 \times 3$  patch is individually compared with median of the gray pixels. The 1 is set to those positions whose gray value is bigger or equal to median gray value else 0 is set. There is the emergence of 9-bit pattern after thresholding. In MBP, as there is the inclusion or exclusion of center bit. In this work, it is fully dropped as a consequence; there is emergence of 8 pattern, which is transformed to MBP code by granting weights. After adopting the equal methodology for all locations (positions), there is the formation of MBP image. The MBP image forms the feature dimension of 256. Figure 5 gives complete MBP example, and Fig. 6 shows map image with histogram. MBP with  $P = 9$  (all gray pixels) is given in Eq. 3, and MBP with  $P = 8$  (gray neighbor pixels) is given in Eq. 4. R designates the radius;  $V_{R,p}$  designates the location-wise gray pixels, and  $V_{\text{median}}$  designates median value.

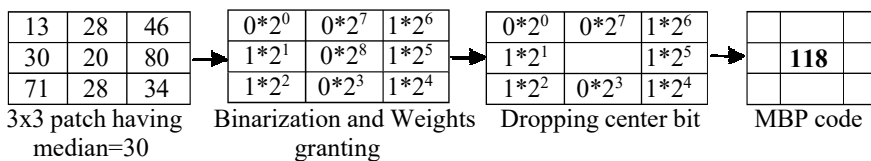


Fig. 5 MBP example

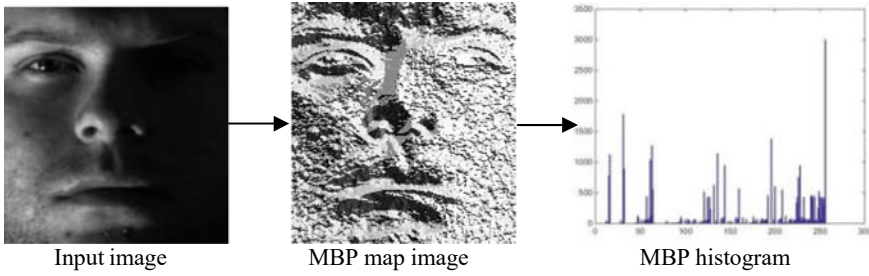


Fig. 6 MBP map image with histogram

$$MBP_{P,R}(x_C) = \sum_{p=0}^{P-1} k(V_{R,p} - V_{median})2^p \quad \text{For } P = 9 \text{ in } 3 \times 3 \text{ patch} \quad (3)$$

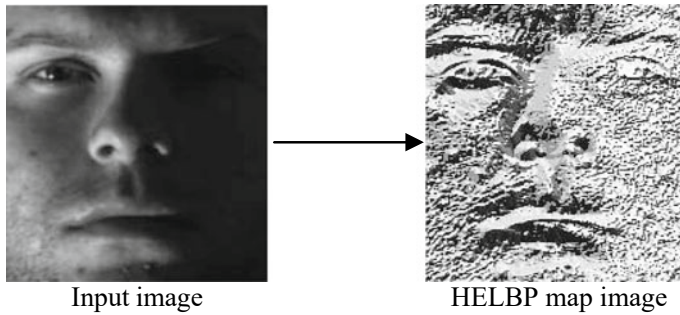
$$MBP_{P,R}(x_C) = \sum_{p=0}^{P-1} k(V_{R,p} - V_{median})2^p \quad \text{For } P = 8 \text{ in } 3 \times 3 \text{ patch} \quad (4)$$

$$k(y) = \begin{pmatrix} 1 & y \geq 0 \\ 0 & y < 0 \end{pmatrix}$$

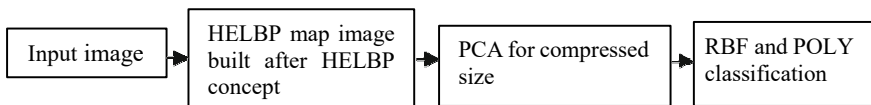
### 3 The Invented Descriptor

#### 3.1 HELBP Map

In this work, it has been analyzed (from the results) and proved that the histogram feature unable to perform well in harsh lightning changes. In contrast to this, the map feature of HELBP performs outstanding in harsh lightning changes. HELBP map is the transformed image size built from HELBP concept. The HELBP map feature also defeats the results of the other implemented descriptors. Results section gives the size details. Figure 7 shows the HELBP map image, and block diagram is conveyed in Fig. 8. The dimension is compacted by the usage of PCA, and then, RBF and POLY are utilized for matching.



**Fig. 7** HELBP map image of corresponding input image



**Fig. 8** Block diagram

## 4 Results

### 4.1 Datasets Explanation

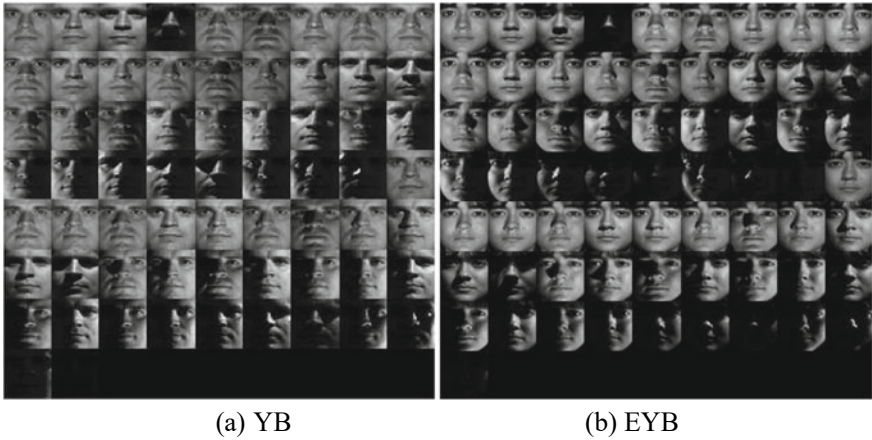
Explanation of two datasets are given as

#### 4.1.1 YB Dataset

YB dataset is developed from 10 individuals. Each individual encompasses 64 lightning variation images in 9 poses. There is the also capturing of one ambient image which sets the total to 5850 images. Those images which are posed frontally are used for the purpose of evaluation. These frontal-posed images are 650. Each one has the resolution of  $192 \times 168$ . Figure 9a portrays one individual images.

#### 4.1.2 EYB Dataset

EYB dataset is developed from 38 individuals. Each individual encompasses the 64 lightning variation images frontally posed. Therefore, this dataset has 2432 images. Each one has the resolution of  $192 \times 168$ . The EYB dataset is left with the 2414 images, as 18 images corrupt in database making. Figure 9b portrays one individual images.



**Fig. 9** One individual dataset images

## 4.2 Descriptors Feature Details

Prior to perform feature extraction, there is the image resizing for computational cost compensation. The size which is set for the feature extraction is  $55 \times 52$ . Then, four descriptors are owned for extracting features. These are LBP hist., HELBP hist., MBP hist., and **HELBP map**. The size of these ones is 256, 256, 256, and **2544**. After PCA, classifier takes the size of 32 and 41, on YB and EYB, respectively. MATLAB R2018a is considered for all the evaluation. Hist. specify histograms in LBP, HELBP, and MBP.

## 4.3 Recognition Rate Estimation

The recognition rate (in %) is estimated by the ratio of the correctly matched samples to all the samples in test dataset. A correctly matched samples are computed by taking the difference between test dataset samples and the incorrect matches.

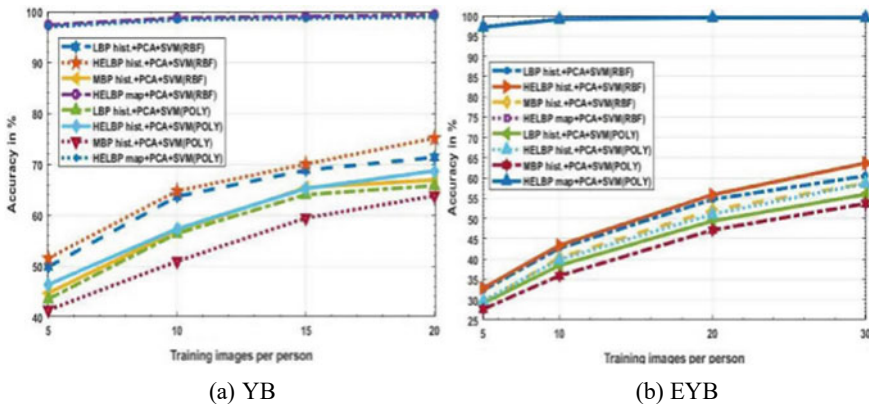
### 4.3.1 YB Dataset

On YB, the training samples/person is (5, 10, 15 and 20), while the test ones are (60, 55, 50 and 45). The finest classifier rate is measured after 25 runs. The **HELBP map** thrashes performance of other histogram-based descriptors. Table 1 and Fig. 10a convey the analysis.



**Table 1** Results investigation on YB and EYB

	Training samples (Z) on YB				Training samples (Z) on EYB			
	Z = 5	Z = 10	Z = 15	Z = 20	Z = 5	Z = 10	Z = 20	Z = 30
All approaches	Recognition rate				Recognition rate			
LBP hist + PCA+RBF	49.83	63.63	68.80	71.33	32.32	42.67	54.71	60.51
HELBP hist.+PCA+RBF	51.50	64.72	70.00	75.11	32.86	43.41	55.86	63.65
MBP hist.+PCA+RBF	44.66	56.90	65.40	66.88	29.18	40.36	52.05	58.71
<b>HELBP map+PCA+RBF</b>	<b>97.33</b>	<b>98.72</b>	<b>99.00</b>	<b>99.33</b>	<b>97.30</b>	<b>99.21</b>	<b>99.51</b>	<b>99.60</b>
LBP hist.+PCA+POLY	43.50	56.36	64.00	65.77	29.09	38.44	49.39	55.96
HELBP hist.+PCA+POLY	46.33	57.27	65.20	68.66	29.72	39.87	51.08	58.63
MBP hist.+PCA+POLY	41.33	50.90	59.40	63.77	27.65	35.88	47.15	53.68
<b>HELBP map+PCA+POLY</b>	<b>97.00</b>	<b>98.36</b>	<b>98.60</b>	<b>98.88</b>	<b>97.16</b>	<b>99.11</b>	<b>99.51</b>	<b>99.52</b>



**Fig. 10** Graph analysis

**4.3.2 EYB Dataset**

On EYB, the training samples/person is (5, 10, 20 and 30), while the test ones are (59, 54, 44 and 34). The finest classifier rate is measured after 20 runs. The **HELBP map** thrashes performance of other histogram-based descriptors. Table 1 and Fig. 10b convey the analysis. In Table 1, hist. specifies the histogram. Best recognition rate is indicated with bold (black).

## 5 Conclusion

This work analyzes the potential of HELBP map than histogram-based descriptors in extreme lightning changes. The three histogram-based descriptors are LBP hist., HELBP hist., and MBP hist. PCA is adopted for all of them to compress the feature length, and matching is done by SVMs. On YB and EYB, datasets results are conducted.

## References

1. Nguyen HT, Caplier A (2012) Elliptical local binary patterns for face recognition. In: ACCV, pp 85–96
2. Ahuja B, Vishwakarma VP (2021) Local binary pattern based ELM for face identification. In: Proceedings of ICAIA, vol 1164, pp 363–369
3. Koley S, Roy H, Bhattacharjee D (2021) Gammadion binary pattern of Shearlet coefficients (GBPSC): an illumination-invariant heterogeneous face descriptor. *Patt Recog Lett* 145:30–36
4. Siddiqui R, Shaikh F, Sannulal P, Lakshmi A (2021) An improved method for face recognition with incremental approach in illumination invariant conditions. In: Proceedings of ICCCE, pp 1145–1156
5. Bindu CHH, Chari KM (2020) Face recognition using improved Co-HOG features. In: ACSN, pp 647–655
6. Kumar D, Garain J, Kisku DR, Singh JK, Gupta P (2020) Unconstrained and constrained face recognition using dense local descriptor with ensemble framework. *Neurocomputing* 408:273–284
7. Tong SG, Huang YY, Tong ZM (2019) A robust face recognition method combining LBP with multi-mirror symmetry for images with various face interferences. *Int J Auto Comp* 16(5):671–682
8. Bhattacharya S, Nainala GS, Rooj S, Routray A (2019) Local force pattern (LFP): descriptor for heterogeneous face recognition. *Patt Recog Lett* 125:63–70
9. Dahmouni A, Moutaouakil KE, Satori K (2018) Face description using electric virtual binary pattern (EVBP): application to face recognition. *Multi Tool App* 77:27471–27489
10. Karanwal S (2021) A comparative study of 14 state of art descriptors for face recognition. *Multi Tool App* 80:12195–12234
11. Alkhatib M, Hafiane A (2019) Robust adaptive median binary pattern for noisy texture classification and retrieval. *IEEE Trans Im Process* 28(11):5407–5418
12. Jia Y, Liu H, Hou J, Kwong S, Zhang Q (2021) Semisupervised affinity matrix learning via dual-channel information recovery. *IEEE Trans Cyber* 1–12
13. Ma Z, Li B (2020) A DDoS attack detection method based on SVM and K-nearest neighbour in SDN environment. *Int J Comp Sci Eng* 23(3):224–234
14. Bao J, Lai Z, Li (2021) Relaxed local preserving regression for image feature extraction. *Multi Tool App* 80:3729–3748

# Determination of the TOE Factors Influencing the Adoption of Internet Banking Services on SMEs in Yemen: A Moderated Mediation Approach



Nabil Hussein Al-Fahim, Rawad Abdulghafor, and Sherzod Turaev

## 1 Introduction

Internet banking services (IBSs) are greatly useful to banks as well as customers. They are useful for banks in terms of safety cost, reach a higher percentage of the population, they are efficient in improving the reputation and better customer service satisfaction of costumers [1, 2]. IBS also provides banks with an unlimited distribution network, thereby offering a competitive advantage. The adoption of IBS renders banks with cost-effective transactions and strengthens the bond between bankers and costumers. Indeed, the advancement of Internet networks and high-value financial services can be offered at lower costs through the Internet [3]. In spite of the fact that, Yemeni banks produce electronic infrastructure and spend tremendous amount of money annually to adopt electronic banking, the adoption of electronic banking and IBS in Yemeni banks is relatively low among small- and medium-sized enterprises (SMEs) [4–6]. Furthermore, the study of [7] has proclaimed IBS as beneficial in that customers are offered access to bank services at a fast rate and with comfort without requiring being present at the bank. Additionally, to ensure being ahead in the progressively competitive financial service market, a comprehensive IBS stratagem is vital.

Although e-banking has a lot of advantages, many banks' consumers do not embrace these advanced of technologically services offered by the banks especially

---

N. H. Al-Fahim · R. Abdulghafor (✉)

Department of Computer Science, Faculty of Information and Communication Technology,  
International Islamic University Malaysia, 53100 Kuala Lumpur, Malaysia

e-mail: [rawad@iiu.edu.my](mailto:rawad@iiu.edu.my)

S. Turaev (✉)

Department of Computer Science and Software Engineering, College of Information Technology,  
United Arab Emirates University, P.O. Box 15551, Al Ain, United Arab Emirates

e-mail: [sherzod@uaeu.ac.ae](mailto:sherzod@uaeu.ac.ae)

in the Arab countries and in Yemen in particular [4, 5, 8, 9]. Internet banking technology in Yemen has not reached its maturity yet, as it is scarcely used by retail or corporate customers and there is a prevalence in the culture that favors face-to-face banking and personal contact. Relatively small empirical study has been implemented to understand the factors that influence the use of Internet banking by SMEs, and its consequential effects, despite this impending strain [10].

For that reason, there should be additional research to comprehend the relevance of IB in Yemen, to recognize areas that are underdeveloped in establishments, primarily in SMEs. To fill a gap in the research, we investigate TOE elements such as technological variables (relative advantages, security), organizational aspects (management support and organization preparedness), and environmental aspects (government support).

This study aims to contribute to the body of knowledge in the field of technology acceptance by expanding our understanding of the factors that influence IB adoption by SMEs in Yemen. As a result, the study's major goal is to look at the factors that influence SMEs' willingness to use Internet banking in Yemen. Furthermore, the study looks into the function of organization preparedness as a mediator between TOE characteristics and SMEs' aims for the usage of Internet banking in Yemen.

## 2 Literature Review

### 2.1 *Internet Banking Services in Yemeni Banks*

Arab Bank, The Yemen Gulf Bank, International Bank of Yemen, Yemen Commercial Bank, and CAC Bank [4, 8, 11] are the five banks that provide IBS in Yemen. Customers can use these banks for services such as balance inquiries, annual statements of account, account-to-account transfers, account-to-account transfers, checkbook requests, and pin code changes.

### 2.2 *SMEs and Yemen Economy*

Yemen's economy is primarily based on oil, and the Yemeni government derives the majority of its revenue from oil-related taxes [4, 12]. The SMEs industry in Yemen is critical to the country's economy. SMEs have long been regarded as the economy's mainstay, and they play an important role in creating jobs, improving human resource quality, and cultivating an entrepreneurial culture, in addition to supporting large-scale businesses and creating new business prospects [4, 5, 13].

Micro, small, and medium companies (MSMEs) employ around 90% of Yemen's workers and contributed 95% of the country's GDP in 2007 [4, 14]. Furthermore, recent estimates put the number of MSMEs at over 378,305. Micro-businesses (those

with fewer than four employees) account for more than 350,138, while small businesses account for 17,248 (with 5–9 employees). However, there are roughly 10,919 medium-sized businesses (having 10–50 employees).

### ***2.3 Internet Banking Adoption***

Internet banking adoption means “the customers use of many services represented in inquiring about an account balance application for a loan, remitting money from one account to another and many other services that are basically carried out online” [11]. The adoption of e-banking technology depends on a number of factors.

Personal touch with banking employees, technology anxiety, a large network of current branches, and limited computer proficiency among customers are all factors impacting client non-adoption of IB [12]. The lack of availability, rather than their avoidance of the technology during its infancy [15], has been determined to be the main cause for customers not embracing ATMs [16]. In general, whether a consumer has access to e-banking technology is determined by whether they are a customer of a bank that offers a number of services, including computer banking. Furthermore, it has been stated that consumers’ perceptions of innovation as well as socioeconomic factors have influenced their adoption of technological innovation.

### ***2.4 Technology–Organization–Environment (TOE) Framework***

Technology–organization–environment (TOE) framework was established by Tornatzky et al. [13]. The TOE framework is a frequently used framework for IT system adoption in organizations. It is one of the most effective alternatives to other IT adoption theories [17]. Furthermore, the TOE framework illustrates how a corporation can accept technology innovation not only owing to technological considerations, but also owing to organizational and environmental factors. These three factors create both challenges and opportunity for technological advancement [18]. They also indicate that the technical context encompasses the organization’s internal and external technologies. The organizational context encompasses the organization’s traits and resources, such as its size and scope, as well as its management structure. Industry features, technology support infrastructure, government regulation, and the firm’s competitors are all part of the environmental context [10]. Although the association between organizational preparedness and SMEs’ desire to embrace electronic data interchange (EDI) is significant, it has been discovered that the effect of organizational readiness on SMEs’ intention to implement EDI is not strong.

## 2.5 Factors Influencing IB Adoption (*Independent Variables*)

The technology–organization–environment framework (TOE) identifies three characteristics to explain a firm’s decision-making behavior in relation to technological innovations [13]. The TOE framework has been used in previous research in many “Information and Communication Technology Innovation (ICTI)” and IS factors. It was used in studying various IS innovations. By examining these factors, a greater understanding of why a company does, or does not, adopt a new technology could appear [14, 19, 20]. In the study [16], the factors have examined with the extent of organizational usage of business intelligence and analytics (BIA) using the TOE model.

To explain a firm’s decision-making behavior in response to technological advancements, the technology–organization–environment framework (TOE) defines three characteristics [18]. Many “Information and Communication Technology Innovation (ICTI)” and IS aspects have been studied using the TOE framework in the past. It was used to research a variety of IT advancements. By evaluating these characteristics, it may be possible to gain a better understanding of why a corporation adopts or does not embrace a new technology [21–23]. Using the technology–organization–environment framework, the factors influencing organizational adoption of business intelligence and analytics (BIA) were investigated in the study [24].

### Technological Context

#### (a) Security

Security is critical to the growth of trust in IB because when clients process information of financial and know that it will be handled securely, they feel at peace, and their trust in the bank grows over time. Security worries have also impeded the rise of mobile shopping by negatively influencing customers’ willingness to accept it, according to the study.

Furthermore, previous studies reveal that privacy and security are crucial factors that influence users’ willingness to use e-based transaction systems.

Many clients are cautious because to the lack of personal touch with the service provider, especially if they believe there will be no recourse if their information or funds are stolen [25, 26]. In other words, a lack of security can restrict fun and, as a result, adoption, whereas a sense of security can promote enjoyment and aid adoption.

#### (b) Relative Advantages

Relative advantage has been found to be essential variable in determining the use of new technologies [15].

Many studies confirmed the benefits of relative advantages are significantly and positively related to IB systems adoption and other related technologies [14, 16, 17]. According to Ammar [18], the finding that perceived relative advantages exert influence on m-banking adoption highlights the importance of innovation characteristics to institutions who are involved in the design and development of m-banking services.

According to previous research [21], the relative advantage of IT innovation is one of the factors frequently used in technological innovation adoption studies. Consumer familiarity with Internet banking, perceived structural assurance, and relative advantage are all important elements in forming initial trust in Internet-only banks, according to [27].

## **Organizational Context**

### **(a) Top Management**

Organizational readiness and top management consider to be factors that impact SMEs' adoption of ES. In [21], it has been reported that without top management commitment, businesses in Malaysia are less likely to adopt e-payment technologies. In many cases, the top management team must mediate between technological and business requirements, as well as settle conflicts of interest among a diverse group of stakeholders.

CEO and top management traits have been proven to be key determinants of technological adoption in SMEs in [28] studies. Technological knowledge, a supportive environment, passion, leadership styles, academic qualifications, and technology awareness through networking are some of these traits. As a result, the authors consider CEO and top management qualities to be a significant factor in the three stages of technological innovation adoption.

### **(b) Organization Readiness**

The level of financial and technical resources available in an organization to implement knowledge management systems is referred to as organizational readiness. The term "perceived organization resources" refers to a company's perceptions of its technical, financial/business, and human resource availability.

Organizations that have reached a certain level of technical competence are typically perceived as highly integrated in terms of computerized processes and as more prepared to absorb IT innovation, allowing them to reap higher benefits [29]. The organizational readiness, according to Tan [22], involves infrastructure, necessary systems, and technical capabilities. Although the literature differs in its definition of organizational readiness, all agreed that organizational readiness has a significant impact on SMEs' technology adoption.

## **Environmental Context**

### **(a) Government Support**

[30] contributed to the research of the influencing elements of online banking adoption by incorporating important antecedents into the TAM model and ultimately discovered that government backing is critical to the trust of online banking products.

Government support, on the other hand, was not found to be helpful in the assimilation of m-banking in the Sudanese microfinance industry, according to. This conclusion could be explained by the fact that the majority of responding MFPs are government-owned or semigovernmental enterprises that receive government assistance by default.

## 2.6 Factors Influencing IB Adoption (Moderating Variable)

The type of organization or the nature of the business of an organization is a moderator element in this study. As a result, the geographic scope of an organization's operations is defined as its business type. It has been used to categorize corporate entities based on their business characteristics [31].

Moreover, it has been found in [22] that service SMEs have expressed quite different views from manufacturing SMEs in Malaysia. Service SMEs opine that IT network infrastructure hinders them from using ICT.

Another study [32] discovered that the type of business had an effect on both the number of Internet users and the perceived benefit level.

On the other hand, investigated industry categories as a moderating factor that may have substantial influence on the adoption of ICT among SMEs in Malaysia in a study of 406 owners and managers.

## 2.7 Conceptual Framework and Hypotheses Testing

- |                    |   |
|--------------------|---|
| H1                 | Security is positively significant influence on Organization Readiness                                      |
| H2                 | Relative Advantages is positively significant influence on Organization Readiness                           |
| H3                 | Top Management is positively significant influence on Organization Readiness                                |
| H4                 | Government Support is positively significant influence on Organization Readiness                            |
| H5                 | Security is positively significant influence on BI to adopt IB  |
| H6                 | Relative Advantages is positively significant influence on BI to adopt IB                                   |
| H7                 | Top Management is positively significant influence on BI to adopt IB  |
| H8                 | Government Support is positively significant influence on BI to adopt IB                                    |
| H9                 | Organization Readiness is positively significant influence on BI to adopt IB                                |
| H10, H11, H12, H13 | Organization Readiness mediates the relationship between TOE factors and BI to adopt IB                     |
| H14                | Organization Type moderates on the relationship between Organization readiness and BI to adopt IB (Fig. 1). |



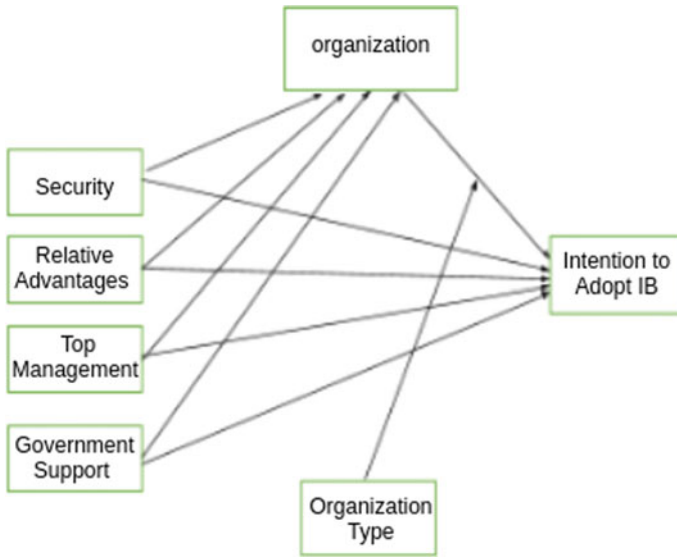


Fig. 1 Conceptual framework and hypotheses testing

### 3 Methodology

For this topic, a quantitative research design is acceptable. This is because it intends to use hypothesis testing to explore the impact of social media use on humanitarian response via social capital as a mediating function, which necessitates the employment of a quantitative methodology to deal with the data.

Data will be collected using a self-administered survey utilizing a stratified random sample method in this project. Managers and owners of SMEs in Yemen are divided into groups based on the number of employees (population of the sample).

#### 3.1 Design of Questionnaire

The data for this paper was gathered through a survey. The questionnaire is divided into two sections. The first section contains demographic information about the respondents, such as their organization’s activities, kind of organization, and types of activities and programs.

The second portion of the survey will ask respondents about the study’s variables of interest, which are (1) intention to use IB and (2) security, (3) comparative advantages, (4) top management, (5) government assistance, and (6) readiness of the company. Factors 2 and 6 are external variables, while factor 1 is an endogenous variable.

**Table 1** Descriptive statistics for all variables

Variable	Code	Min.	Max.	No. items	Mean	%	S.D.
Security	SU	1.00	4.60	5	3.3554	0.67108	0.97025
Relative advantages	RT	1.00	4.60	5	3.3102	0.66204	0.92378
Top management	TM	1.00	4.80	5	3.2325	0.6465	1.0149
Government support	GM	1.00	5.00	4	3.1616	0.63232	1.0670
Organization readiness	OG	1.00	4.67	4	3.2898	0.65796	0.79661
Behavior intention	BI	1.00	4.83	6	3.2192	0.64384	0.81420
Overall		1.00	4.45	29	3.2610	0.6522	0.69621

A total of 352 questionnaires were returned out of 400, representing an approximate 88% response rate. 21 questionnaires were removed from the analysis due to missing values, and 17 cases were outliers; hence, a total of 314 useable questionnaires were included, with a response rate of 73%. For this investigation, a sample size of  $n = 314$  was deemed adequate. The study sample size ( $N = 314$ ) met the 5:1 ratio that was proposed by the researchers [23, 24].

## 4 Analysis of Data

### 4.1 Descriptive Statistics for Variables

The measuring scales' mean and standard deviation (S.D.) were calculated. A five-point Likert scale was utilized in this investigation, with "1" indicating strongly disagreement and "5" indicating strongly agreement.

Table 1 reveals that security and relative advantages had the greatest mean of 3.355 and 3.310 out of a maximum of 5, accounting for 67% and 66% of the total. The organization preparedness score was 3.28, accounting for 65% of the total. Government support, on the other hand, had the lowest mean of 3.16, accounting for around 63% of the total. Furthermore, the total mean (mean of these values) was 3.26 out of 5 or 65% (more than 3). Additionally, the standard deviations (SD) for factors range from 0.796 to 1.067, indicating that the data set has a lot of acceptable variabilities. All variables' descriptive statistics are explained in Table 1.

### 4.2 Assessment of PLS-SEM Path Model Results

As previously stated, SEM has two-stage data analysis method that guides the evaluation of the measurement model and the structural model estimation. The reflective constructs' construct validity, convergent validity, and discriminant validity are

all evaluated as part of the measuring model. In this vein, this study followed the two-stage sequential paradigm by applying PLS-based SEM [33]. We examined the measurement model first, then the structural model analysis and hypothesis testing, as recommended by Hair et al. [34]

The evaluation of convergent validity was a critical component of the measurement model. Constructs have an average variance extracted (AVE) of at least 0.5, and composite reliability (CR) measures of internal consistency reliability are above 0.70 and ranged between 0.906 and 0.956.

The reliability (Cronbach’s alpha) values were greater than 0.70 and varied from 0.866 to 0.940, as shown in Table 2. [34] advised that the AVE of each latent component be 0.50 or higher in order to obtain satisfactory convergent validity. The AVE values in this investigation were greater than 0.50 and ranged from 0.687 to 0.845, suggesting acceptable values suggesting sufficient convergent validity.

Thus, the convergent validity was confirmed in the study (see Table 2).

**Table 2** Loading of items, Cronbach’s alpha, (CR), and AVE

Variables		Loadings	Cronbach’s alpha	CR	AVE
Security	SU1	SU1	0.940	0.954	0.806
	SU2	SU2			
	SU3	SU3			
	SU4	SU4			
	SU5	SU5			
Relative advantages	RT1	0.908	0.937	0.949	0.789
	RT2	0.881			
	RT3	0.882			
	RT4	0.884			
	RT5	0.886			
Top management	TM1	0.908	0.937	0.952	0.799
	TM3	0.876			
	TM4	0.889			
	TM5	0.895			
	TM6	0.900			
Government support	GM1	0.919	0.939	0.956	0.845
	GM2	0.914			
	GM3	0.919			
	GM4	0.930			
Organization readiness	OG1	0.779	0.866	0.906	0.707
	OG2	0.837			
	OG3	0.869			
	OG4	0.879			
Behavior intention	BI1	0.836	0.904	0.926	0.678
	BI2	0.847			
	BI3	0.851			
	BI4	0.849			
	BI5	0.775			
	BI6	0.777			

**Table 3** Discriminant validity for latent variables

	AVE	GM	TM	IB	OG	RT	SU
GM	0.677	0.894					
TM	0.684	0.298	<b>0.919</b>				
IB	0.641	0.529	0.458	<b>0.823</b>			
OG	0.675	0.443	0.396	0.780	<b>0.841</b>		
RT	0.710	0.543	0.366	0.558	0.483	<b>0.888</b>	
SU	0.591	0.518	0.383	0.577	0.528	0.432	<b>0.898</b>

### 4.3 Convergent Validity

The factor loading for the items in this study is greater than 0.50, which is acceptable if the study sample size is greater than 200 people. As a result, convergent validity is sufficiently demonstrated. As a result, all indicators in this study are related to their respective constructs, providing sufficient proof of the model's convergent validity. The convergent validity was confirmed in the study (see Table 2).

### 4.4 Discriminant Validity

The AVE values for each set of constructs are more than the squared correlations, indicating discriminant validity. Furthermore, for any given construct, the square root of the AVE was more than the value of the square of correlation of that variables with any other factor ( $AVE > \text{correlation square}$ ). The square root of the AVE for all variables with correlations bigger than the correlations between the variables and other variables in the model is shown in Table 3.

The degree to which a construct is actually distinct from other constructs is known as discriminant validity. For every concept that exceeds the squared correlation among other components, discriminant validity is assessed using average variance extracted (AVE) [13].

Relative advantages, government support, and top management explained 38.8% of the variance in organization readiness. In addition, security, relative advantages, government support, and top management and organization readiness explained 69% of the variance in behavior intention to adopt IB as shown in Fig. 2 and Table 4.

### 4.5 Assessment of Effects Size ( $F^2$ )

The mean of changes in the  $r^2$  reflects the relative effect of a given exogenous variable on an endogenous latent variable. It is determined as the increase in  $R^2$  of the latent

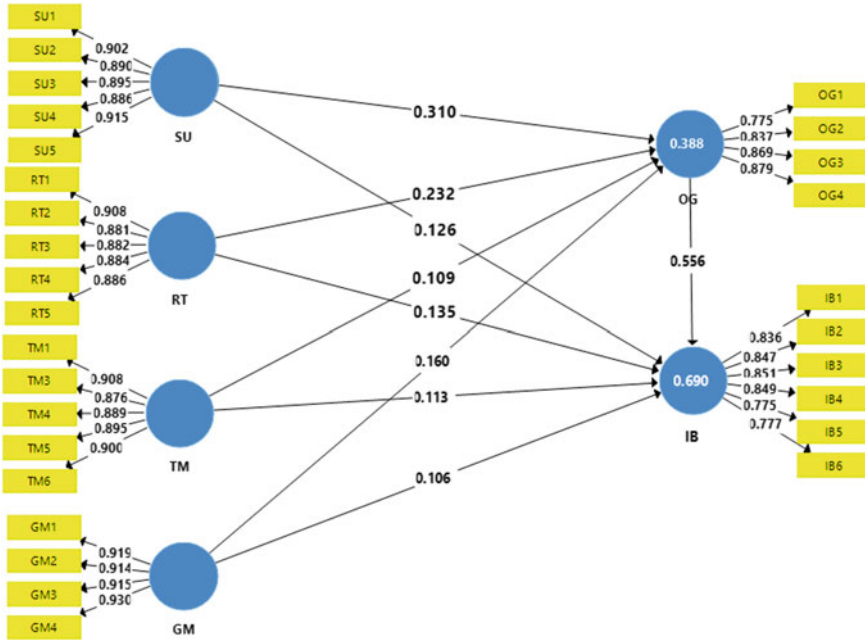


Fig. 2 Measurement model/outer loading and (path coefficient)

Table 4 Coefficient of determination results R<sup>2</sup>

Exogenous construct	Endogenous construct	R <sup>2</sup>	Hair et al. [34]	Cohen [35]	Chin [36]
SU, RT, TM and GM	Organization readiness	0.388	Moderate	Substantial	Moderate 1
SU, RT, TM, GM and OG	Behavior intention	0.690	Moderate	Substantial	Substantial

Source Smart PLS Version 3

variable to which the path is related, divided by the percent of unexplained variance of the latent variable.

The effect size ( $f^2$ ) can be calculated using the following formula proposed by:

$$\text{Effect Size } (f^2) = \frac{R^2_{\text{included}} - R^2_{\text{excluded}}}{1 - R^2_{\text{included}}} \tag{1}$$

According to Chin [25], an effect size of 0.02 is small, 0.15 is medium, and greater than 0.35 is large. Table 5 presents the results of effect size of the exogenous latent variables on endogenous variable with inclusion and exclusion of the mediator.

**Table 5** Determination coefficient results  $R^2$ 

Variable	Effect size $f^z$			
	Organization readiness	Rating	Behavior intention	Rating
SU, RT, TM and GM				
Security	0.104	Small	0.031	Small
Relative advantages	0.057	Small	0.036	Small
Top management	0.012	Small	0.025	Small
Government support	0.034	Small	0.028	Small

#### 4.6 Direct Hypotheses Results

Security had a significant and favorable impact on organization preparedness ( $\beta = 0.310$ ; C.R. = 5.964;  $P = 0.000$ ), indicating that H1 is supported. Following that, relative advantages had a significant and favorable impact on SMEs' intention to adopt IB, indicating that H2 is supported ( $\beta = 0.232$ ; C.R. = 4.733;  $P = 0.000$ ). Furthermore, top management had a significant and favorable influence on organization preparedness ( $\beta = 0.109$ ; C.R. = 2.023;  $P = 0.044$ ), as did government support for IB adoption by SMEs in Yemen ( $\beta = 0.160$ ; C.R. = 2.832;  $P = 0.005$ ); hence, H3 and H4 are supported.

Alternatively, the outcomes of the research show security and relative advantages had a significant and positive effect on intention toward IB adoption. Thus, H5 and H6 are supported.

Furthermore, top management had a positive significant impact on SMEs' intention to use IB ( $\beta = 0.113$ ; C.R. = 3.091;  $P = 0.004$ ), and government assistance had a positive significant impact on SMEs' intention to use IB in Yemen ( $\beta = 0.106$ ; C.R. = 2.878;  $P = 0.001$ ). This H7 and H8 can be used. Finally, organizational readiness had a large favorable impact on behavior intention to adopt IB ( $\beta = 0.556$ ; C.R. = 17.168;  $P = 0.000$ ). The structural model's direct hypothesis outcomes are shown in Table 6.

#### 4.7 The Mediating Role (Indirect Hypotheses Results)

The study conducted a mediation analysis by applying SEM using (PLS) to detect and estimate the mediating effect of training outcome between nature of training, management involvement in training, motivation management, management confidence, management culture, and firm performance.

The indirect influence of an independent variable on a dependent variable is measured using a mediator in a mediation test. The mediation test in this work was based on a PLS bootstrapping strategy, which meant that the hypotheses were examined using the PLS methodology [26]. Management and marketing researchers are increasingly recognizing and accepting of the PLS-SEM approach [34].

**Table 6** Summary of structural model assessment (direct hypotheses)

H	Exogenous variables		Endogenous variables	Estimate (path coefficient)	S.D.	C.R. (t-value)	P-value	Hypothesis result
H1	Security	→	Organization readiness	0.310	0.052	5.964	0.000	Supported
H2	Relative advantages	→	Organization readiness	0.232	0.049	4.733	0.000	Supported
H3	Top management	→	Organization readiness	0.109	0.054	2.023	0.044	Supported
H4	Government support	→	Organization readiness	0.160	0.056	2.832	0.005	Supported
H5	Security	→	Behavior intention	0.126	0.031	4.025	0.000	Supported
H6	Relative advantages	→	Behavior intention	0.135	0.045	2.984	0.003	Supported
H7	Top management	→	Behavior intention	0.113	0.036	3.091	0.002	Supported
H8	Government support	→	Behavior intention	0.106	0.054	3.401	0.001	Supported
H9	Organization readiness	→	Behavior intention	0.556	0.032	17.168	0.000	Supported

The mediating impact in the PLS model was established using bootstrapping analysis in conjunction with the assumptions formulated [34].

Table 8 shows organization readiness (OG) could mediate the link between top management (TM), government support (GM), relative advantages (RT), and security (SU) as exogenous variables and behavior intention (BI) to use IB as endogenous variable. This is due to the fact that the *T* values for four hypotheses were 1.971, 2.790, 4.063, and 5.761 (all greater than 1.960) and the *p*-values were 0.049, 0.006, 0.000, and 0.000 (0.05). As a result, H10, H11, H12, and H13 are recommended. The structural model assessment for indirect hypotheses is summarized in Table 7 and Fig. 3.

**Table 7** Summary of structural model assessment (Indirect hypotheses)

Hypotheses	Relation	Original sample ( $\beta$ )	STDEV	T-value	P-value	Result
H10	TM → OG → IB	0.061	0.031	1.971	0.049*	Supported
H11	GM → OG → IB	0.089	0.027	2.790	0.006*	Supported
H12	RT → OG → IB	0.126	0.032	4.063	0.000**	Supported
H13	SU → OG → IB	0.173	0.030	5.761	0.000**	Supported

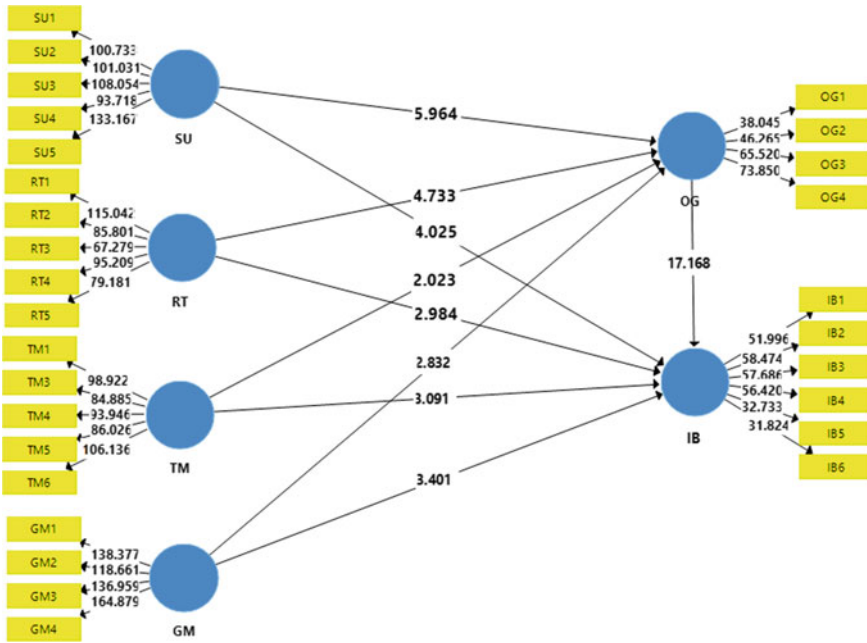
\*Significant at bootstrapping  $p < 0.05$

\*\*Significant at bootstrapping  $p < 0.01$

**Table 8** Summary of structural model assessment (indirect hypothesis)

H	Relationship	Path coefficient ( $\beta$ )	STDEV	t-value	p-value	Result
H14	OG OT... IBA	0.080	0.0321	2.469	0.011*	Supported

\*Significant at bootstrapping p < 0.05



**Fig. 3** Structural model with results of hypotheses

### 4.8 Modeling Effect (Organization Type)

To investigate the role of the organization type (OT) invariance as a moderator between organization readiness and behavior intention to adopt IB a simultaneous analysis of type organization based on trading and services organization grouping was carried out. The study used interaction effect using PLS-SEM to examine the moderating effect.

The study findings show that the hypothesis H14 is supported; OG \* OT... IBA is positively significant. This is because, the path coefficient and t-value were ( $\beta = 0.080$ ;  $t = 2.469$ ;  $p < 0.05$ ). Therefore, organization type is moderator between organization readiness and Internet banking adoption. Table 8 and Fig. 4 show the result of hypothesis testing for moderating effect (OT).



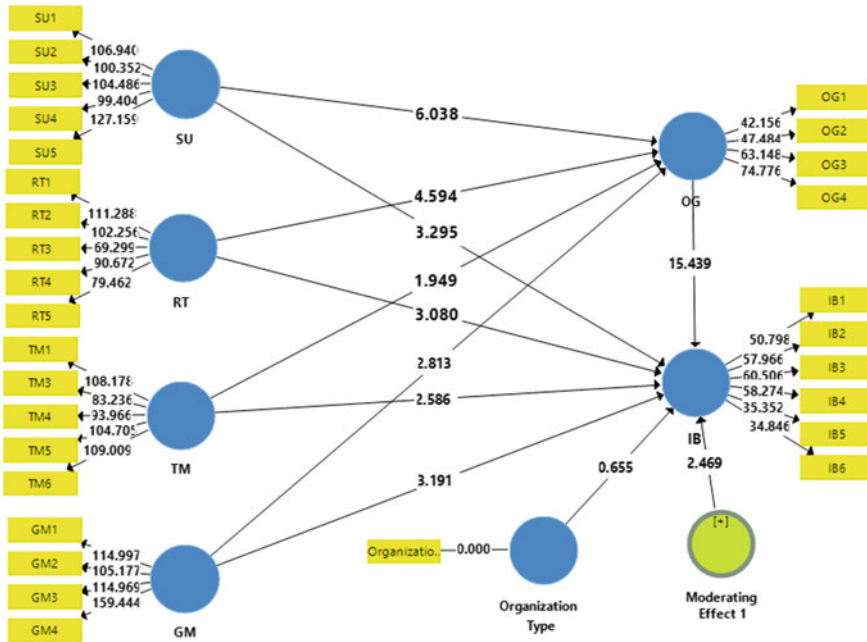


Fig. 4 Moderating effect

## 5 Implications

### 5.1 Practice Implication

This study makes a significant contribution by statistically confirming the parameters that influence SMEs' adoption of IB. As a result, it is reasonable to assume that SMEs with larger relative advantages, security, top management, organizational preparation, and government support will be more likely to adopt IB.

The importance of organizational preparation as a crucial mediator in the adoption of IB was highlighted in this study. This means that in order to employ IB in their commercial transactions, SMEs must be ready in terms of organizational preparedness. The culture, consistent value, and work methods of the organization are all factors in organizational readiness. Furthermore, this research shows that the environment in which a company operates has an impact on its decision to implement IB. To boost the number of IB users among their customers, banks should focus on TOE aspects. This study found that organizational readiness played a role in mediating the relationship between TOE characteristics and the intention to implement IB among Yemeni SMEs. In addition, the association between organization readiness and BI to adopt IB was modified by the kind of organization. As previously said, the goal

of this research is to see if the TOE framework, which was designed in developed countries, can be applied to other non-Western cultures or developing countries.

## **5.2 *Managerial Implication***

To reassure their clients that e-banking services are secure, bank executives must increase the security elements of their systems. Putting a greater emphasis on positive safety elements could help to change negative customer opinions. As a result, banks should send a strong message to customers that present security is more than acceptable, allowing them to be rest-assured when using e-banking services. Bank executives must place a greater emphasis on marketing communications in order to widely disseminate the benefits of adopting e-banking. Banks can employ advertising to alter customers' perceptions toward e-banking services, which affects their willingness to utilize it.

## **5.3 *Policymakers***

Policymakers must create a priority list of technological, organizational, and environmental elements in order to implement IB adoption based on their organizations' actual capacity and capability. As a result, the Yemeni government should enact regulations and rules to safeguard electronic transactions and promote bank client trust. Furthermore, the study proposes that policymakers, bank executives, and practitioners should focus on maximizing TOE variables, which have been shown to influence and contribute to increasing SMEs' intentions toward IBS experimentally.

Additional research might be done because this study does not address all of the factors that influence SMEs' willingness to adopt e-banking services. Other aspects, such as regulatory support, cost, image, or quality of online services, should be studied further. Furthermore, looking into various hypotheses would be extremely beneficial to Yemeni decision-makers and scholars. Comparative studies in other nations with similar situations would greatly expand and improve current understanding.

## **6 Conclusion**

This research might aid bankers in understanding the existing low penetration rate of e-banking and devising strategies to boost e-banking adoption and acceptability by Yemeni SMEs, a nation where e-banking is still regarded an innovation.

It also adds to the literature on technology adoption and acceptance, which many researchers have suggested be expanded to new contexts, particularly in terms of the

TOE's generalizability and applicability in a new context (online banking), with a new user group (SMEs), and in a new cultural setting (Yemen), which is a critical step in moving a theory forward.

According to the findings, managers and owners should be prepared to deliver financial services via the Internet. Yemeni banks would be able to grow their customer base by offering IBS. Finally, as Internet technology becomes more prevalent and necessary, IB adoption will become increasingly more important. This recommends that managers and owners should devote more time and resources to assisting firms in developing their own skills to integrate IB into organizational functions through training programs.

Furthermore, the report proposes that Yemeni banks should offer specific training, workshops, and seminars aimed at assisting businesses in understanding the requirements for launching an online presence. Furthermore, this investigation contributed to the understanding of technology acceptance within the framework of technology acceptance theories study and, optionally, Internet behavior. The report provides information on the state of IB adoption among Yemeni SMEs and serves as a resource for academics, practitioners, and policymakers interested in encouraging enterprises to embrace IB.

**Acknowledgements** The authors would like to thank the United Arab Emirates University for funding this work under Start-Up grant G00003321. Also, the second author's research has been supported by Grant RMCG20-023-0023, Malaysia International Islamic University.

## References

1. Nasri W (2011) Factors influencing the adoption of internet banking in Tunisia. *Int J Bus Manag* 6(8):143–160
2. Khrawish HA, Al-Sa'di NM (2011) The impact of e-banking on bank profitability: evidence from Jordan. *Middle East Financ Econ* 13(1):142–158
3. Al Sukkar A, Hasan H (2005) Toward a model for the acceptance of internet banking in developing countries. *Inf Technol Dev* 11(4):381–398
4. Al-Fahim NH, Jusoh WJW, Abideen A (2014) Factors affecting the intention to adopt internet banking services among small and medium sized-enterprises in Yemen. *J Manag Bus Stud* 3(9):432–440
5. Al-Ajam AS, Nor KM (2013) Influencing factors on behavioral intention to adopt Internet banking service. *World Appl Sci J* 22(11):1652–1656
6. Zolait AHS, Ibrahim AR, Farooq A (2010) A study on the internet security and its implication for e-commerce in Yemen. *Int J Technol Diffus* 1(3):34–47
7. Jayawardhena C, Foley P (2000) Changes in the banking sector—the case of internet banking in the UK. *Internet Res*
8. Zolait AHS, Mattila M, Sulaiman A (2009) The effect of user's informational-based readiness on innovation acceptance. *Int J Bank Mark*
9. Al-Kibsi SM (2010) Customer perception of technology-banking services quality providing by banks operating in Yemen. PhD dissertation, pp 145–155
10. Rawashdeh A, Al-namlah L (2017) Factors influencing electronic data interchange adoption among small and medium enterprises in Saudi Arabia. *AJBA* 10(2):253–280

11. Al-Rawahna ASM, Chen S-C, Hung C-W (2019) The barriers of e-government success: an empirical study from Jordan. Available SSRN 3498847
12. Mols NP, Bukh PND, Nielsen JF (1999) Distribution channel strategies in Danish retail banking. *Int J Retail Distrib Manag*
13. Tornatzky LG, Fleischer M, Chakrabarti AK (1990) Processes of technological innovation. Lexington Books
14. Puklavec B, Oliveira T, Popovič A (2014) Unpacking business intelligence systems adoption determinants: an exploratory study of small and medium enterprises. *Econ Bus Rev* 16(2)
15. Rogers EM (1995) Diffusion of innovations: modifications of a model for telecommunications. In: *Die diffusion von innovationen in der telekommunikation*. Springer, Berlin, pp 25–38
16. Bhushan N et al (2014) Network densification: the dominant theme for wireless evolution into 5G. *IEEE Commun Mag* 52(2):82–89
17. Puklavec B, Oliveira T, Popovič A (2018) Understanding the determinants of business intelligence system adoption stages. *Ind Manag Data Syst*
18. Ammar AA (2017) Factors influencing intention to adopt mobile banking in Sudanese microfinance sector. Multimedia University (Malaysia)
19. Zhu K, Kraemer K, Xu S (2003) Electronic business adoption by European firms: a cross-country assessment of the facilitators and inhibitors. *Eur J Inf Syst* 12(4):251–268
20. Oliveira T, Fraga M (2011) Literature review of information technology adoption models at firm level
21. Sidek N (2015) Determinants of electronic payment adoption in Malaysia: the stakeholders' perspectives
22. Tan KS (2009) An empirical study of internet-based ICT adoption among SMEs in Southern Malaysia. Multimedia University
23. Hair JF (2009) *Multivariate data analysis*
24. Kline RB (2015) *Principles and practice of structural equation modeling*. Guilford Publications
25. Chin WW (1998) The partial least squares approach to structural equation modeling. In: *Quantitative methodology series*, no April, pp 295–336. <https://doi.org/10.1016/j.aap.2008.12.010>
26. Wold HO (1985) Partial least squares. In: Kotz S, Johnson NL (eds) *Encyclopedia of statistical science*. Wiley, New York, pp 581–591
27. Awa HO, Ojiabo OU, Orokor LE (2017) Integrated technology-organization-environment (TOE) taxonomies for technology adoption. *J Enterp Inf Manag*
28. Chatzoglou P, Chatzoudes D (2016) Factors affecting e-business adoption in SMEs: an empirical research. *J Enterp Inf Manag*
29. Gangwar H, Date H, Ramaswamy R (2015) Understanding determinants of cloud computing adoption using an integrated TAM-TOE model. *J Enterp Inf Manag*
30. Marakarkandy B, Yajnik N, Dasgupta C (2017) Enabling internet banking adoption: an empirical examination with an augmented technology acceptance model (TAM). *J Enterp Inf Manag*
31. Goode S, Stevens K (2000). An analysis of the business characteristics of adopters and non-adopters of World Wide Web technology. *Inf Technol Manage* 1(1):129–154
32. Pulakanam V, Voges KE (2010) Adoption of six sigma: review of empirical research. *Int Rev Bus Res Pap* 6(5):149–163
33. Hair JF, Black WC, Babin BJ, Anderson RE (2014) *Multivariate data analysis: Pearson new international edition*, vol 1. Essex Pearson Educ. Ltd., p 2
34. Hair Jr JF, Matthews LM, Matthews RL, Sarstedt M (2017) PLS-SEM or CB-SEM: updated guidelines on which method to use. *Int J Multivar Data Anal* 1(2):107–123
35. Cohen S (1988) Perceived stress in a probability sample of the United States
36. Chin WW (1998) The partial least squares approach to structural equation modeling. *Mod Methods Bus Res* 295(2):295–336

# Technological and Organizational Factors Influencing the Internet Banking Use Among SMES in Yemen: The Mediating Role of Attitude



Nabil Hussein Al-Fahim and Rawad Abdulghafor

## 1 Introduction

Internet banking services (IBS) are greatly useful to banks as well as to customers. They are useful to banks in that they save cost, reach a higher percentage of the population, are efficient and improve bank's reputation, and enable achieving excellent consumer service satisfaction [1]. IBS may involve several services and activities such as electronic fund transfers, point of sale (POS) machines in stores, automated teller machines (ATMs), mobile banking and money, e-statements, e-mails, phone calls, texts, and transaction alerts all of which are executed through secure lines [2].

IBS also provides banks with an unlimited distribution network, thereby offering a competitive advantage to them. As a result of adopting this technology, banks become capable of rendering their services online which lowers the costs of transactions and strengthens the bond between bankers and customers. Internet facilitates banks to provide excellent value-financial services at lower costs [3]. In spite of the fact that Yemeni banks produce electronic infrastructure and spend a humongous amount of money annually to adopt electronic banking, the adoption of Internet banking service is relatively low among small and medium-sized enterprises (SMEs) in Yemen [4–6]. Furthermore, [7] proclaim IBS as beneficial in that customers are offered access to bank services at a fast rate and with comfort without requiring physical interactions with the bank. In addition, establishing a comprehensive set of IBS services and products within banks not only improves their operation and interaction electronically with consumers but also saves cost and time and provides a competitive market advantage.

---

N. H. Al-Fahim · R. Abdulghafor (✉)

Department of Computer Science, Faculty of Information and Communication Technology,  
International Islamic University Malaysia, 53100 Kuala Lumpur, Malaysia

e-mail: [rawad@iium.edu.my](mailto:rawad@iium.edu.my)

Unfortunately, despite all the remarkable advantages of IBS, several consumers of financial institutions have yet to adopt these technologically advanced services proposed by the banking industry especially in the Arab countries and in Yemen in particular [8]. The cultural context of Yemeni people requires face-to-face interaction such as banking and personal contact which hinders the usability and adoption of IBS technology by retail and corporate customers. Despite the potential tension, very few empirical studies have been carried out to examine the factors which may influence the use of IB among SMEs in Yemen as well as the consequent potential impacts of such variables [9]. Thus, tremendous progress and research are required to fully define and understand the influential factors that make the country lags behind particularly in the organization such as SMEs. This study aims to extend the body of knowledge in the IB technology area such as its factors that impact its adoption among Yemeni SMEs. Therefore, the primary goal of this study is to examine the determining technological and organizational factors influencing the use of IB in the Yemeni SMEs sector. In addition, this study aims to investigate the role of attitude to use IB as a mediator between technological, organizational factors, and actual use of IB by SMEs in Yemen.

### ***1.1 Research Objectives***

To determine the technological factors (trust, system quality, ease of use) that could influence the attitude and actual use of IB by SMEs in Yemen.

To investigate the organizational factors (bank support, ICT readiness, customer readiness) that could influence the attitude and actual use of IB by SMEs in Yemen.

To examine the role of attitude towards IB usage as a mediator between (trust, system quality, bank support, ICT readiness, customer readiness, ease of use) and actual use of IB by SMEs in Yemen.

### ***1.2 Research Gap***

Despite the provision of electronic infrastructure by Yemeni banks and spending millions of dollars annually to adopt electronic banking, IB adoption is very low and minimal among (SMEs) in Yemen [10]. Thus, the lack of comprehensive IB research on SMEs and customers is one of the main gaps globally and in Yemen's IB research particularly.

The literature indicates that few studies have utilized the TOE framework (technological and organizational factors) model to discuss the use of IB from SMEs viewpoint in Arab countries region in general and Yemen in particular. Therefore, this study is essential as it enriches the literature with new findings by exploiting the TOE theory to examine the use of IBs among SMEs in Yemen which will be useful

for future academics and researchers to be able to use the variables of TOE theory in their contextual studies.

The researcher also expects that the organizations that depend on IB or transact their products or services via the Internet could benefit from the results of this study after knowing the factors that influence organizations' attitude or behaviour for SMEs to use new technology. Therefore, this study could contribute to help service, manufacturing, and trading organizations such as banks, airlines, insurance companies, and the health sector as well as export and import companies to understand the factors influencing SMEs attitudes regarding the use of technological services. Furthermore, the availability of financial products and information via the Internet is a top priority for these organizations in seeking and planning to do e-commerce in Yemen if the local banks cannot provide online transactions that are necessary for e-business, e-services e-commerce, e-government, and other online activities of these organizations.

## **2 Literature Review**

### ***2.1 IBS Definition***

The literature indicates different definitions of IB in which each definition depends on the scope, type of services, and access limit offered to consumers. Some authors consider the IB as an internal portal that comprises several banking services that are accessible for customers [11]. According to Rotchanakitumnuai and Speece [12], IB is defined as a transactional system that allows customers to interact with online banking activities and services. Other researchers defined IB as a practice to conduct financial transactions by bank consumers over the Internet using a secured online banking website. A wider definition of IB would be the use of technology to access banking activities and services through banks' websites using a private or public network.

Indeed, IB is an online transaction tool that enables customers and banks to carry out their online transactions and offer several benefits such as balance checking, receiving or making bills' payment and money transfer, applying for loans, and making investments. Despite all these benefits of IB, several factors affect the IB services and activities in developing countries such as trust, system quality, ease of use, financial institution support, customer readiness, and ICT in order to get new online SMEs who are Internet users.

## 2.2 Factors Influencing the Use of IB

### 2.2.1 Technological Factors

Technological context is the first component of the TOE model which includes both the external and internal technologies used by the firm. One of the early usages of the TOE model in researches was to study the “technological factors”, where the complexities of IT infrastructure, system development, and management formalization were considered as the important factors to represent the technological context of TOE framework [13]. In [14], the TOE model has been used to determine the information system (IS) adoption among SMEs and to determine the factors influencing the SMEs when deciding to adopt new technology or system within the firm. They found that the technological factors are essential to determine the adoption and implementation of a new system as well as concluded that the use of technological factors could benefit the SMEs internally and externally.

#### Trust

Without any doubt, trust in the system is considered one of the most essential factors in all social activities because it does not only involve uncertainty and dependency but also the degree of uncertainty in online transaction systems such as e-commerce is very high compared to the traditional physical trade or commerce [15]. Therefore, trust in the system or technology is a very important factor that sometimes hinders the intention of SMEs or individuals to use IB and online interactions such as e-banking and e-commerce.

According to Montazemi and Qahri-Saremi [16], investing factors other than trust are effective. They stressed the importance of educating customers about the “usefulness” and “trustworthiness” of online IB activities, services, and interactions. Indeed, trust is an important factor that decides the rejection or adoption of customers to use e-banking or e-commerce technology [17]. This is because significantly positive affects the attitudes of customers as well as their satisfaction with finance-related technology [18]. Therefore, this research hypothesizes a positive relationship as follows:

- Hypothesis 1a (H1a) Trust has a statistically significant and positive impact on attitude to use IB.
- Hypothesis 1b (H1b) Trust has a statistically significant positive impact on the actual use of IB.
- Hypothesis 1c (H1c) Attitude mediates the relationship between trust and actual use of IB.

#### Ease of Use

Ease of use describes the perceived ease of use of difficulty to use technology either by the organization members or by their customers [19]. Therefore, technology should be easy to use; otherwise, the possibility of its adoption and usage will reduce [20]. Conducted a meta-analysis study and found that the ease of use and compatibility factors have a major impact on technology usage, innovation, and adoption. In [21],



the importance of the ease-of-use factor has been highlighted for firms when implementing new technologies. Previous literature has studied ease of use as the most significant construct, influencing CC adoption [21, 22]. It is, therefore, hypothesized that:

- Hypothesis 2a (H2a) Ease of use has a statistically significant and positive effect on attitude to use IB.
- Hypothesis 2b (H2b) Ease of use has a statistically significant and positive influence on actual use of IB.
- Hypothesis 2c (H2c) Attitude mediates the relationship between ease of use and actual use of IB.

### **System Quality**

System quality is considered a significant factor when deciding to implement or use new technology, and it refers to the system processing properties such as speed service delivery, control, ease of use, and enjoyment [23]. System quality measures the desired characteristics of new technology such as IB in terms of its availability, accessibility, functionality, cost, speed, resources, and uniqueness all of which can be perceived by customers and thus influence their decisions to use or adopt new technology. Thus, the following hypothesis is proposed:

- Hypothesis 3a (H3a) System quality has a statistically significant and positive influence on attitude to use IB.
- Hypothesis 3b (H3b) System quality has a statistically significant and positive influence on the actual use of IB.
- Hypothesis 3c (H3c) Attitude mediates the relationship between system quality and the actual use of IB.

### **Organizational Context**

The second element of the TOE model that explains the decision-making behaviour of a firm in relation to technological innovation is the organizational factor. The organizational factors refer to the explicit characteristics and variables of firms. Three important components, namely the size and scope and complexity of firms, are part of the organizational context variables [24]. The top management is one of the earlier researches that were examined using the TOE framework.

According to Puklavec et al. [25], the slack is an important factor of IB system adoption as it generally influences the level of voluntariness and sensitivity to resources availability. Several studies have investigated the organizational factors of the TOE and found that organizational readiness is an essential variable that influences the adoption of IB systems and other technologies and innovations within firms [26, 27].

### **ICT Readiness**

Technological environment, both telecommunication and electronic, where certain firms operate influences ICT adoptions. The lack of information technology infrastructure could act as a barrier to sustaining the growth of e-commerce. ICT is a

combination of information technology and communication technology. The literature indicates that IT readiness has two basic elements which are the personal IT knowledge and technical readiness that are the precursors to IS adoption and innovation [28, 29] studies which postulate the importance of information technology infrastructure in the adoption of e-banking in Ethiopia. Moreover, in [30], it has been found that the development of ICT infrastructure was significantly associated with m-banking development in Bangladesh. Hence, the researcher hypothesizes that

- Hypothesis 4a (H4a) ICT readiness has a statistically significant and positive influence on attitude to use IB.
- Hypothesis 4b (H4b) ICT readiness has a statistically significant and positive influence on the actual use of IB.
- Hypothesis 4c (H5c) Attitude mediates the relationship between ICT readiness and actual use of IB.

### **Bank Support**

Owing to their remarkable roles, banks are very important to the technological facilities as they are considered as the resources in the economic structure which allows them to act as intermediaries between investors and households to establish various types of enterprises [31]. All sorts of either SMEs or large firms depend on banks to get access to the required financial resources and supports to accomplish growth or even survival of their business. In the ever-changing and competitive environment, banks should enhance their performance on a continuous basis to gain sustainable competitive abilities. Therefore, any bank that has online channels aims to provide banking services that require huge investments to make reliable channels for their customers to use these channels [32]. In terms of bank support, in [33] it has been stated that the IB aids both banks and their customers and that IB is illustrated as a sharing wallet for both financial SMEs and institutions. In [34], it has been found that large banks, private ownership, newer banks with higher expenses for fixed assets, and higher deposits were more likely to adopt IB. They also indicate that these banks were attracting more customers through this new channel of delivery. Therefore, the following hypotheses are formulated:

- Hypothesis 5a (H5a) Bank support has a statistically significant and positive impact on attitude to use IB.
- Hypothesis 5b (H5b) Bank support has a statistically significant and positive impact on the actual use of IB.
- Hypothesis 5c (H5c) Attitude mediates the relationship between bank support and actual use of IB.

### **Organization Readiness**

Organization readiness refers to the availability level of both technical and financial resources in the firms to achieve the adoption of knowledge management systems [35, 36]. Therefore, the term “perceived organization resources” denotes firms’ perceptions towards the level of the available technical, financial, and human resources

[37]. Firms that have achieved a level of sophistication with regard to technological readiness are often seen to be extremely highly integrated organizations especially in terms of computerized services and processes and are considered as the firms that are ready to adopt IT innovations with greater benefits [14].

In [38], it has been stressed that organizational readiness includes relevant systems, suitable infrastructure, technical skills, and resources. Although organization readiness has been defined in the literature from different contexts, all agreed that it has a significant and positive impact on SMEs technological adoption. As such, the researcher hypothesizes that:

- Hypothesis 6a (Ha)    Organization readiness has a statistically significant and positive impact on attitude to use IB.
- Hypothesis 6b (H6b)    Organization readiness has a statistically significant and positive impact on the actual use of IB.
- Hypothesis 6c (H6c)    Attitude mediates the relationship between organization readiness and actual use of IB.

### **Attitude Towards IB**

The effect of attitude on behaviour has been showed through behaviour intention in TRA [39]. However, some previous studies suggested that attitude impacts the behaviour in different manners and that behaviour is a strong predictor of attitude than the actual usage. In [40, 41], it has been indicated that the customers' attitude towards system characteristics on adoption of IBS, by Malaysian bank customers, is related to IBSA, and it plays a significant role in Internet banking adoption. Furthermore, a related study [42], which was conducted on 124 respondents, finds a significant relationship between customers' attitudes and IBA.

The literature indicates that the user attitude has a strong and direct impact on the intentions of customers to use and or adopt new technology. For example, in [42], it has been found that social pressure could influence the behaviour of individuals and increase their intentions to use IB. Previous studies indicated that the attitudinal factors influencing the adoption and use of new technology include the perception of customers and their experience and characteristics of the system.

Therefore, the hypothesis is formulated as follows:

- Hypothesis H7 (H7)    Attitude has a statistically significant and positive influence on IB actual use.

## **3 Conceptual Framework**

### **3.1 Methodology and Data Collection**

The primary method used to collect the relevant data in this study is the survey questionnaire method which was designed and divided into two parts. The first part

involves the demographic questions to know the basic information about respondents such as age, gender, experience, education, and income. The second part of the questionnaire asks the participants about the factors of interest in this study which are (1) actual usage, (2) attitude, (3) bank supported, (4) ICT readiness, (5) organization readiness, (6) system quality, (7) ease of use, and (8) trust. The first variable is endogenous, and factors 2 as a mediator as well as from 3 to 8 are exogenous variables.

The total number of distributed questionnaires is 374, and the returned ones are 302 which represents an 80% response rate. From the 302 returned questionnaires, 22 cases were outliers; thus, the total usable number returned questionnaires are 280 which represent a 74% response rate after the outliers were removed. Therefore, the sample size of  $n = 280$  is sufficient for this research. The ( $n = 280$ ) sample size achieved a ratio of 5:1 which means that five respondents are needed per item (indicator) as recommended by Shi et al. [43] and Hair et al. [44]. According to Kline [45], the minimum of at least, the five participants for each estimated variable (estimation of about 40 parameters, the studies on a minimum sample size of 200 participants).

## 4 Analysis and Results

### 4.1 Descriptive Statistics for Variables

Table 1 shows the descriptive statistics for the constructs in this study. It can be observed that the highest mean obtained for bank support with 3.328 and the actual usage with 3.310 out of a maximum of 5 is making up 66%. Furthermore, attitude to use IB and system quality have 3.217 and 3.244 making up 64%. In addition, other variables such as trust, complexity, and customer readiness have a mean of 3.157, 3.182, and 3.161, respectively, which makes a total of 63%, and the overall mean is 3.229 which is more than 3 [46]. In addition, the calculated standard deviation for all

**Table 1** Descriptive statistics for factors

	Variable	Code	No. items	Mean	%	S.D.
	Actual usage	AU	5	3.3102	0.662	0.92378
	Attitude to use IB	ATT	7	3.2179	0.643	0.80555
Technological factors	System quality	QS	5	3.2446	0.648	0.95271
	Ease of use	EOU	4	3.1616	0.632	1.0670
	Trust	TRU	3	3.1571	0.631	0.87449
Organizational factors	Bank support	BS	5	3.3280	0.665	0.85216
	ICT readiness	ICT	5	3.2325	0.646	1.0149
	Organization readiness	OR	4	3.1822	0.636	0.96940
Overall			38	3.229	0.645	0.67324

the variables in this study ranged from 0.805 to 0.1067 which denotes the existence of acceptable variability within the data set.

## **4.2 Validity and Reliability**

### **4.2.1 Convergent Validity**

The items loading in this study are more than 0.5 which is acceptable since the study sample is more than 200 as suggested by Hair et al. [46]. This indeed is sufficient evidence of convergent validity, where the range of the factor loading lies between 0.614 (BS3) and 0.910 (ICT6). Therefore, in this study, all the indicators are related to their specific constructs, and thus, there is satisfactory proof of the convergent validity of the method used in this study. Table 2 shows the items loading, AVE, Cronbach's alpha, and CR.

### **4.2.2 Discriminant Validity**

Table 3 presents the discriminant validity. The discriminant validity denotes the distinctness of the study constructs from other constructs as mentioned by Hair et al. [44]. Average variance extracted (AVE) is an important statistical test that can be used to evaluate the discriminant validity for every construct that may exceed the correlation among other constructs.

In this study, the discriminant validity was validated since the AVE obtained values are more than the correlations squared for each ever variables. Besides, the AVE was greater than the absolute correlation square values of the given constructs with respect to other constructs ( $AVE > \text{correlation square}$ ) as shown in Table 3. The results also indicate that maximum shared variance (MSV) was less than AVE. Table 3 presents discriminant validity for all variables.

## **4.3 Confirmatory Factor Analysis (CFA)**

In this study, items with load loading and high error to achieve the goodness of the fit were excluded from the measurement of model fit because they will affect the goodness of the fit of the model. Modification indices inspection (MI) revealed that the error of items (BS5 and BS3), (CR1 and CR2), (QS4 and QS5), (AU3 and AU5), and (ATT1 and ATT2) should be correlated due to which they are considered to have high correlation as shown in Fig. 1. In this study, all the indicators showed excellent ability to measure the good of the fit as suggested by Hair et al. [48]. The final model showed the ratio of the chi-square to the degree of freedom (normed  $\chi^2$ ) was 1.720, less than 5 shows a good model fit, and the RMSEA is 0.048, less than 0.08 which

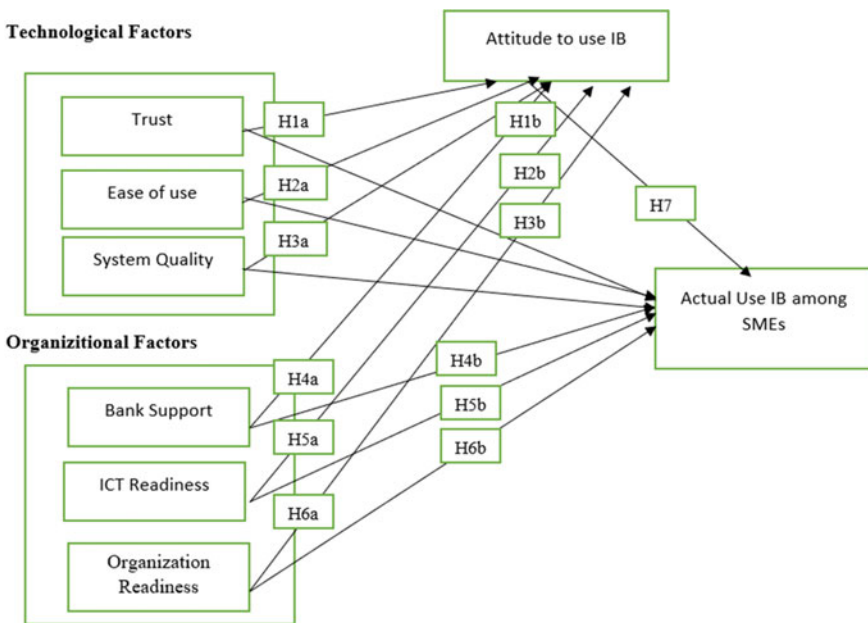
**Table 2** Items loading, Cronbach's alpha, composite reliability, and AVE

Constructs	Items	Factor loading	Cronbach's alpha	Composite reliability	AVE
Actual usage	AU1	0.896	0.933	0.930	0.728
	AU2	0.843			
	AU3	0.830			
	AU4	0.859			
	AU5	0.836			
Attitude	ATT1	0.780	0.914	0.913	0.601
	ATT2	0.788			
	ATT3	0.810			
	ATT4	0.803			
	ATT5	0.756			
	ATT6	0.739			
	ATT7	0.748			
System quality	SQ1	0.886	0.937	0.938	0.752
	SQ2	0.837			
	SQ3	0.858			
	SQ4	0.876			
	SQ5	0.879			
Ease of use	EOU1	0.889	0.939	0.939	0.794
	EOU2	0.884			
	EOU3	0.885			
	EOU5	0.905			
Trust	TRU1	0.878	0.880	0.881	0.713
	TRU2	0.784			
	TRU4	0.868			
Bank support	BS1	0.836	0.884	0.975	0.589
	BS2	0.812			
	BS3	0.614			
	BS4	0.891			
	BS5	0.646			
ICT readiness	ICT1	0.897	0.944	0.944	0.773
	ICT2	0.863			
	ICT3	0.874			
	ICT4	0.850			
	ICT5	0.910			
Organization readiness	OR1	0.908	0.940	0.927	0.761
	OR2	0.875			
	OR3	0.839			
	OR4	0.866			

*Source* Computed data analysis

**Table 3** Discriminant validity for [47]

	AVE	MSV	BS	QS	TRU	EOU	ICT	OR	ATT	AU
BS	0.589	0.255	<b>0.768</b>							
QS	0.725	0.338	0.482	<b>0.867</b>						
TRU	0.713	0.555	0.428	0.469	<b>0.844</b>					
EOU	0.794	0.241	0.373	0.314	0.377	<b>0.891</b>				
ICT	0.773	0.394	0.460	0.575	0.350	0.411	<b>0.879</b>			
OR	0.761	0.394	0.468	0.527	0.389	0.463	0.628	<b>0.872</b>		
ATT	0.601	0.555	0.499	0.581	0.745	0.491	0.492	0.472	<b>0.775</b>	
AU	0.728	0.389	0.505	0.576	0.495	0.396	0.496	0.507	0.624	<b>0.853</b>



**Fig. 1** Conceptual framework and hypotheses testing

is a good fit [46]. Also, other measures revealed GOF of the model to the data (CFI = 0.955, IFI = 0.956, TIL = 0.950) which indicated that the model employed in this paper is a good fit to data [48, 49]. Figure 1 shows the measurement model for exogenous and endogenous variables.

**Determination Coefficient: R<sup>2</sup> Value**

The R<sup>2</sup> value describes the strength or weaknesses of the correlation and amount of variance between the independent and dependent factors. In this study, the values of the R<sup>2</sup> squared multiple correlations of the structural model on attitude and actual

**Table 4** Determination coefficient ( $R^2$ )

Exogenous construct	Endogenous construct	$R^2$	Hair et al. [50]	Cohen [51]	Chin [52]
SQ, EOU, TRU, BS, ICT, OR	Attitude to use IB	0.57	Moderate	Substantial	Moderate
SQ, EOU, TRU, BS, ICT, OR, ATT	Actual usage	0.33	Moderate	Moderate	Moderate

usage of IB were 0.57 and 0.33, respectively. Hence, the results show all exogenous variables: bank support, ICT readiness, customer readiness, system quality, ease of use, and trust explained 57% of the variance in attitude to use IB. In addition, bank support, ICT readiness, customer readiness, system quality, trust, ease of use, and attitude explained 33% of the variance in actual usage of IB as shown in Fig. 1 and Table 4.

#### 4.4 Structural Model

The structure model is explained using the hypotheses testing described in the following sections. This study determines the effect of TOE factors on actual usage of IB via the mediating role of attitude by hypotheses testing. The structural model results showed the model fit indices such as RMSEA is 0.068 which is less than 0.08 and considered a good fit [44, 53]. In addition, the goodness of fit of the model in this study was confirmed using other values of the GOF of the model to the data (CFI = 0.909, IFI = 0.910, TIL = 0.901) which indicates the data. Figure 2 demonstrates the structural model results with the standardized estimate (goodness of fit indices).

#### 4.5 Direct Hypothesis Results

The empirical study tested four main hypotheses that are related to the goal of this study. With regard to technological factors hypotheses, the results show system quality, ease of use, and trust had a significant and positive influence on attitude to use IB ( $\beta = 0.205$ ;  $t = 3.803$ ;  $P = 0.000$ ), ( $\beta = 0.206$ ;  $t = 4.408$ ;  $P = 0.000$ ), and ( $\beta = 0.592$ ;  $t = 8.978$ ;  $P = 0.000$ ), respectively. Therefore, H1a, H2a, and H3a are supported.

According to the organizational factors' hypotheses, bank support and ICT readiness are the important factors which influence attitude to use IB among managers/owners' SMEs in Yemen ( $\beta = 0.109$ ;  $t = 2.343$ ;  $P = 0.019$  and  $\beta = 0.118$ ;  $t = 2.609$ ;  $P = 0.009$ ), respectively. So, H5a and H4a are supported. However, organization readiness was not significantly influenced by attitude to use IB. In addition, bank support, organization readiness, and system quality had a significant and positive effect on actual usage of IB, but trust and ease of use were no significant influence



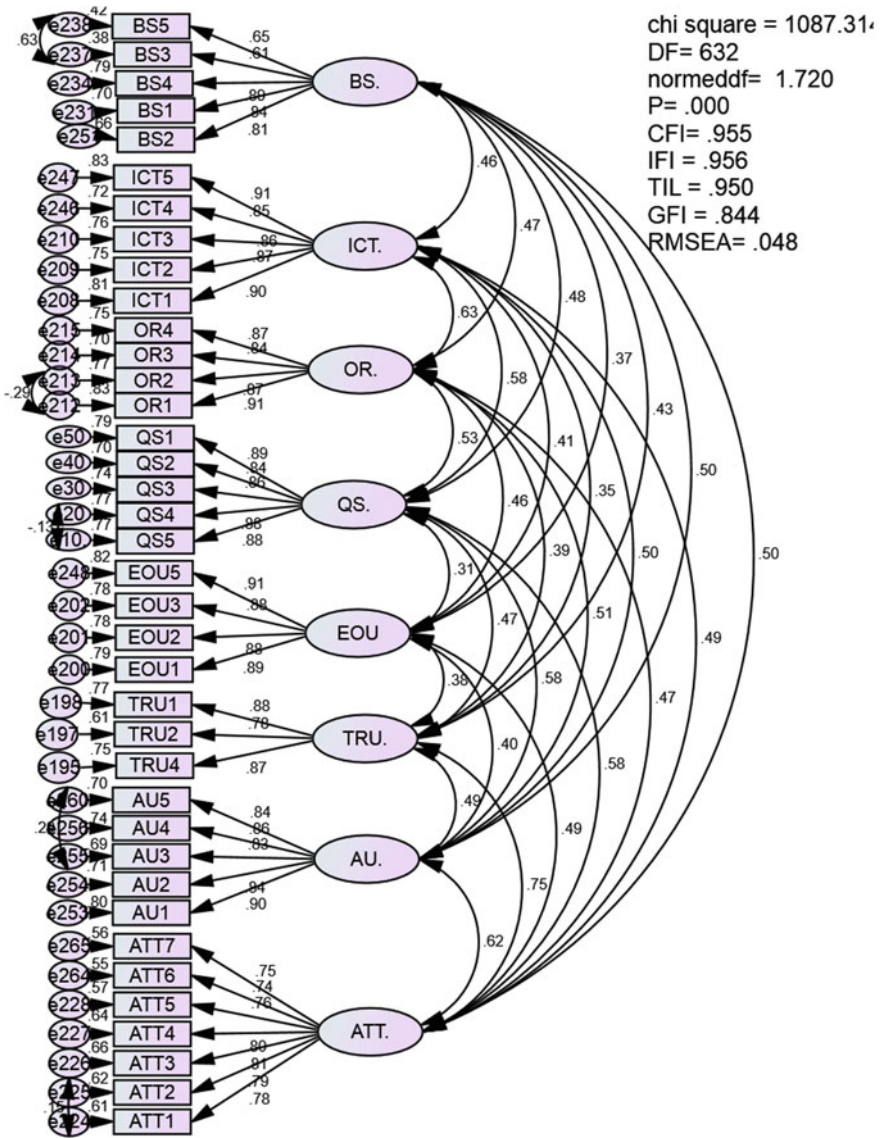


Fig. 2 Full measurement model with standardized estimates

on actual usage of IB among SMEs’ owners and managers in Yemen. Therefore, H5b and H6b are supported, but H1b and H2b were not supported. Finally, attitude had statistically significant and positive effect on actual usage of IB ( $\beta = 0.316$ ;  $t = 3.371$ ;  $P = 0.000$ ). Table 5 shows the results of the direct hypotheses (Fig. 3).

**Table 5** Direct hypotheses testing results of structural model

Hypothesis	Exog.	Endo	Estimated $\beta$	E.S	C.R <i>t</i> -value	<i>P</i> -value	Status	Result
H1a	Trust	Attitude	0.592	0.051	8.978	0.000	Sig.	Supported
H2a	Ease of use	Attitude	0.206	0.031	4.408	0.000	Sig.	Supported
H3a	System quality	Attitude	0.205	0.037	3.803	0.000	Sig.	Supported
H4a	ICT readiness	Attitude	0.118	0.030	2.609	0.009	Sig.	Supported
H5a	Bank support	Attitude	0.109	0.036	2.343	0.019	Sig.	Supported
H6a	Organization readiness	Attitude	0.009	0.030	0.192	0.874	No-sig.	Not supported
H1b	Trust	Actual usage	0.034	0.082	0.396	0.692	No-sig.	No supported
H2b	Ease of use	Actual usage	0.048	0.045	0.856	0.390	No-sig.	No supported
H3b	System quality	Actual usage	0.238	0.053	3.782	0.00	Sig.	Supported
H4b	ICT readiness	Actual usage	0.084	0.043	1.584	0.113	No-sig.	No supported
H5b	Bank support	Actual usage	0.163	0.052	2.998	0.003	Sig.	Supported
H6b	Organization readiness	Actual usage	0.144	0.043	2.772	0.006	Sig.	Supported
H7	Attitude	Actual usage	0.316	0.116	3.371	0.000	Sig.	Supported

### 4.6 Indirect Hypothesis Results—(Mediating Role)

#### 4.6.1 Mediating Influence of Attitude Between Technological, Organization Factors, and Actual Usage of IB

The results in Table 6 showed that the attitude had a full mediation effect on the relationship between organization readiness, ease of use, trust, and actual usage of IB. This is because the path coefficients for indirect effect were 0.037, 0.187, and 0.065, and *p*-values were 0.024 and 0.001 ( $p < 0.05$ ), respectively. Therefore, H6c, H2c, and H3c are supported. In addition, attitude to use IB had partial mediation between bank support, system quality, and actual usage of IB. Therefore, hypotheses HH5c and H3c are supported, but it was no mediation between ICT readiness and actual usage due to which path coefficient and *p*-value were not significant for indirect effect (0.003;  $p = 0.864$ ). H4c is not supported as shown in Table 6. It can be concluded that the attitude of SMEs managers mediates on the relationship between organization

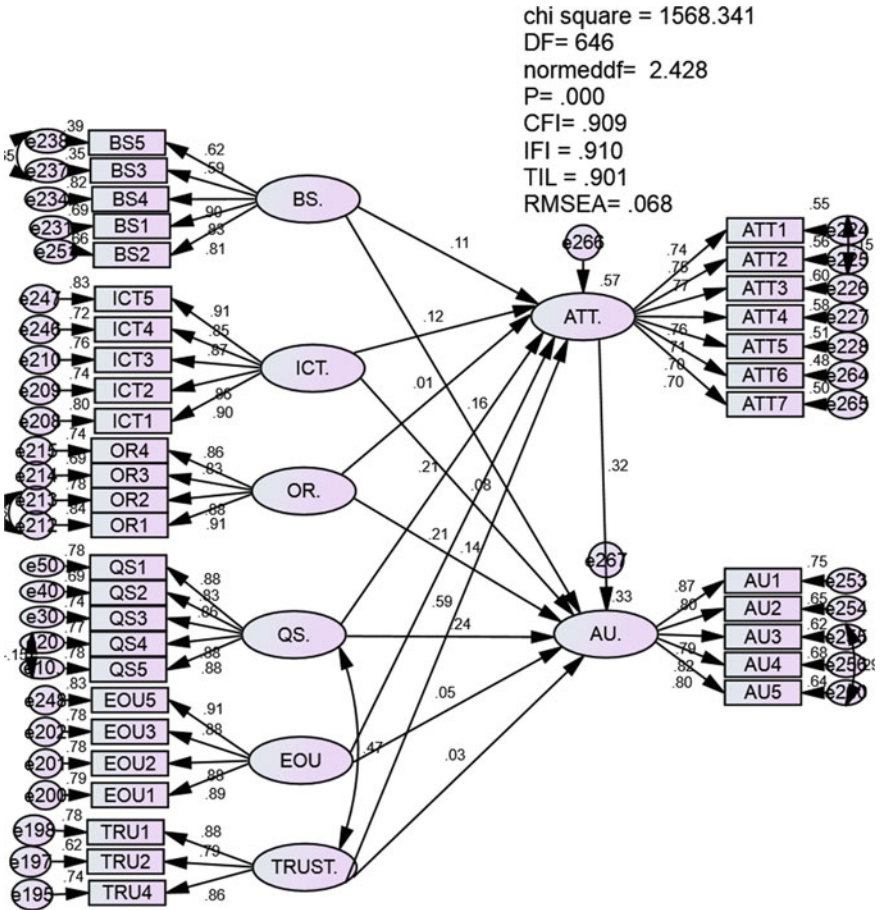


Fig. 3 Structural model with standardized estimated

readiness, ease of use, trust, bank support, system quality, and actual usage of IB by SMEs in Yemen.

## 5 Discussion and Implication

### 5.1 Discussion

The main objective is to investigate and examine the technological and organizational variables that influence the use of IB services by SMEs in Yemen and the role of attitude as a mediator that influences the relationship between TOE factors

**Table 6** Structural parameters of the mediation role of attitude between technological, organizational factors, and actual usage of IB

H	Relationship	Path “a”	Path “b”	Direct effect (c)	Indirect effect (c’)	Significant of indirect effect	Mediation type
H1c	TRU → ATT → AU	0.592	0.316	0.034	0.065	0.001	Full mediation
H2c	EOU → ATT → AU	0.206	0.316	0.048	0.187	0.001	Full mediation
H3c	SQ → ATT → AU	0.205	0.316	0.238	0.065	0.001	Partial mediation
H4c	ICT → ATT → AU	0.118	0.316	0.084	0.003	0.864	No mediation
H5c	BS → ATT → AU	0.109	0.316	0.163	0.034	0.031	Partial mediation
H6c	OR → ATT → AU	0.009	0.316	0.144	0.037	0.024	Full mediation

and actual usage of IB. This study investigates the factors of Internet banking adoption among SMEs’ managers in Yemen, using SEM software (Amos). The results of this study indicated that bank support, ICT readiness, system quality, and ease of use had positively significant effect on attitude to use IB, but ICT readiness was no significant effect. In addition, organization readiness, system quality, and attitude had a significantly influence on the actual use of IB. However, ease of use and trust were not significant influences on using of IB among SMEs in Yemen. Furthermore, attitude mediated the relationship between organization readiness, ease of use, trust, bank support, system quality, and actual usage IB as stated earlier, and this study aims at addressing the applicability of the TOE framework (technological and organizational factors) which was introduced in developed countries to developing countries especially Yemen. Thus, this study proposed the TOE model that accounts for the utilization of the united model within the adoption of Internet banking. This study focusses to address the applicability of the TOE model (technological and organizational factors).

### 5.2 Practical Implication

The main significance of this research is to statistically validate the factors that influence SMEs attitudes and the actual use of IB. Therefore, it was observed that the SMEs with greater system quality, customer readiness, bank support, trust, and attitude are more likely to become users of IB. This study emphasized the role of attitude as a significant mediator in the use of IB. This means that SMEs should be ready in terms of attitude in order to use IB in their business transactions. Attitude

involves the culture, consistent values, and work practices in the organization itself. Moreover, this study also proves that organizational and technological factors are influential in shaping their organizations' attitude to use IB. Therefore, banks should focus on TOE parameters to improve the number of IB users among their customers.

### ***5.3 Managerial Implication***

Bank managers must emphasize the positive features that their e-banking services offer to customers such as safety, security, and trust which improve the perceptions of customers towards using IB services. Therefore, banks should keep their customers well informed by sending them messages to inform them of the current security adequacy to improve the feelings of safety when using IB services. Bank managers should also publicize their messages to reach a wider population by using public newspapers and advisement in social media to influence their customers' attitude to use e-bank services which consequently influences the attitude to use IB.

### ***5.4 Policy-Makers***

Policy-makers are required to establish a list of all the organizational and technological factors that are necessary to implement IB according to the actual capability and capacity of their organizations. Thus, in order to foster and improve trust among IB customers in Yemeni SMEs, the government should implement policies, regulations, and laws to protect electronic transactions and activities. Moreover, this study recommends that the policy-makers, practitioners, and bank managers of SMEs in Yemen to focus more on maximizing the TOE factors which have been proven empirically to contribute and influence the improvement of SMEs attitudes towards the use of IB. In addition, more studies on the topic should also be considered because this study did not cover all the variables that may influence the attitude to use e-banking services by SMEs in developing countries. Therefore, more research could focus on other variables such as ICT and customer readiness, bank support, ease of use, quality system, and quality of online services. In addition, further research should also examine other theories which would further support the decision-making process of SMEs to implement and sustain IB. This study also recommends conducting comparative studies in other countries with similar conditions which would improve the body of knowledge in this area.

## 5.5 Conclusion

This study could be beneficial to bankers and SMEs managers as it provides a benchmark that explains the current low penetration rate of IB to formulate efficient strategies to encourage the attitude of using IB by Yemeni SMEs. In Yemen, the IB is still considered as a new technology that requires more investigations and trust so that it can be fully used and adopted by customers and organizations. This study contributes to the body of knowledge on TOE framework variables (technological and organizational) to examine IB technology adoption, acceptance, and usage which may aid future researchers to extend such variables or use the current model to a new context. TOE framework has been established as an efficient framework to study the technology adoption and acceptance (IB and e-banking) in terms of applicability and generalizability with different SMEs and in different cultural settings (e.g. Yemen). This study indicates that managers/owners should prepare to provide financial services via the Internet and ensure the safety and security of their customer's transactions and online banking activities. Offering IBS would mean that Yemeni banks could expand their customer base, gain a competitive advantage, and sustain their business in the online market. As Internet technology spreads and becomes a necessity, IB adoption becomes even more significant. This suggests to managers/owners to put more effort to assist organizations to develop their own capability to integrate IB into organization functionalities via training programmes. Moreover, this study suggests that Yemeni banks need to provide specialized courses workshops, and seminars targeted at helping organizations understand the prerequisites for launching their Internet presence. Additionally, this work added to understand acceptance of the technology. The study gives and provides points of reference for practitioners, policy-makers, and academics in promoting firms to IB adoption.

## References

1. Nasri W (2011) Factors influencing the adoption of internet banking in Tunisia. *Int J Bus Manag* 6(8):143–160
2. Nuwagaba A, Brighton N (2014) Analysis of e-banking as a tool to improve banking services in Zambia. *Int J Bus Manag Invention* 3(11):62–66
3. Sukkar AA, Hasan H (2005) Toward a model for the acceptance of internet banking in developing countries. *Inf Technol Dev* 11(4):381–398
4. Al-Ajam AS, Nor KM (2013) Influencing factors on behavioral intention to adopt Internet banking service. *World Appl Sci J* 22(11):1652–1656
5. Al-Fahim NH (2015) Factors influencing the intention to adopt internet banking services among small and medium enterprises (SMEs) in Yemen
6. Al-Fahim NH, Jusoh WJW, Abideen A (2016) An examination factors influencing the intention to adopt internet banking among SMEs in Yemen: using an extension of the technology acceptance model (TAM). *J Internet Bank Commer*
7. Jayawardhena C, Foley P (2000) Changes in the banking sector—the case of internet banking in the UK. *Internet Res*
8. Zolait AHS, Ibrahim AR, Farooq A (2010) A study on the internet security and its implication for e-commerce in Yemen. *Int J Technol Diffus (IJTD)* 1(3):34–47

9. Parker C, Castleman T (2007) New directions for research on SME-eBusiness: insights from an analysis of journal articles from 2003–2006. *J Inf Syst Small Bus* 1(1):21–40
10. Nelson P, Richmond W (2007) Internet banking: gold mine or money pit. *Acad Bank Stud J* 6(1):1–25
11. Teo HH, Oh LB, Liu C, Wei KK (2003) An empirical study of the effects of interactivity on web user attitude. *Int J Hum Comput Stud* 58(3):281–305
12. Rotchanakitumnuai S, Speece M (2003) Barriers to internet banking adoption, qualitative study among corporate customers in Thailand. *Int J Bank Mark* 21(6):321–323
13. Chau PY, Tam KY (1997) Factors affecting the adoption of open systems: an exploratory study. *MIS Q* 1–24
14. Gangwar H, Date H, Ramaswamy R (2015) Understanding determinants of cloud computing adoption using an integrated TAM-TOE model. *J Enterp Inf Manag*
15. Zhou RG, Li W, Huan TT, Shen CY, Li HS (2014) An online banking system based on quantum cryptography communication. *Int J Theor Phys* 53(7):2177–2190
16. Montazemi AR, Qahri-Saremi H (2015) Factors affecting adoption of online banking: a meta-analytic structural equation modeling study. *Inf Manag* 52(2):210–226
17. Gupta D, Kamilla U (2014) Cyber banking in India: a cross-sectional analysis using structural equation model. *IUP J Bank Manag* 13(2)
18. Padachi K, Howorth C, Narasimhan MS, Durbarry R (2010) Working capital structure and financing pattern of Mauritian SMEs. In: Oxford business & economics conference program
19. Straub ET (2009) Understanding technology adoption: theory and future directions for informal learning. *Rev Educ Res* 79(2):625–649
20. Gutierrez A, Boukrami E, Lumsden R (2015) Technological, organisational and environmental factors influencing managers' decision to adopt cloud computing in the UK. *J Enterp Inf Manag*
21. Zhang H, Vorobeychik Y (2019) Empirically grounded agent-based models of innovation diffusion: a critical review. *Artif Intell Rev* 52(1):707–741
22. Chong AYL, Lin B, Ooi KB, Raman M (2009) Factors affecting the adoption level of e-commerce: an empirical study. *J Comput Inf Syst* 50(2):13–22
23. Althonayan M, Althonayan A (2017) E-government system evaluation: the case of users' performance using ERP systems in higher education. *Transforming Gov People Process Policy*
24. Tornatzky LG, Fleischer M, Chakrabarti AK (1990) Processes of technological innovation. Lexington Books
25. Puklavec B, Oliveira T, Popović A (2014) Unpacking business intelligence systems adoption determinants: an exploratory study of small and medium enterprises. *Econ Bus Rev* 16(2)
26. Yoon TE, Ghosh B, Jeong BK (2014) User acceptance of business intelligence (BI) application: technology, individual difference, social influence, and situational constraints. In: 2014 47th Hawaii international conference on system sciences. IEEE, pp 3758–3766
27. Senyo PK, Effah J, Addae E (2016) Preliminary insight into cloud computing adoption in a developing country. *J Enterp Inf Manag*
28. Al-Somali SA, Gholami R, Clegg B (2011) Determinants of B2B e-commerce adoption in Saudi Arabian firms. *Int J Digit Soc (IJDS)* 2(2):406–415
29. Gutierrez A, Boukrami E, Lumsden R (2015) Technological, organisational and environmental factors influencing managers' decision to adopt cloud computing in the UK. *J Enterp Inf Manag*
30. Ahad MT (2014) Investigating a framework for providing mobile banking opportunities to rural SMEs in Bangladesh. Doctoral dissertation
31. Al-Swidi AK, Mahmood R (2011) How does organizational culture shape the relationship between entrepreneurial orientation and the organizational performance of banks? *Eur J Soc Sci* 20(1):28–46
32. Ndubis NO, Sinti Q (2006) Consumer attitudes, system's characteristics and internet banking adoption in Malaysia. *Manag Res News*
33. Al Nahian Riyadh M, Akter S, Islam N (2009) The adoption of e-banking in developing countries: a theoretical model for SMEs. *Int Rev Bus Res Pap* 5(6):212–230
34. Malhotra C, Malhotra R, Singh MM, Garg S, Ingle GK (2007) A study of tobacco use among street children of Delhi. *Indian J Community Med* 32(1):58

35. Chwelos P, Benbasat I, Dexter AS (2001) Empirical test of an EDI adoption model. *Inf Syst Res* 12(3):304–321
36. Iacovou CL, Benbasat I, Dexter AS (1995) Electronic data interchange and small organizations: adoption and impact of technology. *MIS Q* 465–485
37. Molla A, Peszynski K, Pittayachawan S (2010) The use of e-business in agribusiness: investigating the influence of e-readiness and OTE factors. *J Glob Inf Technol Manag* 13(1):56–78
38. Zhu K, Kraemer KL (2005) Post-adoption variations in usage and value of e-business by organizations: cross-country evidence from the retail industry. *Inf Syst Res* 16(1):61–84
39. Ajzen I, Fishbein M (1975) A Bayesian analysis of attribution processes. *Psychol Bull* 82(2):261
40. Bentler PM, Speckart G (1979) Models of attitude–behavior relations. *Psychol Rev* 86(5):452
41. Hair JF, Anderson RE, Black WB, Tatham RL (2010) *Multivariate data analysis*. Prentice Hall
42. Tan M, Teo TS (2000) Factors influencing the adoption of Internet banking. *J Assoc Inf Syst* 1(1):5
43. Shi W, Shambare N, Wang J (2008) The adoption of internet banking: an institutional theory perspective. *J Financ Serv Mark* 12(4):272–286
44. Hair JF, Anderson RE, Tatham RL, Black WC (1998) *Multivariate data analysis*, 5th edn. Prentice-Hall, New Jersey
45. Kline RB (2005) *Principle and practice of structural equation modeling*, 2nd edn. Guilford Press, New York
46. Hair JF, Anderson RE, Black WB, Babin B, Tatham RL (2006) *Multivariate data analysis*, 7th edn. Auflage, Upper Saddle River
47. Fornell C, Larcker DF (1981) Evaluating structural equation model with unobservable variables and measurement error. *J Mark Res* 18(1):39–50
48. Hair JF, Ringle CM, Sarstedt M (2013) Partial least squares structural equation modeling: rigorous applications, better results and higher acceptance. *Long Range Plan* 46(1–2):1–12
49. Schumacher RE, Lomax RG (2004) *A beginner's guide to structural equation modeling*. Lawrence Erlbaum Associates Inc., Mahwah, NJ
50. Hair J, Hollingsworth CL, Randolph AB, Chong AYL (2017) An updated and expanded assessment of PLS-SEM in information systems research. *Ind Manage Data Sys*
51. Cohen J (1988) *Statistical power analysis for the behavioral sciences*, 2nd edn. Á/L
52. Chin WW (1998) The partial least squares approach to structural equation modeling. *Mod Methods Bus Res* 295(2):295–336
53. Hair JF, Ringle CM, Sarstedt M (2011) PLS-SEM: indeed a silver bullet. *J Mark Theory Pract* 19(2):139–152



# Markovian Queueing Model with Server Breakdown, Single Working Vacation, and Catastrophe



M. Seenivasan, V. J. Chakravarthy, and R. Abinaya

## 1 Debut

Queueing systems with SWV and consummate service have acquired importance over the most recent twenty years because of large extent uses mainly manufacturing system, service system, telecommunications, and computer system. In numerous genuine queueing circumstances, after assistance fulfillment, if no customer in the queue, a server goes on a vacation epoch. This type is known as vacation queue. Using survey paper by Doshi [1], many researchers introduced queueing model with vacations. Gelenbe et al. [2] studied queues with negative arrivals.

Servi and Finn [3] developed an  $M/M/1$  queueing model upon WV. Wu and Takagi [4] analyzed an  $M/G/1$  queueing model upon MWV. Analysis of  $GI/M/1$  queueing model upon MWV studied by Baba [5]. Server will only take one WV if the queue become null. So, if the queue has no customers when the server comes back from SWV, he will idle on the system and wait for the customers to arrive instead of picking up another WV. A multi-server system with an SWV was proposed by Lin and Ke [6]. The use of inactive epoch a  $M/G/1$  model was studied by Levy and Yechiali [7].

The notion that a disaster is happening randomly leads to the destruction of all the customers there, and it is not uncommon for a service facility to be temporarily disabled until a new customer arrives. Disasters can originate external to the organization or originated on some other queue. In case study, this event takes place

---

M. Seenivasan (✉)

Mathematics Wings-DDE, Annamalai University, Annamalainagar, Tamil Nadu 608002, India  
e-mail: [emseeni@yahoo.com](mailto:emseeni@yahoo.com)

V. J. Chakravarthy

Principal, Arulmigu Kapaleeswarar Arts and Science College, Kolathur, Tamil Nadu 600099, India

R. Abinaya

Department of Mathematics, Annamalai University, Annamalainagar, Tamil Nadu 608002, India

in the computer system, production system, and service system. Gelenbe et al. [2] analyzed the queues with negative arrivals. Stirzaker [8] considered the process with catastrophe. Mostly, the catastrophe occurs only when the systems are full. Chao [9] concentrated with queueing model having catastrophe and item from arrangement. Aissani et al. researched on the single server retrial queue subject to breakdown.

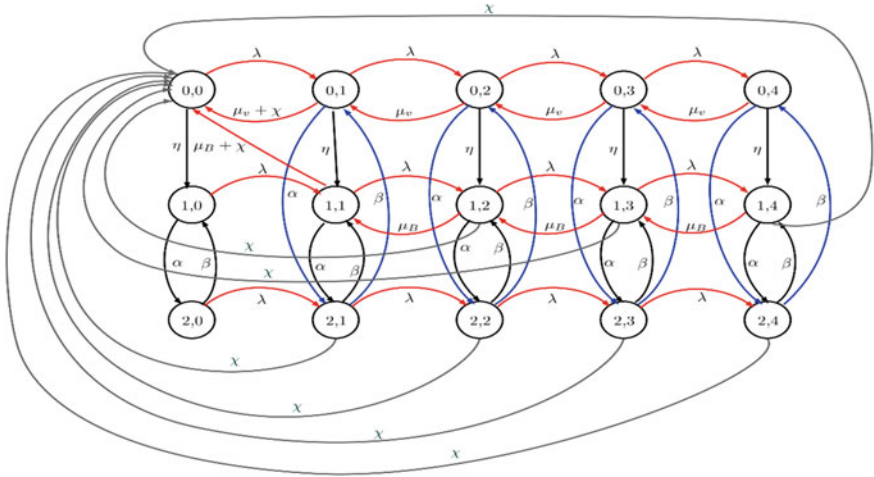
A catastrophic—cum—restorative progression associated contribution for grid-lock created advanced networking administrations have been explored by Direscenzo et al. [10]. An MGM is the best key for finding the difficult queueing obstacle in queueing theory. Neuts [11] thought Markov chains with applications queueing hypothesis, which have a grid mathematical invariant likelihood vector. Neuts [12] determined matrix geometric arrangement in stochastic models. Seenivasan et al. [13] studied performance examination of two heterogeneous servers queueing model with an irregularly reachable server utilizing MGM. Seenivasan et al. [13] investigated a retrial queueing model with two heterogeneous servers using the MGM. The vacation models and the model in which the server may go to breakdowns and repairs are well ascertained in survey papers B. T. Doshi. Gray et al. [14] overworked on multiple  $\varphi$  in queueing theory. Aissaini et al. [15] analysed on single server retrial queue subject to breakdown.

Our study deals with an SWV, server breakdown, and catastrophe in  $M/M/1$  queueing model. In accordance with FCFS principle, customers are served. This model has been analyzed using MGM. The excess of this study designated as follows. We providing construction of model in Sect. 2. The matrix geometric solution derived in Sect. 3. Mathematical illustrations solved in Sect. 4 and brief conclusion in final section.

## 2 Construction of the Model

We consider a slow service, catastrophic, and server breakdown in  $M/M/1$  model. The customers show up in line as per Poisson process with parameter  $\lambda$ . They create a queue dependent on her/his request for appearance. At a normal working period, influx customers served at a service rate  $\mu_b$ , following an exponential assigned. The server starts a WV of arbitrary length at immediate if there is no customer in the system, and it follows an exponential assigned with parameter  $\eta$ . At the time of WV period, influx customers get service with rate  $\mu_v$ , following an exponential assigned. If the queue forms, then the server chops and shifts its rate from  $\mu_v$  to  $\mu_b$ , the normal working interval starts. If the queue not formed after the completion of vacation, the server waits idle. If not, server starts a normal busy period when a customer arrival occurs. When the server may subject to sudden breakdown with rate  $\alpha$  and after, it should be repaired and goes to normal service with rate  $\beta$ .

In this model, when system contains customers according to the Poisson process, the catastrophe occurs with parameter  $\xi$ . This catastrophic event immediately ruins the customers and attacks the system as well. The rate transition diagram is given in Fig. 1.



**Fig. 1** Rate transition diagram

Let  $\{(k(t), n(t)) : t \geq 0\}$  be a MP, where  $k(t)$  and  $n(t)$  represent state of process at time  $t$ , respectively.

- $k(t) = 0$ , when server is on WV,
- $k(t) = 1$ , when server is on normal working epoch, and
- $k(t) = 2$ , when server is on breakdown.
- $n(t)$  denotes total customer presented in the queue.

The QBD process along with state space  $\Omega$  as follow

$$\Omega = (0, 0)U(1, 0)U(2, 0)U(i, j); \quad i = 0, 1, 2 \ \& \ j = 1, 2, \dots, n \geq 1$$

Infinitesimal generator matrix  $Q$  is presented below:

$$Q = \begin{pmatrix} A_0 & A_1 & 0 & \dots & \dots & \dots \\ B_0 & A_2 & A_1 & \dots & \dots & \dots \\ C_0 & A_3 & A_2 & A_1 & \dots & \dots \\ C_0 & 0 & A_3 & A_2 & \dots & \dots \\ C_0 & \dots & \dots & A_3 & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \end{pmatrix}$$

where

$$A_0 = \begin{pmatrix} -\lambda - \eta & \eta & 0 \\ 0 & -\alpha - \eta - \lambda & \alpha \\ 0 & \beta & -\beta - \lambda \end{pmatrix}; \quad A_1 = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix};$$

$$\begin{aligned}
 B_0 &= \begin{pmatrix} \mu_v + \xi & 0 & 0 \\ \mu_b + \xi & 0 & 0 \\ \xi & 0 & 0 \end{pmatrix}; \quad C_0 = \begin{pmatrix} \xi & 0 & 0 \\ \xi & 0 & 0 \\ \xi & 0 & 0 \end{pmatrix}; \\
 A_2 &= \begin{pmatrix} -(\lambda + \eta + \alpha + \mu_v + \xi) & \eta & \alpha \\ 0 & -(\alpha + \lambda + \mu_b + \xi) & \alpha \\ \beta & \beta & -(2\beta + \lambda + \xi) \end{pmatrix}; \\
 A_3 &= \begin{pmatrix} \mu_v & 0 & 0 \\ 0 & \mu_b & 0 \\ 0 & 0 & 0 \end{pmatrix}
 \end{aligned}$$

### 3 Matrix Geometric Solution

We define  $p_{ij} = \{k = i, n = j\} = \lim_{t \rightarrow \infty} k(t) = i, n(t) = j$ , where  $j$  represents total customers presented in the queue, and  $i$  represents the server state. Probability vector is defined as  $P = (P_0, P_1, P_2, \dots)$  and  $P_j = (p_{0j}, p_{1j}, p_{2j})$ ,  $j = 0, 1, 2, 3, 4, \dots$ . The static probability row matrix is represented by using  $PQ = O$ .

$$P_0[A_0 + R[B_0 + R(1 - R)^{-1}C_0]] = 0 \tag{1}$$

$$P_0A_1 + P_1A_2 + P_2A_3 = 0 \tag{2}$$

$$P_1A_1 + P_2A_2 + P_3A_3 = 0 \tag{3}$$

$$\begin{aligned}
 &\cdot \\
 &\cdot \\
 &\cdot
 \end{aligned}
 \tag{4}$$

$$P_iA_1 + P_{i+1}A_2 + P_{i+3}A_3 = 0$$

and

$$P_i = P_0R^i \text{ for } i \geq 1 \tag{5}$$

Let us denote rate matrix be  $R$ . Using matrix equation [16], it is irreducible non-negative solution.

$$R^2A_3 + RA_2 + A_1 = 0 \tag{6}$$

The balancing mathematical statement is

$$P_0(1 - R)^{-1}e = 1 \tag{7}$$

Here, “ $e$ ” is unit matrix.

From Eq. (6),

$$R = -[A_1 + R^2A_3]A_2^{-1} \tag{8}$$

Rate matrix  $R$  is mathematically found by iterative method with  $R(0) = 0$

$$R_{n+1} = -[A_1 + R_n^2A_3]A_2^{-1} \tag{9}$$

Using the above equation to solve  $R$  matrix and check all values of  $R$  will monotonically increasing. And  $R$  will converge to  $-A_2^{-1}$  and  $[A_1 + R_n^2A_3]$  are positive.

By considering matrix geometric procedure, the steady state is obtained.  $P$  partitioned as  $P = (P_0, P_1, P_2)$  is probability row vector of the (reducible) generator matrix  $M = A_1 + A_2 + A_3$ . Then,  $M$  can be written as

$$M = \begin{pmatrix} -\gamma_0 & \eta & \alpha \\ 0 & -\gamma_1 & \alpha \\ \beta & \beta & -\gamma_2 \end{pmatrix}$$

where

$$\begin{aligned} \gamma_0 &= \eta + \alpha + \xi \\ \gamma_1 &= \alpha + \xi \\ \gamma_2 &= 2\beta + \lambda + \xi \end{aligned}$$

$M$  is clear and row vector  $P$  can be shown to be solitary such that

$$PM = 0 \text{ and } Pe = 1 \tag{10}$$

We have

$$P_0 = \left[ \alpha + \left( 1 + \frac{\eta + \gamma_0}{\gamma_1} \right) - \frac{\gamma_0\gamma_1}{\beta} \right] \tag{11}$$

$$P_1 = \frac{\eta + \gamma_0}{\gamma_1} P_0 \tag{12}$$

$$P_2 = \frac{\gamma_0}{\beta} P_0 \tag{13}$$

The static condition of such a QBD (see Neuts [12]) can be obtained by the drift condition  $PA_1e < PA_3e$ .

We get the static condition takes format

$$\lambda[P_0 + P_1 + P_2] < \mu_b P_2 + \mu_v P_1 \tag{14}$$

### 3.1 Evaluation of the Results

Performance measures have been found using steady-state probabilities as given below

If server is idle

$$E(I) = P_0 \tag{15}$$

If server is WV, then mean number of customers presented

$$E(V) = \sum_{j=0}^{\infty} j p_{0j} \tag{16}$$

If the server is busy epoch, then mean number of customers presented

$$E(B) = \sum_{j=1}^{\infty} j p_{1j} \tag{17}$$

If the server is breakdown, then mean number of customers presented

$$E(BD) = \sum_{j=1}^{\infty} j p_{2j} \tag{18}$$

Mean number of customers presented

$$E(N) = E(I) + E(B) + E(V) + E(BD) \tag{19}$$

Total output gained is

$$\delta = \mu_v \sum_{j=1}^{\infty} j p_{0j} + \mu_b \sum_{j=1}^{\infty} j p_{1j} \tag{20}$$

Delay time is given by

$$Dt = \frac{E(N)}{\delta} \tag{21}$$

### 4 Mathematical Study

Here, we make mathematical calculation for model given by the segment above. Our goals are to show effect of a parameter on system features. By modifying  $\lambda$  and  $\mu_b$ , eight illustrations are presented in these sections.

The parameter  $\lambda$  value varies, and all other argument values are stabled. Illustration 1 to Illustration 4 are presented below.

**Illustration 1** We take  $\lambda = 0.05, \mu_b = 0.5, \mu_v = 0.4, \alpha = 0.2, \eta = 0.15, \beta = 0.3$  and  $\xi = 0.1$ .

$$R = \begin{pmatrix} 0.0667 & 0.0244 & 0.0218 \\ 0.0086 & 0.0729 & 0.0196 \\ 0.0330 & 0.0440 & 0.0832 \end{pmatrix}$$

By swapping rate matrix  $R$  in Eq. (1), vector  $P_0$  is obtained and normalization equation  $P_0[I - R]^{-1}e = 1$  for the mathematical argument selected previously, row vector  $P_0$  is granted by  $P_0 = (0.2283, 0.3516, 0.2960)$ . More, the balance vector  $P_j$ 's gained from  $P_j = P_0R^j, j = 1, 2, 3, \dots$  and are shown in Table 1. Column 2, 3, and 4 contain the three elements of  $P_j, j = 0, 1, 2, \dots$ . Final column constitutes the total of three elements. Total probability was confirmed to be  $0.9999 \approx 1$ .

**Illustration 2** We take  $\lambda = 0.1, \mu_b = 0.5, \mu_v = 0.4, \alpha = 0.2, \eta = 0.15, \beta = 0.3$  and  $\xi = 0.1$ .

$$R = \begin{pmatrix} 0.1305 & 0.0523 & 0.0372 \\ 0.0178 & 0.1462 & 0.0337 \\ 0.0657 & 0.0941 & 0.1516 \end{pmatrix}$$

**Table 1** Probability vectors for  $\lambda = 0.05$

$P_j$	$p_{0j}$	$p_{1j}$	$p_{2j}$	Total
$P_0$	0.2283	0.3516	0.2960	0.8759
$P_1$	0.0280	0.0442	0.0365	0.1087
$P_2$	0.0035	0.0055	0.0045	0.0135
$P_3$	0.0004	0.0007	0.0004	0.0015
$P_4$	0.0001	0.0001	0.0001	0.0003
Total				<b>0.9999</b>

**Table 2** Probability vectors for  $\lambda = 0.1$

$P_j$	$p_{0j}$	$p_{1j}$	$p_{2j}$	Total
$P_0$	0.3297	0.2792	0.1593	0.7682
$P_1$	0.0585	0.0730	0.0458	0.1773
$P_2$	0.0119	0.0180	0.0116	0.0415
$P_3$	0.0026	0.0044	0.0028	0.0098
$P_4$	0.0006	0.0010	0.0007	0.0023
$P_5$	0.0001	0.0002	0.0002	0.0005
$P_6$	0.0000	0.0001	0.0000	0.0001
Total				<b>0.9997</b>

By swapping rate matrix  $R$  in Eq. (1), vector  $P_0$  is obtained and normalization equation  $P_0[I - R]^{-1}e = 1$  for the mathematical argument selected previously, row vector  $P_0$  is granted by  $P_0 = (0.3297, 0.2792, 0.1593)$ . More, the balance vector  $P_j$ 's gained from  $P_j = P_0R^j, j = 1, 2, 3, \dots$  and are shown in Table 2. Column 2, 3, and 4 contain the three elements of  $P_j, j = 0, 1, 2, \dots$ . Final column constitutes the total of three elements. Total probability was confirmed to be  $0.9997 \approx 1$ .

**Illustration 3** We take  $\lambda = 0.15, \mu_b = 0.5, \mu_v = 0.4, \alpha = 0.2, \eta = 0.15, \beta = 0.3$  and  $\xi = 0.1$ .

$$R = \begin{pmatrix} 0.1907 & 0.0834 & 0.0484 \\ 0.0268 & 0.2192 & 0.0440 \\ 0.0969 & 0.1491 & 0.2091 \end{pmatrix}$$

By swapping rate matrix  $R$  in Eq. (1), vector  $P_0$  is obtained and normalization equation  $P_0[I - R]^{-1}e = 1$  for the mathematical argument selected previously, row vector  $P_0$  is granted by  $P_0 = (0.3676, 0.2262, 0.0762)$ . More, the balance vector  $P_j$ 's gained from  $P_j = P_0R^j, j = 1, 2, 3, \dots$  and are shown in Table 3. Column 2, 3, and 4 contain the three elements of  $P_j, j = 0, 1, 2, \dots$ . Final column constitutes the total of three elements. Total probability was confirmed to be  $0.9999 \approx 1$ .

**Illustration 4** We take  $\lambda = 0.2, \mu_b = 0.5, \mu_v = 0.4, \alpha = 0.2, \eta = 0.15, \beta = 0.3$  and  $\xi = 0.1$ .

$$R = \begin{pmatrix} 0.2469 & 0.1166 & 0.0566 \\ 0.0353 & 0.2913 & 0.0517 \\ 0.1255 & 0.2074 & 0.2583 \end{pmatrix}$$

By swapping rate matrix  $R$  in Eq. (1), vector  $P_0$  is obtained and normalization equation  $P_0[I - R]^{-1}e = 1$  for the mathematical argument selected previously, row vector  $P_0$  is granted by  $P_0 = (0.3684, 0.1854, 0.0236)$ . More, the balance vector  $P_j$ 's gained from  $P_j = P_0R^j, j = 1, 2, 3, \dots$  and are shown in Table 4. Column 2, 3, and



**Table 3** Probability vectors for  $\lambda = 0.15$

$P_j$	$p_{0j}$	$p_{1j}$	$p_{2j}$	Total
$P_0$	0.3676	0.2262	0.0762	0.6700
$P_1$	0.0835	0.0916	0.0437	0.2188
$P_2$	0.0226	0.0336	0.0172	0.0734
$P_3$	0.0069	0.0118	0.0062	0.0249
$P_4$	0.0022	0.0041	0.0021	0.0084
$P_5$	0.0007	0.0014	0.0007	0.0028
$P_6$	0.0003	0.0005	0.0003	0.0011
$P_7$	0.0001	0.0002	0.0001	0.0004
$P_8$	0.0000	0.0001	0.0000	0.0001
Total				<b>0.9999</b>

**Table 4** Probability vectors for  $\lambda = 0.2$

$P_j$	$p_{0j}$	$p_{1j}$	$p_{2j}$	Total
$P_0$	0.3684	0.1854	0.0236	0.5774
$P_1$	0.1005	0.1019	0.0365	0.2389
$P_2$	0.0330	0.0490	0.0204	0.1024
$P_3$	0.0124	0.0223	0.0097	0.0444
$P_4$	0.0051	0.0100	0.0044	0.0195
$P_5$	0.0022	0.0044	0.0019	0.0085
$P_6$	0.0009	0.0019	0.0008	0.0036
$P_7$	0.0004	0.0008	0.0004	0.0016
$P_8$	0.0002	0.0004	0.0002	0.0008
$P_9$	0.0001	0.0002	0.0001	0.0004
$P_{10}$	0.0000	0.0001	0.0000	0.0001
Total				<b>0.9976</b>

4 contain the three elements of  $P_j, j = 0, 1, 2, \dots$ . Final column constitutes the total of three elements. Total probability was confirmed to be  $0.9976 \approx 1$ .

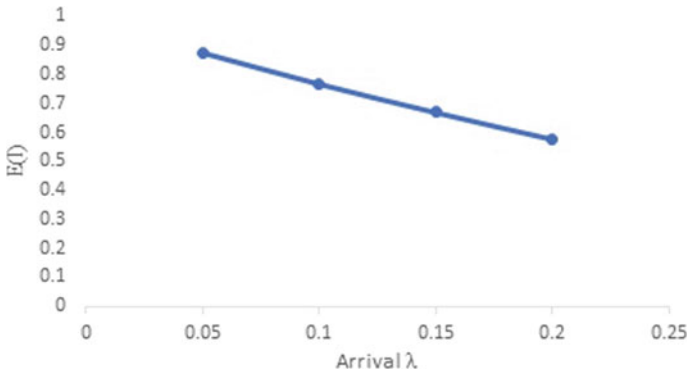
### 4.1 Evaluation of the Results

Varying the parameter  $\lambda = 0.05-0.2$ , the performance measures shown in Table 5.

Out of the above figures, we derived few performance measurements with the effect of  $\lambda$  such as mean no. of customer if server is idle, mean no. of customer if server is on working vacation, mean no. of customer if server is on busy period,

**Table 5** Performance measures for  $\lambda$

$\Delta$	0.05	0.1	0.15	0.2
$E(I)$	0.8759	0.7682	0.6700	0.5774
$E(B)$	0.0577	0.1278	0.2228	0.3518
$E(V)$	0.0366	0.0930	0.1642	0.2458
$E(BD)$	0.0471	0.0812	0.1111	0.1436
$E(N)$	1.0173	1.0702	1.1681	1.3186
$\Delta$	0.0435	0.1011	0.1771	0.2742
$Dt$	23.3862	10.5856	6.5957	4.8089



**Fig. 2** Arrival versus server is in idle

mean no. of customer if server is on breakdown, mean no. of customers throughout system, and output gained from the system and delay time, respectively. From Fig. 2 shows that arrival rate increases, mean number of customer if server is idle decreases, Figs. 3, 4, 5, 6, and 7 show arrival rate increases, mean no. of customer if server is WV, busy period, breakdown, mean no. of customers throughout system, and total output gained increases. Finally, if the arrival increases, delay time of the system decreases as shown in Fig. 8.

## 5 Conclusion

In this article, we have studied a different manner of service rates, server breakdown, and catastrophes. We have obtained the static probability row vector by MGM, and also, we derived so performance measures with the effect of  $\lambda$  such as mean number of customer if server waits idle, mean number of customer if server is on working vacation, mean no. of customer if server is on busy period, mean number of customer

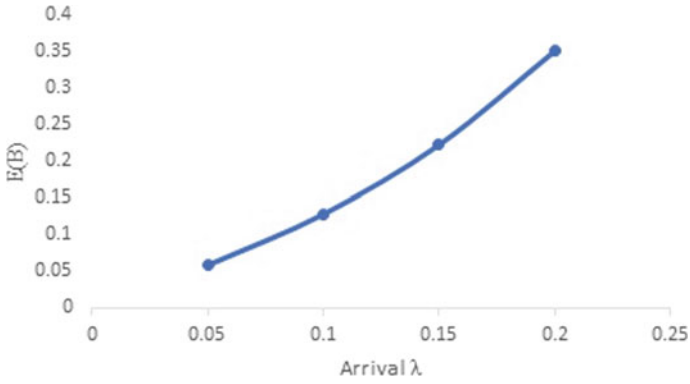


Fig. 3 Arrival versus mean number of customer during busy period

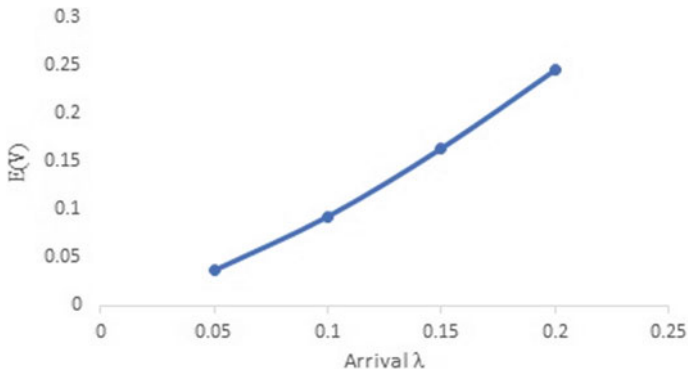


Fig. 4 Arrival versus mean number of customer during working vacation

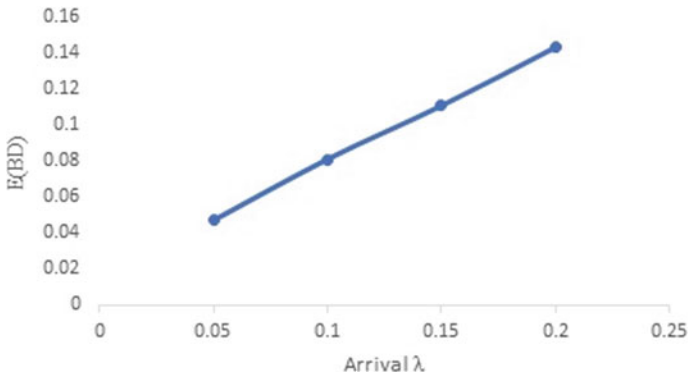


Fig. 5 Arrival versus mean number of customer during breakdown

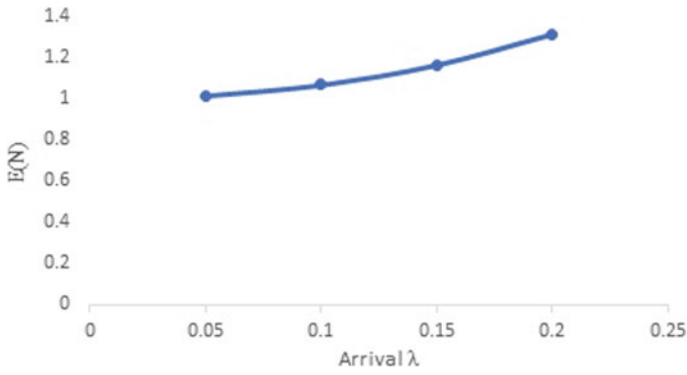


Fig. 6 Arrival versus mean number of customer in the system

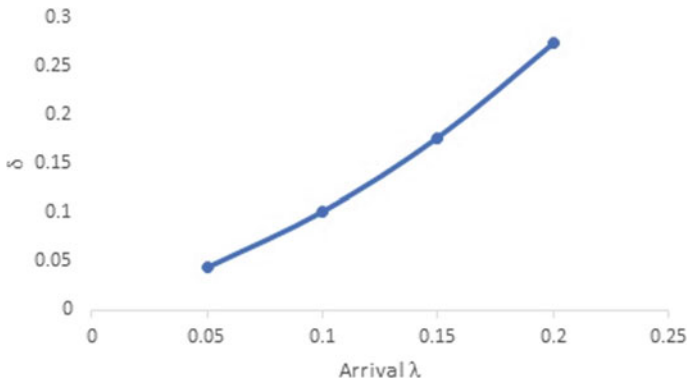


Fig. 7 Arrival versus total output from the system

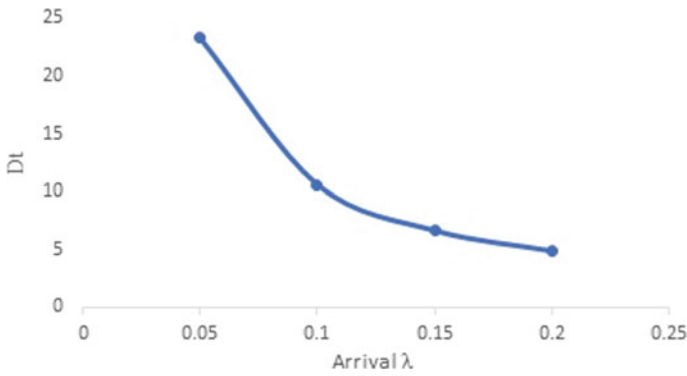


Fig. 8 Arrival versus delay time of the system

if server is on breakdown, mean number of customers throughout system, and output gained from the system and delay duration, respectively.

## References

1. Doshi BT (1986) Single-server queue with vacation a survey. *Queueing Syst* 1:29–66
2. Gelenbe E, Glynn P, Sigman K (2003) Queues with negative arrivals. *Appl Prob* 28:245–250
3. Servi LD, Finn SG (2002) M/M/1 queues with working vacation (M/M/1/WV). *Perform Eval* 50:41–52
4. Wu D, Takagi H (2003) M/G/1 queue with multiple working vacation. In: *Proceedings of the queueing symposium, stochastic models and their application, Kakegawa*, pp 51–60
5. Baba Y (2005) Analysis of a GI/M/1 queue with multiple working vacations. *Oper Res Lett* 33:201–209
6. Lin CH, Ke JC (2009) Multi-server system with single working vacation. *Appl Math Model* 33:2967–2977
7. Levy Y, Yechiali U (1975) Utilization of idle time in an M/G/1 queueing system. *Manag Sci* 22:202–211
8. Stirzaker D (2006) Processes with catastrophe. *Math Sci* 31:107–118
9. Chao X (1995) A queueing network model with catastrophes and product form solution. *Oper Res Lett* 18:75–79
10. Dicrescenzo A, Giorno V, Nobile AG, Ricciardi LM (2003) On the M/M/1 queue with catastrophes and its continuous approximation. *Queueing Syst* 43:329–347
11. Neuts MF (1978) Markov chains with applications queueing theory, which have a matrix geometric invariant probability vector. *Adv Appl Prob* 10:185–212
12. Neuts MF (1981) Matrix-geometric solutions in stochastic models. In: Vol. 2 of *Johns Hopkins's series in the mathematical series*. Johns Hopkins University Press, Baltimore, USA
13. Seenivasan M, Rajarajan G, Indumathi M (2021) Retrial queueing model with two heterogeneous servers using the matrix geometric method. *Mater Sci Eng* 1070
14. Gray WJ, Patrick Wang P, Scott M (2004) A queueing model with multiple types of server breakdowns. *Qual Technol Quant Manag* 1:245–255
15. Aissani A, Artalejo JR (1998) On the single server retrial queue subject to breakdown. *Queueing Syst* 30:309–321
16. Seenivasan M, Indumathi M, Chakravarthy VJ (2021) Performance analysis of two heterogeneous server queueing model with intermittently obtainable server using matrix geometric method. *Phys Conf Ser* 1724

# Lessening Spectrum Sensing Data Falsification Attack by Weighted Fuzzy Clustering Means Using Simulation Annealing in Cognitive Radio Networks



L. Thulasimani and A. Hyils Sharon Magdalene

## 1 Introduction

In 1999, Joseph Mitola founded the Cognitive Radio Network (CRN) for accommodating mobile communications [1]. The component finds the frequency band as possible. But when it is not available, it investigates the release band and use it. Then, it deviates skillfully from one band to another [2], rather than stick to a particular band. One, who consumes the licensed band, noted as primary users (PUs) [3] and of those, who use the unlicensed band called as the secondary users (SUs).

The channel being consumed named as ‘busy’, which causes a frequency fully engaged by the PU. The channel being not consumed named as ‘release’, which guides the PU to non-usage of the spectrum.

This causes undesirable interferences to the PU and extended to annual the SU to receive the channel. This is called as general SSDF/Byzantine attack [4]. CSS [2] is an advanced choice to raise the execution of spectrum sensing, and is clearly defined in Fig. 1.

## 2 Related Work and Contributions

Aida [5] implements a weighted sequential probability ratio test (WSPRT) for the best improvement of disseminated spectrum sensing against the SU terminus. Reference [6] depicts about the binary strategy that strains decrease the SSDF attack. The paper [7] is based on the repute of the SUs. They have a binary recount approach

---

L. Thulasimani (✉) · A. Hyils Sharon Magdalene  
Department of Electronics and Communication Engineering, PSG College of Technology,  
Coimbatore 641004, India  
e-mail: [ltm.ece@psgtech.ac.in](mailto:ltm.ece@psgtech.ac.in)

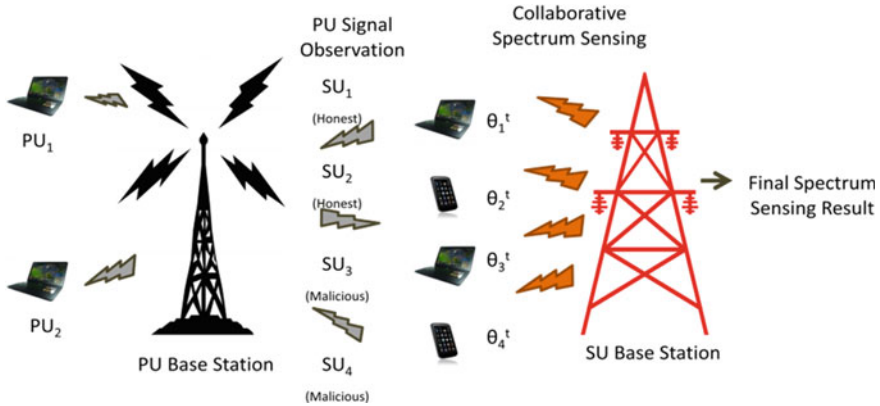


Fig. 1 Performance of collaborative sensing along with SSDF attack

that strives to reduce SSDF attack. Hyils Sharon Magdalene and Thulasimani in [8] proposed another CSS strategy to treat SSDF attack, in which each user has alloted a weight. Saud in [9] separates the theory of fuzzy sets used in fuzzy logic, and is an associated theory. Lotfi in [10] recounts the range that requires for measurable factor of fuzzification to produce a much more overlapping clusters. Ehab et al. in [11], assure that the scanning can be performed through the entrance stage for the intruders by means of fuzzy based intrusion detection system. In Li and Han [12], another data mining strategy is applied to relief the Sybil attack that identities the harmful characters.

### 3 Existing Method

#### 3.1 Multiple Trust Factor

There are four factors to be defined in multiple trust factor [13] to mitigate SSDF attack. In preceding-based Trust Factor, the entail level of trust  $Se_n$ , considering the level of trust is calculated as,

$$TL[Se_n] = \begin{cases} \sum_{i=1}^t TL[Se_n^i] * g(t) & T_m \neq 0 \\ 0.5 & T_m = 0 \end{cases} \quad (1)$$

In dynamic factor,  $O[Se_n^t]$  be the available duration of the flow of sensing time period 't' and  $S[Se_n^t]$  be the entire sensing time that is flowing up to  $Se_n$  at time 't' which merges the interlinked system. Then it calculates as,

$$DF[Se_n] = \frac{S[Se_n^t]}{O[Se_n^t]} \quad (2)$$

In motivation factor, the factor dissemination performs the harmful users to give a penalty by weakening the ground trust. This can be evaluated as:

$$MF[Se_n] = 1 - \gamma[Se_n] \quad (3)$$

where  $\gamma$  represents the penalty factor. In consistency factor,  $TL[Se_n] = TL[Se_n^1, Se_n^2, \dots, Se_n^t]$  suggests the point position of the trust for SU  $S_n$  and  $T$  indicates the slithering time duration. The asymmetry of trust SU  $S_n$ , is denoted as  $S^2[Se_n]$  and is defined as

$$S^2[Se_n] = \sum_{i=1}^t \frac{(TL[Se_n^i] - TL[Se_n^-])^2}{T_m - 1} \quad (4)$$

where  $TL[Se_n^-]$  is the trust assessing average value.

### 3.2 Evidence Theory

In Evidence Theory [14], the frame is defined as,  $\{H_1, H_0, \Omega\}$ , where  $\Omega$  denotes the hypothesis is true. The sample number of the  $X_i(t)$  is large enough to obey the normal distribution. So, the normal form of membership function is applied to get  $m_i(H_1)$ ,  $m_i(H_0)$ ,  $m_i(\Omega)$  in two cases of  $H_1$  and  $H_0$ , respectively, and is described as,

$$m_i(H_0) = \int_{x_i(t)}^{+\infty} \frac{1}{\sqrt{2\pi}\delta_{0i}^2} \exp\left(-\frac{(x - \mu_{0i})^2}{\delta_{0i}^2}\right) dx \quad (5)$$

$$m_i(H_1) = \int_{-\infty}^{x_i(t)} \frac{1}{\sqrt{2\pi}\delta_{1i}^2} \exp\left(-\frac{(x - \mu_{1i})^2}{\delta_{1i}^2}\right) dx \quad (6)$$

$$m_i(\Omega) = 1 - m_i(H_0) - m_i(H_1) \quad (7)$$

where  $\mu_{0i}$ ,  $\mu_{1i}$ ,  $\delta_{0i}^2$ ,  $\delta_{1i}^2$  denotes the mean and variance of the  $X_i(t)$  under  $H_0$  and  $H_1$ , respectively. The above formula shows the credibility of each cognitive user as,  $m_i(H_0)$  and  $m_i(H_1)$ .



### 3.3 *Balanced Cluster-Based Redundancy Check Aggregation (BCRCA) Mechanism [15]*

This is a method that searches the exponential family of a maximum entropy solution. It is expressed as,

$$P(x_{ji}) = \prod_i a_j^{f_j(x_{ji})} \quad (8)$$

where  $a_j$  is unknown. Then the maximum likelihood was limited to the restriction  $f_j(x_{ji}) \in \{0, 1\}$ . The method starts with arbitrary values of  $a_j^{(0)}$  values, which define the initial probability estimate:

$$P^{(0)}(x) \stackrel{\text{def}}{=} \prod_j a_j^{(0)f_j(x_{ji})} \quad (9)$$

### 3.4 *XOR Distance Analysis*

The XDA designs [16] the trust mechanism to get the CSS, which is free from SSDF attack. In this existing method, first the XOR distance analysis between any two SUs is calculated. Second the lightweight algorithm is implemented to detect SSDF attackers. Third the ideal trust mechanism is applied to correct the SSDF attackers trust value.

The above constituents of the preceding methods (from section A to D) are helped to simulate in the computer simulation results. The estimation taken is based on 5–30 attackers.

### 3.5 *Proposed Contribution*

#### **Improved Weighted FCM Algorithm**

In the traditional FCM algorithm [17], the final clustering result of each sample node is same in each iteration and it is hard to determine. According to the classification of data, the problem can be solved by using a weighted processing toward the target function and the cluster center of the FCM clustering algorithm. This will improve the classification in an accurate number as possible. The weighted FCM clustering procedure technique is shown elaborated in Sect. 5.

## 4 Strategy Model

### 4.1 Collaborative Spectrum Sensing

In CSS [18], a CR user makes the network and executes their spectrum sensing separately. The local sensing data must be settled between the two hypotheses tests [13] and is followed.

$$X_i(t) = \begin{cases} w_i(t) & H_0 \\ h_i S(t) + w_i(t) & H_1 \end{cases} \tag{10}$$

where

- $H_0$  PU is absent,
- $H_1$  PU is present.

At time ‘ $t$ ’, the received signal for  $i$ th user is  $X_i(t)$  and  $S(t)$  is the PU signal. The thermal noise is ‘ $w_i(t)$ ’. Then,  $\theta_i$  is the local sensing data of  $i$ th user at  $t$ th slot and is expressed in binary diverge as follows:

$$\theta_i^t = \begin{cases} 0, & H_0 \\ 1, & H_1 \end{cases} \tag{11}$$

Consider of  $K$  primary users and  $N$  secondary users. Then,  $H$  is the group of honest user that contains a ‘ $K$ ’ PUs and  $M$  is the group of harmful SUs that contains in a ‘ $N$ ’ SUs. Consider some of them are harmful in  $N$  SUs. The group of  $N$  SUs in the network is,  $SU_n = \{\theta_1, \theta_2, \dots, \theta_n\}$ . Hence, time slot  $\tau$  consists of two similar sized replaced time windows ( $t_0, t_1, t_2, \dots, t_n$ ). Each contains a sensing slot and the directed sensing data of all the SUs in  $t$ th sensing slot is carried as:  $\theta[t] = \{\theta_1^t, \theta_2^t, \theta_3^t, \dots, \theta_n^t\}$  [19], where each  $\theta_i^t \in \{0, 1\}$ . Once the completion of one sensing slot is concluded, the slithering window descends the oldest time events and adds the new sensing time observations. The SUs sensing period from 1 to  $t$  is found as  $\theta_n = \{\theta_n^1, \theta_n^2, \theta_n^3, \dots, \theta_n^t\}$ . The past  $\tau$  sensing time had been always present and is expressed in Fig. 2 [13].

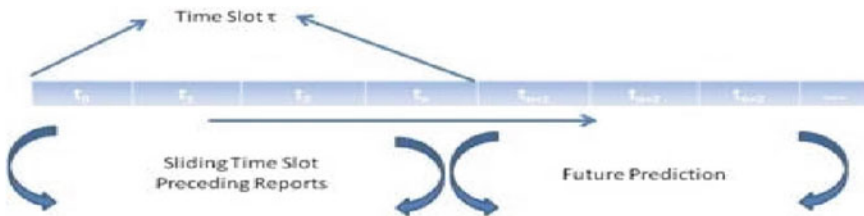


Fig. 2 Future behavior prediction graph

## 5 III-Will User Detection Technique

The suggested technique consists of three processes, namely: Strainer, Lessen, and Secure.

### 5.1 Process-I: ‘Strainer’ Process

#### Improved Weighted FCM Algorithm Based on Simulated Annealing (SA): Suggested Strategy

For the suspicious value calculation, the historical part of the SU report was taken and the assumption is made as follows:

$$\{S_i(t)\} \square P(\tau_i = M | F_t) \tag{12}$$

This is the suspicious phase for node  $i$  and node  $j$  at time ‘ $t$ ’. The instance of a node is PU or SU, which is denoted as,  $\tau_n(= H)$  or  $\tau_n(= M)$ . Every recordings of a time schedule from 1 to  $t$  is found as  $F_t$ . Then, ‘ $\square$ ’ means that every ‘ $F_t$ ’ data is taken as a sensing report in the time frame  $S_i(t)$ .

#### Steps to Filter III-Will Users

*Step 1.* Assume ‘ $l$ ’ clusters and classify the sensing bit to ‘ $l$ ’ clusters. Set the value of  $S = [S_{ij}]$  matrix, and the centroid for fuzzy via.,  $S(0), S(1), \dots, S(l)$ , are allotted to each of the clusters,  $l_1, l_2, \dots, l_n$ .

*Step 2.* Compute the span among the relevant cluster center as,  $\text{dist} = |S_j - S|$ . The minimization of an impartial function from the SUs span is calculated as,

$$S_{\min}^{i,j} = \sum_{i=1}^m \sum_{j=1}^n S_{ij}^m ||S_j - S||^2 \tag{13}$$

where  $1 < m < \infty$ .

Where ‘ $m$ ’ is the ‘membership’, which is greater than one. The level of rank of  $S_i$  in the  $j$ th cluster is  $S_{ij}$ , Let  $\|*\|$  is the mean exhibiting the similarity of any sensing report and the cluster center. Then, set the threshold  $e_t$  which is always greater than zero ( $>0$ ) to a very small value.

*Step 3.* Modify the fuzzy membership function and the sensing report is from ‘1’ to ‘ $t$ ’. The above Eq. (13) is found as,

$$S_{ij} = \frac{1}{\sum_{t=1}^n \left( \frac{\|S_j - S\|}{\|S_j - S_i\|} \right)^{\frac{2}{m-1}}} \tag{14}$$

*Step 4.* In earlier, at ‘ $t$ ’ step, the centroid is computed as,

$$S = \frac{\sum_{j=1}^n \sum_{i=1}^N S_{ij}^m S_i(t+1)}{\sum_{i=1}^N S_{ij}^m} \quad (15)$$

For the weighted process  $W_i$  with the  $i$ th node, and define a variable  $E_i$  to measure  $S_{ij}$ , which means the validity of the membership degree of node  $i$  that belongs to the  $j$ th node. This makes the clustering evaluation as [17],

$$E_i = \sum_{j=1}^m S_{ij} \log_2 S_{ij} \quad (16)$$

$$W_i = \frac{1 + E_i}{n + \sum_{i=1}^n E_i} \quad (17)$$

The iteration of the algorithm will be reduced by the term,  $\sum_{i=1}^m W_i = 1$ . The paper now proposes a new FCM objective function and a clustering center function. Then it is implemented to the FCM clustering algorithm as follows [17]:

$$S = \frac{\sum_{i=1}^m w_i^{W_i} S_{ij}^m S_i(t+1)}{\sum_{i=1}^N w_i^{W_i} S_{ij}^m} \quad (18)$$

Calculate the probability for  $S_{ij}$ , and find out the probability for the total Sensing report. The number of changeover probability of a state ‘ $i$ ’ to all other positions must be 1. So the expression is made as,

$$\sum_{j=1}^l P(s_{i,j}) = 1 \quad (19)$$

A fixed probability transmission ‘ $\pi$ ’ is determined as the set of and its probability is defined as  $\{1, 2, \dots, n\}$ . This is associated with the eigenvalue 1, i.e.,  $\pi P = \pi$ . The vector is individual by analyzing any ‘ $i$ ’ vector which is having the following limit:

$$\lim_{i \rightarrow j} (P^\tau)_{i,j} = \pi_\tau \quad (20)$$

where  $\pi_\tau$  is the  $\tau$ th element of the row vector  $\pi$ . Thus, states that the long-run probability of a state ‘ $j$ ’ is not associated with the initial state ‘ $i$ ’. Thus the suspicious value calculation of each cluster centroid is:

$$S_i(t) = \frac{\prod_{\tau=1}^t \theta_i(\tau)}{\sum_{j=1}^N \prod_{\tau=1}^t \theta_i(\tau) \theta_j(\tau) \dots \theta_n(\tau)} \quad (21)$$

$$S_i(t+1) = \frac{\prod_{\tau=1}^{t+1} \theta_i(\tau)}{\sum_{j=1}^N \prod_{\tau=1}^{t+1} \theta_j(\tau) \theta_j(\tau) \dots \theta_n(\tau)} \quad (22)$$

The time slot is  $\tau$  and the number of SU is  $N$  and is presented in the cluster and the probability of report is  $\theta_i(\tau)$  at time  $t$ . This is expressed as,

$$\theta_i(\tau) = P_{d,i} + P_{f,i} \quad (23)$$

$$\theta_j(\tau) = P_{d,j} + P_{f,j} \quad (24)$$

In the probability of true detection  $P_d$ , the signal is available in the existing band and the detection is successful. In the probability of false alarm  $P_f$ , the detection algorithm expresses that the frequency is occupied [18]. The signal differentiation is as follows:

$$P_d = \Pr(D > \lambda | H_1) \quad (25)$$

$$P_f = \Pr(D > \lambda | H_0) \quad (26)$$

The decision metric  $D$  is compared again to the threshold value, to examine whether the signal is existing or not. Then the computation is found at,

$$D = \sum_{n=1}^N |X_i(t)|^2 \quad (27)$$

Time slot  $\tau$  is the SU count that is available in the cluster  $N$  and  $\theta_i(\tau)$  is a probability of the data and is found in Eq. (28). Time slot  $\tau$  indicates that the node  $n$  is harmful.

$$\theta_i(\tau) = P_{d,i} + P_{f,i} \quad (28)$$

*Step 5.* The objective function is calculated as,

$$\begin{aligned} \Delta S_{ij} &= F(S, c) = \sum_{i=1}^n \sum_{j=1}^m \left( w_i^{W_i} S_{ij}^m || S_i(t+1) - S_i(t) || \right) \\ \Delta S_{ij} &< \text{Threshold} \end{aligned} \quad (29)$$

If the increment of the objective function is found decreased than the threshold value ' $e_t$ '. The algorithm halt and return to step 3.

**Table 1** Threshold range of SUs

Stage	Threshold set	Entailing
1	$\Delta S_{i,j} < e_t$	Honest user
2	$\Delta S_{i,j} \geq e_t$	Harmful user
3	Otherwise $\Delta S_{i,j} = 0$	Suspicious user

## 5.2 Process-II: ‘Lessen’ Process

### Cluster Configuration

The largest complete homogeneous of sensing stage with the other Sensing level [20] is preferred as a cluster core and is computed as,

$$S_i = \sum_{\substack{j \neq i \\ j \in S}} S_{ij} \quad (30)$$

In this technique, a ratio of difference-to homogeneous (DHP) is computed for each Sensing data point is declared by the following expression as,

$$DHP_i = \sum_{\substack{j \neq i \\ j \in S}} \frac{1 - S_{ij}}{S_{ij}} \quad (31)$$

The stage having the least possible DHP is taken as a cluster core.

After the cluster configuration, the attack has been reduced by using Weighted FCM resolutions.

## 5.3 Server-III: ‘Secure’ Process

Every cluster centroid is computed by averaging all the sensing levels in each cluster. WFCM re-utter the cluster center to the ‘right’ positions of the data set. The contrast was shown in Table 1.

# 6 Results

## 6.1 Simulation Setup

The collaborative sensing of ‘N’ SUs are considered as 70 and ‘K’ PUs are considered as 5. More of N SUs, lessening is made by raising the number of harmful users from

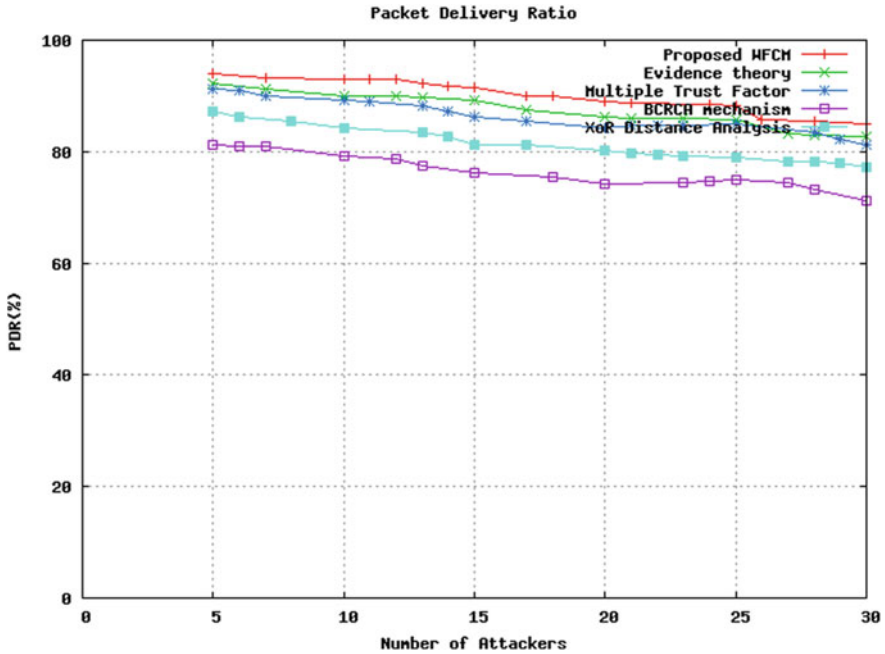


Fig. 3 Performance of suggested and existing defense for packet delivery ratio

5 to 30 attackers in the network. The leftover affairs are announced as honest users. For sensing time duration,  $\tau = 160$  time slots are taken.

### 6.2 Simulation Results

The suggested weighted FCM give better performance than all the other preceding methods, which is shown in graph results.

Figures 3, 4, 5 and 6 represents the execution of packet delivery ratio, packet loss ratio, throughput, and the delay for the existing and the WFCM methods.

The observation depicts that the packet delivery ratio for the FCM in the existing of 30 attackers is 85% which is greater than the existing defense.

## 7 Conclusions

The proposed work reduces the number of malicious user nodes available in the CRN network. When compared with the existing works the results obtained with the proposed work is better. In the preceding work, lessening of SSDF attack was

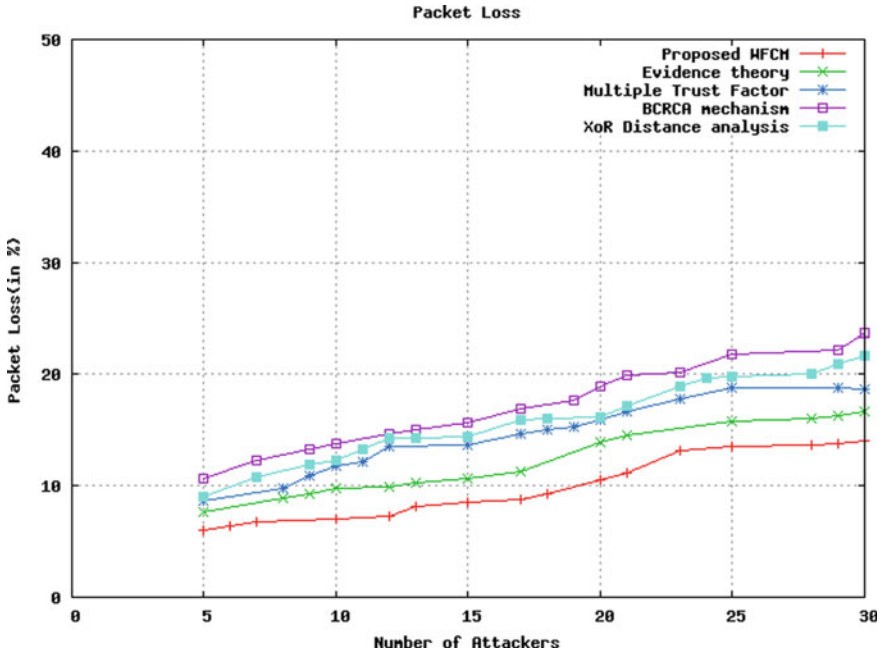


Fig. 4 Performance of suggested and existing defense for packet loss ratio

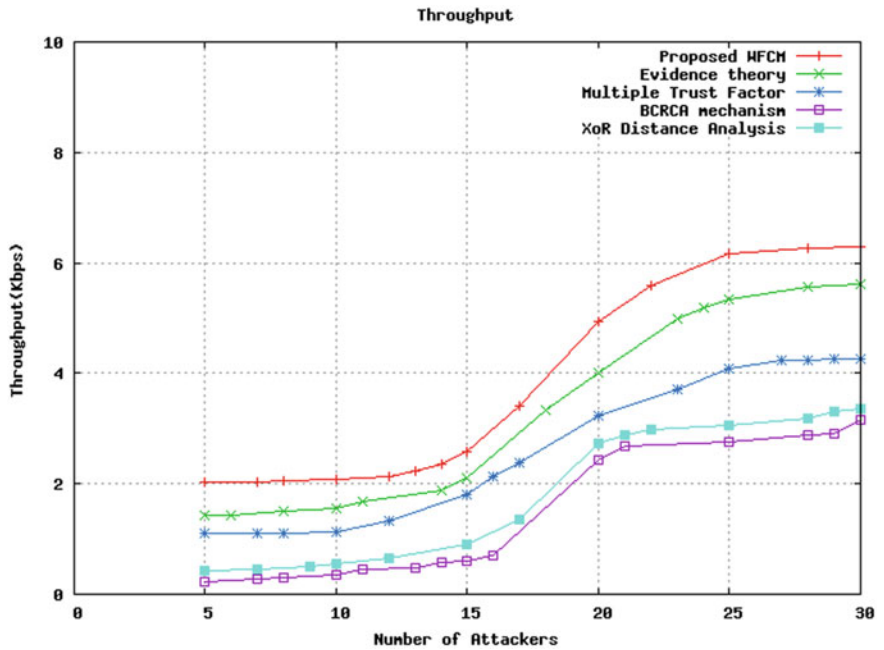


Fig. 5 Performance of proposed and existing defense for throughput



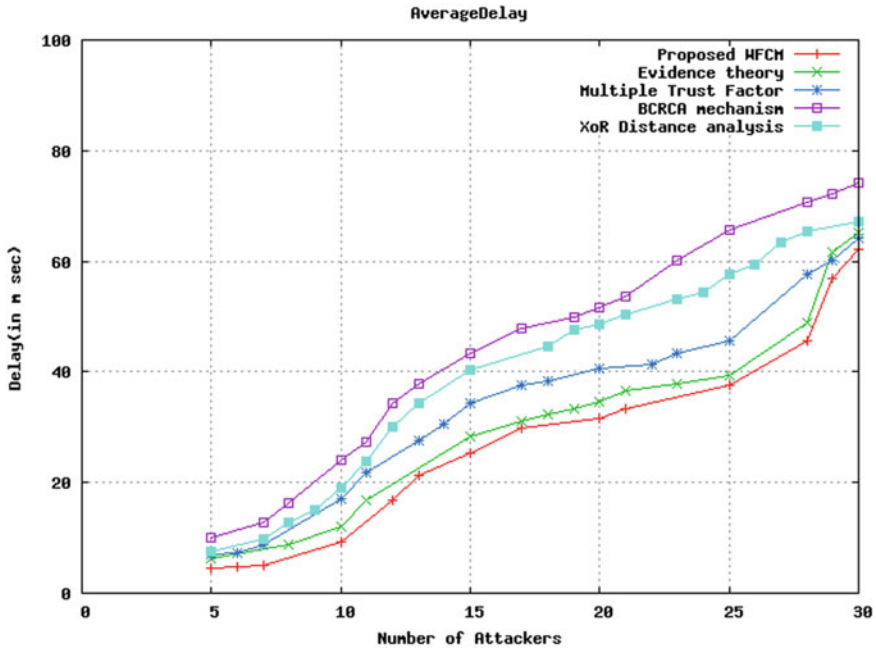


Fig. 6 Performance of suggested and existing defense for delay

performed by applying the different methods. The suggested WFCM schema [20] also renders beneficial function in decentralized network. The execution was done by NS-3 software to get a better result.

### 8 Future Work

The future work is analyzed by modeling a new security protocol for preventing the same SSDF attack in Cognitive radio based Internet of Things (IoT) concepts.

**Acknowledgements** The part of the proposed work was presented in IEEE ICCIC 17' conference on Dec. 2017. It was held in Tamil Nadu College of Engineering, Coimbatore-59, India. The same will be published in the IEEE Xplore digital library on 8 Nov. 2018. In this article, the proposed work is integrated with an another method to get an enhanced performance, when compared with the different existing methods.

### References

1. Li H (2009) Restless watchdog: selective quickest spectrum sensing in multichannel cognitive radio systems. EURASIP J Adv Signal Process 2009(6):1–12

2. Ningrinla M, Roshni R, Suchismita BB, Amar T (2015) Dynamic decision rule for cooperative spectrum sensing in cognitive radio networks. In: International conference on electrical computer and communication technologies. IEEE, India, pp 1–5
3. Kennong R, Nath S, Ningrinla M, Taggu A (2016) Can clustering be used to detect intrusion during spectrum sensing in cognitive radio networks? *IEEE Syst J* 12(1):938–947
4. Shwetal V, Thakare MN, Vaidya SR (2015) Design approach for cross layer attacks defence in cognitive radio. *Int Res J Eng Technol* 2(4):763–766
5. Aida V (2016) Robust distributed cooperative spectrum sensing for cognitive radio ad hoc networks. PhD thesis, Rice University, Texas, US
6. Wang W, Li H, Sun Y et al (2009) Securing collaborative spectrum sensing against untrustworthy secondary users in cognitive radio networks. *EURASIP J Adv Signal Process* 2010(4):1–15
7. Shameek B, Shamik S, Mainak C (2013) Vulnerabilities in cognitive radio networks: a survey. *Comput Commun* 36(13):1387–1398
8. Hyils Sharon Magdalene A, Thulasimani L (2017) Fuzzy clustering means (FCM) for mitigating spectrum sensing data falsification (SSDF) attack in cognitive radio networks. In: 2017 IEEE International conference on computational intelligence and computing research (ICIC)
9. Saud A, Marco Di R, Fabrizio G (2014) Robust algorithm against spectrum sensing data falsification attack in cognitive radio networks. In: IEEE Vehicular technology conference. IEEE, South Korea, pp 1–5
10. Lotfi AZ (1965) Fuzzy sets. *Inf Control* 8(3):338–353
11. Ehab A, Mahamod I, Rosdiadee N, Nor FA (2017) Beamforming techniques for massive MIMO systems in 5G: overview, classification, and trends for future research. *Front Inf Technol Electr Eng* 18(6):753–772
12. Li W, Han J, Pei J (2001) CMAR: accurate and efficient classification based on multiple class-association rules. In: Proceeding of the 2001 international conference on data mining (ICDM'01), San Jose, CA, pp 369–376
13. Suchismita B, Raiping K, Ningrinla M (2016) Association rule mining for detection of colluding SSDF attack in cognitive radio networks. In: International conference on computer communication and informatics. IEEE, India, pp 1–6
14. Sumit K, Srinivas S, Ramesh KS (2017) A multi-factor trust management scheme for secure spectrum sensing in cognitive radio networks. *Wirel Pers Commun* 97(2):2523–2540
15. Wang H, Li Y, Chang TC (2018) An enhanced cooperative spectrum sensing scheme for anti-SSDF attack based on evidence theory. *Microsyst Technol* 24(6):2803–2811
16. Duggineni C, Chari KM (2018) Mitigation strategy against SSDF attack for healthcare in cognitive radio networks. *Int J Biomed Eng Technol* 27(1/2):33–49
17. Jingyu F, Man Z, Yun X, Hongzhou Y (2018) Securing cooperative spectrum sensing against collusive SSDF attack using XOR distance analysis in cognitive radio networks. *Sensors* 18(370):1–14
18. Liu P, Duan L, Chi X, Zhu Z (2013) An improved fuzzy C-means clustering algorithm based on simulated annealing. In: 10th International conference on fuzzy systems and knowledge discovery. IEEE, China, pp 39–43
19. Linyuan Z, Guoru Z, Qihui W, Zou Y, Han Z, Wang J (2015) Byzantine attack and defense in cognitive radio networks: a survey. *IEEE Commun Surv Tutor* 17(3):1342–1363
20. Luella F (2011) Applying localized realized volatility modeling to futures indices. Senior thesis, Claremont McKenna College, California, USA

# Performance Analysis of Thinning Algorithms for Offline-Handwritten Devanagari Words



Sukhjinder Singh and Naresh Kumar Garg

## 1 Introduction

Character recognition systems contribute for the advancement of the automation process, and can improve the interaction between man and machine in many applications of image processing and pattern recognition such as office/bank automation, cheque verification, writer identification, script identification [1]. Thinning is generally used as preprocessing step in such type of applications including fingerprint recognition and shape description [2] to refine the pattern/word shape. It also helps to reduce large memory requirement for storing structural information of a pattern/word for further processing such as feature extraction and classification. Thinning is a process of skeletonization of a binary image of pattern/word to extract the skeleton of the shape of interest considering topology and shape preservation [3]. Many character recognition methods also use some thinning approach to facilitate shape analysis and stroke identification. There are several thinning algorithms reported in the literature of character recognition [4]. Like Odia and Bangla scripts, Devanagari script has been originally derived from Brahmi script and accompanies top horizontal line. In Fig. 1, Handwritten Devanagari Word (HDW)—“बठिंडा” (Bathinda) before and after applying the thinning algorithm, has been depicted as an example to illustrate the impact of thinning.

Researchers developed and modified various thinning algorithms for many general purpose applications, in the past few years. In many such applications, sometimes static images of handwritten script need to be processed rather than dynamic information which may cause deterioration in system performance.

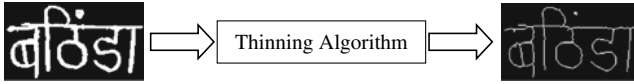
---

S. Singh (✉)

ECED, Giani Zail Singh Campus CET, MRSPTU, Bathinda, India  
e-mail: [er.ssrakhra@gmail.com](mailto:er.ssrakhra@gmail.com)

N. K. Garg

CSED, Giani Zail Singh Campus CET, MRSPTU, Bathinda, India



**Fig. 1** A thinning example (**Left:** input HDW “बठिंडा” (Bathinda); **Right:** resultant thinned HDW)

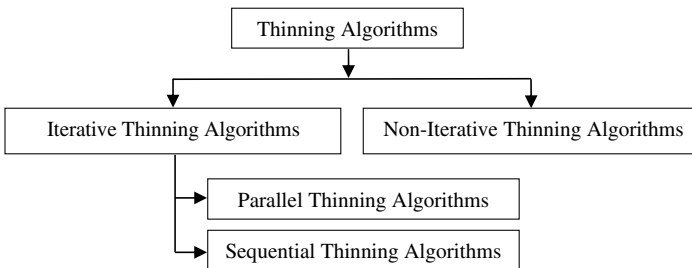
Therefore, it become more challenging to deal with static images rather than dynamic information. A key preprocessing step in the processing of static scripts is thinning or skeletonization where the script is reduced to a collection of lines. In this paper, the performance of three thinning algorithms namely, Zhang-Suen [ZSu] [5], Guo-Hall [GHa] [6] and Lee-Kashyab-Chu [LKC] [7] have been analyzed in terms of various objective and subjective performance metrics. Such algorithms may play an important role in subsequent recognition process of Handwritten Devanagari Words (HDWs). Skeletonization algorithms compute the centerlines skeletons from image boundaries, whereas thinning algorithms remove outer layers of an image while simultaneously preserving the image connectivity [8]. The broad classification of thinning algorithms are presented in the following subsection.

### 1.1 Classification of Thinning Algorithms

Thinning algorithms are broadly classified into two categories namely, iterative and non-iterative [9] as depicted in the Fig. 2.

#### 1.1.1 Iterative Thinning Algorithms

As their name implies, iterative algorithms repeat a predefined procedure or loop on image pixels preferably to obtain unit-pixel thickness/wide skeleton of a pattern/word. These types of algorithms are based on iteratively peeling off the outer boundary layer by layer while moving inside a pattern/word image from outside.



**Fig. 2** Classification of thinning algorithms

Based on way of pixel examination, these algorithms shall further be classified as parallel and sequential-iterative thinning algorithms [2]. These have been briefly explained below.

### Parallel Thinning Algorithms

In parallel thinning algorithms; firstly, desired pixels are identified and thereafter, undesired pixels are removed after identifying all the desired pixels in iterative manner considering results of previous iteration. Such algorithms have gained much attention of researchers as these algorithms are capable of reducing computation time requirements. In this algorithm, all pixels are operated on simultaneously. [ZSu], [GHa] and [LKC] thinning algorithms considered for this work are the examples of parallel thinning algorithm.

### Sequential Thinning Algorithm

In sequential thinning algorithms; undesired pixels are removed while identifying all the desired pixels in iterative manner. The decision to either remove or keep the pixel is governed by previously processed pixels within the same iteration. Earlier thinning algorithms were mainly based on sequential approach later on with the more advancements, researchers preferred parallel approach. To cope up with the problem of preserving the connectivity of the original pattern, many researchers partitioned their algorithms into separate sub-iterations and adopted partial sequentially for these sub-iterations.

#### 1.1.2 Non-Iterative Thinning Algorithms

Non-iterative thinning algorithms do not perform pixel to pixel examination unlike in iterative thinning algorithms. Generally, their implementation is based on some mathematical transformations like distance transforms. These algorithms convert a pattern/word into a set of polygons (which may or may not be regular) based on border of the pattern/word. These set of polygons, help to extract the skeleton of the pattern/word.

## 1.2 Motivation

The purpose of this paper is to take Handwritten Devanagari Words (HDWs) scanned images as input, recognize the handwritten words and modify the handwriting such that it is a beautified thinned version of the input. Thus this work comprises of thinning of HDW for their beautification and further in future

recognition. Thinning is commonly used to reduce font variability in handwritten word recognition [10]. Moreover, it also helps to minimize multiple open end and junction points in a word or character before feature extraction phase [11–12]. In pattern recognition-based applications such as handwritten word recognition, thinning algorithms play an important role due to:

- (i) The thinning algorithms generally provide important information about word or character shape by simplifying its shape and hence thus reducing the amount of data which is required for processing.
- (ii) It shall also help to reduce the processing-time requirements.
- (iii) It shall be used for extraction some features as end points, junction points and connection among various character-components.
- (iv) In pattern recognition-based systems, some algorithms like vectorization also require unit-pixel wide lines as input.

### ***1.3 Requirements of a Good Thinning Algorithm***

Mainly, the requirements of good thinning algorithms are application specific [13]. Variety of thinning algorithms are available in literature whose methodology may or may not be complicated and resultant skeleton depends upon the nature of input pattern. However in general, the requirements of the same are briefly summarized below [9, 14, 15]:

- (i) The resultant skeleton should be of one-pixel wide without redundancies.
- (ii) The resultant skeleton should be centered inside the component and maintain or preserve the component connectivity.
- (iii) There should be no excessive deletion or erosion of pixel points.
- (iv) The resultant skeleton should compress the image data.
- (v) It should take minimum execution time.
- (vi) It should preserve the pattern/word topologically and geometrically which shall be important component for its classification or recognition.
- (vii) It should be robust against noise.

### ***1.4 Challenges***

The thinning algorithms pay a vital role in HDW recognition systems. Therefore, it must be chosen or designed in such a way that system does not affected badly. The various challenges that must be considered while designing or selecting the thinning algorithm for HDW have been briefly discussed below [8, 16–18]:

- (i) *Filtering and preserving the dots:* To remove noise from scanned hand-written word images or samples, filtering such as Median or Gaussian is generally applied before applying thinning algorithms. In handwritten Devanagari words, there may be dots at the top (अं क मात्रा or बिंदी) or bottom (चुकत्ता) of the characters or vowels. Therefore, it become more challenging to identity and remove noise from HDW image samples rather removing or filtering than these character dots.
- (ii) *Retaining the handwritten character connectivity and structural properties:* Handwritten Devanagari words are constructed from the combination of various characters including consonants and vowels modifiers (e.g., अ + ङ + स + न + स + ळ + ल = आसनसोल). The effective HDW-thinning algorithms must retain the connectivity of characters along with structure properties; otherwise it may affect the recognition accuracy of HDW recognition systems due to disconnected or unstructured words.
- (iii) *Unit-pixel width:* It is required that skeleton of HDW should be of minimum thickness preferably having unit-pixel width. Because thickness of word itself is not considered as a main feature for HDW recognition. Thickness of hand-written word may be considered as the complex function of various parameters viz. tip size and type of writing instrument or pen (ball or gel), pressure inserted by writer during writing and paper material. Hence, it becomes a challenge to the researchers while developing a thinning algorithm for HDW.
- (iv) *Crossover or intersection artifacts:* It is desirable to preserve the crossover or intersection points in HDW while extracting skeleton during thinning by taking correct pixels and paths. If these crossover points in HDW are not properly preserved, it may result distorted word or character and hence affect the recognition accuracy of HDW recognition systems.
- (v) *Bogus branches issue:* During the thinning process of HDW, there may be a possibility of producing bogus branches or tails (that was not the part of input word patterns) due to varying character or word thickness. These bogus branches may badly affect the feature extraction and classification stages of HDW recognition systems and hence decreasing the recognition accuracy.

Rest of the paper is organized as follows: Sect. 2 overviews the literature available on thinning algorithms. Implementation of various thinning algorithms is presented in Sect. 3. Performance metrics considered for this word are given in Sect. 4. Section 5 provides the results obtained after implementing [ZSu], [GHa] and [LKC] thinning algorithms. Finally, conclusions are drawn in the Sect. 6 along with future scope.

## 2 Related Works

A study on various thinning algorithms available in literature has been done which shall be helpful to analyze the existing algorithms to do further research in the area of pattern recognition. Singh and Amin obtained image-skeletons of hand-printed

Latin characters using parallel thinning algorithm [1]. They firstly, build a binary tree thereafter applied feature extraction and nonlinear classification approach for recognition of hand-printed Latin characters and achieved average recognition rate 86% using cross-validation. Shimizu et al. produced a stable the skeleton of a digital binary pattern by developing a parallel Hilditch algorithm considering two kinds of masks [19]. They tested their algorithm on 432 kinds of commonly used character fonts and claimed that their algorithm gives better results than other thinning algorithms. To ease the process of feature extraction, Pervouchine et al. extracted the skeleton of handwritten characters (150 images of grapheme “th” written by different writers) in three steps [15]. The produced skeletons were close to human perception of the original pen tip trajectory. Nel et al. analyzed the static handwritten scripts to produce its skeleton that has desirable characteristics [8]. Authors produced skeletons of the handwritten script that are nearly identical to the pen trajectory. Chatbri and Kameyama used space filtering to make thinning algorithm robust against noise to thin binary sketch images [14]. They used adaptive preprocessing to achieve best trade-off between noise removal and shape distortion. They proposed an algorithm to estimate the optimal amount of filtering automatically, so as to achieve good thinning results. Pujari et al. analyzed various thinning algorithms available in literature to check their suitability to skeletonize Odia characters [20]. To preserve vital features of the script and overcome the deficiencies of previous thinning algorithms, authors also proposed another parallel thinning algorithm. Zhang et al. produced a training set of skeleton images from offline/online handwriting samples using vectorization [10]. They also built a handwriting recognizer by taking IAM benchmark databases (achieved error rate reduction of 8%) and online handwriting corpus (achieved error rate reduction of 31%). This recognizer was based on Deep Bidirectional Long Short-Term Memory (DBLSTM) and Hidden Markov Model (HMM).

Goyal and Dutta analyzed the performance of three thinning algorithms for offline-handwritten Devnagri numeral script [9]. On the basis of various performance parameters viz. pixel count, compression ratio, connectivity, information loss and topology preservation; authors gathered that that adaptive thinning algorithm based on Gaussian filtering gives better parametric results for noisy images. To obtain the skeleton of Arabic text (20 image samples), Saudagar and Mohammed implemented Zhang-Suen thinning algorithm using various platforms namely, OpenCV-Java, MATLAB and C++ [4]. Based on the computational time, they concluded that OpenCV-Java-based implementation of Zhang-Suen thinning algorithm is faster as compared with other two platforms viz. MATLAB and C++ . Neelima et al., developed a rank-based skeleton extraction algorithm using MPEG7 dataset [21]. They also compared the performance of their algorithm with Zhang-Suen, Lu-Wang and Modified Zhang-Suen skeletonization algorithms on the basis of various performance parameters namely, elapsed time, thinning rate/speed and connectivity. Authors concluded that the rank-based skeletonization method is superior to other mentioned skeletonization algorithms.

Literature survey reveals that with the advancement of technology, this area has gained some momentum; however, a lot of work still needs to be done for



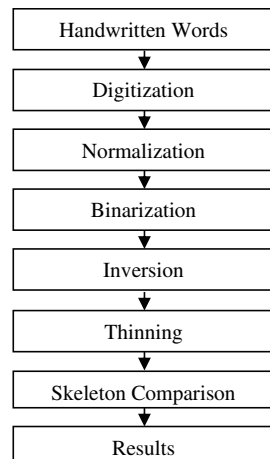
optimization of skeleton algorithms for many applications including handwritten character/word recognition.

### 3 Implementation of Thinning Algorithms

Generally, thinning is used as preprocessing step in the Devanagari character or word recognition systems. Thinning algorithms are designed to simplify the character or word shape which further helps to reduce the amount of data to represent/store a character or word. The resultant skeleton of handwritten Devanagari character or word obtained through thinning algorithms can be used for other stages of character or word recognition such as detection of baseline, word segmentation and feature extraction for classification. Hence, choosing/designing the effective thinning algorithm plays an important role for handwritten Devanagari character or word recognition. In this work, performance analysis of [ZSu], [GHa] and [LKC] thinning algorithms have been carried out for offline-HDWs in Python using Scikit-learn based on various objective and subjective quality metrics. Implementation steps for this work have been depicted in the Fig. 3 and summarized below:

- (i) *Digitization*: In the process of digitization, paper-based handwritten Devanagari word samples collected from various writers are converted into electronic form by scanning at 300 dots per inch resolution.
- (ii) *Normalization*: This is used to normalize the handwritten words into same or uniform size. Here each handwritten word is normalized into  $256 \times 64$  window size due to horizontal writing approach of Devanagari words.
- (iii) *Binarization*: Handwritten Devanagari word samples obtained in step-ii above, are converted into black (binary 0) and white (binary 1) pixels using

**Fig. 3** Implementation steps for this work



Ostu's threshold selection method from gray-level histograms [22]. It reduces the computational complexity of the algorithm due to two colors for processing [23].

- (iv) *Inversion*: Thereafter, binary HDW samples are inverted which shall turn white pixels into black and vice-versa.
- (v) *Thinning*: Finally, thinning algorithms namely, [ZSu], [GHa] and [LKC] have been implemented for offline-HDW.
- (vi) *Skeleton Comparison*: The performance of above mentioned thinning algorithms has been analyzed using various objective and subjective quality metrics.

### **3.1 A Brief Overview of [ZSu] Algorithm**

In 1984, Zhang and Suen proposed a thinning algorithm which shall be capable of thinning digital patterns in an efficient way [5]. Their algorithm removes the pixels on the object boundary by making successive passes within the entire image pixels until no more pixels can be removed. It algorithm is based on two sub-iterations. In the first sub-iteration: south-east boundary points and the north-west corner points shall be deleted. Whereas in the second sub-iteration: north-west boundary points and the south-east corner points shall be deleted while preserving the end points and pixel connectivity. In this way, after performing several iterations, only a similar look skeleton of the pattern or handwritten word shall remain with unitary thickness.

### **3.2 A Brief Overview of [GHa] Algorithm**

Guo and Hall algorithm is based on the removing of the border pixels at each iteration until none other pixel can be removed without shifting the connectivity. In this way, it produces a relatively thicker skeleton of digital pattern. Their algorithm is based on two sub-iteration approaches. In first approach: alternatively north and east boundary pixels, thereafter south and west boundary pixels shall be deleted. While in second approach: alternately a thinning operator shall be applied to one of two subfields. This algorithm shall result a very thin medial curves along with preserving image connectivities [6].

### **3.3 A Brief Overview of [LKC] Algorithm**

In 1994, Lee et al. [7] proposed another parallel thinning algorithm which can also handle 3-D pattern/object images. It was developed for extracting both medial surfaces and axes of binary image. It was based on iterative approach. In every

iterative-loop; firstly, it moves over the all pixels of a pattern and thereafter, it shall remove the undesired pixels until the pattern stops altering. In order to preserve the local-connectivity of a pattern in a better way, its every iterative-loop generally consists of two phases. In the first phase, it gathers the list of undesired pixels to be removed. Secondly, in the next phase, it rechecks the shortlisted undesired pixels of first phase to ensure the preserved connectivity of the pattern.

The goal of this work is analyze the performance of above mentioned thinning algorithms to check their suitability to thin or skeleton offline-handwritten Devanagari words.

## 4 Performance Metrics

Performance metrics play a vital role in the development of various algorithms such as thinning, feature extraction and classification. These can be used to evaluate the performance of such algorithms in terms of quality of the processed images. In literature, there are basically two broad categories of performance metrics namely, objective and subjective performance metrics or measurements.

### 4.1 Objective Performance Metrics

Some available objective performance metrics to test the performance of thinning algorithms viz. reduction rate, sensitivity measurement and thinness measurement considered for this work, are briefly defined in the following sub-sections [9, 14, 24]:

#### 4.1.1 Reduction Rate (RR)

Reduction rate is calculated on the basis of foreground pixels present in the original image and resultant skeleton of the image. Mathematically reduction rate is defined as given below:

$$\text{Reduction Rate(RR)} = \left[ \frac{(fgps - fgpst)}{fgps} \right] \times 100$$

where

$fgps$  = foreground pixels in the original image

$fgpst$  = foreground pixels in the skeleton image

Ideally, reduction rate should be 100%. Practically, it should be high as possible.

### 4.1.2 Sensitivity Measurement (SM)

It is another qualitative metric that helps to determine whether the thinning algorithm has selected the best thinned image from the available scale space. The total number of cross-points present in an image can be used for the measurement of sensitivity. It is expressed by the following mathematical equation:

$$\text{Sensitivity Measurement}(SM) = \sum_{i=0}^n \sum_{j=0}^m S(P[i][j])$$

where

$$S(P[i][j]) = \begin{cases} 1, & \text{if Trans}(P[i][j]) > 2 \\ 0, & \text{otherwise} \end{cases}$$

Lower value of SM, indicates the skeleton image contains less artifacts, redundant branches and lines caused by noise.

### 4.1.3 Thinness Measurement (TM)

The Thinness Measurement (TM) parameter measures the extent or degree to which a pattern or handwritten word present in the scanned image is thinned. Mathematically, TM can be calculated as:

$$\text{Thinness Measurement}(TM) = \left( 1 - \frac{TM_1}{TM_2} \right)$$

where

$$TM_1 = \text{Total triangle count in thinned image} = \sum_{i=0}^n \sum_{j=0}^m \text{triangle\_count}(P[i][j])$$

$$TM_2 = 4 \times [\max(\text{height, width}) - 1]^2 = 4 \times [\max(m, n) - 1]^2$$

Its value range is of  $[0, 1]$ .  $TM = 1$  indicates that pattern or handwritten word is completely unit-pixel wide.

## 4.2 Subjective Performance Metrics

The performance analysis of thinning algorithms is also definitive measure of the quality of the skeleton image. The thinned image quality may be specified by Mean Opinion Score (MoS), which is the result of the perception-based subjective

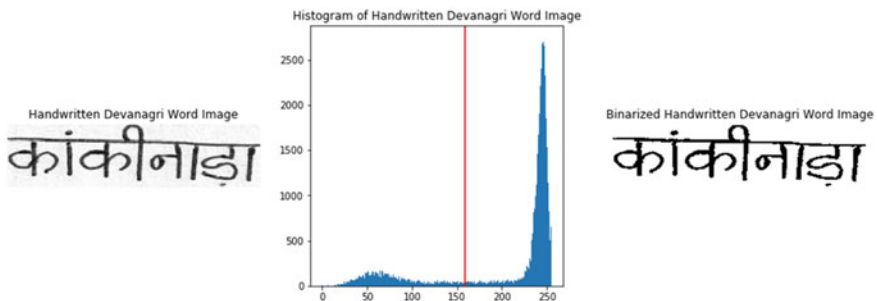
evaluation. The 5-level grading scores of MoS (i.e., 5-excellent, 4-good, 3-acceptable, 2-poor quality and 1-unacceptable) have been considered for this work too. In the following sections the simulation results have been represented for various thinning algorithms based upon above performance metrics.

## 5 Results and Evaluation

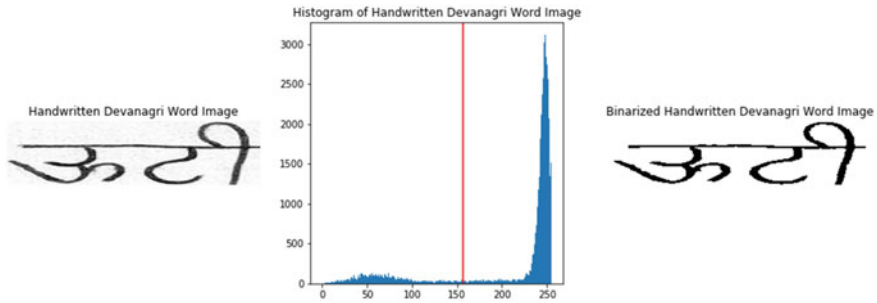
### 5.1 Dataset Description and Preprocessing

Performance analysis of various thinning algorithms has been carried out on a common set of HDW samples (15-word-classes) collected from hundreds of writers as mentioned in [25–29] so as to check the suitability of the algorithms for the same. The 15 word-classes taken for this work are namely, “आसनसोल”(Asansol), “औरंगाबाद” (Aurangabad), “कांकीनाडा” (Kakinada), “कपूरथला” (Kapurthala), “खजुराहो” (Khajuraho), “ऋषिकेश” (Rishikesh), “नैनीताल” (Nainital), “चौरंगी” (Chowringhee), “त्रिवेणी” (Triveni), “वाराणसी” (Varanasi), “हुगली” (Hooghly), “मैसूर” (Mysuru), “छपरा” (Chapra), “मेरठ” (Meerut), “ऊटी” (Ooty). For each word, the quality of resultant thinned word image is evaluated using objective as well as subjective performance metrics. Firstly, each HDW samples were collected using A-4 sized paper, then scanned, normalized and thereafter, converted into binary image using Ostu’s threshold selection method [22]. Results for two handwritten word-classes namely, “कांकीनाडा” (Kakinada) and “ऊटी” (Ooty) are depicted before and after applying Ostu’s threshold along with their histograms in the Figs. 4 and 5, respectively.

Corresponding binarized HDWs are further processed and various thinning algorithms are applied on them as discussed in the following section.



**Fig. 4** Handwritten Devanagari word “कांकीनाडा” (Kakinada) before and after applying Ostu’s method along with its histogram



**Fig. 5** Handwritten Devanagari word “ऊटी” (Ooty) before and after applying Ostu’s method along with its histogram

## 5.2 Performance Analysis and Comparison

Before applying various thinning algorithms to Ostu’s thresholded handwritten word samples (binarized image), the each samples of binary HDW are inverted which shall turn white pixels into black and vice-versa. After that thinning algorithms namely, [ZSu], [GHa] and [LKC] have been applied/implemented on a common set of handwritten Devanagari word samples (15 word-classes) using Scikit-learn library available in Python. Pictorial representation of two HDW samples (2 word-classes) namely, “कांकीनाड़ा” (Kakinada) and “ऊटी” (Ooty) are presented in the following Figs. 6 and 7 for reference.

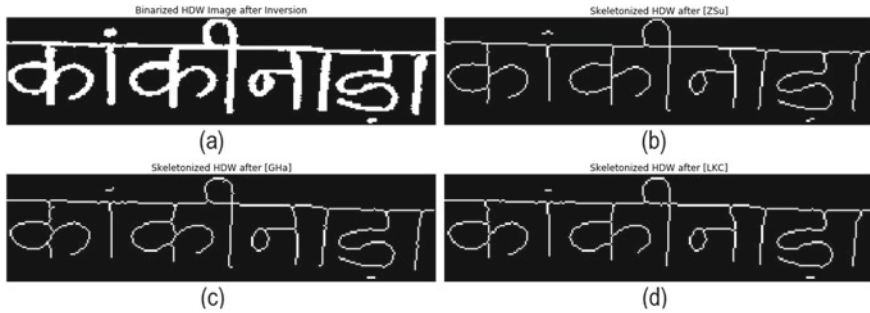
To check the suitability of these algorithms to thin or skeleton offline-HDW, various available objective performance metrics namely, reduction rate, sensitivity measurement and thinness measurement are calculated for various HDW samples of 15 different word-classes. The average measurements of these performance metrics are presented in the following Tables 1, 2 and 3, respectively.

Table 1 shows that [LKC] algorithm has achieved higher average reduction rate as compared with [ZSu] and [GHa] algorithms for HDWs. Moreover, both [ZSu] and [GHa] algorithms resulted nearby figures in terms of reduction rate. Graphical representation of the same has been depicted in the following Fig. 8. This shows ability of thinning algorithms to reduce the foreground pixels in original pattern/word.

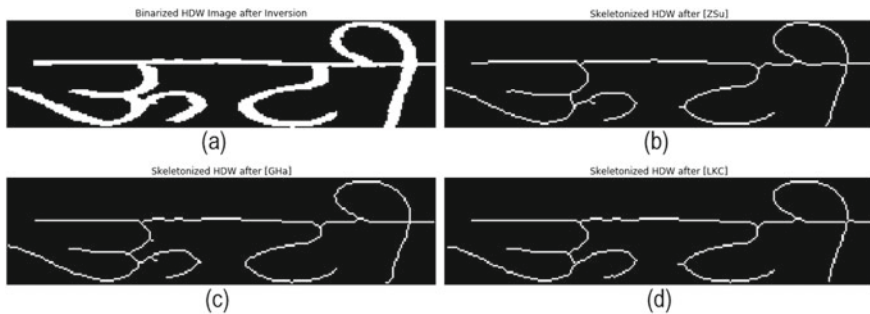
Table 2 shows that both [ZSu] and [GHa] algorithms has achieved lower average sensitivity measurement as compared with [LKC] algorithm for HDWs.

Graphical Representation of Sensitivity Measurement (SM) has been depicted in the below Fig. 9. Therefore, it can be summarized that resultant thinned image by [LKC] algorithm may contain some artifacts, redundant branches and lines caused by noise.

Table 3 shows that [LKC] and [GHa] algorithms has achieved higher average thinness measurement as compared with [ZSu] algorithm for HDWs.



**Fig. 6** Handwritten Devanagari word “कांकीनाड़ा”(Kakinada) and its resultant skeleton **a** after Ostu’s threshold and inversion **b** after Zhang and Suen [5] **c** after Guo and Hall [6] **d** after Lee et al. [7]



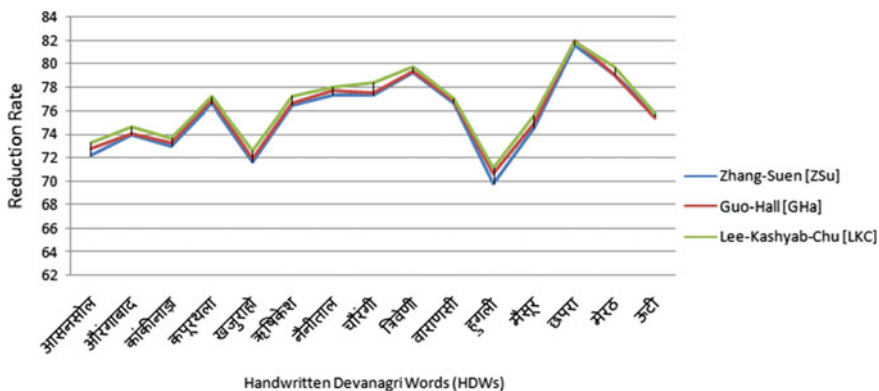
**Fig. 7** Handwritten Devanagari word “ऊटी”(Ooty) and its resultant skeleton **a** after ostu’s threshold and inversion **b** after Zhang and Suen [5] **c** after Guo and Hall [6] **d** after Lee et al. [7]

Graphical representation of the thinning measurement for various thinning algorithms has been depicted in the following Fig. 10. This shows ability of thinning algorithms to produce unit-pixel wide skeleton of original pattern/word.

Above representation shows that Lee-Kashyab-Chu [LKC] thinning algorithm perform slightly better thinning of collected HDW samples in terms of reduction rate and sensitivity measurement as compared with Zhang-Sue [ZSu] and Guo-Hall [GHa] algorithms. In terms of subjective performance metrics, i.e., MoS, skeletons produced by Lee-Kashyab-Chu [LKC] algorithm achieved higher MoS grade-4 (good) as compared with other two thinning algorithms. Moreover, both Guo-Hall [GHa] and Zhang-Suen [ZSu] algorithms achieved MoS grading score-3 (acceptable) for the HDWs database considered for this work.

**Table 1** Reduction rate (RR)

HDWs	Zhang-Sue [ZSu]	Guo-Hall [GHa]	Lee-Kashyab-Chu [LKC]
आसनसोल	72.21705426	72.80620155	73.30232558
औरंगाबाद	73.95287958	74.05104712	74.67277487
कांकीनाडा	73.01197421	73.28830212	73.62603623
कपूरथला	76.60914818	76.93782525	77.18433306
ऋषिकेश	71.6117851	71.95840555	72.6169844
ऋषकिश	76.45631068	76.69902913	77.21143474
नैनीताल	77.32476962	77.71572187	78.02289863
चौरंगी	77.30870712	77.54324245	78.39343301
त्रिवेणी	79.24880128	79.43526905	79.70165157
वाराणसी	76.66281087	76.83632157	77.06766917
हुगली	69.76058932	70.68139963	71.12338858
मैसूर	74.53838678	74.99190152	75.5749919
छपरा	81.56970913	82.02106319	81.87061184
मेरठ	79.04389657	79.07396272	79.70535177
ऊटी	75.73721538	75.47592385	75.73721538
Average reduction rate	75.67026921	75.96770777	76.38740672

**Fig. 8** Graphical representation of reduction rate (RR)

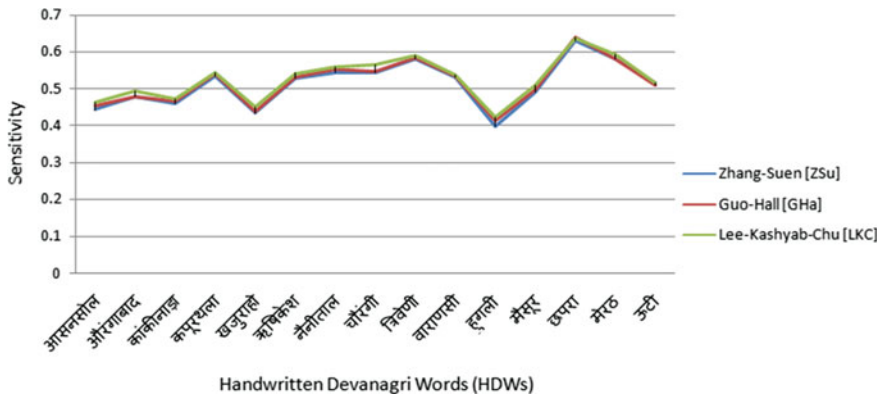
## 6 Conclusions and Future Scope

In this work, performance analysis of thinning algorithms namely, Zhang-Suen [ZSu], Guo-Hall [GHa] and Lee-Kashyab-Chu [LKC] has been carried out to check their suitability to skeleton offline-handwritten Devanagari words. These algorithms have been tested using a common dataset of offline-HDWs with 15-word-classes written by hundreds of writers. Finally, conclusions are drawn on the basis of



**Table 2** Sensitivity measurement (SM)

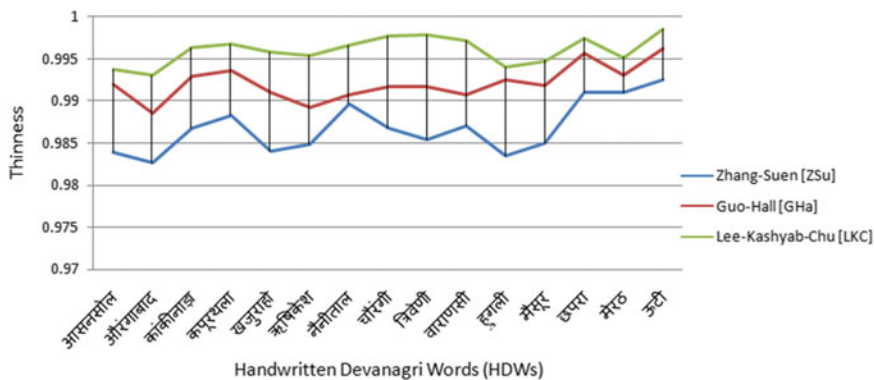
HDWs	Zhang-Sue [ZSu]	Guo-Hall [GHa]	Lee-Kashyab-Chu [LKC]
आसनसोल	0.441860465	0.453023256	0.463875969
औरंगाबाद	0.478075916	0.480039267	0.492473822
कांकीनाडा	0.45901136	0.464844949	0.471292601
कपूरथला	0.532456861	0.5393043	0.543960559
खजुराहो	0.432235702	0.43847487	0.451299827
ऋषिकेश	0.526968716	0.531823085	0.54180151
नैनीताल	0.544540631	0.552080424	0.558503211
चौरंगी	0.54353562	0.548226327	0.565523307
त्रिवेणी	0.581513053	0.585242408	0.590570059
वाराणसी	0.530075188	0.533545402	0.537883169
हुगली	0.39558011	0.413627993	0.422099448
मैसूर	0.488824101	0.498218335	0.509556203
छपरा	0.630892678	0.639919759	0.636910732
मेरठ	0.579374624	0.580276609	0.592603728
ऊटी	0.51399776	0.509145203	0.514371034
Average sensitivity measure	0.511929519	0.517852813	0.526181679

**Fig. 9** Graphical representation of sensitivity measurement (SM)

various objective and subjective performance metrics. From this work, it has been gathered that Lee-Kashyab-Chu [LKC] algorithm performed better than other two mentioned algorithms in terms of Reduction Rate (RR), Thinness Measurement (TM) and Mean Opinion Score (MoS). However, as this algorithm also achieved slightly-higher value of Sensitivity Measurement (SM) as compared with other

**Table 3** Thinness measurement (TM)

HDWs	Zhang-Sue [ZSu]	Guo-Hall [GHa]	Lee-Kashyab-Chu [LKC]
आसनसोल	0.983855945	0.991927973	0.993790748
औरंगाबाद	0.982651391	0.988543372	0.993126023
कांकीनाडा	0.986797667	0.992938287	0.996315628
कपूरथला	0.988209487	0.993693447	0.996709624
ऋषिकेश	0.984049931	0.990984743	0.995839112
ऋषिकेश	0.984889369	0.989206692	0.995412844
नैनीताल	0.989667691	0.990784697	0.996648981
चौरंगी	0.986788021	0.991779213	0.997651204
त्रिवेणी	0.985348961	0.991742142	0.99786894
वाराणसी	0.986982933	0.990743419	0.997107318
हुगली	0.983539095	0.99251777	0.994014216
मैसूर	0.985026042	0.991861979	0.994791667
छपरा	0.990966123	0.995734003	0.99749059
मेरठ	0.990977444	0.993082707	0.99518797
ऊटी	0.99252895	0.996264475	0.99850579
Average thinness measurement	0.986818603	0.992120328	0.99603071

**Fig. 10** Graphical representation of thinness measurement (TM)

mentioned algorithms. Thus, it also depicts that resultant thinned image of this algorithm may contain some artifacts, redundant branches and lines caused by noise as compared with others algorithms considered for this work. Thus, thinning scheme could be used as a useful method of preprocessing of handwritten Devanagari word recognition. In future, some other thinning algorithms along with few more performance metrics may also be considered to test their performance on offline-HDW samples.

## References

1. Singh S, Amin A (1999) Neural network recognition of hand-printed characters. *Neural Comput Appl* 8:67–76
2. Park SH, Kim SK, Kim HJ (1996) A fully parallel thinning algorithm using a weighted template. In: *Proceeding of IEEE TENCON—digital signal processing applications*. pp 300–303
3. Liu L, Chambers EW, Letscher D, Ju1 T (2010) A simple and robust thinning algorithm on cell complexes. *Pacific Graph* 29:2254–2260
4. Saudagar AKJ, Mohammed HV (2016) OpenCV based implementation of Zhang-Suen thinning algorithm using java for arabic text recognition. In: *Information systems design and intelligent applications, advances in intelligent systems and computing*, vol 435. pp 265–271
5. Zhang TY, Suen CY (1984) A fast parallel algorithm for thinning digital patterns. *Commun ACM* 27(3):236–239
6. Guo Z, Hall RW (1989) Parallel thinning with two sub-iteration algorithms. *Image Process Comput Vision* 32(3):359–373
7. Lee TC, Kashyap RL, Chu CN (1994) Building skeleton models via 3-D medial surface/axis thinning algorithms. *CVGIP: Graph Model Image Process* 56(6):462–478
8. Nel EM, Preez JA, Herbst BM (2009) A pseudo-skeletonization algorithm for static handwritten scripts. *Int J Docum Anal Recogn* 12:47–62
9. Goyal G, Dutta M (2016) Experimental approach for performance analysis of thinning algorithms for offline handwritten devnagri numerals. *Indian J Sci Technol* 9(30):1–10
10. Zhang X, Wang HM, Wang L, Huo Q, Haifeng L (2015) Building handwriting Recognizers by leveraging skeletons of both offline and online samples. In: *Proceeding of IEEE 13th international conference on document analysis and recognition (ICDAR)*. pp 406–410
11. Arora S, Bhattacharjee D, Nasipuri M, Basu DK, Kundu M (2008) Combining multiple feature extraction techniques for handwritten devnagari character recognition. In: *Proceeding of IEEE region 10 and the third international conference on industrial and information systems (ICIINFS)*, pp 1–6
12. Qaroush A, Jaber B, Mohammad K, Washaha M, Maali E, Nayef N (2019) An efficient, font independent word and character segmentation algorithm for printed arabic text. *J King Saud Univers-Comput Inform Sci* 1–15
13. Gonzalez Woods R (1992) In: *Digital image processing*. Addison-Wesley Publishing Company
14. Chatbri H, Kameyama K (2014) Using scale space filtering to make thinning algorithms robust against noise in sketch images. *Pattern Recogn Lett* 42:1–10
15. Pervouchine V, Leedham G, Melikhov K (2005) Handwritten character skeletonisation for forensic document analysis. In: *Proceeding of ACM symposium on applied computing (SAC)*, pp 754–758
16. Chouinard C, Plamondon RJ (1992) Thinning and segmenting handwritten characters by line following. *Mach Vis Appl* 5:185–197
17. Yu K, Wu J, Zhuang Y (2008) Skeleton-based recognition of chinese calligraphic character image. In: Huang YMR et al. (eds) *Advances in multimedia information processing—PCM 2008*. *Lecture Notes in Computer Science*, Springer, Berlin, Heidelberg, vol 5353. pp 228–237
18. Cowell J, Hussain F (2001) Thinning arabic characters for feature extraction. In: *Proceeding of IEEE fifth international conference on information visualization*, pp 181–185
19. Shimizu M, Fukuda H, Nakamura G (2000) A thinning algorithm for digital figures of characters. In: *Proceeding of the 4th IEEE southwest symposium on image analysis and interpretation*, pp 83–87
20. Pujari AK, Mitra C, Mishra S (2014) A new parallel thinning algorithm with stroke correction for odia characters. In: *Advanced computing, networking and informatics, volume 1, smart innovation, systems and technologies*, vol 27. pp 413–419

21. Neelima N, Srikrishna A, Rao KG (2020) Rank based approach for extracting unit pixel width skeleton. In: Proceeding of international conference on recent advancements in engineering and management (ICRAEM), pp 1–10
22. Ostu N (1979) A threshold selection method from gray-level histograms. *IEEE Trans Syst Man Cybern* 9(1):62–66
23. Sahare P, Dhok SB (2018) Multilingual character segmentation and recognition schemes for Indian document images. *IEEE Access* 6:10603–10617
24. Ng GS, Zhou RW, Quek C (1994) A novel single pass thinning algorithm. *IEEE Trans Syst Man Cybernet* 1267–1275
25. Shaw B, Bhattacharya U, Parui SK (2014) Combination of features for efficient recognition of offline handwritten devanagari words. In: Proceeding of IEEE 14th international conference on frontier in handwritten recognition, pp 240–245
26. Shaw B, Bhattacharya U, Parui SK (2015) Offline handwritten devanagari word recognition: information fusion at feature and classifier level. In: Proceeding of IEEE 3rd IAPR asian conference on pattern recognition, pp 720–724
27. Shaw B, Parui SK, Shridhar M (2008) A segmentation based approach to offline handwritten devanagari word recognition. In: Proceeding of IEEE international conference on information technology (ICIT), pp 256–257
28. Shaw B, Parui SK, Shridhar M (2008) Offline handwritten devanagari word recognition: a holistic approach based on directional chain code feature and HMM. In: Proceeding of IEEE international conference of information technology, pp 203–208
29. Shaw B, Parui SK (2010) A two stage recognition scheme for offline handwritten Devanagari words. *Stat Sci Interdisciplinary Res Mach Interpretation Patterns* 11:145–165

# Application of Digital Image Correlation Technique to Study Strain Localization in Fe–Ni–W Alloy



Kavita Tewari and Ramesh Kulkani

## 1 Introduction

Deterioration of engineering structures, such as bridges, roads, dams, tunnels, and buildings, due to aging and corrosion, accidents, or overuses is one of the main concerning issues. To avoid any catastrophic failure, these engineering structures are kept under constant vigilance. Such monitoring provides valuable information that helps in estimating the structural integrity, durability, and ensures that the structures are in proper condition. This monitoring mechanism can also be used as a preventive system where before any catastrophic failure, signals can be identified by focusing on areas that are most likely to fail and identifying early signs of failures, like the initiation of cracks. The most likely regions in these structural materials are the points of the highest strain or stress triaxiality, initiating the formation of cracks [1]. Therefore, to have an anticipatory accurate prediction of the life of a structural component, many times, it becomes necessary to have prior knowledge of regions of stress concentrations or distribution of the stress in the structural components.

In this regard, digital image correlation (DIC) techniques have tremendous scope as these techniques could be used to identify the regions of stress concentrations in a material undergoing deformation. For example, the deformation of a specimen can be monitored by estimating displacements of pixels of digitally acquired images of the surface. DIC technique was developed to measure the full-field displacement (FFD) and displacement gradient of a specimen under load [2, 3]. Since then, the DIC

---

K. Tewari (✉)

Department of Electronics Engineering, VESIT, University of Mumbai, Mumbai, India

e-mail: [kavita.tewari@ves.ac.in](mailto:kavita.tewari@ves.ac.in)

R. Kulkani

Department of Electronics and Telecommunication Engineering, VESIT, University of Mumbai, Mumbai, India

e-mail: [ramesh.kulkarni@ves.ac.in](mailto:ramesh.kulkarni@ves.ac.in)

method has been optimized and improved in terms of accuracy and computational complexity by various researchers. DIC has been used to understand the behavior of systems in biological materials [4, 5], metallic alloys [6], and polymers [7], to name a few. DIC has also been used to understand the nonlinear behavior of materials during the formation of slip bands [8] and crack tips [9].

Typically, by analyzing microscopic images prior to and subsequent to deformation DIC can provide semi-quantitative to quantitative information in terms of displacement and strain fields [10]. DIC technique, without any presumption, provides accurate estimation of strain locally as well as globally and that too of a real system. In addition, DIC being a non-contact technique is superior to other contact-based techniques in providing more accurate estimate of full-field displacement (FFD) and strain.

In current scanning electron microscope (SEM) with the capability of nanometer resolution, it is possible to estimate deformation at submicron level. Application of 2D-DIC techniques on high spatial resolution images obtained under such SEM finds applications in broader areas [11]. In this paper, a 2-D DIC system that makes use of an open-source MATLAB-based 2D DIC software, Ncorr, has been used to compute the distribution of FFD as well as strain [12] in the samples which were subjected to in situ deformation under an SEM. For this purpose, intermittently images were recorded during deformation. In this study, microstructural features within samples were utilized to study the deformation, and later on, the study was extended to featureless samples to prove the wider application of the technique. In addition, a brief description of the algorithm used to process the images has also been presented.

## 2 Methodology

### 2.1 Fundamentals of DIC

2D-DIC involves tracking of the image points on the digital images obtained prior and subsequent to deformation. The DIC approach has two successive stages:

- (1) First stage involves recording of the images of the same region of the specimen. In order to have high accuracy, intermittently images of the same region of the sample during deformation were recorded.
- (2) In second stage, DIC software is applied to attain the required results as per the desired analysis. The image taken prior to deformation is taken as reference and images obtained after deformation are compared with respect to this reference image to identify the displacements of pixels within a given region.

To implement DIC analysis, the undeformed image is divided into smaller portions called subsets. It is assumed that within each subset, deformation is homogenous, and the relationship between the reference and the deformed images is linear (Eq. 1). It is also assumed that the transformation of the coordinate of the points of the

undeformed image to the points of deformation uses a first-order shaping function that involves translation and rotation only.

$$\begin{aligned} x'_i &= x_i + u + Du + u_x Dx_i + u_y Dy_i \quad \text{and} \\ y'_i &= y_i + v + Dv + v_x Dx_i + v_y Dy_i \end{aligned} \tag{1}$$

where  $x_i, y_i$  are the coordinates of a pixel in the reference image which after deformation become  $x'_i$  and  $y'_i$  in deformed images,  $u$  and  $v$  are the integer pixel displacements, and  $\Delta u$  and  $\Delta v$  are subpixel displacements in the  $x$  and  $y$  directions, respectively.

To find out deformation within a subset, the extrema of a correlation function are established using DIC algorithm. To initiate the computation, a seed point is located at an integer position. Similarity, between the reference and deformed subset is compared through a pre-defined correlation criterion. This is achieved by DIC algorithms by finding the extremum of a correlation cost function. The normalized cross-correlation (NCC) [4, 10] could be expressed as:

$$C_{\text{NCC}} = \sum_{i=-M}^M \sum_{j=-M}^M \left[ \frac{f(x_i, y_j)g(x'_i, y'_i)}{\bar{f}\bar{g}} \right] \tag{2}$$

In this equation,  $f(x_i, y_i)$  and  $g(x'_i, y'_i)$  are the grayscale intensity functions of the reference image and deformed image at a specified location  $(x_i, y_i)$ , respectively, and  $\bar{f}, \bar{g}$  are determined by using functions given in Eqs. 3 and 4:

$$\bar{f} = \sqrt{\sum_{i=-M}^M \sum_{j=-M}^M [f(x_i, y_j)]^2} \tag{3}$$

$$\bar{g} = \sqrt{\sum_{i=-M}^M \sum_{j=-M}^M [g(x'_i, y'_j)]^2} \tag{4}$$

Lagrangian strains can be obtained using the following equations [4, 12]:

$$E_{XX} = \frac{1}{2} \left( 2 \frac{\partial u}{\partial x} + \left( \frac{\partial u}{\partial x} \right)^2 + \left( \frac{\partial v}{\partial x} \right)^2 \right) \tag{5}$$

$$E_{XY} = \frac{1}{2} \left( \frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} + \frac{\partial u}{\partial x} \frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} \frac{\partial v}{\partial y} \right) \tag{6}$$

$$E_{YY} = \frac{1}{2} \left( 2 \frac{\partial v}{\partial y} + \left( \frac{\partial u}{\partial y} \right)^2 + \left( \frac{\partial v}{\partial y} \right)^2 \right) \tag{7}$$

Strain contour maps computed by the DIC method were used to calculate local true strain using the following equation:

$$\varepsilon_{\text{eng}} = -1 + \sqrt{2 \times \varepsilon_{\text{Lag}} + 1} \quad (8)$$

$$\varepsilon_{\text{true}} = \ln(1 + \varepsilon_{\text{eng}}) \quad (9)$$

where subscripts “eng” stands for engineering, “Lag” stands for Lagrangian, respectively.

To investigate relationships between displacements and between strains Pearson correlation coefficient “r” was calculated using following relationship [13]:

$$r = \frac{1}{n-1} \sum \frac{(x_i - \bar{X})(y_i - \bar{Y})}{s_x s_y} \quad (10)$$

where n represents number of samples, x and y represent variables,  $\bar{X}$  and  $\bar{Y}$  represent their mean values,  $s_x$  and  $s_y$  represent their standard deviations, respectively. The value of “r” attains values between +1 and -1. In the case, variables do not bear any relationship, there will be no systematic pattern in  $(x_i - \bar{X})$  and  $(y_i - \bar{Y})$  and summation of  $(x_i - \bar{X})(y_i - \bar{Y})$  would cancel each other, reducing r to a negligible value. However, if the variables are positively associated, i.e., increment in one brings increment in other, the positive values of  $(x_i - \bar{X})$  will coincide with the positive values of  $(y_i - \bar{Y})$  and negative values of  $(x_i - \bar{X})$  with negative values of  $(y_i - \bar{Y})$  and summation of  $(x_i - \bar{X})(y_i - \bar{Y})$  will yield a positive value of r. In the case variables are negatively associated, the summation of  $(x_i - \bar{X})(y_i - \bar{Y})$  will give negative value of r.

When the distribution is not completely known, Pearson’s coefficient, r, does not provide the proper relationship between the variables, in such cases, the Spearman rank correlation coefficient test can be performed to cross-check the validity of correlation between the two variables. The Spearman rank correlation can be expressed as [13]:

$$r_s = 1 - 6 \sum \frac{d^2}{N(N^2 - 1)} \quad (11)$$

where N is the number of samples, and d is the difference in statistical rank of corresponding variables (u and v). Because of the use of formulation ranks, computation of the Spearman rank correlation coefficient is easier in comparison to the Pearson coefficient [13].



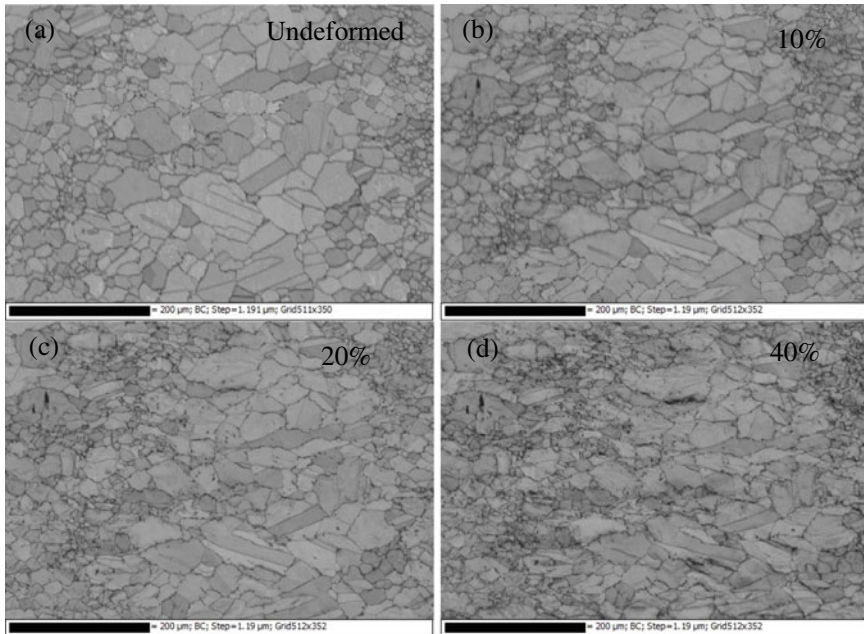
## 2.2 Materials

In the present study, deformation in the Fe–Ni–W-based alloy was studied using the DIC technique. The Fe–Ni–W alloy is generally used as a single-phase structural material for high-temperature applications. Specimens of the alloys were ground using successive grades of emery papers up to grade 2000. Finally, these samples were mirror-polished using colloidal silica to reveal their microstructures. Samples of the Fe–Ni–W alloys were subjected to in situ deformation, and images were recorded using the secondary electron detector. Images of the sample were intermittently recorded to observe the progression of deformation in the sample. These images were later subjected to DIC analysis to find out the displacements and distribution of the stress in the material.

## 3 Results and Discussion

### 3.1 Computation of Strain and Statistical Analysis in Fe–Ni–W Alloys

Images at successive deformations of the Fe–Ni–W alloy are shown in Fig. 1. From these images, a bimodal distribution of grains with small grains ( $\sim 1\mu\text{m}$ ) and large grains ( $\sim 10\mu\text{m}$ ) could be noticed. In this study, the region of the large grain was selected, as coarse grains require lower stress for plastic yielding [14], and therefore, the major portion of deformation is expected to confine mainly to this region. In addition, the large grain regions surrounded by difficult to deform small grain regions, deform under constrained rigid boundaries; because of which the displacement of grains will not remain uniform. Due to this bimodal distribution of grain, as may be noticed from Fig. 2,  $u$  and  $v$  displacements along  $x$  and  $y$  directions show band structures. The relevant computed parameters and their ranges are tabulated in Table 1. By comparing images of samples (Fig. 2) deformed 10, 20, and 40%, a gradual change in the microstructure due to deformation could be noticed. The clear-cut bands of displacements noticed in Fig 2a and b became relatively diffused in successive deformation (compare Fig 2 a and b with Fig 2 a and f). It may be noticed from the color chart representing the band of displacement that the band is effectively increasing with deformation. It is worth mentioning here that these bands should not be considered in the absolute sense. For example, in the case of a 10% deformed sample, minimum value of  $u$ -displacement is  $-1.542\mu\text{m}$ , and the maximum value is  $55.06\mu\text{m}$ . In the case of  $v$  displacement, minimum and maximum displacements were  $23.43\mu\text{m}$  and  $76.48\mu\text{m}$ . The mean values of  $u$  and  $v$  displacements are  $21.65\mu\text{m}$  and  $52.34\mu\text{m}$ , respectively and their standard deviations were computed as 11.12 and 7.75, respectively. In the case of the 40% deformation frame, the minimum and maximum displacements of  $u$ -displacement became  $-4.63\mu\text{m}$  and  $158.46\mu\text{m}$ , and for  $v$  displacement, these values are  $-9.6$  and  $105.48$ , respectively. These values

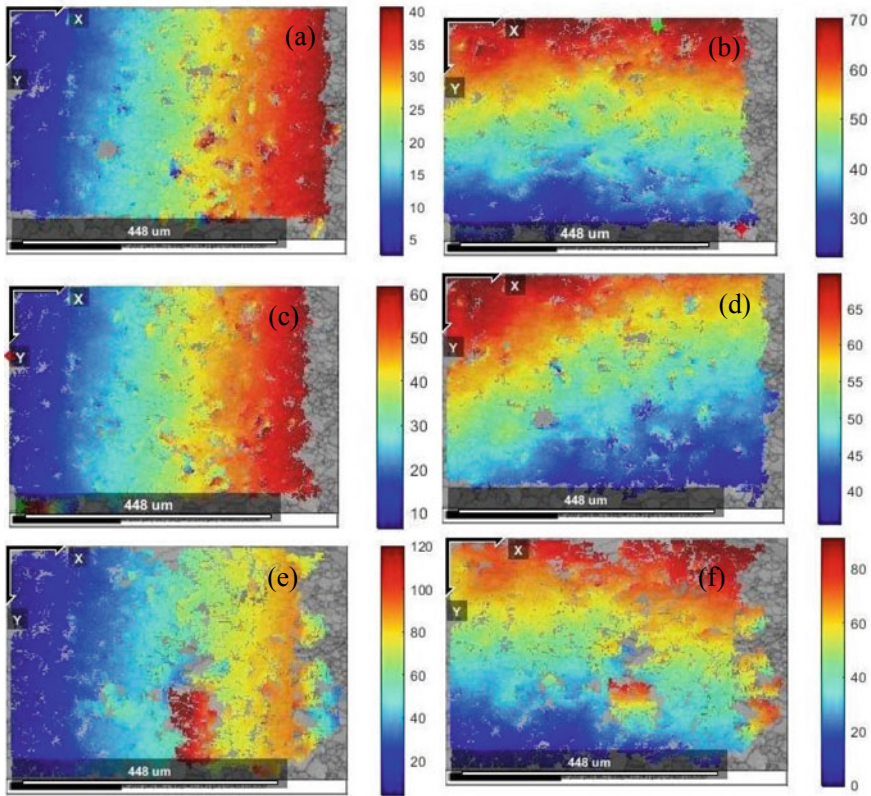


**Fig. 1** showing images of the same region under (a) undeformed, (b) 10% deformed, (c) 20% deformed, and (d) 40% deformed

are summarized in Table 1. It can, therefore, be inferred that with increasing deformation, the range of displacements, means, and deviations are increasing, which is consistent with the process of deformation. This analysis, therefore, shows that DIC is providing consistent results.

This analysis has also shown that the  $v$  displacement is more pronounced as compared to  $u$ -displacement which could be attributed to the orientation of boundary of regions comprising smaller grains. In the images Fig 2 a and f representing  $u$  and  $v$  displacements, it may be noticed that localization of large displacements visible in both images indicating initiation of strain concentration regions. However, before making comments, it is necessary to have a detailed statistical analysis to find out the correlation between  $u$  and  $v$  displacements.

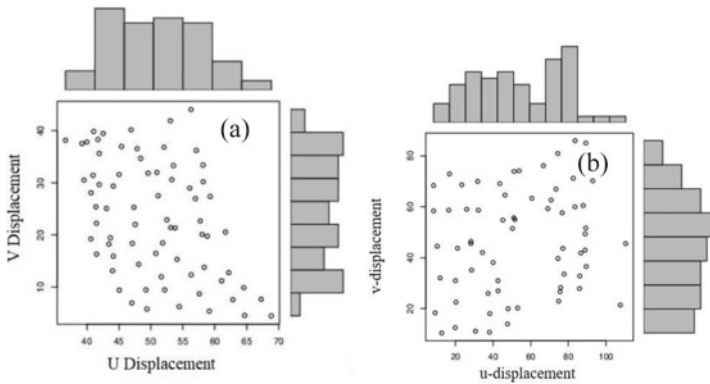
The scattered plots for 10 and 40% of the Pearson's function are shown in Fig. 3. By comparing Figs. 3 a and b, it can be seen that the relationship between the displacements is nonlinear in nature and with increasing deformation, the correlation between the two displacements becomes weaker. Equation 10 has used to compute Pearson's correlation coefficients which came out to be  $-0.44$  and  $0.182$  for the case of 10% deformation and 40% deformation, respectively. The value of  $-0.44$  suggests that if the value of one of the displacements increases, the value of the other reduces. The magnitude of  $0.44$  suggests that the strength of the relationship is medium in nature. Such correlation indicated that the displacements are not independent and



**Fig. 2** showing images of the u-displacements (a) 10%, (c) 20%, (e) 40% deformed and v displacements (b) 10%, (d) 20%, and (f) 40% deformed samples

**Table 1** Values of displacements under successive deformation

Displacement	Min (in $\mu\text{m}$ )	Median (in $\mu\text{m}$ )	Max (in $\mu\text{m}$ )	Mean (in $\mu\text{m}$ )	Standard deviation (S)	Standard error (Se)
U—Displacement (10% deformation)	-1.54	21.65	55.07	23.19	11.12	1.31
V—Displacement (10% deformation)	23.43	52.34	76.49	50.65	7.75	0.91
U—Displacement (40% deformation)	-4.63	46.29	158.47	53.33	28.17	3.47
V—Displacement (40% deformation)	-9.6	49.74	105.48	46.32	20.58	2.53

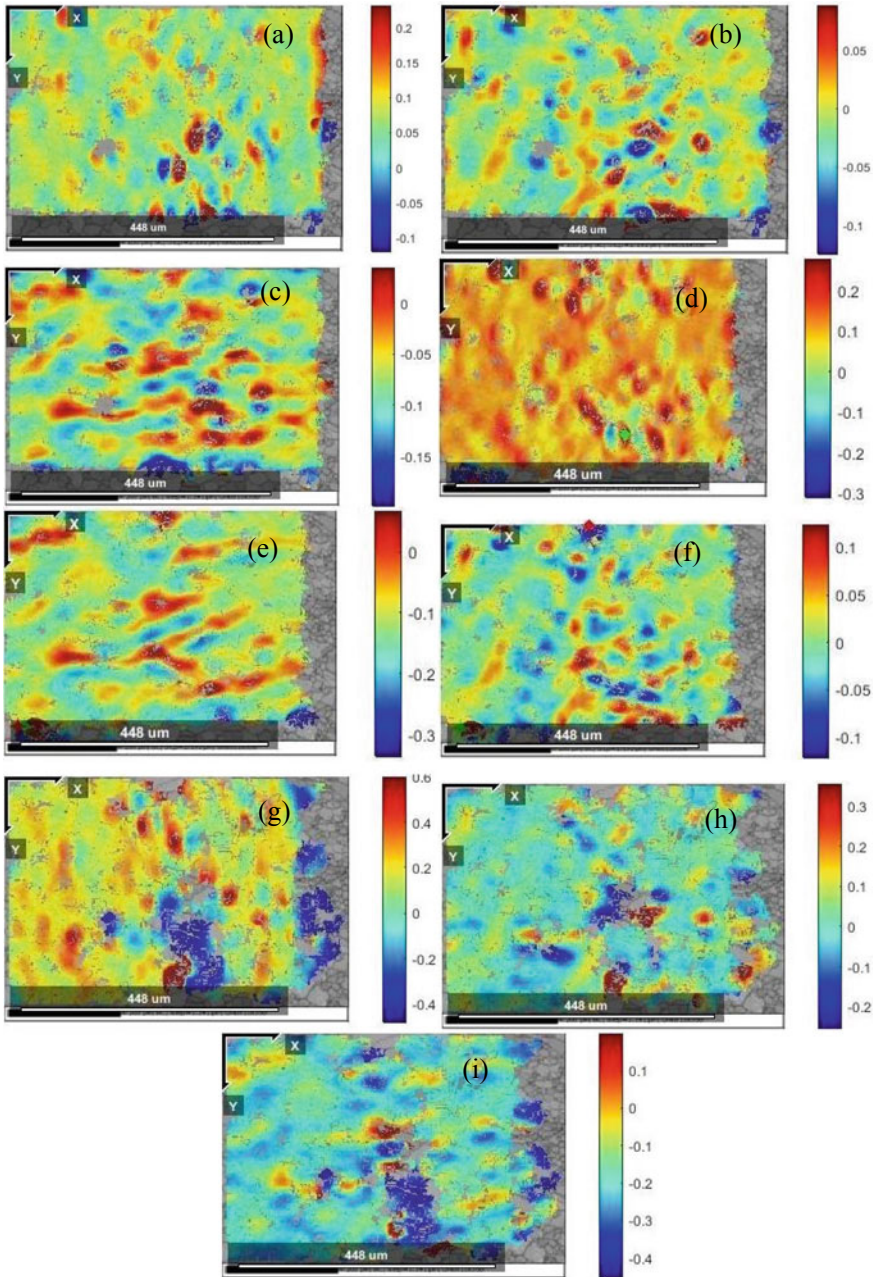


**Fig. 3** Pearson's function between  $u$  and  $v$  displacements (a) 10% and (b) 40% deformed samples

due to the constraints imposed by the boundary, an increase in one displacement is partially compensated by the reduction of the other. With increasing deformation, the correlation becomes relatively weak which indicates that the restrictions imposed by the boundaries have reduced considerably. However, detailed information could only be extracted by converting these displacements into strains.

Using Eqs. 5–7, all the three components of the strains ( $E_{xx}$ ,  $E_{yy}$ ,  $E_{xy}$ ) were estimated. Figure 4 shows the computed  $E_{xx}$ ,  $E_{xy}$ , and  $E_{yy}$  strains of samples deformed by 10, 20 and 40%, respectively. In addition, detailed statistical analysis of values of strains has also been carried out. The values of relevant parameters and their ranges are tabulated in Table 2. It may be noted that all the components of strains increase with increasing deformation. In all deformed samples,  $E_{xx}$  is mainly positive in nature (Fig. 4a, d and g; Table 2) means  $E_{xx}$  strain is mainly tensile in nature.  $E_{xy}$  remains nearly insignificant as in most of the region its value is close to zero (Fig. 4 b, e, and h). In contrast,  $E_{yy}$  is mainly negative in all the deformed samples (Fig. 4c, f, and i). Such behavior arises due to the uniaxial deformation generating elongation on the  $x$ -direction and as reactionary compressive stress is generated in the  $y$ -direction. It may also be noted that with increasing deformation, localization of strain started and regions with large positive strains are always having regions of large negative strains in their vicinity. Up to 10% deformation, the strain was mainly confined to the region of large grains, and the region of small grain remained more or less undeformed. With deformation reaching 40%, the presence of strain in the region of small grains could be noticed (Fig. 4g). Because of the transfer of strain into the regions of small grains, fluctuations in the strain observed in 20% deformation get reduced (compare Fig. 4 d and g). Detailed statistical analysis to find out the correlation between the components of strains has been carried out and presented in Fig. 5, and the value of the coefficient of Pearson correlation is given in Table 3.

As the bimodal distribution of the grain's constraints the distribution of the strains in the matrix, the elongation strains ( $E_{xx}$ ,  $E_{yy}$ ) show a weak but positive correlation with the shear strain ( $E_{xy}$ ) during early deformation (10%), which means that increase



**Fig. 4** Strain distribution in the deformed samples. (a)  $E_{xx}$ , (b)  $E_{xy}$ , (c)  $E_{yy}$  in 10% deformed sample, (d)  $E_{xx}$ , (e)  $E_{xy}$ , (f)  $E_{yy}$  in 20% deformed sample, and (g)  $E_{xx}$ , (h)  $E_{xy}$ , (i)  $E_{yy}$  in 40% deformed sample

**Table 2** Values of strain under successive deformation

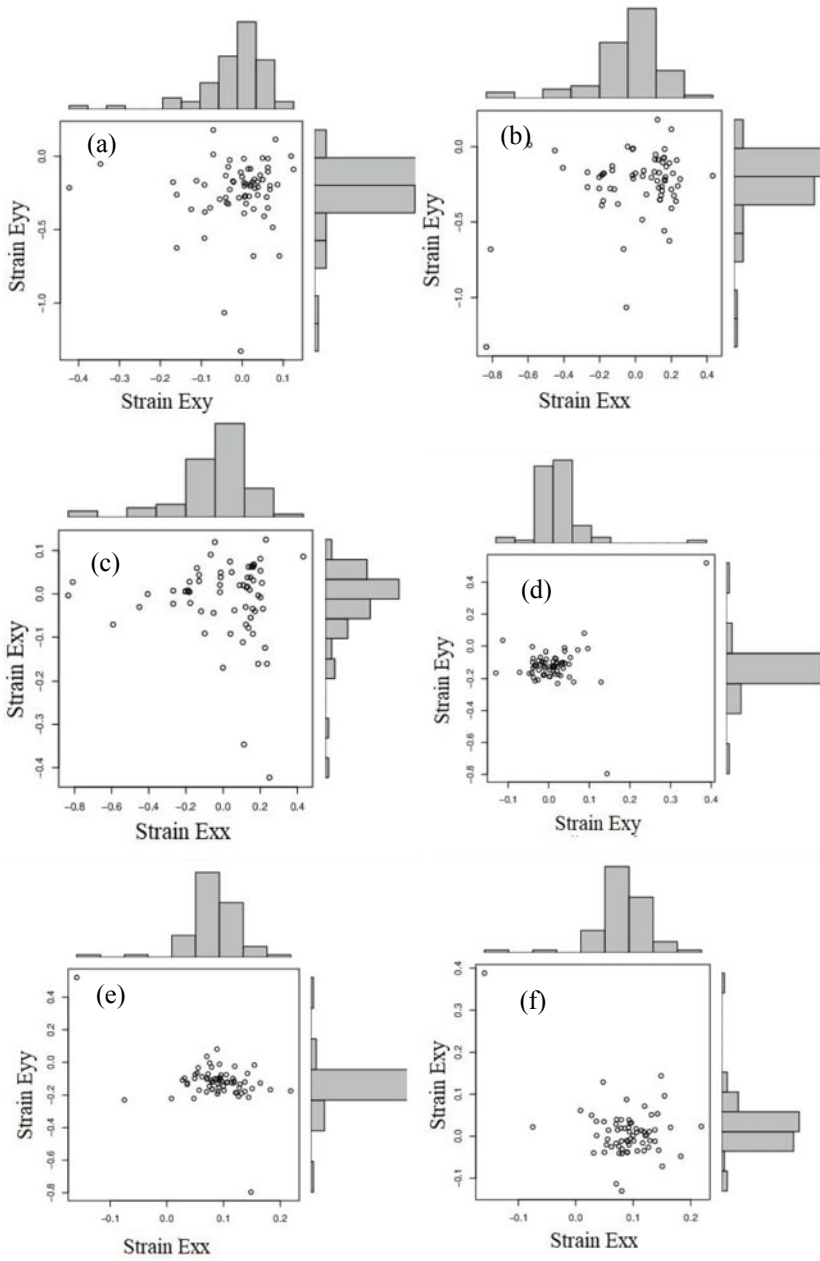
Deformation (%)	Strain	Min	Median	Max	Mean (N)	Standard deviation (S)	Standard error (Se)
10	$E_{xx}$	-0.466	0.07	0.97	0.06	0.05	0.00774
	$E_{xy}$	-0.346	-0.006	0.50	-0.01	0.02	0.00322
	$E_{yy}$	-0.448	-0.068	0.27	-0.08	0.11	0.00609
20	$E_{xx}$	-0.494	0.104	0.49	0.09	0.06	0.00668
	$E_{xy}$	-0.299	0.004	0.75	0.01	0.06	0.00756
	$E_{yy}$	-0.497	-0.112	2.72	-0.12	0.13	0.01611
40	$E_{xx}$	-0.50	0.1688	4.64	0.01	0.23	0.02941
	$E_{xy}$	-1.675	0.019	1.59	-0.01	0.09	0.01143
	$E_{yy}$	-0.496	-0.187	1.5	-0.24	0.23	0.02921

in one helps in increasing the magnitude of the other. Due to restrictions imposed by the small grain regions, the elongation in the region of large grain is associated with the shear deformation ( $E_{xy}$ ). This conjecture is supported by the fact that the correlation between the elongation strains ( $E_{xx}, E_{yy}$ ) is weakly negative, which means an increase in the magnitude of one reduces the others. In another word, any tensile strain generated in one direction promotes compressive strain in another which is the manifestation of the rigid body constrained condition.

With increasing strain (20% deformation), correlation between the elongation strains ( $E_{xx}, E_{yy}$ ) becomes more negative suggesting relatively more severe constraints imposed by the system. In addition, the correlation between the  $E_{xx}$  and  $E_{xy}$  also becomes negative. However, for elongation along with perpendicular directions  $E_{yy}, E_{xy}$  shows a medium correlation. It can be inferred that the rigid body restriction imposed by small grains confined most of the deformation in the y-direction. Between 20 and 40% deformation, the accumulated strain in the regions of large grains induces deformation in the regions of small grains. This helps in increasing the deformation. As a consequence of this induced deformation in the region of small grains, the correlation between the shear strain ( $E_{xy}$ ) and elongation strains ( $E_{xx}, E_{yy}$ ) becomes weak, but a moderate correlation between the elongation strains is established.

## 4 Conclusion

In the present work, a single-phase Fe–Ni–W alloy was studied for microstructural changes imparted by deformation. Images under SEM were recorded when the sample was deformed from 5–50%. Bimodal distribution of grains introduced heterogeneity in the displacement and strains as shown by the DIC of images obtained before and after deformation. Detailed Pearson correlation analysis showed that the displacements,  $u$  and  $v$ , determined using first-order shape functions, depending



**Fig. 5** Pearson correlation scattered graphs (a–c)  $E_{yy}$  for 10% deformed sample, (d–f) for 20% deformed sample, and (g–i) for 40% deformed sample

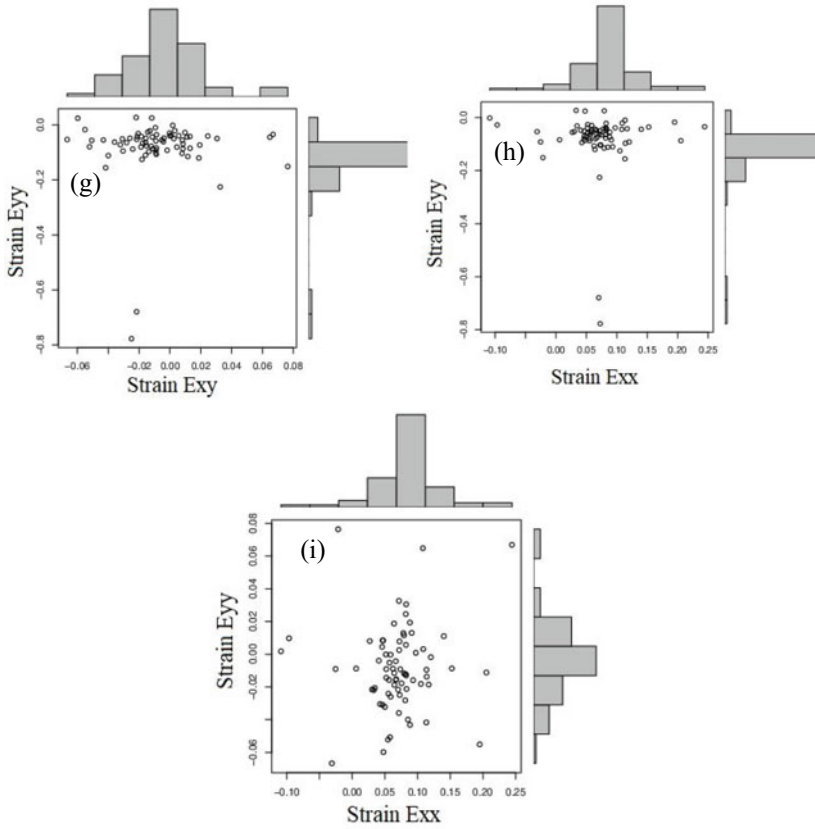


Fig. 5 (continued)

Table 3 Pearson correlation coefficient of strains

	$E_{xx} - E_{xy}$	$E_{xx} - E_{yy}$	$E_{yy} - E_{xy}$
10%	0.07	-0.01	0.029
20%	-0.41	-0.44	0.29
40%	-0.078	0.326	0.048

upon the amount of deformation showed medium to the weak correlation between them. Similar correlations among various components of strains were changing with the deformation. These changes were attributed to the heterogeneous distribution of grains which led to forming two separate regions, one containing large grains and the other with smaller grains. It has been shown that till the deformation is confined to the region of large grains, the correlation between shear deformation and elongation strains were existing, but as the deformation initiated in the regions of small grains also, the correlation between the two became weak.



## References

1. Vanniamparambil PA, Bartoli I, Hazeli K, Cuadra J, Schwartz E, Saralaya R, Kontsos A (2012) An integrated structured health monitoring approach for crack growth monitoring. *J Intell Mater Syst Struct* 23(14):1563–1573. <https://doi.org/10.1177/1045389X12447987>
2. Peters WH, Ranson WF (1981) Digital imaging techniques in experimental stress analysis. *Opt Eng* 21:427–431
3. Chu TC, Ranson WF, Sutton MA (1985) Applications of digital-image-correlation techniques to experimental mechanics. *Exp Mech* 25:232–244
4. Dickinson AS, Taylor AC, Ozturk H, Browne M (2011) Experimental validation of a finite element model of the proximal femur using digital image correlation and a composite bone model. *J Biomech Eng* 133:014504
5. Franck C, Hong S, Maskarinec SA, Tirrell DA, Ravichandran G (2007) Three-dimensional full-field measurements of large deformations in soft materials using confocal microscopy and digital volume correlation. *Exp Mech* 47:427–438
6. Rehr C, Kleber S, Antretter T, Pippan R (2011) A methodology to study crystal plasticity inside a compression test sample based on image correlation and EBSD. *Mater Charact* 62:793–800
7. Bernasconi A, Carboni M, Ribani R (2020) On the combined use of digital image correlation and micro computed tomography to measure fibre orientation in short fibre reinforced polymers. *Compos Sci Technol* 195:108182
8. Zdunek J, Brynk T, Mizera J, Pakieła Z, Kurzydłowski KJ (2008) Digital image correlation investigation of Portevin–Le Chatelier effect in an aluminium alloy. *Mater Charact* 59:1429–1433
9. Tracy J, Waas A, Daly S (2015) Experimental assessment of toughness in ceramic matrix composites using the J-integral with digital image correlation part I: methodology and validation. *J Mater Sci* 50:4646–4658
10. Elliot DA (Ed) (1980) Applications of digital image processing to astronomy. *Int Soc Opt Photon* 264
11. Gonzalez RC, Wood RE (2002) In: Digital image processing. Pearson Educations Pvt. Ltd.
12. Lee S-H, Lee M, Kim H-J (2014) Anatomy-based image processing analysis of the running pattern of the perioral artery for minimally invasive surgery. *Br J Oral Maxillofac Surg* 52(8):688–692
13. Kothari CR, Garg G (2015) In: Research methodology. 3rd edn, New Age. pp 136–140
14. Fan GJ, Choo H, Liaw PK, Lavernia EJ (2006) Plastic deformation and fracture of ultrafine-grained Al–Mg alloys with a bimodal grain size distribution. *Acta Mater* 54:1759–1766

# Design and Construction of Analogue Signal Processing System for Distributed Fibre Optic Sensors



Kadambari Sharma , J. M. Nair, and P. P. Vaidya

## 1 Introduction

Technological advances in optoelectronics along with the development of optical fibres with minimum attenuation losses and high sensitivity to variations in optical properties of light propagating through it led to the emergence of fibre optic sensors (FOSs). FOS can be used to measure large number of physical parameters like temperature, pressure, vibrations, humidity, chemical analysis, displacement and magnetic fields. The optical fibre acts as a sensor to one such physical parameter, and any change in its value causes variation in the properties of light propagating through it. The deviations of light are measured by converting it into an electrical signal and further processed for retrieving the information of the physical parameter. FOS is subject of interest as they offer various advantages over conventional sensors like high immunity to electromagnetic interference, easy implementation in hazardous areas and lightweight.

The most important building blocks of the FOS system are the light source (LASER, LASER diode), an optical fibre, a photodetector (photodiode, photo-multipliers etc.) and an excellent signal processing system (Fig. 1). FOS system involves various operating principles like modulation of intensity, wavelength, phase, polarisation and incorporating periodic variation in refractive index technique for the

---

Supported by organisation VESIT.

---

K. Sharma (✉) · J. M. Nair · P. P. Vaidya  
VES Institute of Technology, Affiliated to University of Mumbai, Mumbai, India  
e-mail: [kadambari.sharma@ves.ac.in](mailto:kadambari.sharma@ves.ac.in)

J. M. Nair  
e-mail: [principal.vesit@ves.ac.in](mailto:principal.vesit@ves.ac.in)

P. P. Vaidya  
e-mail: [pp.vaidya@ves.ac.in](mailto:pp.vaidya@ves.ac.in)

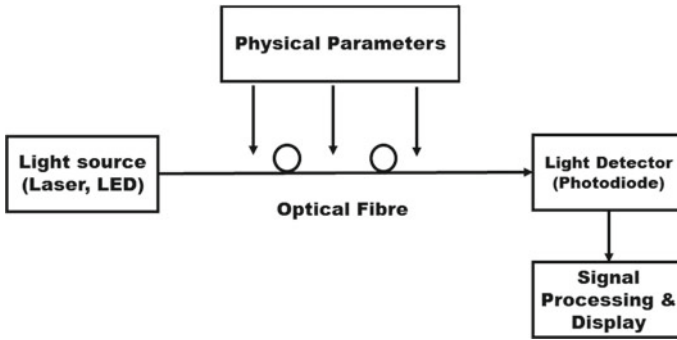


Fig. 1 Block diagram of fibre optic sensor (FOS) system

measurement of various physical parameters. One such principle employed for measurement of temperature, vibrations and strain is based on sensing the backscattered light from the fibre. The sensor is known as distributed fibre optic sensors (DFOSs), and the technique is termed as distributed fibre optic sensing (DFS). In this technique, the sensing can be done over the entire fibre length; parameter change that occurs at any point of the fibre can be sensed and detected. It provides a continuous measurement over the entire length.

DFOS is based on optical time domain reflectometer (OTDR) [1, 2] (Fig. 2) in which the short- and high-peak power optical pulses are launched into the fibre of known length. The propagating light loses its power due to the various attenuation factors like absorption, bending, splices, connectors and scattering. The fraction of the scattered light that falls within the angle of acceptance of the fibre in the reverse direction is recaptured and guided back towards the launching end, called the backscattered light. The backscattered light is a weak signal as it loses power through attenuation whilst returning. At the launching end, the backscattered light is focussed on the receiver, which is further electronically processed, digitised and displayed. [2, 3]. The amplitude of the backscattered signal is very weak and depends on the optical

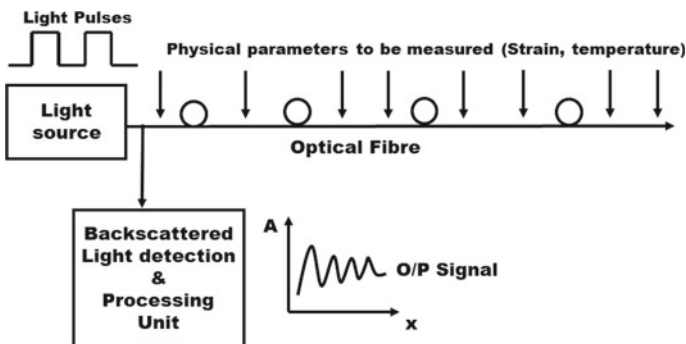


Fig. 2 General block diagram of distributed fibre optic sensor (DFOS) system

power of the launched pulse. The pulse peak power cannot be increased beyond a threshold as high power introduces undesired non-linearities in the fibre that results in erroneous output. The backscattered signal after experiencing attenuation reaches the receiver with an optical power which is much less than the launching power.

One of the important parameters of the DFOS system is the spatial resolution which is defined as the smallest length of the fibre over which the parameter variation can be sensed. This spatial resolution depends on the pulse width of the light pulses. High spatial resolution requires narrow pulse width. Thus, the typical value of the light pulse width is around 10 ns (ns) which gives a spatial resolution of 1 m. Special computer-controlled light sources are designed for obtaining such narrow light pulses [4].

The accuracy of the DFOS system depends on the performance of its detection circuit. The detector circuit should be such that it converts the whole of the backscattered weak optical power pulse into electrical signal without any dispersion.

This paper describes the design of optical signal processing system that offers adequate bandwidth for the detection of such high-speed and low-amplitude light pulses.

## 2 Optical Signal Processing System

The captured weak backscattered light is focussed on the photodetectors which are either photodiodes, photon counters or photomultiplier tubes. Many experiments favour photomultiplier tubes as they offer large surface area and large gain. Avalanche photodiodes are also employed as they offer high sensitivity. However, with recent technological advancements in the field of photodiodes, it is possible to detect high-speed, weak light signals with positive intrinsic negative (PIN) photodiodes which are highly cost-effective. This paper describes signal processing using the PIN photodiode. However, the same system with slight modifications in the component values can be utilised to detect signal from any type of photodetectors.

A photodetector in its proper operating mode along with an appropriate and sensitive signal processing circuit constitutes an optical signal processing (OSP) system.

The overall OSP system performance is governed by the proper selection, design and construction of the photodetector and its associated circuitry. Thus, selection of an appropriate detector is a very important step in the design of an OSP system. The important criteria to be considered in its selection are high sensitivity at the operating wavelength of the light source, high quantum efficiency, linear response to light, wide bandwidth, fast response time, less-noise contribution and stability in operation [5]. The photodiode capacitance should be as small as possible to reduce the charging time constant causing an increase in the bandwidth. Small active area and high shunt resistance of the photodiode aids in the signal-to-noise ratio improvisation [6].

The photodiode (PD) generates photocurrent proportional to the incident light. In order to attain the fastest response and widest of the bandwidth, the PD should

be operated in the photoconductive mode. In this mode, an external reverse bias is applied to the PD.

The PD, in photoconductive mode, converts the optical power to corresponding current signal. The photodetectors require a front-end transimpedance amplifier (TIA) to convert the weak current signal to a low-magnitude voltage signal which can be further amplified using cascaded amplification stages to achieve the desired output [7].

## 2.1 Transimpedance Amplifier Design Criteria

The TIA is the foremost and critical stage of the OSP system. This stage converts input weak current to an amplified weak output voltage. It also helps in controlling the overall system performance in terms of bandwidth and noise figure [8].

The basic parameters that characterise the TIA (Fig. 3) are the transimpedance gain and the input impedance. The transimpedance gain of the amplifier is decided by the feedback resistor  $R$  which implies that the gain is given by the feedback impedance for sufficiently large values of the op-amp open loop gain.

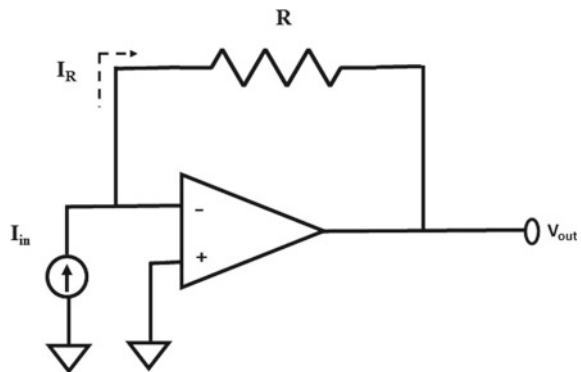
$$I_{in} = \frac{0 - V_{out}}{R} \quad (1)$$

$$I_{in} = \frac{-V_{out}}{R} \quad (2)$$

$$V_{out} = -RI_{in} \quad (3)$$

Designing a high-resolution detection circuit that incorporates photodiodes is a challenging mission as the factors like gain, bandwidth and input-referred noise that all needs to be considered together.

**Fig. 3** Basic transimpedance amplifier



The performance of the TIA is affected by various factors like voltage noise, current noise, input capacitance, bias current and finite bandwidth [9]. In general, field effect transistors-based (FETs) operational amplifiers (op-amps) offer high input impedance, low bias current that results in low DC error voltages and low noise due to reduced input current noise [10, 11].

The input bias current affects the TIA circuit by summing together with the weak PD current. Both the current flows through the feedback resistor to generate the output voltage as given by Eq. 4

$$V_{\text{out}} = -R(I_{\text{in}} + I_{B-}) \quad (4)$$

where  $I_B$ —is the op-amp bias current.

The magnitude of the input bias current limits the smallest value of the actual current signal to be detected accurately. This is one of the most important consideration in the DFOS systems as the amount of current generated by the sensitive detector will be very small in magnitude in the range of picoamperes to nanoamperes. This small value of current requires an op-amp having very low bias current in the range of femtoamperes. This is due to the fact that if the signal value is less than the op-amp bias current, then the signal would be left undetected as per Eq. 4. Thus, low-bias current amplifiers are required.

In DFOS system, the light is sent in the form of pulses having a narrow width of the order of few nanoseconds to few microseconds as per the spatial resolution requirements. As mentioned, the typical value employed is 10 ns. All these narrow light pulses should be processed by the OSP system without any degradation in their rise and fall times so as to obtain the accurate information. Such short-duration pulse detection requires a very high gain bandwidth amplifier, for example op-amps having a Gega Hertz bandwidth.

In addition to this, the amplitude of the backscattered light is very small in the range of pico- to nanoamperes. A single-stage TIA amplification would require a very high gain to obtain a satisfactory analogue output. This would require a high-value feedback resistor (example few hundreds of mega ohms).

This high-value resistor imposes a restriction on the desired wide bandwidth of the TIA due to which it may lead to unsatisfactory results. Also, the large feedback resistor increases its thermal noise, as the thermal noise voltage is given by Eq. 5

$$e_{\text{out}} = \sqrt{4KTBR} \quad (5)$$

where  $e_{\text{out}}$  is the noise voltage ( $V_{\text{rms}}$ ),  $K$  the Boltzmann's constant equal to  $1.38 \times 10^{-23}$  J,  $T$  the temperature in Kelvin,  $B$  the noise bandwidth in Hertz and  $R$  the feedback resistor in ohms [6].

The signal-to-noise ratio improves by a factor of  $\sqrt{R}$ . Thus, there is a trade-off between high gain, bandwidth and noise. Hence, there is a need to select an optimum TIA gain that provides the required bandwidth. The overall gain of the system can

then be further increased by utilising wide frequency range-cascaded amplification stages.

In order to achieve better voltage noise and high-gain bandwidth products, decompensated amplifiers are preferred [11]. Such op-amps are utilised for high-gain applications.

Thus, proper choice of the operational amplifier is very critical in the OSP system as it will optimise for the selected bandwidth, gain and signal-to-noise ratio. With several cascaded stages, it is desirable to select the maximum gain at the earlier stage that will minimise the noise contribution as it has the highest impact on signal-to-noise ratio. But, the transimpedance gain should be such that it allows the required optimum bandwidth. Thus, for the first stage that is the transimpedance stage, the best op-amp that offers all the desired qualities of high bandwidth, less current and voltage noise and best transimpedance gain possible for the needed frequency range should be selected.

Considering all the above important criteria, we have selected the op-amp LTC6268-10 for the OSP system designed for DFOS. It is FET input, decompensated op-amp (gain of 10 stable). It has the required low bias current, low input capacitance, low voltage noise and high-gain bandwidth. It can operate on a single supply of 5 V which reduces the power consumption.

## 2.2 Circuit Design

The design requirements of this OSP system are very high bandwidth (up to 400 MHz) and low supply voltage ( $\leq 5$  V). The input current to the receiver is very low (in the range of pA-nA); the final signal at the receiver output is designed to be few millivolts for this application.

The optical input signal to the PD is in the form of pulses having a duration not less than 5 ns, amplitude in the range of pA-nA, and the frequency of these pulses is decided depending on the required spatial resolution and on the fibre length.

The circuit was configured for the two different bandwidth requirements of 360 and 200 MHz, respectively. The photodiode employed was an OSI model FCI125G-006HRL. The important parameters of the photodiode are mentioned in Table 1

**Table 1** Photodiode parameters

Sr. no	Parameters	Values
1	Reverse voltage	5 V
2	Dark current at 5 V	30 pA
3	Diode capacitance at 5 V	650 fF
4	Shunt resistance	166 Gohms
5	Rise time	35 picoseconds

As the op-amp selected is LTC6268-10, the total input capacitance is 1.2pF; the GBWP of the IC is 4 GHz. The equations needed to obtain the values of the feedback resistor and the feedback capacitor for the TIA circuit are given by Eqs. 6 and 7 [12, 13]:

$$f_{-3db} = \sqrt{\frac{GBWP}{2\pi RC_{in}}} \tag{6}$$

$$C_F > \sqrt{\frac{C_{IN}}{\pi GBWP}} \tag{7}$$

where  $f_{-3db}$  is the desired bandwidth, GBWP is the gain bandwidth of the op-amp,  $R$  is the feedback resistor,  $C_{IN}$  is the total capacitance including the op-amp capacitance and diode capacitance,  $C_F$  is the feedback capacitor.

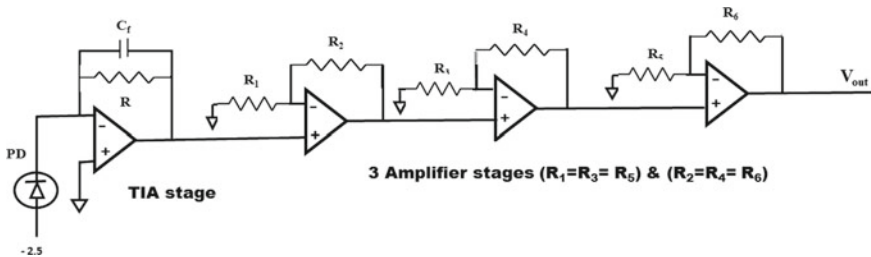
In high-speed circuits, the input stray capacitance ( $C_{IN}$ ) reduces the operating bandwidth of the op-amp and introduces oscillations in the circuit. To increase the stability of the op-amp, a feedback capacitor ( $C_F$ ) is added in the circuit.

For the two required frequencies of 360 and 200 MHz, the optimal values of the  $R$  and  $C_F$  for the TIA circuit were obtained as shown in Table 2:

Also, Eq. 6 indicates that the  $-3$  dB bandwidth of the TIA is inversely related to the feedback resistor. Therefore, if the bandwidth is the crucial parameter, then the best approach is to have a reasonable transimpedance gain stage followed by a broad-band voltage gain stage. Thus, the TIA stage was cascaded with three noninverting amplification stages each of gain 11 (Fig. 4).

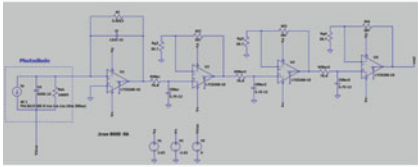
**Table 2** Value table

Frequency $f_{-3db}$	R ohms	$C_F$
360 MHz	3.65 k	131fF
200 MHz	12.1 k	126fF

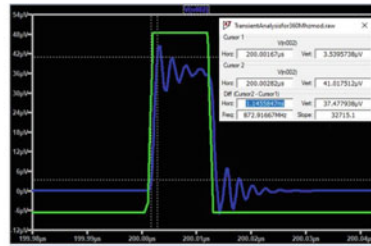


**Fig. 4** Circuit diagram for OSP system

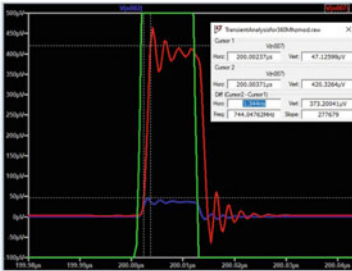




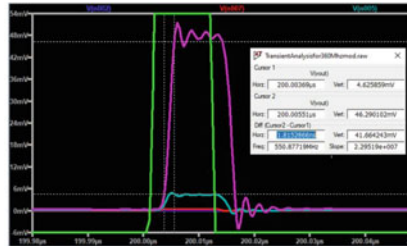
(a) Receiver circuit Diagram(TIA + 3 amplifiers)



(b) TIA output



(c) 2nd Amplifier output



(d) 3rd Amplifier output

Fig. 5 Simulation output for 360 MHz

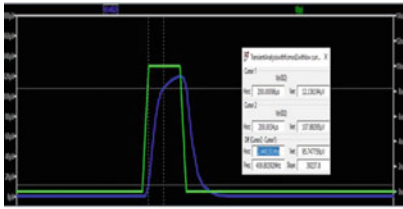
### 2.3 Circuit Simulation

The circuit was simulated before the actual conception of the hardware since the simulation helps in a better behavioural understanding of the circuit. The simulation was carried out using LTSpice, the circuit diagram, and the output waveforms obtained for the two bandwidth requirements are shown in Figs. 5 and 6.

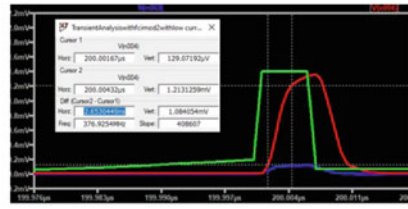
The rise time of the output pulses obtained for 360 MHz and that of 200 MHz is 1.77 ns and 2.44 ns, respectively, and the amplitude of the pulses is 100 millivolts. The rise time is of concern in this design as the input pulses are of very short duration in the range of 5 ns or more, so a rise time of less than 5 ns confirms that the short-duration pulses will never go undetected. Also, the small rise time confirms the wide bandwidth offered by the circuit which is the requirement of the DFOS.

### 2.4 Hardware Design

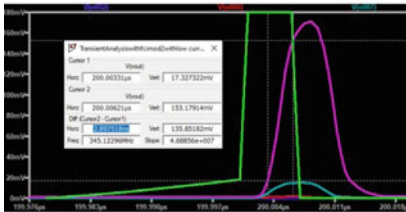
**Important Considerations for PCB Layout** As the DFOS detection circuit deals with small current values, the circuit layout needs to be prepared with great care. The three most important factors to be considered whilst designing the layout are the leakage currents, noise sources, and the stray capacitance. A good printed circuit



(a) TIA output



(b) 2nd Amplifier output



(c) 3rd Amplifier output

Fig. 6 Simulation output for 200 MHz

board design (PCB) should minimise these factors and obtain a balance between optimal performance and product manufacturability [14].

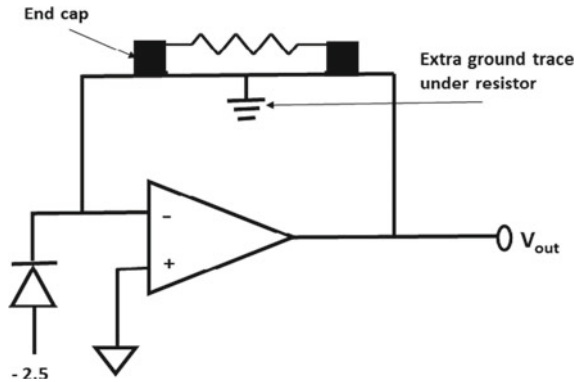
Proper grounding is very crucial in this board. The bottom layer of the double-sided PCB was made the ground plane. A ground plane acts as a common reference voltage, provides shielding and reduces stray inductance. The stray inductance is reduced by cancellation of the magnetic field between the conductor and the ground plane [15]. This plane had breaks in PCB to avoid voltage drops due to residual impedances. All these planes then were connected to a common ground point. This single point is termed as the star ground.

Noise effects arising due to the power supply were reduced by connecting the supply terminals of the op-amps with bypass capacitors. Typically, a 10 $\mu$ Farad tantalum capacitor in parallel with 0.1 $\mu$ Farad ceramic capacitor is connected between the supply line and the ground. This provides a low AC impedance path to ground at the power supply pins for a wide range of frequencies, thus, preventing unwanted noise from entering the op-amp. The capacitor ground was connected to the ground plane with the shortest trace length and was placed near to the amplifier load to minimise disturbances between the rails and ground [15]. As four op-amps were connected to the same power supply lines, the supply lines were made short and wide; also, each op-amp had its individual bypass capacitors.

Noise effects were further reduced by implementing surface-mount devices (SMD); track length was minimised; ground plane was around the perimeter of analogue section.

The active electronic component employed in the OSP system was the wide bandwidth op-amp LT6268-10. To preserve the stability of the op-amp without sacrificing

**Fig. 7** Ground trace to feedback resistor



on its bandwidth, certain techniques were implemented whilst preparing the board layout.

At high frequencies, the impedance of the feedback resistor is decreased by its own parasitic capacitance which limits the operating bandwidth. In order to increase the bandwidth range, the parasitic capacitance needs to be reduced to a very low value. This was achieved by providing a ground trace between the feedback resistor pads near the TIA output end (Fig. 7).

This ground trace protects the output field from interacting with the summing node end of the resistor and bypasses the field to ground. The ground trace extended under the resistor dielectric [9, 13]. This technique was used to obtain a wide operating bandwidth.

The critical nodes in the TIA circuit are the amplifier's inverting node and its feedback elements. The accuracy of the circuit is dependent on the extent of leakage current flowing through this node. Any unexpected sources like adjacent signal on the PCB both on the same layer or different layers, other components on the signal path, any type of board contamination and environmental factors may cause leakage current in the circuit [13]. Several factors were considered in order to avoid leakage currents from affecting the circuit.

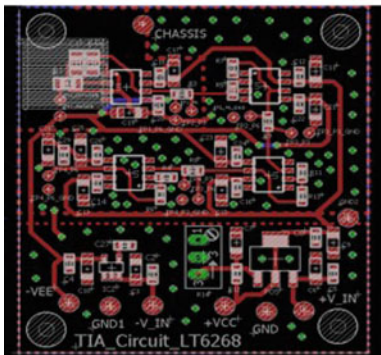
The SOIC package of LTC6268-10 is best suited for low input bias currents [13]. The high-impedance signal path on the PCB was made as short as possible.

In addition to this, guard ring was placed around the node. The concept of guarding is to surround the high impedance conductor with another conductor called guard that is driven to the same voltage potential [16]. Guards are important as they cancel the input leakage currents because the high impedance conductor and the guard are at the same potential. A guard ring is a structure used to implement the guarding technique on the surface of the PCB. It is a filled copper shape that completely surrounds the high impedance trace from the photodiode connection to the noninverting terminal. It also extends around the passive components in the feedback path to ensure that the entire high impedance node is surrounded by the guard [16].

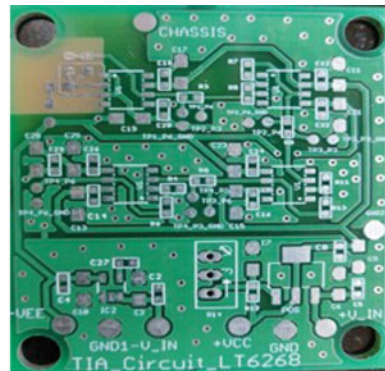
### 2.5 Hardware Results

The PCB layout was designed in Eagle, and the PCB was fabricated. The final board layout is as shown in Fig. 8a, the fabricated PCB is shown in Fig. 8b. The board after component assembly and for preliminary testing is as shown in Fig. 8c. The board was tested using the same photodiode and for the same frequency requirements, namely 360 and 200 MHz, and following results were obtained (Fig. 9).

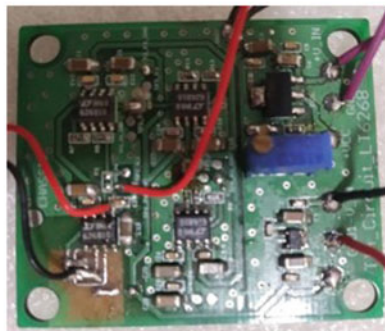
The hardware results also confirm the rise time of the pulses as 2.013 nanoseconds for 360 MHz requirement and 2.927 ns for 200 MHz requirement that are similar to the simulation results. Thus, the hardware is successfully able to meet the design criteria of the DFOS system.



(a) TIA Board Layout

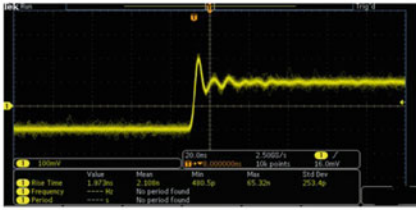


(b) TIA Board

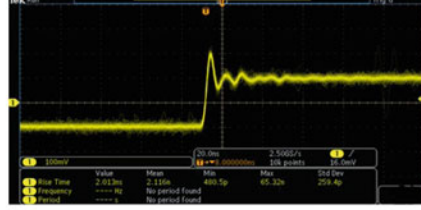


(c) Testing Board

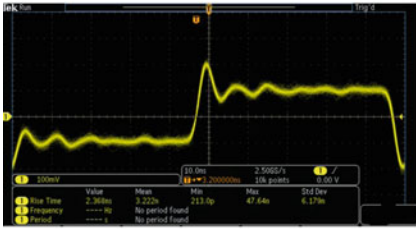
Fig. 8 PCB



(a) TIA output for 360MHz requirement



(b) Amplifier output for 360MHz requirement



(c) TIA output for 200MHz requirement



(d) Amplifier output for 200MHz requirement

Fig. 9 Hardware results

### 3 Conclusion

The paper describes the design considerations for optical signal processing for DFOS. The simulation and hardware results ascertain the design of OSP system utilising the IC LTC6268-10 meet the stringent requirements of measurement of low-intensity (carrying a current in the range of pA-nA) light pulses of duration as small as 5 ns. The OSP system developed is very economical as compared to the systems available in the market. Initial testing of the hardware was done by using the double-layer PCB, but the system performance can be further enhanced by implementing multilayer PCB and associated modifications in the circuit layout.

### References

- Hartog A (1983) A distributed temperature sensor based on liquid-core optical fibers. *J Lightwave Technol* 1.3:498–509
- Hartog AH (2018) In: An introduction to distributed optical fiber sensors. 2nd edn. CRC press
- Bao X, Chen L (2012) Recent progress in distributed fiber optic sensors. *Sensors* 12(7):8601–8639
- Kadambari S, Vaidya PP, Nair JM (2018) A research grade computer controlled light source. In: 2018 International conference on smart city and emerging technology (ICSCET). IEEE
- Hui R, Sullivan MO (2009) Fiber optic measurement techniques. Elsevier Inc

6. Instruments Texas. "Designing photodiode amplifier circuits with opa128." Application Note [Z]
7. Xue Y et al. (2019) Enhanced bandwidth, high gain, low noise transimpedance amplifier for asynchronous optical sampling systems. *Rev Scientific Instrum* 90.6: 063103
8. Bansal U, Gupta M (2018) High bandwidth transimpedance amplifier using FGMOS for low voltage operation. *Integration* 60:153–159
9. Glen B (2015) Op amp combines femtoamp bias current with 4GHz gain band-width product, shines new light on photonics applications. *LT J Analog Innov* 25.1
10. Caldwell J (2014) 1 mhz single-supply photodiode amplifier reference design. TI Application Note TIDU 535:1–19
11. Instruments, Texas (2009) Transimpedance considerations for high-speed amplifiers. Application Report No. SBOA122
12. Hashemi H (2017) Texas instruments, SNOA942A. Transimpedance amplifiers (TIA): choosing the best amplifier for the job. Nov 2015 [Rev 2017]
13. "LTC6268–10 Datasheet and Product Info—Analog Devices"
14. Grohe P (2012) Design femtoampere circuits with low leakage, part1. Texas Instruments
15. Ardizzoni J (2005) A practical guide to high-speed printed-circuit-board layout. *Analog Dialogue* 39(9):1–6
16. "AD4530–1 Datasheet and Product Info—Analog Devices"

# Design of Versatile Reconfigurable ADC for Wide Range of Resolution and Conversion Time



Jayamala Adsul, P. P. Vaidya, and J. M. Nair

## 1 Introduction

The different types of conventional ADCs (flash, pipeline, successive-approximation, subranging, delta-sigma, etc.) have fixed resolution and fixed conversion time, so these conventional ADCs won't be useful for extensive range of applications [1–4]. Also, researchers have designed different types of reconfigurable ADCs whose resolution and operating speed can be varied, but its reconfigurability factor is small, i.e. the range of application is narrow [5–10].

This reconfigurability factor can be improved by designing a single ADC [11] for a wide range of applications (high resolution, low speed, medium resolution, medium speed, and low resolution high speed). Figure 1 shows many high-speed applications (satellite communication, radar processing, video, IF sampling, software radio, etc.) in the range of high hundreds MHz where flash ADC is useful and high-resolution applications (data acquisition, measurement, instrumentation, etc.) in the range of 8–20 bit where subranging and interpolation ADCs are useful. Generally, in this range of resolution, successive-approximation (SAR), sigma-delta ADCs are useful. So to achieve conversion for extensive range of applications (high speed low resolution–high resolution low speed), a single ADC which is reconfigurable can be designed. The flash ADC is a basic block in the subranging ADCs. This flash ADC consists of a resistive string, comparators, and priority encoders, while subranging ADC consists

---

J. Adsul (✉)

Department of Electronics Engineering, VESIT, University of Mumbai, Mumbai, India  
e-mail: [jayamala.adsul@ves.ac.in](mailto:jayamala.adsul@ves.ac.in)

P. P. Vaidya · J. M. Nair

Department of Instrumentation Engineering, VESIT, University of Mumbai, Mumbai, India  
e-mail: [pp.vaidya@ves.ac.in](mailto:pp.vaidya@ves.ac.in)

J. M. Nair

e-mail: [principal@ves.ac.in](mailto:principal@ves.ac.in)

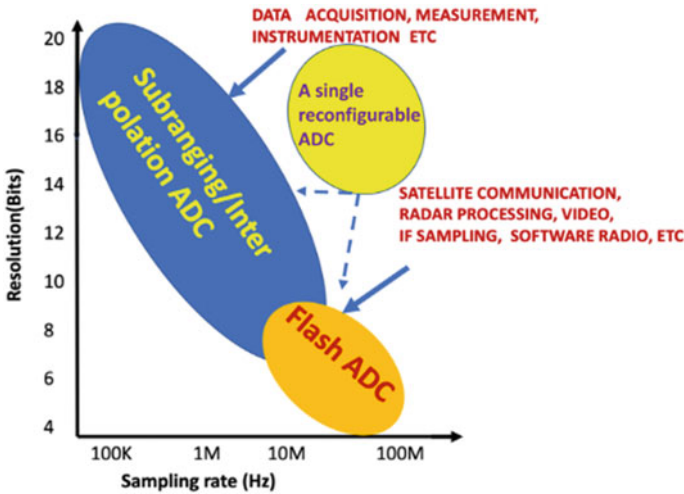


Fig. 1 Application of flash and sub-ranging ADC [1–4, 12, 13]

of flash ADCs, digital to analog converter (DACs), and residue amplifiers. The same building blocks are configured in different ways to achieve the operation of flash ADC, sub-ranging ADC, and interpolation ADC.

## 2 A New Proposed Design

In this paper a new ADC architecture has been introduced that makes use of the same hardware in a flash and sub-ranging ADC to capably adapt the operation at the architectural level, to obtain ADC with optimum resolution, and conversion time as required for the specific application. The novel hybrid flash/sub-ranging/interpolation ADC is designed and simulated using NI Multisim 14.1, and results are presented.

A new ADC design has been conceptualized, which is shown in Fig. 2. This would be useful for extensive range of applications as its resolution can be adjusted as 4, 6, 8, 10, 12 and 16-bit with different conversion times. A single reconfigurable ADC consists of basic components such as 16 comparators, 15 residue amplifiers (gain 2), a 16-bit DAC, priority encoders, and a resistive string of 16 resistors. These components could be arranged and interconnected to obtain the required resolution and conversion time. This would behave as 4-bit flash ADC, 6-bit sub-ranging ADC, 8-bit recirculating ADC, 10-bit sub-ranging ADC, 12-bit interpolation ADC, 16-bit sub-ranging ADC, and 20-bit interpolation ADC.

Generally, conventional reconfigurable designs are implemented using discrete components in hardware which are reconfigured manually to achieve the different



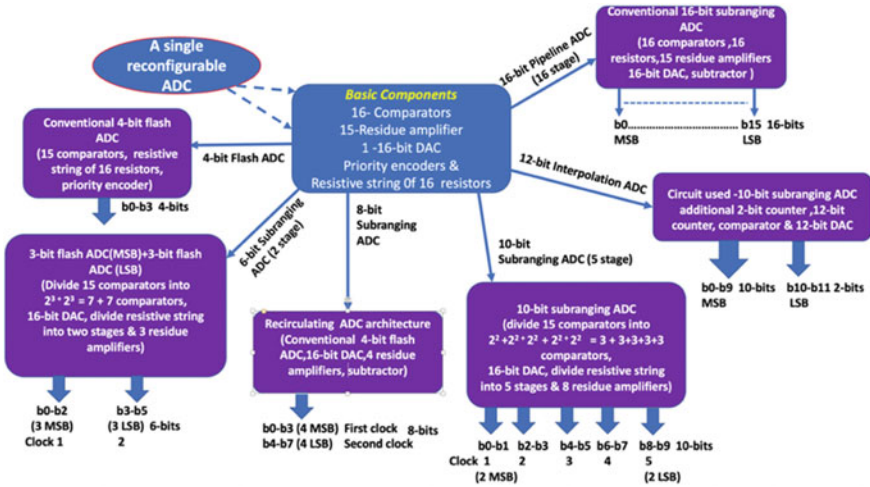


Fig. 2 Reconfiguration of ADC

configurations of ADC. But in this proposed design, interconnection is done automatically through programmable devices to achieve the different configurations of ADC, as shown in Fig. 2.

### 2.1 Reconfiguration for 4-bit Flash ADC

This is an architectural reconfiguration of ADC. For 4-bit flash, ADC as shown in Fig. 3, the components requirement generally are  $2^4 - 1$  comparators (15), 16 resistors, and priority encoder. So these components are interconnected, as shown in Fig. 3, which converts the input signal with a 4-bit resolution and lower conversion time. The resistive string is formed with 16 resistors to generate reference voltage (ref1 to ref15) for comparators (C1–C15). The output of each comparator (LSB1 to LSB15) is given as an input to the priority encoder to obtain 4-bit (b0 to b3) digital signal. So this is like a conventional flash ADC which provides low-resolution and high-frequency conversion. The conversion time for this 4-bit resolution is T which is nearly equal to comparator response time + priority encoder propagation delay.

### 2.2 Reconfiguration for 6-bit Subranging ADC

For 6-bit subranging ADC, the basic components mentioned in Fig. 2 are interconnected as two 3-bit flash ADC as shown in Fig. 4. It consists of two 3-bit flash ADC,

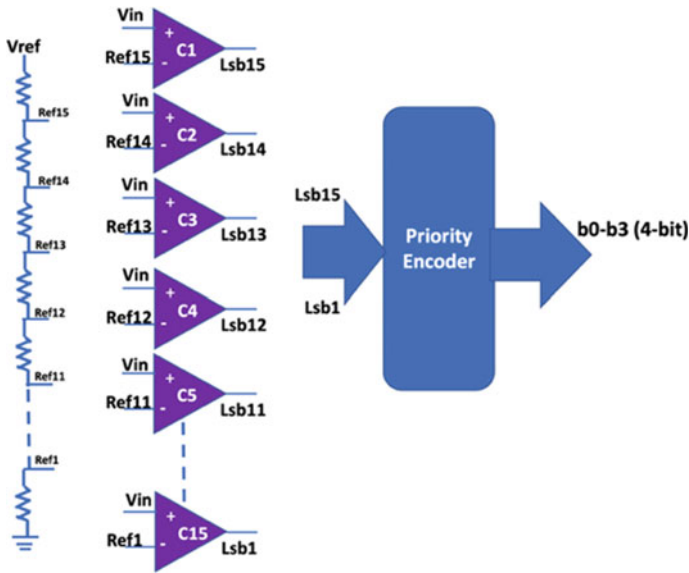


Fig. 3 4-bit flash ADC

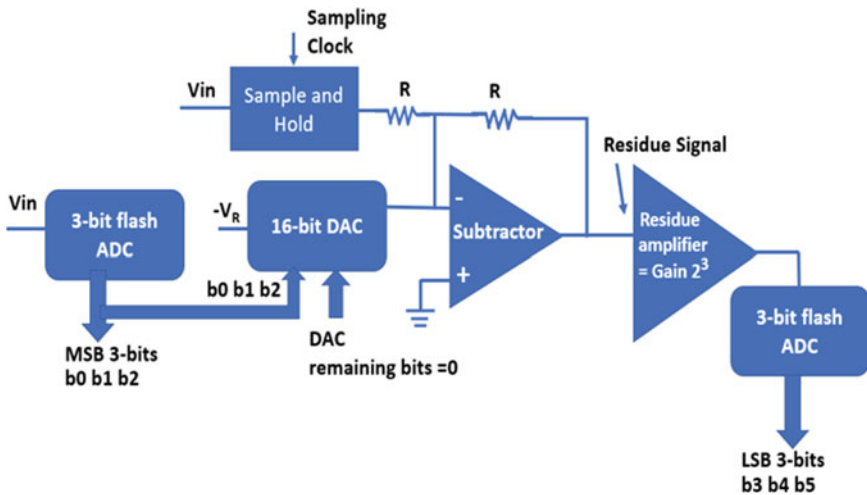


Fig. 4 Two stage 6-bit subranging ADC

subtractor, residue amplifier, and a 16-bit DAC. So each 3-bit flash consists of a resistive string of eight resistors,  $2^3 - 1$  comparators (7), and a priority encoder.

The resistive string is divided into two strings to generate reference voltage for each 3-bit flash ADC comparator. In the first clock,  $V_{in}$  input signal is converted by the first 3-bit flash ADC to obtain 3 MSB bits ( $b_0$  to  $b_2$ ), and this digital output is

given as an input to a 16-bit DAC (MSB bits of DAC), while the remaining lower bits of DAC are kept at 0 logic. This DAC output is subtracted from  $V_{in}$  input signal to extract the remainder (residue) signal. This remainder signal is amplified by a gain of 8. This gain of 8 is achieved by connecting three residue amplifiers (gain 2) in a cascade. The residue has been amplified to obtain the residue signal in the full scale range. Then, in the second clock, the remaining three LSB bits ( $b_3$ – $b_5$ ) are obtained by converting the amplified residue signal by the second 3-bit flash ADC. So the conversion is complete for the current sample, and again, this 6-bit subranging ADC is ready to accept the new sample for conversion. The conversion time for this 6-bit resolution is  $2 T + 16\text{-bit DAC settling time} + 4\text{opamps setting time}$ .

### 2.3 Reconfiguration for 8-bit Recirculating Type ADC

For 8-bit recirculating type ADC, the basic components are interconnected as 4-bit flash ADC followed by a 16-bit DAC and residue amplifiers, as shown in Fig. 5. This 4-bit flash ADC internal circuit will be the same as the circuit shown in Fig. 3 with additional components, subtractor, latch and sample, and hold amplifier. In the first clock,  $V_{in}$  input signal is converted by 4-bit flash ADC to obtain four MSB bits ( $b_0$ – $b_3$ ), and it is latched. This digital output is given as an input to a 16-bit DAC (MSB bits of DAC) while the remaining lower bits of DAC are kept at 0 logic. This DAC output is subtracted from  $V_{in}$  input signal to extract the remainder signal. This remainder signal is amplified by a gain of 16. This gain of 16 is achieved by connecting four residue amplifiers (gain 2) in a cascade.

The residue has been amplified to obtain the residue signal in the full scale range. Then, in the second clock, the remaining four LSB bits ( $b_4$ – $b_7$ ) are obtained by converting the amplified residue signal by the same 4-bit flash ADC. Then, the conversion is complete for the current sample, and again, this 8-bit recirculating type

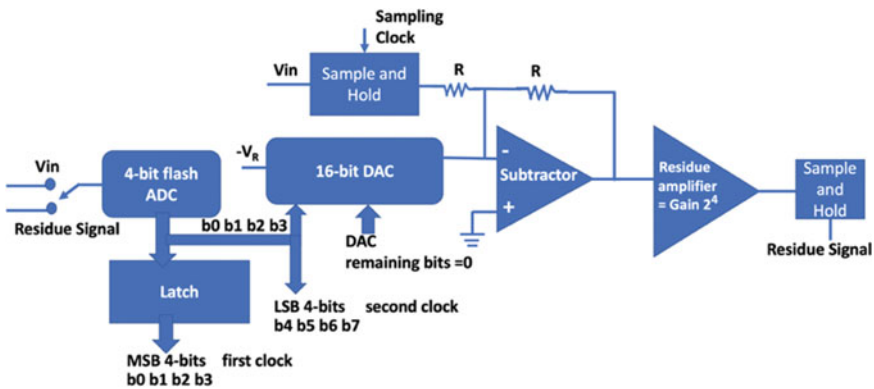


Fig. 5 8-bit recirculating ADC

ADC is ready to accept the new sample for conversion. The conversion time for this 8-bit resolution is  $2 T + 16\text{-bit DAC settling time} + 50\text{pamps setting time}$ .

### 2.4 Reconfiguration for 10-bit Subranging ADC

For 10-bit subranging ADC, the basic components mentioned in Fig. 2 are interconnected as five 2-bit flash ADC as shown in Fig. 6. It consists of five 2-bit flash ADC, subtractor, residue amplifier, and a 16-bit DAC. So each 2-bit flash consists of a resistive string of four resistors,  $2^2 - 1$  comparators (3), and a priority encoder.

The resistive string is divided into five strings to generate reference voltage for each 2-bit flash ADC comparator. In the first clock,  $V_{in}$  input signal is converted by the first 2-bit flash ADC to obtain two MSB bits ( $b_0$ – $b_1$ ), and this digital output is given as an input to a 16-bit DAC (MSB bits of DAC) while remaining lower bits of DAC are kept at 0 logic. This DAC output is subtracted from  $V_{in}$  input signal to extract the remainder signal. This remainder signal is amplified by a gain of 4. This gain of 4 is achieved by connecting two residue amplifiers (gain 2) in a cascade. The residue has been amplified to obtain the residue signal 1 in the full scale range. Then, in the second clock, the next two bits ( $b_2$ – $b_3$ ) are obtained by converting the amplified residue signal 1 by second 2-bit flash ADC, and this digital output is given as an input to a 16-bit DAC (next 2bits of DAC) while remaining lower bits of DAC are kept at 0 logic. This DAC output is subtracted from  $V_{in}$  input signal to extract the remainder signal. This remainder signal is amplified by a gain of 16. This gain of 16 is achieved by connecting four residue amplifiers (gain 2) in a cascade. The

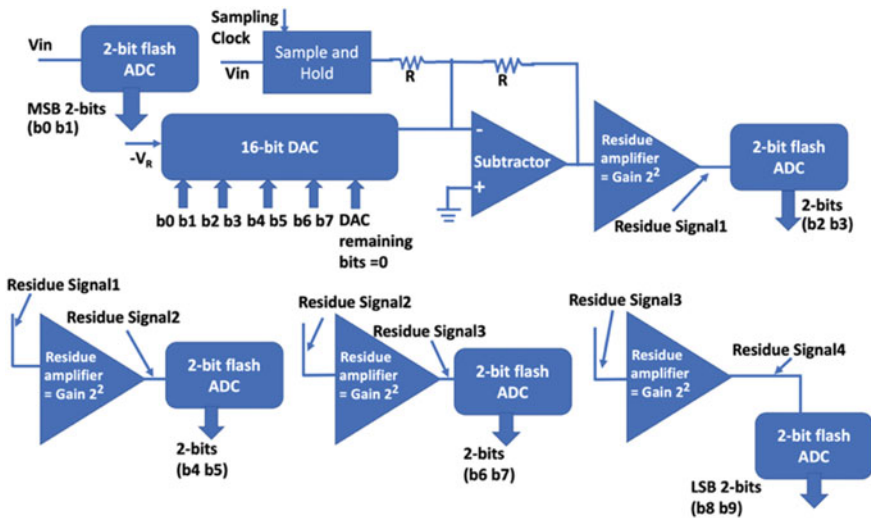


Fig. 6 10-bit subranging ADC

residue has been amplified to obtain the residue signal 2 in the full scale range. Then, in the third clock, the next two bits (b4–b5) are obtained by converting the amplified residue signal 2 by the third 2-bit flash ADC, and this digital output is given as an input to a 16-bit DAC (next 2 bits of DAC) while remaining lower bits of DAC are kept at 0 logic. This DAC output is subtracted from  $V_{in}$  input signal to extract the remainder signal. This remainder signal is amplified by a gain of 64. This gain of 64 is achieved by connecting six residue amplifiers (gain 2) in a cascade. The residue has been amplified to obtain the residue signal 3 in the full scale range.

Then, in the fourth clock, the next two bits (b6–b7) are obtained by converting the amplified residue signal 3 by the fourth 2-bit flash ADC, and this digital output is given as an input to 16-bit DAC (next 2 bits of DAC) while remaining lower bits of DAC are kept at 0 logic. This DAC output is subtracted from  $V_{in}$  input signal to extract the remainder signal. This remainder signal is amplified by a gain of 256. This gain of 256 is achieved by connecting eight residue amplifiers (gain 2) in a cascade. The residue has been amplified to obtain the residue signal 4 in the full scale range. Then, in the fifth clock, the LSB 2 bits (b8–b9) are obtained by converting the amplified residue signal 4 by the fifth 2-bit flash ADC. The conversion is complete for the current sample, and again, this 10-bit subranging ADC is ready to accept the new sample for conversion. The conversion time for this 10-bit resolution is  $5 T_{16\text{-bit DAC settling time}} + 9\text{opamps setting time}$ .

### 2.5 Reconfiguration for 12-bit Interpolation ADC

For 12-bit interpolation ADC, the subranging ADC, which has been designed for 10-bit resolution, is used to convert the signal with a 12-bit resolution by adding circuits such as 12-bit DAC, analog adder, latches, counter, and a digital comparator. Figure 7 shows the block diagram of extension of 10-bit resolution into 12-bit.

The 12-bit DAC output is incremented by LSB equivalent to 12-bit, i.e. 2.44 mV, and is added with an analog input (to be converted into 12-bit digital output). Then the output of the adder is converted into 10-bit, which are MSBs of 12-bit (b0–b9), while

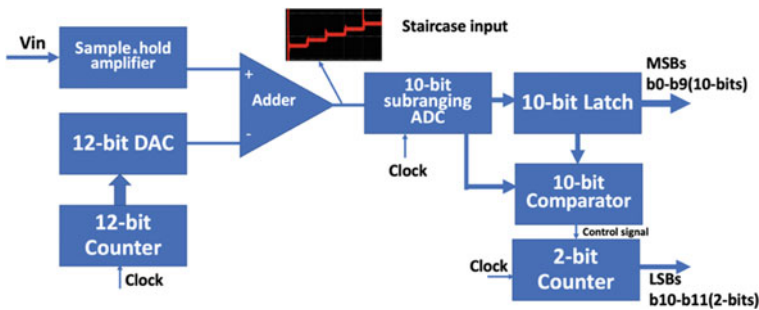


Fig. 7 12-bit interpolation ADC

the 10-bit data, which was initially obtained with only analog input, was latched and compared with the next samples and the counter which was initially loaded with 11 data, it starts counting down. If the next sampled 10-bits are greater than the latched 10-bits, then the counter stops counting, and the last count is the last 2LSB bits of 12-bit ADC. Here, MSB 10-bits are  $b_0$ – $b_9$ , and LSB 2-bits are counter output ( $b_{10}$ – $b_{11}$ ). This concept can be extended to more number of bits at the cost of speed of operation. The conversion time for this 12-bit resolution is  $5T + 16\text{-bit DAC settling time} + 9\text{opamps setting time} + 12\text{-bit DAC settling time} + \text{propagation delay of digital comparators and 2-bit counter}$ .

### 2.6 Reconfiguration for 16-bit Subranging ADC

For 16-bit subranging ADC, the basic components mentioned in Fig. 2 could be interconnected as 16 1-bit flash ADC as shown in Fig. 8. It consists of 16 1-bit flash ADC (1 comparator), subtractor, residue amplifier, and a 16-bit DAC. So each 1-bit flash consists of  $2^1 - 1$  comparators(1). The resistive string is made of two resistors to generate reference voltage for each 1-bit flash ADC comparator. The working principle of this 16-bit ADC is the same as 10-bit ADC.

Here, in every clock, 1-bit is obtained, so it takes 16 clocks to obtain the 16-bit digital output. This ADC does not require any priority encoder; the output of the comparator itself is the digital output. The conversion time for this 16-bit resolution is  $16\text{ comparators settling time} + 16\text{-bit DAC settling time} + 17\text{opamps setting time}$ . For 20-bit interpolation ADC, the subranging ADC, which has been designed for 16-bit resolution, is used to convert the signal with the 20-bit resolution by adding circuits such as 20-bit DAC, analog adder, 16-bit latch, 20-bit counter, 16-bit digital

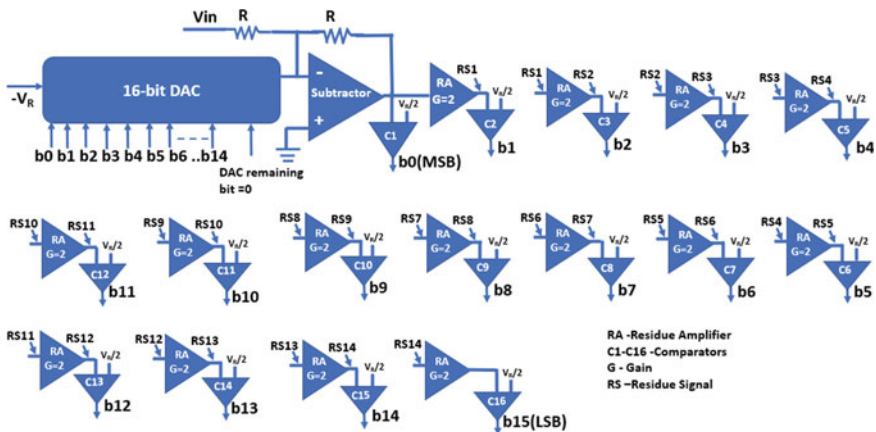


Fig. 8 16-bit subranging ADC

comparator, and 4-bit counter. The block diagram for 20-bit interpolation ADC would be similar to 12-bit interpolation ADC, as shown in Fig. 7.

### 3 Design of the Switches to Configure Resistive String, Input Signal, and Residue Signals

The reconfigurability of ADC is achieved automatically by configuring the interconnection among the components through different switches. A reconfigurable ADC has been designed and simulated using NI Multisim 14.1 for giving a resolution of 4, 6, 10 and 12-bit. So, in this section, configuration of switches for a resistive string and an input signal for comparators (4, 6, and 10-bit resolution) is described.

#### 3.1 Arrangement of Basic Components

Figure 9 shows the arrangement of 15 comparators and a resistive string of 16 resistors (additional few resistors for reconfigurability) in five stages. Each stage (1–5) includes three comparators and four resistors. Users can program the resolution and conversion time of the proposed ADC by changing the control signals named as *4 or 6 bit*, and *EN* mentioned in Table 1.

For 4-bit resolution, the user can give control signal *4 or 6 bit* = 0 and *EN* = 0. For 6-bit resolution, *4 or 6 bit* = 1 and *EN* = 0. For 10-bit resolution, *4 or 6 bit* = 0 and *EN* = 1. For 12-bit resolution, *4 or 6 bit* = 1 and *EN* = 1.

#### 3.2 Switches for Resistive String Configuration for 4-bit Flash ADC

Figure 10 shows analog switches used to configure the connection of resistive string for 4-bit flash ADC according to a conventional 4-bit flash ADC. So, according to this conventional 4-bit flash ADC, these 16 resistors are connected to each other in series to generate the reference voltages *ref1* (LSB) to *ref15* (MSB).

So, to obtain the connectivity of 16 resistors in series, Table 2 shows the interconnection among the stages (1 to 5) through different switches (SW1 to SW6).

Now, all the resistors are connected in series through switches (SW1–SW6) to provide the 15 reference voltages for all 15 comparators like conventional 4-bit flash ADC.

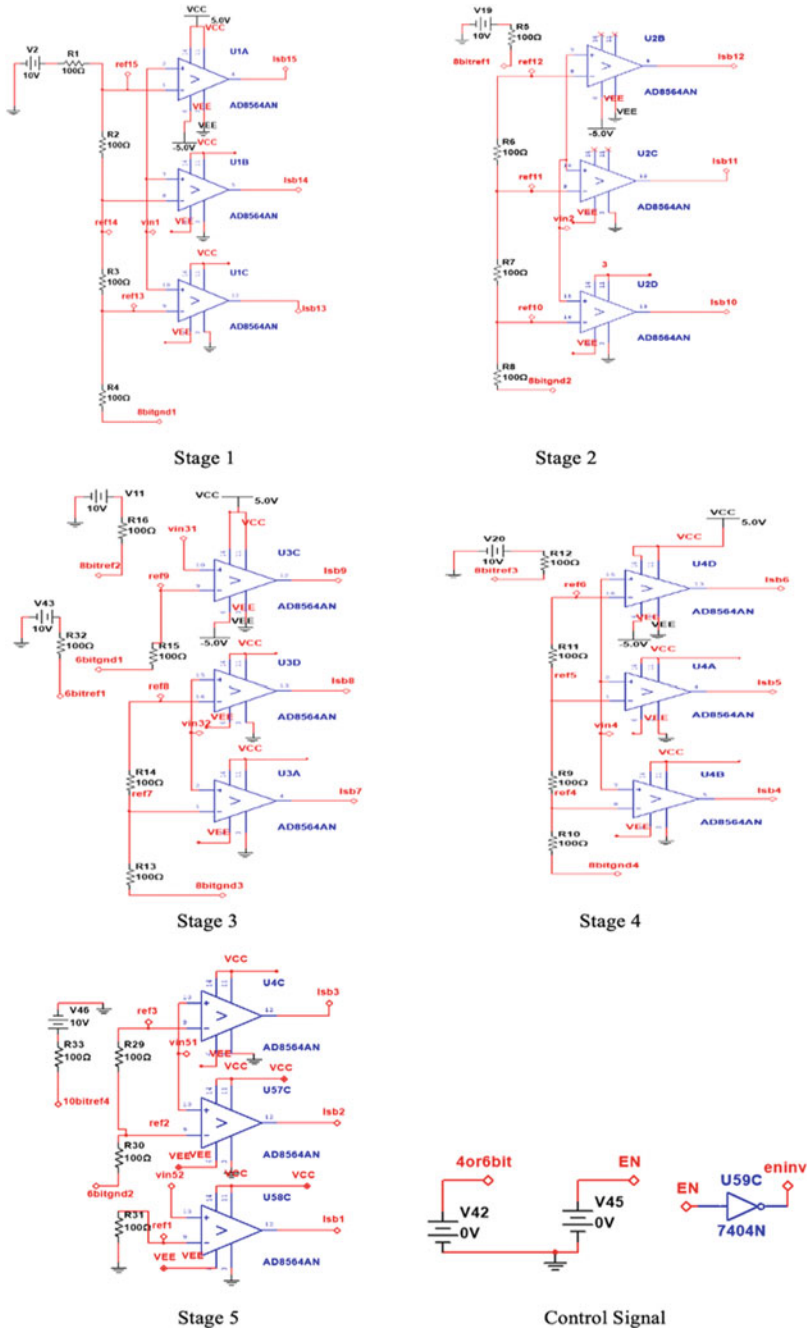
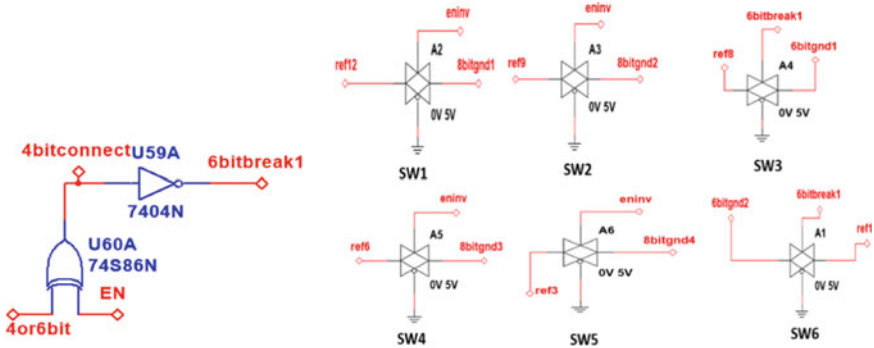


Fig. 9 Arrangement of comparators and resistive string in five stages



**Table 1** Selection of control signal

<i>EN</i>	<i>4 or 6 bit</i>	Resolution
0	0	4-bit
0	1	6-bit
1	0	10-bit
1	1	12-bit



**Fig. 10** Control signals and switches for resistive string configuration for 4-bit flash ADC

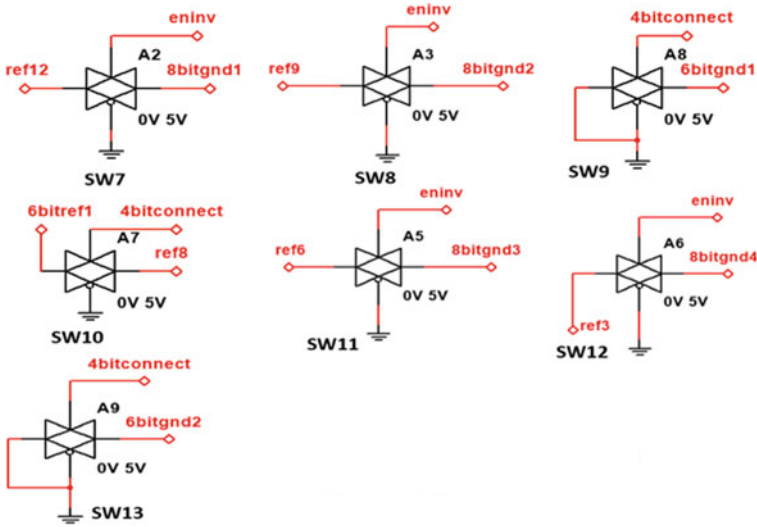
**Table 2** Interconnection of different signals for 4-bit flash ADC resistive string

Switch	Control signal	Signal1	Signal2
SW1	<i>eninv</i>	<i>ref12</i> of stage2	<i>8bitgnd1</i> of stage1
SW2	<i>eninv</i>	<i>ref9</i> of stage3	<i>8bitgnd2</i> of stage2
SW3	<i>6bitbreak1</i>	<i>ref8</i> of stage3	<i>6bitgnd1</i> of stage3
SW4	<i>eninv</i>	<i>ref6</i> of stage 4	<i>8bitgnd3</i> of stage3
SW5	<i>eninv</i>	<i>ref3</i> of stage 5	<i>8bitgnd4</i> of stage4
SW6	<i>6bitbreak1</i>	<i>6bitgnd2</i> of stage5	<i>ref1</i> of stage 5

### 3.3 Switches for Resistive String Configuration for 6-bit 2 Stage Subranging ADC

Figure 11 shows analog switches used to configure the connection of resistive string for 6-bit two-stage subranging ADC similar to conventional 6-bit subranging ADC. Each stage consists of 3-bit flash ADC. According to this conventional 6-bit two-stage subranging ADC, these 16 resistors are divided into two strings (8 + 8) for each 3-bit flash ADC.

So, to obtain the connectivity of (8 + 8) resistors in two stages, Table 3 shows the interconnection among the stages (1–5) through different switches (SW7 to SW13) where some of the switches are common (SW7,SW8,SW11, and SW12).



**Fig. 11** Switches for resistive string configuration for 6-bit two-stage subranging ADC

**Table 3** Interconnection of different signals for 6-bit two-stage subranging ADC resistive string

Switch	Control signal	Signal1	Signal2
SW7(SW1)	<i>eninv</i>	<i>ref12</i> of stage2	<i>8bitgnd1</i> of stage1
SW8(SW2)	<i>eninv</i>	<i>ref9</i> of stage3	<i>8bitgnd2</i> of stage2
SW9	<i>4bitconnect</i>	<i>ground</i>	<i>6bitgnd1</i> of stage3
SW10	<i>4bitconnect</i>	<i>6bitref1</i> of stage3	<i>ref8</i> of stage 3
SW11(SW4)	<i>eninv</i>	<i>ref6</i> of stage 4	<i>8bitgnd3</i> of stage3
SW12(SW5)	<i>eninv</i>	<i>ref3</i> of stage 5	<i>8bitgnd4</i> of stage4
SW13	<i>4bitconnect</i>	<i>ground</i>	<i>6bitgnd2</i> of stage5

### 3.4 Switches for Resistive String Configuration for 10-bit 5 Stage Subranging ADC

Figure 12 shows analog switches used to configure the connection of resistive string for 10-bit five stage subranging ADC similar to a conventional 10-bit subranging ADC. Each stage consists of 2-bit flash ADC. According to this conventional 10-bit,

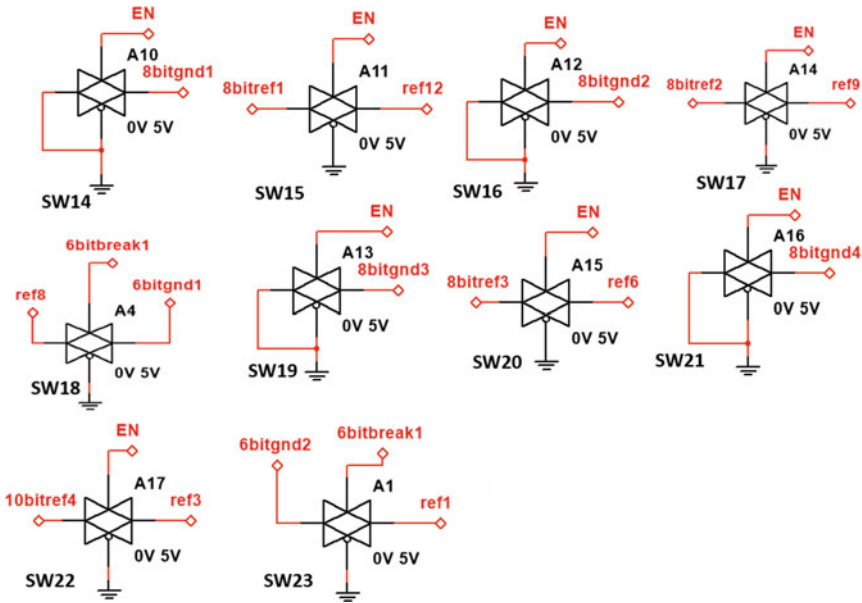


Fig. 12 Switches for resistive string configuration for 10-bit 5 stage subranging ADC

five stage subranging ADC, these 16 resistors are divided into five strings (4 + 4 + 4 + 4 + 4) for each flash ADC.

So, to obtain the connectivity of (4 + 4 + 4 + 4 + 4) resistors in five stages, Table 4 shows the interconnection among the stages (1–5) through different switches (SW14 to SW23) where switch SW23 is common.

### 3.5 Switches for Comparator Input Signal Configuration (4-bit Flash ADC)

Figure 13 shows switches (SW24 to SW 30) to apply the input for 4-bit flash ADC.

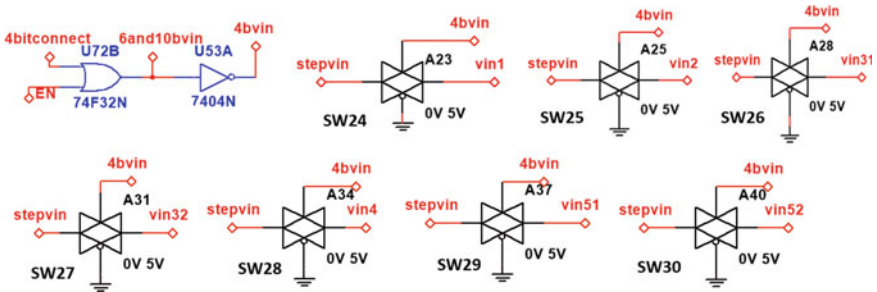
Table 5 shows the connection between input signals through switches (SW24 to SW 30) for 4-bit flash ADC.

### 3.6 Switches for Comparator Input Signal Configuration (6-bit Subranging ADC)

Figure 14 shows switches (SW31–SW36) to apply input for 6-bit two-stage subranging ADC.

**Table 4** Interconnection of different signals for 10-bit five-stage subranging ADC resistive string

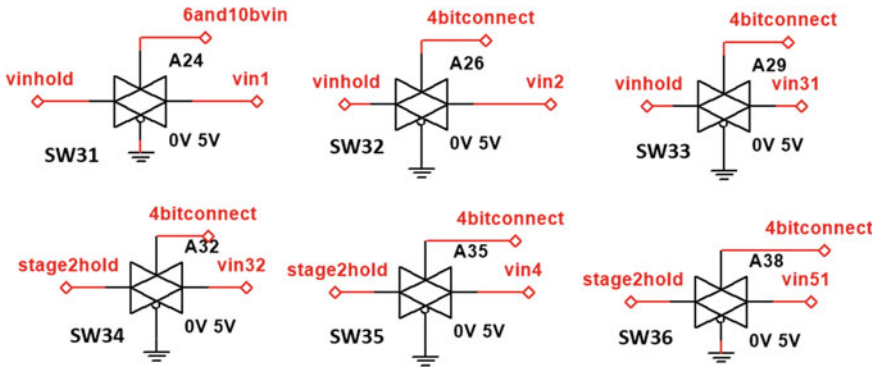
Switch	Control signal	Signal1	Signal2
SW14	EN	ground	8bitgnd1 of stage1
SW15	EN	8bitref1(Ref voltage)	ref12 of stage2
SW16	EN	ground	8bitgnd2 of stage2
SW17	EN	8bitref2(Ref voltage)	ref9 of stage 3
SW18	6bitbreak1	ref8 of stage 3	6bitgnd1 of stage3
SW19	EN	ground	8bitgnd3 of stage3
SW20	EN	8bitref3 (Ref voltage)	ref6 of stage4
SW21	EN	ground	8bitgnd4 of stage4
SW22	EN	10bitref4 (ref voltage)	ref3 of stage5
SW23(SW6)	6bitbreak1	6bitgnd2 of stage5	ref1 of stage5



**Fig. 13** Switches to configure input for comparators for 4-bit flash ADC

**Table 5** Interconnection of different signals for 4-bit flash ADC comparators input signal

Switch	Control signal	Signal1	Signal2
SW24	4bvin	stepvin(input signal)	vin1(input to comparators of stage1)
SW25	4bvin	stepvin(input signal)	vin2(input to comparators of stage2)
SW26	4bvin	stepvin(input signal)	vin31(input to comparators of stage3)
SW27	4bvin	stepvin(input signal)	vin32(input to comparators of stage3)
SW28	4bvin	stepvin(input signal)	vin4 (input to comparators of stage4)
SW29	4bvin	stepvin(input signal)	vin51(input to comparators of stage5)
SW30	4bvin	stepvin(input signal)	vin52(input to comparators of stage5)



**Fig. 14** Switches to configure input for comparators for 6-bit two stage subranging ADC

**Table 6** Interconnection of different signals for 6-bit two-stage subranging ADC comparators input signal

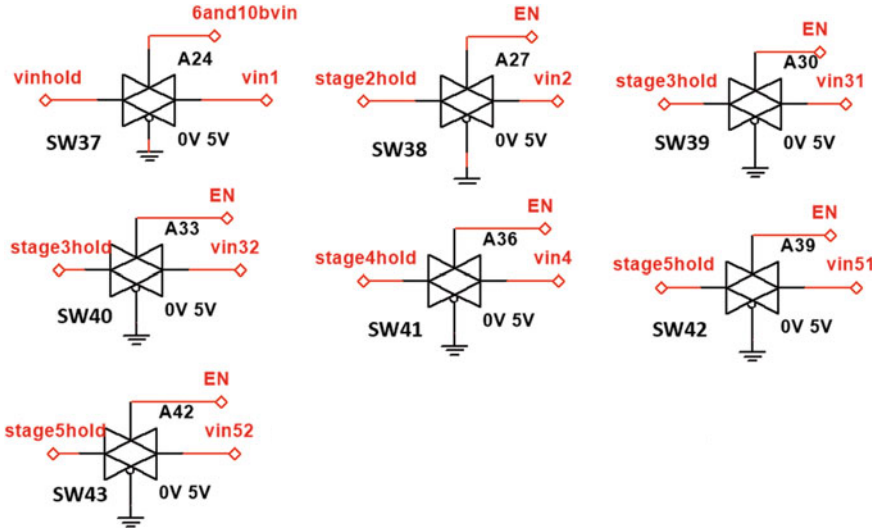
Switch	Control signal	Signal1	Signal2
SW31	6 and 10bvin	vinhold (sampled input signal)	vin1 (input to comparators of stage1)
SW32	4bitconnect	vinhold (sampled input signal)	vin2 (input to comparators of stage2)
SW33	4bitconnect	vinhold (sampled input signal)	vin31(input to comparators of stage3)
SW34	4bitconnect	stage2hold (residue signal)	vin32 (input to comparators of stage3)
SW35	4bitconnect	stage2hold (residue signal))	vin4 (input to comparators of stage4)
SW36	4bitconnect	stage2hold (residue signal)	vin51(input to comparators of stage5)

Table 6 shows connections among different terminals through switches (SW31 to SW 36) for 6-bit two-stage subranging ADC.

### 3.7 Switches for Comparator Input Signal Configuration (10-bit Subranging ADC)

Figure 15 shows switches (SW 37 to SW 43) to apply input for 10-bit five-stage subranging ADC.

Table 7 shows connection among different terminals through switches (SW37 to SW43) for 10-bit five-stage subranging ADC.



**Fig. 15** Switches to configure input for comparators for 10-bit five-stage subranging ADC

**Table 7** Interconnection of different signals for 10-bit five-stage subranging ADC comparators input signal

Switch	Control signal	Signal1	Signal2
SW37	6 and 10bvin	vinhold (sampled input signal)	vin1 (input to comparators of stage1)
SW38	EN	stage2hold (residue signal)	vin2 (input to comparators of stage2)
SW39	EN	stage3hold (residue signal)	vin31 (input to comparators of stage3)
SW40	EN	stage3hold (residue signal)	vin32 (input to comparators of stage3)
SW41	EN	stage4hold (residue signal))	vin4 (input to comparators of stage4)
SW42	EN	stage5hold (residue signal)	vin51(input to comparators of stage5)
SW43	EN	stage5hold (residue signal)	vin52(input to comparators of stage5)

Similarly, the configuration of switches could be designed for the remaining resolution (8, 16 and 20-bit) for a wide range of applications which has not been simulated here.

### 4 Timing Analysis of Reconfigurable ADC

Figure 16 shows timing analysis for the various resolution of ADCs such as 4, 6, 8, 10, 12 and 16-bit. It shows as resolution increases, the conversion time of reconfigurable ADC also increases.

Table 8 shows the conversion time for each type of ADC.

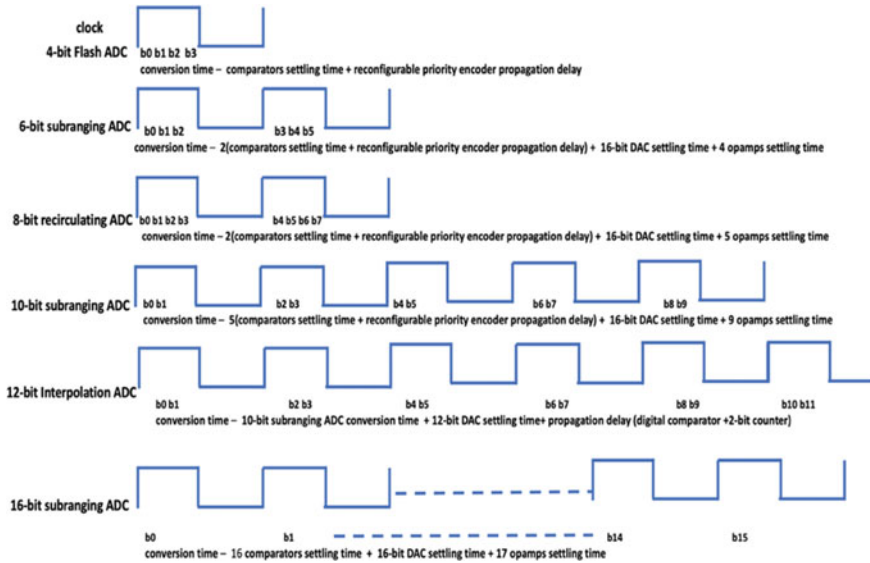


Fig. 16 Timing analysis of reconfigurable ADC

Table 8 Resolution and conversion time of ADCs

Resolution	Conversion time
4-bit	comparator settling time + priority encoder propagation delay = $T(<ns)$
6-bit	$2 T + 16\text{-bit DAC settling time}(T1) + 4 \text{ opamps settling time}(4T2)$
8-bit	$2 T + T1 + 5T2$
10-bit	$5 T + T1 + 9T2$
12-bit	$5 T + T1 + 9T2 + 12\text{-bit DAC settling time} + \text{propagation delay (digital comparator and 2-bit counter)}$
16-bit	$16 \text{ comparators settling time} + T1 + 17 T2$

## 5 Validation

Tables 9, 10, 11 and 12 shows validation for 4-bit flash ADC, 6-bit two stage subranging, 10-bit five-stage subranging ADC and 12-bit interpolation ADC for different input signals with 1LSB error, respectively.

**Table 9** Validation of 4-bit flash ADC (b0–b3)

b0	b1	b2	b3	b4	b5	b6	b7	b8	b9	Vin (V)
0	0	0	1	0	0	0	0	0	0	0.625
0	0	1	0	0	0	0	0	0	0	1.25
0	0	1	1	0	0	0	0	0	0	1.88
0	1	0	0	0	0	0	0	0	0	2.5
1	0	0	0	0	0	0	0	0	0	5
1	0	0	1	0	0	0	0	0	0	6.25

**Table 10** Validation of 6-bit two-stage subranging ADC (b0–b5)

b0	b1	b2	b3	b4	b5	b6	b7	b8	b9	Vin(V)
0	0	0	0	0	1	0	0	0	0	0.15625
0	0	0	0	1	0	0	0	0	0	0.3125
0	0	0	0	1	1	0	0	0	0	0.46875
0	0	0	1	0	0	0	0	0	0	0.625
0	0	0	0	1	1	0	0	0	0	1.25
0	0	1	1	1	1	0	0	0	0	2.5

**Table 11** Validation of 10-bit five-stage subranging ADC (b0–b9)

b0	b1	b2	b3	b4	b5	b6	b7	b8	b9	Vin(V)
0	0	0	0	0	0	0	0	0	1	0.0097
0	0	0	0	0	0	0	0	1	0	0.0195
0	0	0	0	0	0	0	1	0	0	0.0293
0	0	1	0	0	0	0	0	0	0	1.25
0	0	1	0	0	0	0	0	0	1	1.26
0	0	1	0	0	0	0	0	0	1	1.27



**Table 12** Validation of 12-bit interpolation ADC (b0–b11)

b0	b1	b2	b3	b4	b5	b6	b7	b8	b9	b10	b11	Vin(V)
0	0	0	0	0	0	0	0	1	0	1	0	25 m
0	0	0	0	0	0	0	1	0	1	1	0	55 m
0	0	0	0	1	1	0	0	1	1	1	1	0.508
0	0	0	1	1	0	0	1	1	1	0	1	1.01
1	0	0	0	0	0	0	0	0	0	0	0	5.00
1	0	0	0	0	0	0	0	0	0	0	1	5.003

## 6 Figure of Merit

Figure of merit of a reconfigurable ADC is defined in different ways by various researchers depending upon the targeted applications. For example, some researchers [9, 14] have defined FOM as:

$$FOM = P_{ADC}/2 \times BW \times 2^{ENOB} \tag{1}$$

Their application is targeted towards optimization of power for achieving reconfigurability.

Where  $P_{ADC}$  is the total power consumption, BW is the input signal bandwidth, and ENOB is the effective number of bits of the ADC. Similarly, other researchers have defined FOM [6] as:

$$FOM = 2^{2N} \times (\text{Data Rate})/\text{Power}. \tag{2}$$

Also, according to researcher, the ADC flexibility can be measured by its wide resolution, bandwidth, and it is defined as reconfigurability factor (RF) [15] which is proposed here:

$$R_F = (2^{ENOB,max}/2^{ENOB,min})/(f_{s,max}/f_{s,min}). \tag{3}$$

In most of the reconfigurable ADC research papers, the power is optimized while reconfiguring resolution and bandwidth by keeping unused stages in standby mode and/or designing low power components used in ADCs. But in this proposed design, the main objective is to reconfigure the resolution and conversion time to meet a wide range of applications. So here, a new figure of merit is proposed:

$$FOM = (f_{s,max} \times \text{No. of Possible Resolutions}). \tag{4}$$

Where  $f_{s,max}$ —highest sampling rate.

The  $f_{s, \max}$  can be as high as 100 MHz, lowest sampling rate can be as low as DC, maximum resolution can be as large as 20-bit and min can be 1-bit. Also, in this proposed design,  $P_{\text{ADC}}$  can be optimized while reconfiguring resolution and bandwidth by keeping unused stages in the standby mode, then the FOM can be defined as:

$$\text{FOM} = (f_{s, \max} \times \text{No. of Possible Resolutions}) / P_{\text{ADC}}. \quad (5)$$

## 7 Conclusion and Future Scope

The new hybrid flash-subranging ADC combines features of the conventional 4-bit flash ADC and the sub-ranging ADC topologies to achieve flexibility of operation well suited to many high-speed and high-resolution applications. This design has been simulated and validated in NI Multisim 14.1, which proves the functionality of the proposed concept. This reconfigurable ADC utilizes all comparators to achieve 4, 6, 10 and 12-bit resolution instead of keeping those comparators in standby mode. So hardware utilization while reconfiguring the parameters is optimum. But to meet the specific high-speed and high-resolution applications, it is desirable to implement in VLSI.

## References

1. MT-020 (2021) ADC architectures I: the flash converter. Analog Devices by Walt Kester. <https://www.analog.com/media/en/training-seminars/tutorials/MT-020.pdf> (Accessed on 21 June 2021)
2. MT-024 (2021) ADC architectures V: pipelined subranging ADCs, Analog devices by Walt Kester, <https://www.analog.com/media/en/training-seminars/tutorials/MT024.pdf> (Accessed on 21 June 2021)
3. MT-021 (2021) ADC Architectures II: Successive Approximation ADCs. Analog devices by Walt Kester. <https://www.analog.com/media/en/training-seminars/tutorials/MT-021.pdf> (Accessed on 21 June 2021)
4. MT-022 (2021) ADC Architectures III, sigma-delta ADC basics. analog devices by Walt Kester, <https://www.analog.com/media/en/training-seminars/tutorials/MT-022.pdf> (Accessed on 21 June 2021)
5. Yoo J, Lee D, Choi K, Kim J (2002) A power and resolution adaptive flash analog-to-digital converter. In: Proceedings of the international symposium on low power electronics and design. <https://doi.org/10.1109/LPE.2002.146744>
6. Gulati K, Lee H-S (2001) A low-power reconfigurable analog-to-digital converter. IEEE J Solid-State Circuit 36(12):54–55. <https://doi.org/10.1109/4.972140>
7. Kun C, Mason A, Chakrabartty S (2005) A dynamic reconfigurable A/D converter for sensor applications. Sensors IEEE 4. <https://doi.org/10.1109/ICSENS.2005.1597926>
8. Mostafa S, Qu W, Islam SK, Mahfouz M (2010) A calibration circuit for reconfigurable smart ADC for biomedical signal processing. In: IEEE computer society annual symposium on VLSI (2010), pp 185–189. <https://doi.org/10.1109/ISVLSI.2010.85>

9. Taherzadeh-Sani M, Hamoui AA (2012) A reconfigurable and power-scalable 10–12 Bit 0.4–44 MS/s pipelined ADC With 0.35–0.5 pJ/Step in 1.2 V 90 nm digital CMOS. *IEEE Trans Circuits Syst I*. <https://doi.org/10.1109/TCSI.2012.2215712>
10. Stojcevski A, Singh J, Zayegh A (2002) Scalable pipeline analog-to-digital converter for ultra-tdd mobile station receiver. In: 9th international conference on neural information processing, computational intelligence for the E-Age. pp 19–21. <https://doi.org/10.1109/SMELEC.2002.1217844>
11. Gulati K (2007) Low power reconfigurable analog-to-digital converters: cambridge analog technologies. In: Sixth IEEE dallas circuits and systems workshop. Inc. Bedford, Massachusetts, pp 1–8. <https://doi.org/10.1109/DCAS.2007.4433193>
12. Shriram KV, Venkateshwaran K, Vaithyanathan V, Subashri V (2009) Design of analog to digital converter for SDR: international conference for internet technology and secured. *Transactions*. <https://doi.org/10.1109/ICITST.2009.5402602>
13. Ghetmiri S, Salama CAT (2009) An 8-bit 12.5-GSample/s SiGe ADC: analog integrated circuits and signal processing. <https://doi.org/10.1007/s10470-009-9422-7>
14. Fateh S, Schönle P, Bettini L, Rovere G, Benini L, Huang Q (2015) A reconfigurable 5-to-14-bit SAR ADC for battery-powered medical instrumentation. *IEEE Trans Circuits Syst I*. <https://doi.org/10.1109/TCSI.2015.2477580>
15. Yousry R, Chen M-S, Chang M-CF, Yang C-KK (2013) An architecture- reconfigurable 3b-to-7b 4GS/s-to-1.5GS/s ADC using subtractor interleaving. In: IEEE Asian solid-state circuits conference. <https://doi.org/10.1109/ASSCC.2013.6691038>

# Mini Interconnect IC Protocol and Camera Command Set Controller for Camera Communication in Mobile Phones



Utsav Malviya, Gopal Kumar, Champalal Lalani, Sampad Barik,  
and P. S. Shrivastava

## 1 Introduction

CSI-3 is a high-speed, low power serial communication protocol described in [1–4] and [5]. CSI-3 used in mobile devices for real-time camera data transmission from Camera controller to baseband processor inside SoC. CSI-3 also uses D-PHY (5Gbps) [4] or faster version C-PHY (11.4Gbps) [2] for camera data communication. CSI-3 uses subset Camera Control Interface (CCI) [1] protocol for commands communication and control camera from SoC using fast mode I2C communication and 7-bit slave address. Liu et al. [6] explain High-speed I2C bus interfacing using a single I2C-master and multiple I2C slaves and synchronization controls. Valdez and Becker [7] describes I2C protocol controls, format, timing, frequency, current, voltage, and power ratings. Lokapure et al. [8] implement the I2C module on FPGA and interface Temperature, Pressure, and Humidity (TPH) sensors on NIOS-II soft-core processor on FPGA/SoC Board. SMIA camera command set [9] for less-featured single camera device control. SMIA is not good for State-of-the-art multiple camera mobile phones. MIPI developed a new Camera Command Set (CCS) [10] for multi-camera featured modern mobile phones. [10] and [11] describe Camera Command Set (CCS) and functionality in state-of-the-art multiple camera devices based on mobile phones. Kushwaha and Kapse [11] define the Camera Control Interface (CCI) protocol using

---

U. Malviya (✉) · G. Kumar · C. Lalani  
Department of Electronics and Telecommunication, Government Engineering College, Bilaspur,  
CG, India  
e-mail: [utsav00706@gmail.com](mailto:utsav00706@gmail.com)

S. Barik  
Department of Electrical Engineering, Government Engineering College, Bilaspur, CG, India

P. S. Shrivastava  
Department of Electronics and Communication, Guru Ghasidas Vishwavidyalaya, Bilaspur, CG,  
India

Basic I2C (400 kHz) for SMIA single-byte commands communication. All modules of CSI IP design [4, 12, 13] before 2017 support the SMIA camera command register set hence cannot be used in the State-of-the-art CCS camera command supported modern mobile phones. CSI module design [14] supports the CCS command set for modern mobile phones, [14] uses the I2C for CCI. [3, 4, 12, 13] uses a I2C core module for CCI with 2.4 MHz operations, a 7-bit slave address, and an SMIA command set [15]. I2C is a general-purpose protocol which can be used in many applications.

### 1.1 Camera Serial Interface Protocol

MIPI Alliance defines mobile phone standards protocols for interfacing different peripheral device controllers with mobile phone SoC. MIPI uses the Camera Serial Interface (CSI-3) [12, 16] for communication between the SoC and camera shown in Fig. 1. CSI-3 transmits very fast N-Lane encoded pixel data from the camera

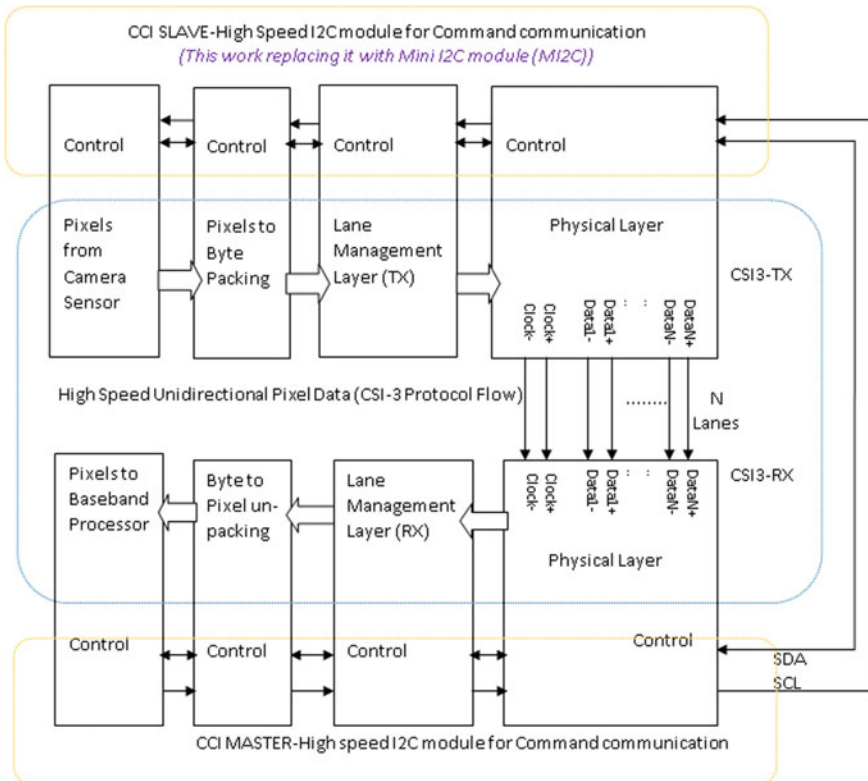


Fig. 1 MIPI CSI layers and CCI signalling

controller to the mobile phone SoC. The data on N-lanes are transmitted serially and synchronized with the system clock. CSI-3 transmits pixel data from the camera to the processor in one direction to maintain very high speed and, for camera commands communication CSI-3 uses bidirectional relatively slow subset CCI protocol [12, 16]. CCI is developed on top of the I2C protocol. This work is a less area and high-speed design named MI2C-slave and at Camera controller. MI2C-slave is fully compatible with I2C. MI2C at camera controller communicates with I2C-master of SoC on SDA and SCL lines.

## 1.2 Camera Command Register Set (CCS) Memory

State-of-the-art mobile phone's camera device has different command registers defined in CCS specification [17]. CCS memory stores the commands from the SoC processor through I2C signals and Camera controller accesses the CCS memory using Advance High-speed Bus (AHB). Camera commands are used for controlling and setting up the camera controller. The CCS Command registers have different sizes, 8-bit CCS registers are general setting registers, 16-bit CCS registers are frame-length, exposure values registers, 32-bit CCS registers are high accuracy setting registers and 64-bit CCS registers are multi-sensor setting and control registers.

## 1.3 Mini Interconnect IC Protocol (MI2C)

The data transfer in the MI2C protocol is the same as the I2C standard [17]. MI2C has a 7-bit slave address. MI2C supports an 8-bit index or 16-bit index with 8-bit, 16-bit, 32 bit, or 64-bit command/status data.

Figure 2a shows single 8-bit command data write operation on the CCS register from mobile phone SoC. In Fig. 2, 'S' is the start state, 'P' is the stop state, 'A' is a positive-acknowledgement from the receiver,  $\bar{A}$  is a negative acknowledgement from the receiver. After selection of camera controller using 7-bit MI2C slave address (0111100b) followed by ( $R/\bar{W} = 0$ ), the I2C-master sends three consecutive 8-bit data to MI2C slave. The MI2C-slave first receives two data bytes (index address of CCS), then each data bytes are camera commands for that index address of CCS. After each successful byte transfer, PACK has to send on the 9th clock from the receiver. Figure 2b shows multiple bytes write started from any random location. Figure 2c and d shows the current location command write operation, in which index addressing does not require.

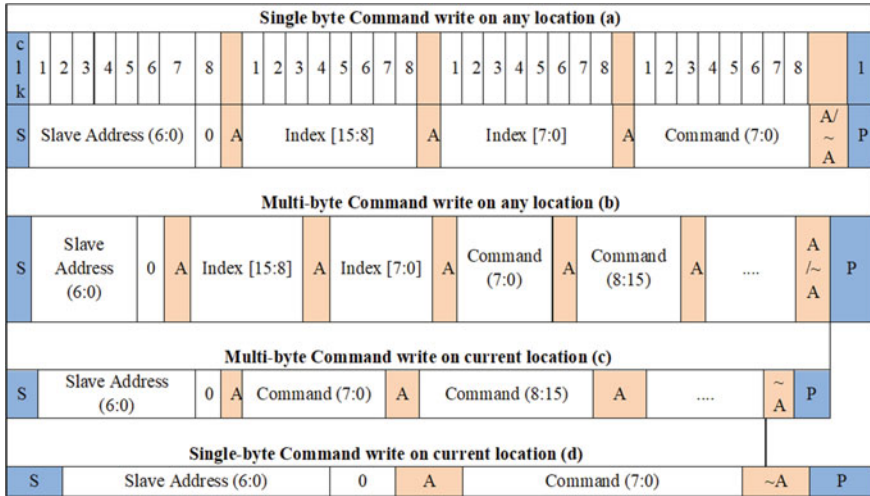


Fig. 2 Command data write operations from I2C-master at mobile SoC to MI2C slave in the camera controller

## 2 Methodology

Process flow (A) is the Camera command transfer method used in available systems [3, 4, 12–14]. Process flow (B) is the Camera command transfer method used in this work. There are three modifications done (i) Replaced I2C-slave with MI2C-slave (ii) A Camera Set Bridge (CSB) signal Set develop for the CCS controller interface and (iii) FSM-based new CCS controller design for simultaneous interfacing with Dual-Port RAM for CCS registers set.

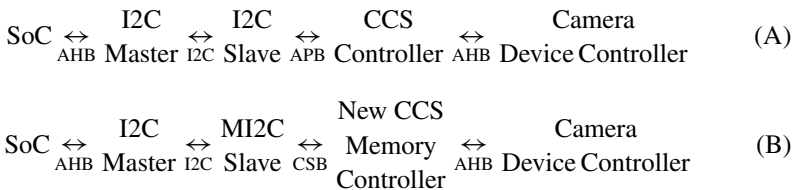


Figure 3 is the top RTL module of CCI in CSI-3, where core and its signal interfacing at the top-level CCI Signals are as follows:

I2C-Master communicates with MI2C-slave using SDA and SCL signals only. MI2C-slave must acknowledge the I2C-master after every time it receives the correct index address of the command. MI2C-slave writes the camera commands inside CCS on the index address. The MI2C slave module identifies the requirement of atomic operation performs atomic operations accordingly. MI2C-slave also interfaces the APB bus from the camera controller to update the slave registers. CCS controller notifies the MI2C-slave for I2C stretching in SDA and SCL.

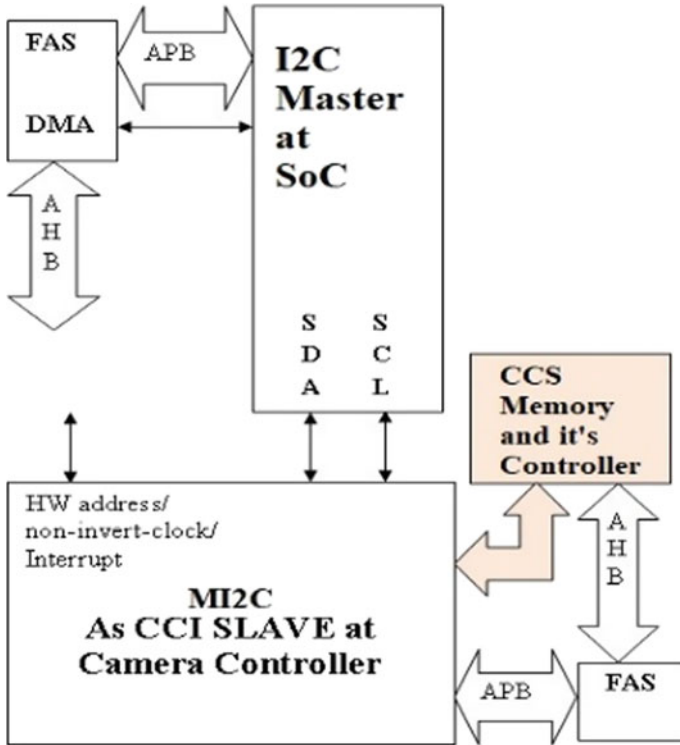


Fig. 3 CCI IP Top module Interfacing

### 2.1 MI2C Module

I2C-master is an essential module in mobile phone SoC it interfaces with multiple sensors like camera sensors [18], but the same fully fledged I2C module at the slave side of the camera controller is not essential hence a mini version of I2C (MI2C) is used on the Slave side [19]. Figure 4 shows all internal modules and control registers of a fully fledged I2C module. CSI-3 uses the same fully fledged I2C module at the mobile phone’s SoC and in-camera controller. Figure 5 shows the MI2C-slave module used in proposed work instead of the fully fledged I2C of Fig. 4 used in this work. This work is a mini design of I2C which kept only essential modules and necessary control registers of I2C [20]. Proposed MI2C is an application-specific design rather than a general-purpose design. MI2C excluded the DMA interfacing module and DMA registers. MI2C does not need a multi-master mode of I2C hence registers related to the master module are also excluded. AHB to APB and vice-versa communication done using FPGA AMBA Subsystem (FAS) [2], FAS reduces FIFO size in the MI2C module. An 8-byte temporary register is used for atomic operations. In atomic operation MI2C-slave module can send/receive different size (8, 16, 32 and 64 bytes) commands to/from mobile phone SoC’s I2C master with communication



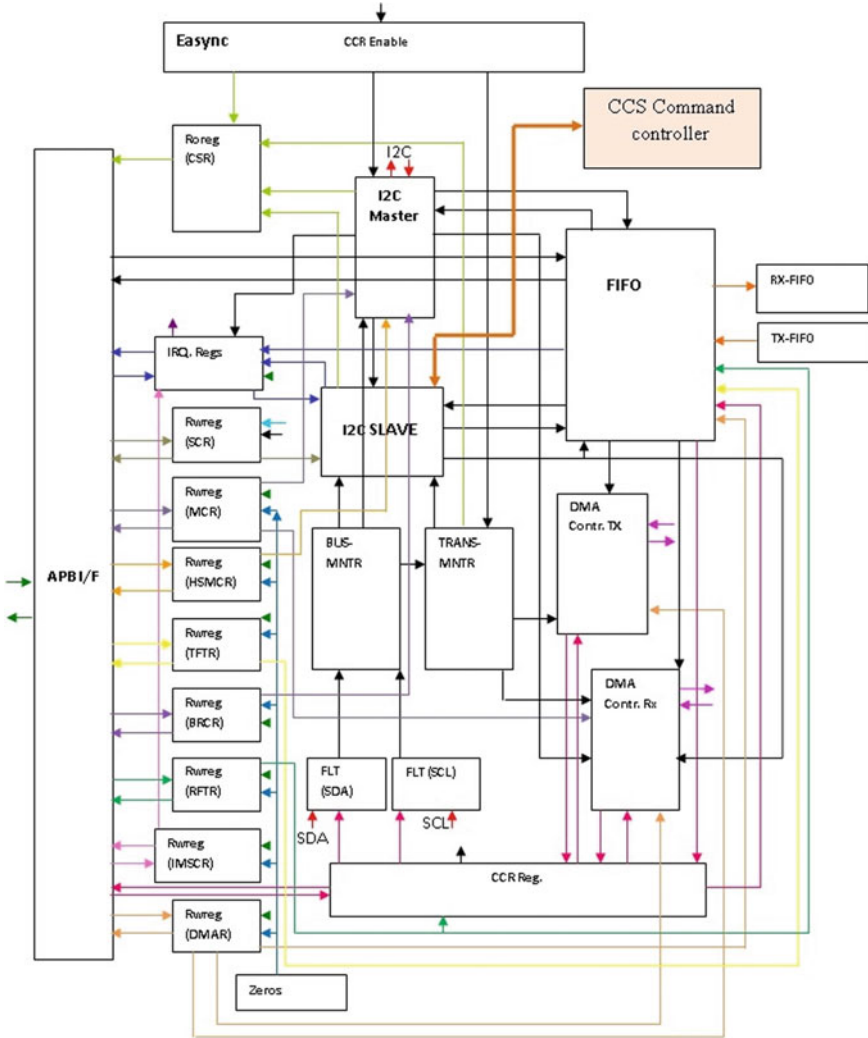


Fig. 4 I2C internal modules

break [20]. The configuration and control registers in the MI2C-slave module can be configured from the camera controller using the APB bus. CCS module contains CCS memory and CCS controllers. The slave module is also interfacing with the CCS module using the proposed signals set named Camera set bridge (CSB).

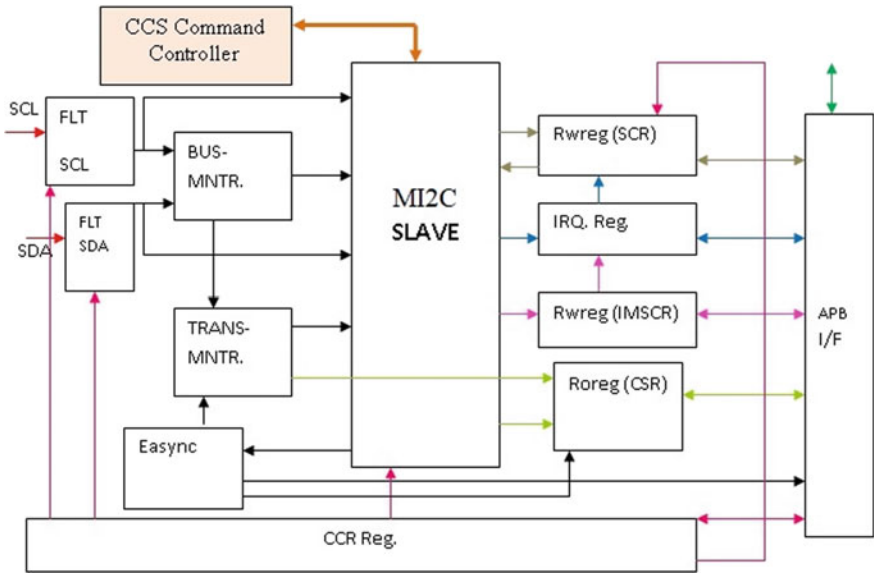


Fig. 5 MI2C internal modules

## 2.2 Camera Command Set (CCS) Module

This CCS module shown in Fig. 6 contains a CCS controller and two DPRAM's of 16 kb each for storing Camera Commands. CCS module interface with MI2C using new signal set developed in this work named Camera control bridge (CSB) explained in Table 1.

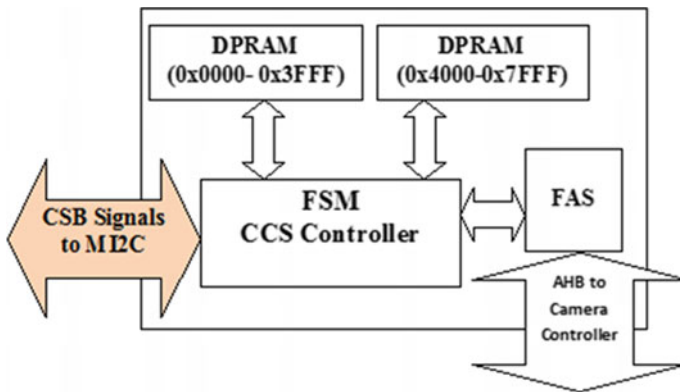


Fig. 6 Camera command set module

**Table 1** CSB Signals between MI2C slave and CCS module

Signal name	Function
I/P: rx_drdy	SDA = 1, SCL = 1 when the Camera controller writes status into CCS via AHB
O/P: rx_push	Write command from MI2C to the CCS
O/P: index_ls	Store LSB of CCS index address
O/P: index_ms	Store MSB of CCS index address
O/P: tx_pop	Read command/Status from CCS to MI2C
I/P: tx_drdy	SDA = 1, SCL = 1 when Camera controller read command from CCS via AHB
O/P: rx_data_in(7:0)	Command byte from MI2C to CCS
I/P: tx_data_out(7:0)	Status/command byte from CCS to MI2C

### 2.2.1 CCS Controller

This FSM-based controller developed for interfacing the camera controller with CCS memory using AHB. This controller also interfaces MI2C-slave with CCS memory using CSB signal set. The multi-byte operation must be atomic to avoid wrong command interpretation by the camera controller or wrong status interpretation at SoC. The proposed work CCS controller has especially taken care of atomic operations by using additional control signals in CSB and re-timing rules. According to the re-timing when a Write operation, a specific set of register updating must suspend until the last bit of multi-byte command received. CCS Controller gives high priority to the camera controller operations, so when AHB interface with CCS, the CCS controller must put hold (SCL = 1, SDA = 1) on the I2C bus. FSM performs the following task:

- (1) Synchronize MI2C data with I2C clock and Camera controller data with System clock.
- (2) I2C-master must hold I2C bus communication and buffer received data when the Camera controller requires access to CCS.
- (3) During writing from current location operation isolates 16-bit index address from MI2C received data also increment index address by one after every push.
- (4) Write command on index address of CCS using push signal.
- (5) Write command/status into CCS when the camera controller sends information on the AHB bus.
- (6) Read command/status from the index address of CCS using a pop signal.
- (7) The camera controller reads commands from CCS when on the AHB bus.
- (8) Identify the command more than one byte and buffer the same until all bytes received. Avoid multi-byte corruption using the atomic cycle for reading or write commands.

- (9) In AHB read generate AHB ready-out after generating I2C line stretching signal.
- (10) Deny write operation if index address from the I2C master received wrong.
- (11) Generates start pulse by writing a command  $0 \times 02$  on CCS address is  $0 \times 0010$ .
- (12) Generates stop pulse by writing a command  $0 \times 01$  on CCS address is  $0 \times 0010$ .

### 2.2.2 Camera Command Set Memory (CCS)

CCS register set (CCS memory) are two Dual-port RAM (DPRAM) IP of coarse-grained Zynq-7000 FPGA [21] developed using Core-Generation. CCS holds the command registers for the camera controller and status registers information for mobile SoC. DPRAM can simultaneously read/write the data. DPRAM can interface with the camera controller and MI2C slave at the same time. Two physically separate memory allows memory banking for 16-bit parallel data access when required.

## 3 Results and Discussion

The individual is modules developed with combinational and behavioural modelling with the system (AMBA) and I2C clock synchronization. RTL entries and synthesis of the proposed work are done on the Xilinx Vivado 2018 EDA tool. This work has tested on the System clock (AHB Clock) that equals 200 MHz and I2C high-speed mode clock with 3.4 MHz speed. Transaction level model (TLM 2.0) testing environment used to create a test suit for I2C Master. A total of 76 possible test cases have generated and verified successfully. Few test cases successfully simulated on the Cadence NCsim simulator for the proposed work are below:

**Test 1**—I2C signal (SCL and SDA) states decoded correctly.

**Test 2**—MI2C slave module addressed correctly by the I2C-master.

**Test 3**—MI2C Message is properly isolated indexed address (16-bit) and commands in:

- (a) Single write on a random location.
- (b) Sequential write on a random location.
- (c) Single write on current location.
- (d) Sequential write on current location operations.

**Test 4**—MI2C Message isolated indexed address (16-bit) and command/status in:

- (a) Single read from random location operation.
- (b) Sequential read from a random location.
- (c) Single read from the current location.
- (d) Sequential read from current location operations.

**Test 5**—MI2C writes 1 byte, 2byte, 4 byte, and 8-byte commands at CCS with the atomic operation.

**Test 6**—MI2C reads 1 byte, 2byte, 4 bytes, and 8-byte status/command register from CCS with the atomic operation.

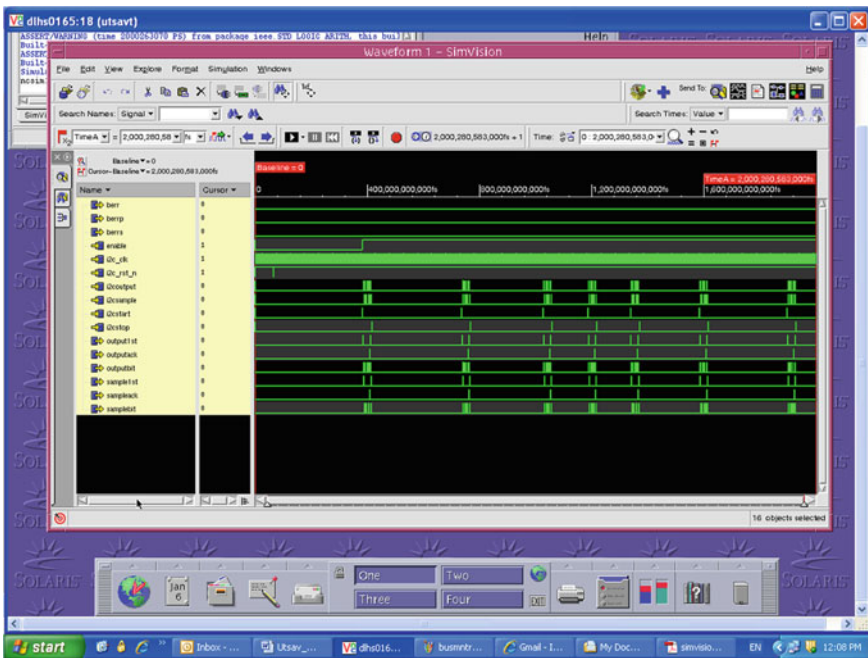
**Test 7**—No read or write operation from MI2C slave when wrong index address sends by the I2C-master.

**Test 8**—Correct single-byte and multi-byte status data write on CCS from AHB Bus at the correct address.

**Test 9**—Correct single-byte and multi-byte command data read from the correct index address of CCS on AHB Bus.

In Fig. 7, MI2C Message properly isolated indexed address (16 bit) in and correct interpretation of reading operation with repeated slave address in sequential read operation from random CCS location operations. The 16-bit index address buffered in the register and latched correctly in CCS. Also, the 4-byte Status data from the CCS specified index address send from MI2C- slave to I2C-master and received proper acknowledgement from I2C-master. In Fig. 8, MI2C correctly writes 4-byte command on CCS with the atomic operation after the 16-bit address of CCS and write operation selection by I2C mater.

Synthesizable RTL entries made for the design using the balance design goal. The floor planning did for routing with reducing the path delay. Synthesis report shows area utilization, time delay, and operating frequency obtained for the design. MI2C-slave obtained from a synthesis report generated on the Vivado design suite. The summary of the synthesis report is shown in Table 2. The dynamic power consumption obtained is 11mw for MI2C. The number of slices (area on-chip) is 160. The proposed



**Fig. 7** CCS read operation from CCS to I2C via MI2C

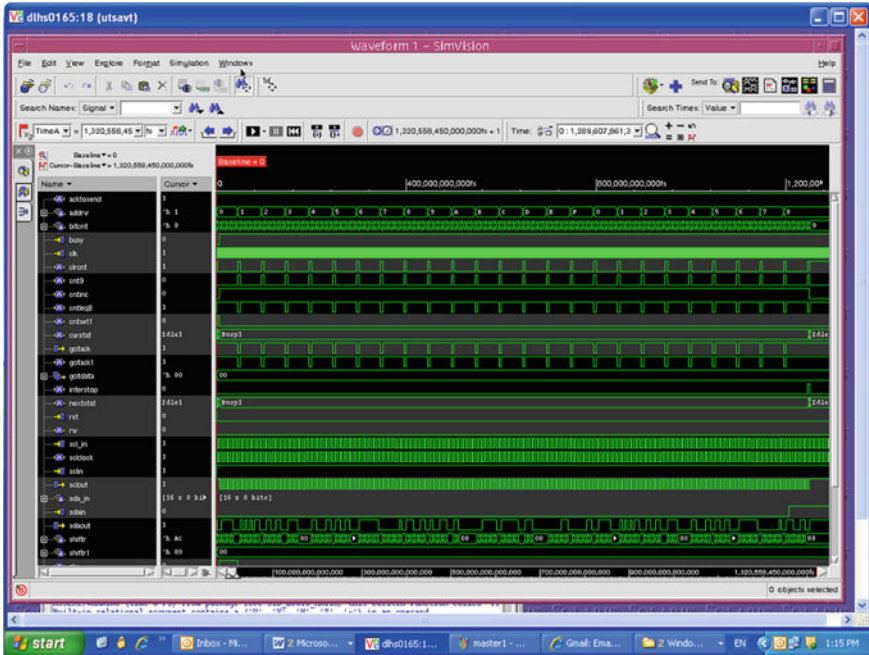


Fig. 8 CCS write operation from I2C to CCS via MI2C

Table 2 Device utilization summary

MI2C-HDL synthesis report summary	
Selected device	Zynq 7000
Slices	160
Slice flip flops	133
4-input LUTs	299
bonded IOBs	90
GCLKs	2
Speed grade	-5
Minimum period	6.397 ns
Maximum frequency:	156.333 MHz
Maximum combinational path delay	7.066 ns
Power consumption X-power analyzer	11 mw at 25 degrees centigrade

is compatible with high-speed I2C and AMBAv3 because the maximum operating frequency for the design is 156.33 MHz.

Figure 9 shows the FPGA prototype experimental setup for the proposed design. Figure 10 shows the Emulation environment created for testing. For implementation and validation of the proposed work Zybo Z7(2 nos.), FPGA Development Boards are used. I2C Master with TLM test environments are programmed on the first FPGA Development Boards [21]. CCS module (CCS controller & CCS), Camera controller AMBA (APB & AHB) bridge interface, MI2C-slave programmed on the second FPGA development board. Both FPGA are connected with only two I2C signal wires. Onboard oscillator clock of 3.4 MHz is used for SCL and 125 MHz external clock connected on FPGA board K17 is used for a system clock (AHB clock). RAM at the master FPGA board is used for testing commands and status. MI2C-slave has CCS to store camera commands and status. Camera traffic emulator RTL is also implemented for AMBA signals interfacing with the CCS controller and



Fig. 9 Experimental setups for validation of MI2C and CCS controller

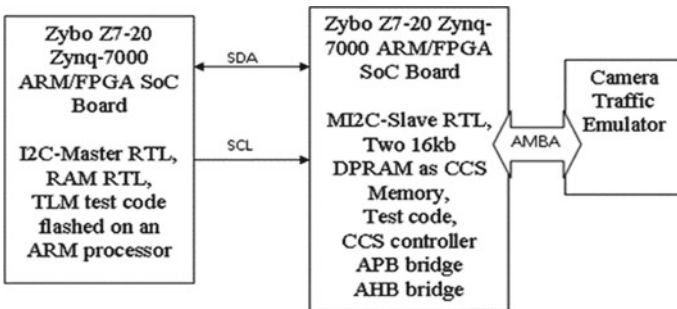


Fig. 10 Emulation environment created for testing

**Table 3** Comparative analysis

Parameter	Proposed design	Kaushik and Singh [3]	Xilinx inc. [14]
Design	CCI slave using MI2C and new CSB signal set, new FSM based CCS Controller	CCI slave using I2C and SMIA interface using APB	CCI slave using I2C and CCS interface using APB
Platform and deigning style	35 nm 7th generation Zynq-Z7 FPGA and Semi-custom design	30 nm CMOS based full-custom design	35 nm 7th generation Vertex-7 FPGA based Semi-custom design
Camera command standard	CCS Standard	SMIA Standard	CCS Standard
Maximum frequency	156.33 MHz	125 MHz	153.21 MHz
Slices on FPGA	160	–	179
Time delay	6.397 ns		11.235 ns

MI2C. Chip-scope pro used to verify the results [21]. TLM software's written in C language is flashed on the ARM processor and created Zybo FPGA development board to host Processor I2C IP.

Table 3 shows the comparative results. The proposed design has fewer slices and better maximum frequency than other similar application designs. The MI2C-slave communicate with 32 kb CCS camera command registers set controller. A CCS command support itself makes the work better than [3, 4] where the old camera command SMIA [15] was in use. MI2C-slave design has more area and speed optimized than I2C-slave [3, 14].

## 4 Conclusion

This work designs an area and speed optimized application-specific MI2C-Slave for CSI-3 protocol. An additional set of signals named CSB is developed for interfacing CCS with MI2C-slave. A new FSM-based CCS controller is developed for atomic operation and simultaneous interfacing with the camera device using AHB and MI2C using CSB. Mix modelling style and VHDL language are used for RTL entries. Simulations are done on Cadence NC-Sim. Synthesis and RTL entries on Xilinx Vivado. Two Zynq-7000 FPGA Development Boards used for the testing. The chip area for the MI2C module observed is less than the I2C module.



## References

1. Lim K, Kim GS, Kim S, Baek K (2010) A multi-lane MIPI CSI receiver for mobile camera applications. *IEEE Trans Consum Electron* 56(3):1185–1190
2. Lu Y, Chen Z, Chang P (2013) Low power multi-lane MIPI CSI-2 receiver design and hardware implementations. In: *IEEE international symposium on consumer electronics (ISCE)*. pp 199–200
3. Kaushik R, Singh P (2015) Study and implementation of CSI-2 receiver and lane merging. *Int J Adv Res Eng Technol Sci* 2(3):56–64
4. Lee P, Lee H, Kim Y, Hong H, Jang Y (2017) A 10-Gbps receiver bridge chip with deserializer for FPGA-based frame grabber supporting MIPI CSI-3. *IEEE Trans Consum Electron* 63(3):209–215
5. Feng A, Knieser M, Rizkalla M, King B, Salama P, Bowen F (2012) Embedded system for sensor communication and security. *IET Inf Secur* 6(2):111–118
6. Liu Q, Meng Q, Liao T, Bao X, Xu C (2019) A flexible hardware architecture for slave device of I2C bus. In: *International conference on electronic engineering and informatics (EEI)*. pp 309–313
7. Valdez J, Becker J (2015) Application report on understanding the I2C bus. *Texas Instruments Incorporated-SLVA704*. pp 1–45
8. Lokapure A, Mandar NS, Satyanarayana B, Raees PCM (2016) Interfacing of digital TPH sensors with FPGA using I2C interface. In: *IEEE bombay section symposium (IBSS)*. pp 1–5
9. Ailus S, MIPI Camera Command Set v1.1 public version, 2019; 1–98.
10. Muukki K (2018) Introduction to MIPI camera command set (CCS) version 1.0. pp 1–28. <https://www.mipi.org/sites/default/files>
11. Kushwaha M, Kapse V (2012) A new subset I2C protocol for interfacing camera module with a baseband processor. *Int J Adv Res Technol* 1(7):1–8
12. MIPI Alliance Standard for Camera Serial Interface (2014) MIPI CSI-3 V1.1 Specification. pp 1–150. <https://www.mipi.org/specifications/csi-3>
13. Anusha P, Ravinder K (2014) CMOS camera implementation with CSI-2 based on ARM processor. *Int J Ethics Eng Managem Educ* 1(9):58–61
14. Xilinx inc. (2020) Product guide of vivado design suite. MIPI CSI-2 Transmitter Subsystem v2.1. PG260. pp 1–58. <https://www.xilinx.com/support/documentation/Ipdocumentation>
15. Nicol RL (2006) Image sensor characterization using SMIA standard. *Photon Multimedia Proc SPIE* 6196(0E):1–12
16. McKeone P, Iepure B, Agili S, Morales A (2019) Time domain reflectometer measurements on MIPI D-PHY protocol for signal integrity analysis. In: *IEEE international conference on consumer electronics (ICCE)*. pp 1–6
17. Bharath KB, Kumaraswamy KV, Swamy RK (2016) Design of arbitrated I2C protocol with DO-254 compliance. In: *IEEE international conference on emerging technological trends (ICETT)*. pp 1–5
18. Pagliari DJ, Macii E, Poncino M (2017) Zero-transition serial encoding for image sensors. *IEEE Sens J* 17(8):2563–2571
19. Paresbhai SR, Jagtap S (2016) Design and verification of serial low-power inter-chip media bus (SLIMbus) device controller supporting MIPI standard. In: *IEEE international conference on recent trends in electronics, information and communication technology (RTEICT)*. pp 932–936
20. Malviya UK, Swain A, Kumar G (2020) Tiny I2C protocol for camera command exchange in CSI-2: a review. In: *International conference on inventive computation technologies (ICICT)*. pp 149–154
21. Cai H, Yu H, Xu JK, Wang G (2013) Application camera links on xilinx FPGA. In: *10th international conference on fuzzy systems and knowledge discovery (FSKD)*. pp 1121–1125

# Automatic Estimation of Multiplicity in Partial Discharge Sources Using Machine Learning Techniques



Lekshmi Kaimal and Ramesh Kulkarni

## 1 Introduction

Insulation breakdown is the foremost cause of malfunctions in high-voltage equipment. Partial discharge (PD) is a major indicator of insulation failures. Partial discharges are low-energy signals that develop at the most stressed of the insulator. It does not cause any damage to the insulation immediately. However, it causes irreversible damage of physical, chemical, and dielectric and other electrical properties in that stressed part of the insulator (ageing). This damage triggers more such stressed sites and in turn PD occurrence. With each PD, the insulating material deteriorates further. This causes rapid degradation of insulation leading to electrical breakdown.

PD analysis is extensively used for the diagnosis of the health of the insulation of HV components. PD analysis includes identification of the defect, its severity assessment, and its localization. Partial discharge occurring in different insulation materials had different characteristics depending upon the deterioration mechanism. It is essential to recognize the cause of the discharge seen in high-voltage equipment so that the appropriate maintenance measure can be taken to avoid further damage to the insulation.

A lot of research has been undertaken in the field of automatic PD source classification. In practical cases, more than one discharge source may be present and active within the electrical equipment. These PD sources can be either of the same type or

---

Supported by Organization VES Institute of Technology, Mumbai, India.

---

L. Kaimal (✉)

VES Institute of Technology, (Affiliated to University of Mumbai), Mumbai, India  
e-mail: [phd14.ajeshlekshmi@ves.ac.in](mailto:phd14.ajeshlekshmi@ves.ac.in)

R. Kulkarni

Department of Electronics and Telecommunications, VES Institute of Technology,  
(Affiliated to University of Mumbai), Mumbai, India  
e-mail: [ramesh.kulkarni@ves.ac.in](mailto:ramesh.kulkarni@ves.ac.in)

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2022  
T. Sengodan et al. (eds.), *Advances in Electrical and Computer Technologies*,  
Lecture Notes in Electrical Engineering 881,  
[https://doi.org/10.1007/978-981-19-1111-8\\_39](https://doi.org/10.1007/978-981-19-1111-8_39)

different discharge types, but at different sites and of varying dimensions. Hence, for effective health diagnosis of in-service power assets, it is vital to first detect the number of active PD phenomena and then recognize the corresponding sources. *The objective of this paper is to describe a tool that can accurately estimate the number of active PD sources and separate the data into clusters for each source.* The information obtained from this tool can then be used with PD classification algorithms to recognize the corresponding sources.

### 1.1 Literature Survey on PD Source Classification

Figure 1 shows the typical PD spectroscopy workflow which forms an integral part of condition monitoring system.

It consists of three major stages. Data acquisition and pre-processing is the first step in which experimental PD pulse data is acquired using the digital PD detectors and filtered. Then, in the next stage, an identification set or PD fingerprint data pattern or feature vector is generated from the pulse data acquired from the experiments. This is followed by a pattern recognition classification algorithm for defect separation and classification, severity assessment, and localization.

The research on automatic PD source classification has been carried out over many years now [1–3]. There is correlation between the PD fault and their corresponding phase resolved partial discharge (PRPD) patterns. This has been used along with the pattern recognition algorithms to categorize according to the PD sources. However, these methods are effective in recognizing the defect if only one single source is present and active at a time. However, in most of the practical scenarios, more than one active PD source is seen in insulation. In such a case, PRPD patterns are often seen to be partially overlapped [4]. The automatic PD classification system needs to be enhanced such that the multi-source PDs can also be successfully classified. The methodology adopted in the instance of multiple active sources is to first separate the clusters associated with the different sources. A feature vector needs to be created from the raw data acquired from the experiments or from any other analysis done on the raw data so that the clustering is done accurately. Then, the cluster information can be further analysed for PD classification and localization.

The first multi-source PD signals defect recognition method was developed using the mixed Weibull distribution fit to the amplitude distribution (charge of the PD vs. count plot) [5]. The pattern classification algorithm is implemented based on the PRPD patterns of a multi-source system in [6]. A two-step logistic regression (LR) algorithm is then performed. The single-source PRPD patterns data is used

**Fig. 1** Block diagram of PD source classification



to train the two-step LR algorithm while the multi-source PRPD pattern sample is used to test the algorithm. Finally, kernel support vector machine (KSVM) is used for classification of new samples [6]. Other than PRPD patterns, PD pulse waveform data can be used to identify the fault causing the discharges and also to separate out noise sources. The features—power ratio (PR) maps [7] and three-dimensions maps [8] can be used to filter the noise from the PD signal. Feature vector based on signal processing of PD signals, like fast Fourier transform [7], pulse waveform comparison [9, 10], power spectrum density [11], and time-frequency analysis techniques [4], have also been used. The wavelet decomposition [8], S-transform [12], and mathematical morphology [4] are the other approaches that have been employed. Lately, novel characteristics such as the cumulative energy function [13], bag of words (BoW) [12], entropy [14], a high-order cumulant called bispectra [15] and separation algorithms like blind signal separation (BSS) [16, 17], and improved density peak clustering (IDPC) [18] have also been employed. New evaluation techniques like the employment of horizontal visibility graph spectral analysis (HVGSA) have also been explored [19]. The linear prediction analysis (LPA) algorithm generates a 12-dimensional feature vector from the raw data. Principal component analysis (PCA) then reduces this feature vector to two dimensions. To facilitate noise removal, an extra attribute called height score was added based on the isolation forest (IF) algorithm. This universal separation methodology can be utilized with the data acquired using different PD detection techniques [20].

In practical scenarios, multiple sources of PDs can be present and active. It is not possible to guess the number of active sources. It is even more difficult to know the associated PD types that could be seen in advance. Thus, identifying the number of PD sources in data acquired from the field is an important step in the multi-source PD classification. One such effort in this direction is done based on the BSS method via eigenvalue decomposition. The excess kurtosis value is used for estimating the number of sources [17]. Since the specific type of PD sources expected are not known, unsupervised algorithms need to be used for separating the data based on the source. Usually, most of these algorithms need to be provided with the exact number of clusters anticipated.

*The objective of this paper is to use an unsupervised clustering method that can pick the number of active components automatically given the input regarding the maximum limit on the number of sources. This technique will efficiently estimate the multiplicity of sources in the data acquired from the field and separate them into clusters. Each of these clusters can be then used with a PD source classification algorithm to identify the corresponding source of each cluster. This can be followed by subsequent analysis such as localization.*

The paper is organized as follows: Sect. 2 presents the proposed methodology for clustering. A brief description of the different clustering algorithms that can be used on the data set is presented in Sect. 3. The data set of Weibull distribution PDHD is generated for analysis, the details of which are discussed in Sect. 4. The proposed method is applied to the data, and its effectiveness in estimating the number of probable active PD phenomena is evaluated and further clustering is done. These results are discussed in Sects. 5 followed by conclusion in Sect. 6.

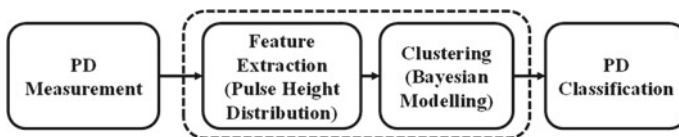
## 2 Proposed Methodology

In multiple, simultaneously activated PD source scenarios, it would be difficult to categorize each data point into a particular label. This can affect the accuracy of the PD classification algorithm. The different clusters corresponding to the various PD sources needs to be first identified. Each of these clusters can be fed to the PD classification model to categorize based on the PD sources accurately [20, 21]. However, the specific PD sources that could be present in unknown data acquired from the field are not known in advance. Hence, unsupervised learning algorithms need to be used to evaluate and categorize the unlabeled data sets.

Figure 2 shows the proposed methodology for a multi-source system. An unsupervised algorithm is implemented in two stages. In the first step, the characteristic feature that can accurately categorize the different PD sources is identified. A clustering algorithm is employed to categorize the clusters based on the selected parameter [22]. These separated clusters can then be applied to source classification algorithms to identify the PD phenomenon associated with each cluster. This work is limited to feature extraction and clustering phases of the workflow as explained in Fig. 2.

The PDs caused by different sources exhibit different PDHD patterns [5, 23]. The Weibull distribution which is widely used for lifetime distributions in reliability engineering is applied to the PDHD patterns. The shape parameter of the Weibull distribution is seen to be associated with the different PD sources. In the case of overlapping PD sources, the mixed Weibull model can be used. A mixed Weibull function can be applied to divide the PDHD pattern into sub-distributions relevant to each PD source [23]. In this paper, Weibull distributed PD pulse height is chosen as the characteristic feature for clustering.

In a practical case, it is difficult to predict the number of PD sources that would be simultaneously active. However, most of the unsupervised clustering algorithms ought to be provided with the exact number of clusters anticipated, whereas Bayesian mixture modelling automatically chooses the suitable number of effective components. The maximum number of sources anticipated is given as the prior information to start with in the BM algorithm. The paper also compares the clustering results obtained using the k-means, Gaussian mixture (GM), and Bayesian mixture (BM) modelling. For the benefit of the interest of the researchers, a brief explanation of the various clustering algorithms implemented in the paper is given in Sect. 3.



**Fig. 2** Block diagram of PD multi-source classification—proposed methodology

### 3 Clustering Algorithms for PD Sources Discrimination

Clustering analysis separates a set of data into meaningful groups, where there is some similarity between the observations within each subset [24]. The clustering algorithms can be of two types: model-based and model-free. A distance (or a similarity measure) is used to cluster the data into labels in the model-free type of approach. In cases where each cluster of the data set is seen to follow a probability distribution, a model-based clustering algorithm can be used to model the distribution of each subset [24].

#### 3.1 *K-Means*

K-means clustering is a model-free clustering technique where the data is divided into 'k' clusters [25]. The objective is to find a centroid or centre of the cluster, which can characterize all the points within a sub-population and need not be real data points. The clusters are finalized based on the minimization of a well-defined cost function.

Let  $D = \{x_1, x_2, x_3, \dots, x_N\}$ , of  $N$  unlabelled data points consisting of  $k$  clusters, where  $k$  is chosen by the user. The following cost function is minimized.

$$V = \sum_{i=1}^k \sum_{x_j \in C_i} (x_j - \mu_i)^2 \quad (1)$$

Here,  $C_i$  represents each of the  $k$  clusters and  $\mu_i$  are the corresponding cluster centroids. The algorithm randomly selects the  $k$  centroids initially and assign each data point to the closest centroid. It then recalculates the cluster centre based on the new allocations of the data points to the clusters iteratively and reassigns until convergence.

#### 3.2 *Finite Mixture Model*

The finite mixture model is being extensively being employed in various fields like biology, medicine, finance, astronomy, where each of the data clusters follows a particular distribution. This model facilitates identifying the individual distributions and also their mixing probability. The mixture of various distributions such as normal, Weibull, Poisson, Gamma, and exponential can be fitted by a finite mixture model. The determination of a correct number of sub-components in the data is a major task in the implementation of mixture models [26].

The description of a mixture of distributions is given by:

$$\sum_{i=1}^K p_i f_i(x), \text{ where } \sum_{i=1}^K p_i = 1, k > 1 \quad (2)$$

where  $f_i$  is a mixture. In most cases, the  $f_i$ 's are from a parametric family, with unknown parameter  $\theta_i$  [27]. Thus, now the mixture is modelled as follows:

$$\sum_{i=1}^K p_i f(x|\theta_i) \quad (3)$$

### 3.3 Gaussian Mixture Modelling

The major challenge of using the k-means clustering technique for finite mixture models is that it does not account for variance. GM overcomes this limitation and can be visualized as generalized k-means clustering which incorporates covariance information of the data and means of the normal distribution. The k-means clustering performs hard classification, i.e. the information on which data point belongs to which cluster is provided. However, it does not provide the probability with which a given data will belongs to each of the probable sub-population. Such a feature is essential in cases of mixture distributions where there would be overlaps between the boundaries of each distribution. This information is given by GM models which perform soft classification [28].

Gaussian mixture clustering models the data set as a linear combination of multivariate Normal distributions. GM non-parametric method uses an expectation-maximization algorithm, which maximizes a quality measure called log-likelihood to find the distribution parameters.

Suppose,  $\{x_1, x_2 \dots x_N\}$  are data samples made from  $k$  Gaussian distributions, i.e.  $k$  distinct classes [28]. Every distribution is multiplied by a weight  $\pi$ .  $P(z_i = k|\Theta)$  is the likelihood that the  $i_{th}$  sample came from Gaussian  $k$ .  $\Theta$  represents parameters defining each distribution—mean, covariance, and weight, as given in (4).  $P(x_i|z_i = k, \Theta)$  is the likelihood of observing a data point that belongs to the distribution number  $k$ . It is also given as,

$$P(x_i|z_i = k, \Theta) = \pi_k \quad (4)$$

where,  $\Theta = \mu_1, \mu_2, \dots \mu_K, T_1, T_2, \dots T_K, \pi_1, \pi_2, \dots \pi_K$

To take into account all the different distributions, the likelihood of observing the sample  $i$  is given by sum of likelihoods of observing the sample given that it belongs to each of the probable distribution. It is given by  $\sum_{j=1}^k P(x_i|z_i = k, \Theta)P(z_i = k|\Theta)$ .

In order to extend this to all  $N$  samples in the data set, the likelihood function is given by,

$$\begin{aligned} L(\Theta) &= P(x|\Theta) = \prod_{i=1}^N \sum_{j=1}^k P(x_i|z_i = k, \Theta) P(z_i = k|\Theta) \\ &= \prod_{i=1}^N \sum_{j=1}^k N(x_i|\mu_k, T_k)\pi_k \end{aligned} \tag{5}$$

$$\log(P(x|\Theta)) = \prod_{i=1}^N \log\left(\sum_{j=1}^k N(x_i|\mu_k, T_k)\pi_k\right) \tag{6}$$

The major advantage of the GM algorithm is its speed of learning the mixture models. However, since this algorithm involves the estimation of the covariance matrices of each cluster, a large data set with sufficient data per mixture is required. The GM algorithm requires precise information regarding the number of component distributions are present in the mixture since it uses all the components it has access to.

### 3.4 Bayesian Mixture Modelling

Bayesian mixture modelling is a soft clustering method and can also give the number of components in each cluster [25]. Bayesian inference is based on Bayes' theorem [29]. Bayes' theorem allows the use some knowledge that is already present (the 'prior') along with the learning from the observation (the 'likelihood') to estimate the probability of an associated event, (the updated knowledge—the 'posterior') [29].

Assuming that the prior knowledge about the parameter  $\Theta$  is known and that prior is described by probability distribution  $p(\Theta)$ , i.e. how probable are each value of the parameter before any observation. Then, when data  $x$  is observed, the prior knowledge about this parameter can be updated using the Bayes theorem as follows

$$p(\Theta|x) = \frac{p(x|\Theta)p(\Theta)}{p(x)} \tag{7}$$

$p(x|\Theta)$  denotes the likelihood function, i.e. how probable are the observed data for this parameter value while  $p(x)$  represents the probability of the evidence which is constant, how probable is it to observe these particular data [30].  $p(\Theta|x)$  gives the posterior distribution. It gives the updated knowledge of how probable are each value of the parameter given the observed data. An inference about the number of active components can also be made from the data.

The Bayesian model is an extension of the Gaussian model wherein prior knowledge of the parameters—mixing coefficients ( $\pi$ ) and the parameters of the component distributions ( $\mu, T$ ), as well as over the number of components in the mixture—are expressed in terms of probability distributions [26].



### 3.4.1 Bayesian Mixture Algorithm

In BM, we also consider latent variables  $s$  which are discrete distributions of mixing coefficients  $\pi$ . Maximum likelihood method is used for obtaining the optimum values of the mixing coefficients  $\pi$  [26]. The joint distribution of the random variables given the mixing coefficients is given by

$$p(x, \pi, T, s|\pi) = p(x|\pi, T, s)p(s|\pi)p(\mu)p(T) \quad (8)$$

The advantage of this method is that the variational Bayesian mixture model normally fixes some of the mixture weights close to zero. This helps the model to select an appropriate number of sub-populations automatically. For this, the model needs to be supplied with the maximum threshold to this number. However, the extra inputs required by variational inference algorithm, make the inference model slower.

## 4 Data Set Generation Using Weibull Distribution

Having identified BM as the suitable clustering algorithms for estimating the multiplicity of PD sources, the efficiency of the algorithm can be verified by implementing it on the data set of PDHD pattern from known number of sources. For this purpose, we have generated a data set using Weibull distribution [23]. It has been noted that Weibull distribution can also be used for identification of PD phenomena in practical insulation systems such as rotating machine stators, particularly in the presence of simultaneously acting PD sources have been.

The PDHD pattern resulting from a single PD source can fit a single Weibull function. The Weibull cumulative distribution function,  $F(q)$  is defined as [5, 16, 17].

$$F(q) = 1 - \exp\left[-\left(\frac{q}{\alpha}\right)^\beta\right] \quad (9)$$

Here,  $q$  is the charge of the PD pulses. The shape parameter  $\beta$  and  $\alpha$ , the scale parameter together characterizes the Weibull distribution.

In the presence of more than one active PD phenomena, a mixed Weibull distribution is applied. The cumulative  $n$ -parameter mixed Weibull distribution for  $n$  active sources is given by,

$$F(q) = \sum_{i=1}^n p_i F_i(q) = \sum_{i=1}^n p_i \left(1 - \exp\left[-\left(\frac{q}{\alpha_i}\right)^{\beta_i}\right]\right) \quad (10)$$

Here,  $F_i(q)$  gives the cumulative Weibull distribution of the sub-populations which describes each PD source.  $p_i$  is the probability of occurrence of the corresponding sub-population  $F_i(q)$  with  $\sum_{i=1}^n p_i = 1$ .  $F_i(q)$  is given by Eq. 9. Thus, the model

**Table 1** Specimen generation using Weibull parameters as per [23]

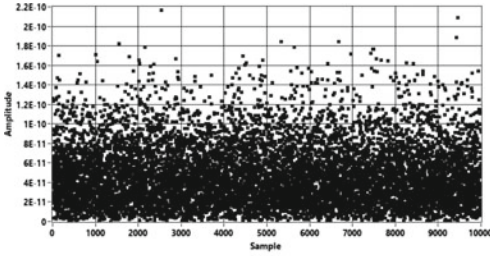
Specimen		Weibull model parameters								
Name	PD source	$p_1$	$\alpha_1$	$\beta_1$	$p_2$	$\alpha_2$	$\beta_2$	$p_3$	$\alpha_3$	$\beta_3$
A1	PMMA sheet	1	50.51	1.63	–	–	–	–	–	–
B1	Spherical void (2mm dia) in 3.5 mm thick Ciba Geigy epoxy resin	0.09	658.77	5.57	0.9	571.03	56.98	–	–	–
A1B1	A1 & B1 in parallel	0.05	658.76	5.00	0.37	570.99	68.05	0.58	50.06	1.62
A2	PMMA sheet	1	64.08	1.60	–	–	–	–	–	–
B2	Spherical void (1.5 mm dia) in 3.5 mm thick Ciba Geigy epoxy resin	0.72	141.37	23.58	0.28	149.02	14.39	–	–	–
A2B2	A & B2 in parallel	0.15	140.71	23.52	0.12	149.69	11.98	0.72	60.99	1.65

for two simultaneously active PD sources is characterized by 5 parameters, i.e.  $\alpha_1$ ,  $\beta_1$ ,  $\alpha_2$ ,  $\beta_2$  and  $p$  [5].

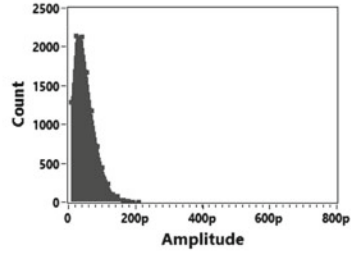
In this paper, Weibull distributed PD pulse height as in [23] is used for clustering. In [23], experiments were conducted on samples of epoxy resin containing more than one internal void under varying test conditions. Tests were carried out on the two specimens, one with a flat air gap and another with a spherical void, individually. These two specimens were then connected in parallel and tested. Weibull parameters were fitted into the amplitude distributions obtained from experiments for analysis [23]. These parameters as mentioned in [23] (summarized in Table 1) were used to generate the data required for analysis in this work. It is noted in [23] that the Weibull fitting is implemented either on the positive discharges or on negative ones. Here, in this work, the Weibull parameters are used to generate positive polarity discharges.

Specimen A consists of Cigre II with a polymethylmethacrylate (PMMA) sheet. The second specimen is a cell with an embedded spherical void in the middle of a Ciba Geigy epoxy resin dielectric with a thickness of 3.5 mm. This is named specimen B. Two different versions of specimen B were used—B1 with a 2 mm diameter spherical cavity while B2 with a diameter of 1.5 mm spherical cavity. Two samples of specimen A were used in experiments, with each subversion- B1 and B2. The Weibull parameters fitted to both of these samples had slight variations. Hence, here, in this work, the samples of specimen A are taken to be A1 and A2. Each of the two versions of B were connected in parallel with the A specimen samples and tested, giving A1B1 and A2B2 groups, respectively.

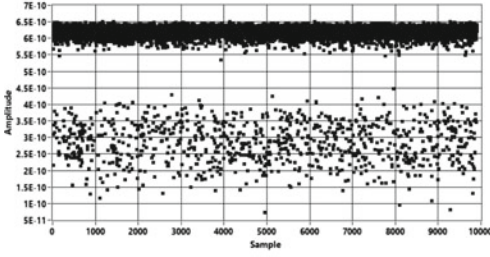
Figure 3 shows the scatter and the amplitude distribution plots of the specimens used. Specimen A sees the lower amplitude pulses with values less than 200 pC. It is seen that amplitude values in B1 (around 600 pC) are greater than those in B2 (around 200 pC). This difference is because of the variation in the diameter of the spherical void in the two specimens. This causes the distributions of the two components in A2B2 to be overlapped. This specimen will thus help in evaluating the efficiency of the proposed design to separate overlapping distributions.



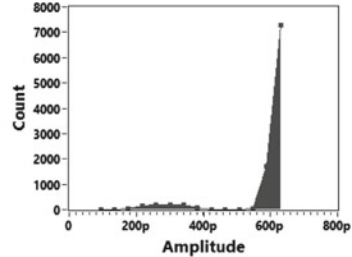
(a) Specimen A - Scatter plot



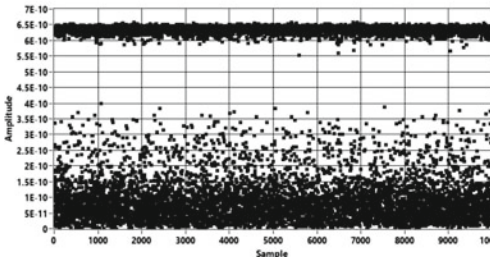
(b) Specimen A - histogram plot



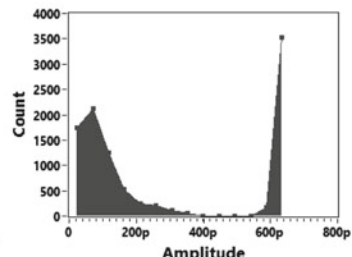
(c) Specimen B1 - Scatter plot



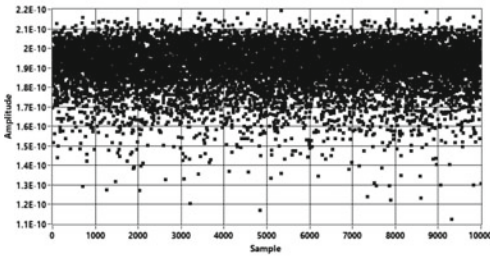
(d) Specimen B1 - histogram plot



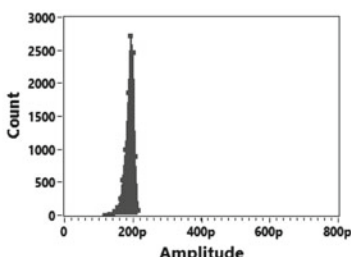
(e) Specimen A1B1 - Scatter plot



(f) Specimen A1B1 - histogram plot



(g) Specimen B2 - Scatter plot



(h) Specimen B2 - histogram plot

Fig. 3 Scatter plot and histogram of the specimens

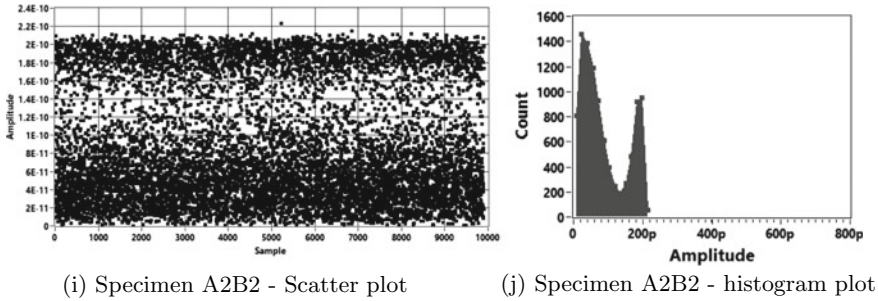


Fig. 3 (continued)

## 5 Results and Discussion

The data set for each specimen generated is first applied to BM to estimate the number of sub-components. This number is then used while clustering the data using three different algorithms—k-means, GM, and BM models.

The GM model fits the mixture data by applying the expectation-maximization (EM). It is noted that the clustering on the data set generated is more effective when it is assumed that all the sub-populations have the same general covariance matrix. The estimation algorithm used to implement BM model is variational inference. The weights, the means, and the covariances are initialized randomly while the prior for the weight's distribution in the mixture model is assumed to follow the Dirichlet process in this work for effective clustering. The BM also estimates the number of components in each specimen. The upper bound of 10 components was given as an input.

The BM model returns the weights related to each of the components. Table 2 gives the active components as estimated by BM for each specimen and also the time taken by the algorithm to arrive at the estimate. It can be seen that the number of active components has been detected accurately. The weight of each component esti-

**Table 2** Active components as per BM for each specimen

Specimen	Weight of each component Using BM	Active components	Time taken (s)
B1	0.091, 0.909, 0, 0, 0, 0, 0, 0, 0, 0	2	24
A1B1	0.558, 0.369, 0.072, 0, 0, 0, 0, 0, 0, 0	3	77
B2	0.799, 0.201, 0, 0, 0, 0, 0, 0, 0, 0	2	21
A2B2	0.551, 0.281, 0.168, 0, 0, 0, 0, 0, 0, 0	3	60

mated also matches with the probability of each weibull sub-component ( $p_1, p_2, p_3$ ) provided to generate the original data set (seen in Table 1).

Figure 4 gives the scatter plot of the specimens after clustering. It is seen that in B2 and A1B1, the amplitudes values of two clusters are very close, and there is no proper demarcation between the clusters while, in A2B2, all the three sub-components are overlapping. Hence the best way to find if the clustering is proper is to use the ‘number of datapoints in each cluster’ as the parameter. If the number of data points seen under each label after clustering is same or close to the ones seen in the input original data set, it can be said that the clustering is accurate.

The number of samples detected under each class is noted, when different clustering algorithms were used. Table 3 gives the expected number of samples under each cluster and in the output after clustering. From this table, it is seen that in B1, all three methods give good results, since there is no overlapping between clusters. However, in the case of overlapping clusters—B2 and A1B1, the size of each cluster as seen by BM is closest to the expected quantity of data points, while k-means and GM give a decent accuracy. This can be seen from the scatter plot of Specimens B2 and A1B1 (Fig. 4i, f, respectively). For the specimen A2B2, it can be seen from Table 3 and Fig. 4j, k, l, that the k-means clustering fails to properly differentiate the datapoints into correct cluster, while BM and GM are able to cluster with a better accuracy, given by the amount of data points in each cluster.

Thus, it can be seen that BM applied to the amplitude distribution does a good job of estimating the number of components active in the sample and also cluster accurately, even in case of overlapping distributions.

## 6 Conclusion

The practical scenario mostly involves more than one simultaneously activated PD source. A method has been proposed in this paper to accurately estimate the number of active PD sources and separate the data into clusters for each source. This information can then be used with PD classification algorithms to recognize the corresponding sources. BM algorithm can estimate the number of clusters in the data set on its own, which is unlike other clustering algorithms that need input regarding the number of components in the sample. The PDs from different sources is seen to exhibit different PD pulse height distribution (PDHD) patterns. The PDHD pattern associated with a single PD phenomenon can fit a single Weibull function while in presence of more than one active PD phenomenon, a mixed Weibull distribution can be used.

In this paper, Weibull distributed PD pulse height data set was generated for different specimens. In the first step, the BM model estimates the weight of each component in the data set of each specimen. The weights of each component are estimated to match the probability of each Weibull sub-component provided to generate the original input data set. Then, the input data was subjected to three different clustering algorithms—k-means, Gaussian mixture, and Bayesian mixture models. In some specimens, the amplitudes values of two or more clusters are similar and the

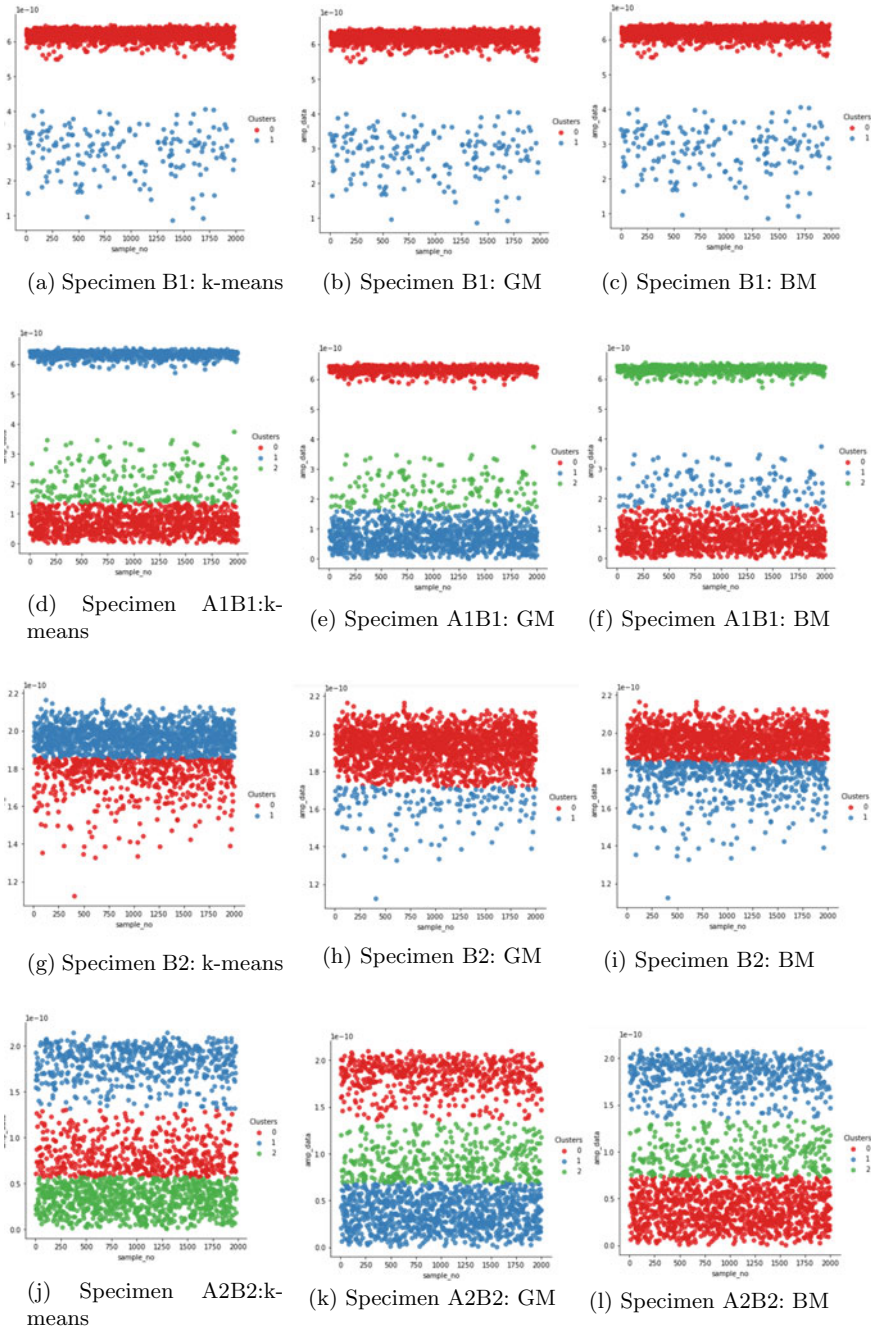


Fig. 4 Scatter plot after clustering for all specimens

**Table 3** Output results for specimen

Specimen	No. of data points in each cluster											
	Expected			Using k-means			Using GM			Using BM		
	Clr 1	Clr 2	Clr 3	Clr 1	Clr 2	Clr 3	Clr 1	Clr 2	Clr 3	Clr 1	Clr 2	Clr 3
B1	1780	178	–	1800	180	–	1800	180	–	1800	180	–
A1B1	1160	740	100	1026	740	234	1100	740	160	1110	740	150
B2	1440	560	–	1382	618	–	1836	164	–	1439	561	–
A2B2	1425	249	238	871	574	535	1063	564	373	1116	564	320

distributions overlap. Hence, the number of data points seen under each label after clustering is compared with the ones seen in the input original data set. This is used to estimate the efficiency of the clustering. K-means clustering gives good accuracy in the specimen where there is no overlapping in distribution, while in the case of overlapping distributions, it fails to cluster accurately. The GM and BM clustering is more efficient even in presence of overlapping distributions. The size of each cluster as seen by BM is closest to the expected quantity of data points considering all the specimens.

Thus, it can be seen that BM applied to the amplitude distribution can be used to estimate the number of components active in the sample and also for divide the data set into clusters associated with each PD source. Thus, the method described in this paper will be a valuable tool to recognize the cause of the discharge seen in high-voltage equipment, even in presence of the simultaneously activated PD phenomena. This is crucial for scheduling the appropriate maintenance measure to be taken to avoid further damage to the insulation and the electrical equipment.

## References

1. Janani H, Kordi B (2018) Towards automated statistical partial discharge source classification using pattern recognition techniques. *High Voltage* 3(3):162–169
2. Chan JC, Saha TK, Ekanayake C (2013) Pattern recognition techniques and their applications for automatic classification of artificial partial discharge sources. *IEEE Trans Dielectrics Electrical Insulation* 20(2):468–478
3. Peng X, Zhou C, Hepburn D et al (2013) Application of K-means method to pattern recognition in on-line cable partial discharge monitoring. *IEEE Trans Dielectrics Electrical Insulation* 20(3):754–761
4. Chan JC, Ma H, Saha TK (2015) Time-frequency sparsity map on automatic partial discharge sources separation for power transformer condition assessment. *IEEE Trans Dielectrics Electrical Insulation* 22(4):2271–2283
5. Cacciari M, Contin A, Montanari GC (1995) Use of a mixed-Weibull distribution for the identification of PD phenomena. *IEEE Trans Dielectrics Electrical Insulation* 2(6):1166–1179
6. Janani H, Kordi B, Jozan MJ (2017) Classification of simultaneous multiple partial discharge sources based on probabilistic interpretation using a two-step logistic regression algorithm. *IEEE Trans Dielectrics Electrical Insulation* 24(1):54–65

7. Ardila-Rey JA, Martínez-Tarifa JM, Robles G, Rojas-Moreno MV (2013) Partial discharge and noise separation by means of spectral power clustering techniques. *IEEE Trans Dielectrics Electrical Insulation* 20(4):1436–1443
8. Hao L, Lewin PL, Hunter JA, Swaffield DJ, Contin A, Walton C, Michel M (2011) Discrimination of multiple PD sources using wavelet decomposition and principal component analysis. *IEEE Trans Dielectrics Electrical Insulation* 18(5):1702–1711
9. Contin A, Pastore S (2009) Classification and separation of partial discharge signals by means of their auto-correlation function evaluation. *IEEE Trans Dielectrics Electrical Insulation* 16(6):1609–1622
10. Pinpart T, Judd MD (2010) Differentiating between partial discharge sources using envelope comparison of ultra-high-frequency signals. *IET Sci Measurement Technol* 4:256–267
11. Robles G, Parrado-Hernández E, Ardila-Rey J, Martínez-Tarifa JM (2016) Multiple partial discharge source discrimination with multiclass support vector machines. *Expert Syst Appl* 55:417–428
12. Firuzi K, Vakilian M, Darabad VP, Phung BT, Blackburn TR (2017) A novel method for differentiating and clustering multiple partial discharge sources using S transform and bag of words feature. *IEEE Trans Dielectrics Electrical Insulation* 24:3694–3702
13. Zhu MX, Liu Q, Xue JY, Deng JB, Zhang GJ, Shao XJ et al (2017) Self-adaptive separation of multiple partial discharge sources based on optimized feature extraction of cumulative energy function. *IEEE Trans Dielectrics Electrical Insulation* 24:246–258
14. Mitiche I, Morison G, Nesbitt A, Hughes-Narborough M, Stewart BG, Boreham P (2018) Classification of partial discharge signals by combining adaptive local iterative filtering and entropy features. *Sensors* 18:14
15. He Z, Hou H, Zang Y, Sheng G, Jiang X (2020) A method for separating multisource partial discharges in a substation based on selected bispectra of UHF signals. *Appl Sci* 10(11):3751
16. Boya C, Robles G, Parrado-Hernandez E, Ruiz-Llata M (2017) Detection of partial discharge sources using UHF sensors and blind signal separation. *Sensors* 17:13
17. Au M, Agba BI, Gagnon F (2015) Fast identification of partial discharge sources using blind source separation and kurtosis. *Electron Lett* 51(25):2132–2134
18. Li Y, Li Z (2020) Improved density peak clustering for separation of multiple source partial discharge in large generators. *Measurement Sci Technol* 31(1):15013
19. Roy SS, Chatterjee S (2021) Partial discharge detection framework employing spectral analysis of horizontal visibility graph. *IEEE Sensors J* 21(4):4819–4826
20. Wang Y, Chang D, Qin S, Fan Y, Mu H, Zhang G (2020) Separating multi-source partial discharge signals using linear prediction analysis and isolation forest algorithm. *IEEE Trans Instrumentation Measurement* 69(6):2734–2742
21. Wu M, Cao H, Cao J, Nguyen N, Gomes JB, Krishnaswamy SP (2015) An overview of state-of-the-art partial discharge analysis techniques for condition monitoring. *IEEE Electrical Insulation Mag* 31(6):22–35
22. Long J, Wang X, Zhou W, Zhang J, Dai D, Zhu G (2021) A comprehensive review of signal processing and machine learning technologies for UHF PD detection and diagnosis (I): preprocessing and localization approaches. *IEEE Access* 9:69876–69904
23. Schifani R, Candela R (1999) A new algorithm for mixed Weibull analysis of partial discharge amplitude distributions. *IEEE Trans Dielectrics Electrical Insulation* 6(2):242
24. Rodríguez CE, Núñez-Antonio G, Escarela G (2020) A Bayesian mixture model for clustering circular data. *Comput Stat Data Anal* 143:106842
25. Heller A (2007) Efficient Bayesian methods for clustering. Katherine PhD thesis, University of London
26. Corduneanu A, Bishop C (2001) Variational Bayesian model selection for mixture distributions. In: *Proceedings eighth international conference on artificial intelligence and statistics*, pp 27–34
27. Marin J-M, Mengersen K, Robert C (2005) Bayesian modelling and inference on mixtures of distributions. In: *Handbook of statistics*, vol 25. [https://doi.org/10.1016/S0169-7161\(05\)25016-2](https://doi.org/10.1016/S0169-7161(05)25016-2)



28. Maklin C (2020) Gaussian mixture models clustering algorithm explained. <https://towardsdatascience.com/gaussian-mixture-models-d13a5e915c8e>. Last accessed 1 June 2020
29. Brooks-Bartlett J (2020) Probability concepts explained: Bayesian inference for parameter estimation. <https://towardsdatascience.com/probability-concepts-explained-bayesian-inference-for-parameter-estimation-90e8930e5348>. Last accessed 1 June 2020
30. Phoong S-Y, Ismail MT (2014) A study of finite mixture model: Bayesian approach on financial time series data. In: AIP conference proceedings, 1605, 805

# Comparison of Encryption Techniques to Encrypt Private Parts of an Image



Nisha P. Shetty, Balachandra Muniyal, Rithish Reddy Kaithi,  
and Sarath Chandra Reddy Yemma

## 1 Introduction

Along with everyday advances in modern technology, the usage of image data among the numerous amount of systems and devices has increased. Today, systems like the IoTs, CCTVs, and drones having built-in imaging devices can obtain images of public and automatically store and share them. Thus, increasing use of the Internet by the public and the availability and sharing of public and private digital data have led industry experts and researchers to give special attention to information security. And there is a need to protect this private information from exponentially growing unauthorized access and attacks. Cryptography is about building, studying, and analyzing protocols that prevent third parties or the public from accessing private information. Main objective of cryptography techniques is to make the content of a message incomprehensible to unauthorized people. Every enciphering/encryption and deciphering/decryption process have two main aspects: the scheme of algorithm and the usage of keys for enciphering/encryption and deciphering/decryption. However, the security of the cryptographic process is maintained by the key used to encrypt and decrypt the information. There are two kinds of mechanisms in cryptography: Symmetric key cryptographic mechanism, in which equal key is used to encode and decode. In the asymmetric key cryptographic mechanism, two distinct keys are used to encrypt and decrypt. The algorithm for symmetric key is much faster, easier to implement, and requires less computing power compared to the asymmetric key algorithm. Steganography techniques are supported by concealing the existence of data by embedding the key message in another cover medium. A map displaying some kind of chaotic behavior is a chaotic map. A chaotic map is a difference equation describing a chaotic dynamical system which is discreet. A chaotic map is

---

N. P. Shetty (✉) · B. Muniyal · R. R. Kaithi · S. C. R. Yemma  
Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal 576104, India  
e-mail: [Nisha.pshetty@manipal.edu](mailto:Nisha.pshetty@manipal.edu)

normally a discrete map that has sensitive dependence on initial conditions. Due to their appealing features like sensitivity to the initial conditions and random spreading out behavior, chaotic maps are adopted for several applications to protect the data. The primary target of this paper is to execute a picture encryption and decoding utilizing both methods and compare the experimental results and security analysis.

## 2 Related Work

Amal Hafsa et al. [3] propose an encryption scheme that includes the “advanced encryption standard (AES)” and “elliptic curve cryptography (ECC)” for encrypting medical images, combining the advantages of symmetric AES for the acceleration of encryption of data and asymmetric ECC for the secure exchange a symmetric session key. The analysis results prove the efficiency, speed, and better security of this algorithm. Arab et al. [4] propose a combination of both modified AES algorithm and chaotic sequence to develop a novel image encryption algorithm. Modified AES scheme is used to encrypt the original image using the encryption key, which is generated with the help of Arnold chaos sequence, and chaotic sequence produced round keys implementation. It is observed from the results from simulation that when small changes are introduced into the original image and cipher key, it results in a consequential change in the enciphered image making the original image inaccessible. Alsaffar et al. [5] compare “advanced encryption standard (AES)” and “Rivest-Shamir-Adleman (RSA)” algorithms for image encryption. Results from tests reveal that image encryption quality is better with AES algorithm, with uniform distribution of pixels in histogram. Moreover, coefficient of correlation is closer to 0 for AES algorithm. The overall results of this study conclude that the AES algorithm is a better algorithm than RSA algorithm for image encryption. Farah et al. [6], to enhance the performance of encryption algorithms, proposed a novel hybrid chaotic map and a distinct way of utilizing optimization technique. In the proposed algorithm, S-boxes based on their respective nonlinearity score are generated accordingly using chaotic Jaya optimization algorithm to carry out a new optimized S-box dependent substitution phase. Results from security analysis reveal that the proposed encryption scheme is resistant to various unauthorized attacks. Ping et al. [7] proposed discrete Henon map, attaining a novel two-point diffusion techniques. The key stream for substitution is generated and made based on the original image to obtain the greater capability of showing resistance against chosen-plaintext attack or known-plaintext attack. The results from security analysis reveal that this encryption scheme provides a superior security, meanwhile, it is observed that faster encryption speed is achieved compared to various image encryption schemes from the time complexity analysis. Shadangi et al. [8] propose a novel CBC-AES algorithm to encrypt images based on Arnold scrambling having several encryption levels. In this method, at first, the original image is shifted circularly and Arnold scrambled together with a bit-wise shuffle operation, after which each bit is complemented. Finally, CBC mode of AES encryption scheme was employed to attain the resulting cipher image.

Furthermore, it is evident from the results of security analysis that this algorithm has ability of resisting various attacks like differential attacks, statistical attacks, and entropy attacks. Shaktawat et al. [9] propose a hybrid approach for image encryption by combining AES, a standard cryptography algorithm, along with splitting and block permutation. Comparison of various parameters with and without block permutation confirms the superiority of the proposed method with regard to better results after using splitting and permutation functions. Wang et al. [10] proposed chaotic encryption of image with a novel one-time pad design which is based upon the multiple mixed combinations of hash functions mixed and cyclic-shift function. By using both the data from the original image and the chaotic sequences, which are calculated using hash algorithms such as SHA1 and MD5 hash, the initial value is generated. The logistic map and the nonlinear equations generate the scrambling sequences. Experimental results and security analysis prove better security provided by this scheme and are protected from common attacks. Hua et al. [11] present a “two-dimensional (2D) logistic-sine-coupling map (LSCM)”. Furthermore, they propose an algorithm to encrypt image based on 2D-LSCM (LSCM-IEA) to adopt classical confusion-diffusion structure using the proposed 2D-LSCM. To interchange pixels of image to various rows and columns, a permutation algorithm is devised and to impart changes of original image to encrypted image, a diffusion algorithm is designed. The results from analysis reveal that better security performance than various algorithms is obtained by LSCM-IEA. Hua et al. [12] proposed an encryption scheme that uses highly efficient scrambling to split adjoining pixels and employing substitution in a random order to impart a small variation in the original image to each and every pixel of the encrypted image. The security analysis reveals that security provided by this encryption scheme is better than various modern image encryption algorithms.

### 3 Methods Used

#### 3.1 Combination of AES, RSA, LSB

- In December 2001, The National Institute of Standards and Technology (NIST) published “advanced encryption standard (AES)” as FIPS 197 (Federal Information Processing Standard). AES is a symmetric block cipher. The algorithm was developed by Joan Daemen and Vincent Rijmen, Belgian Researchers. The AES algorithm is easy to implement and uses keys of different length (128, 192, or 256 bits) according to number of rounds (10, 12, or 14, respectively) used in the implementation of the algorithm.

Each round of AES has 4 steps:

1. SubByte:

It uses the substitution box (s-box) to substitute the values of each byte. Each byte is interpreted as 2 digit hexadecimal value.

2. **ShiftRows:**

The permutation step of the method, each row of the matrix is shifted cyclically to the left.

3. **MixColumns:**

Each column of the matrix is multiplied by a constant to transform it into a new column.

4. **AddRoundKey:**

The step that uses the key, transforms each column by adding it to the key.

The process of encryption is done after the completing all the rounds (10 in our case).

- RSA algorithm named after its publishers Ron Rivest, Adi Shamir, and Leonard Adleman. It is a public key, asymmetric encryption algorithm where only the receiver has the private key. Both sender and receiver have the value of “n”. The public key is {e, n} which is used for encryption, and the private key is {d, n} used for decryption.

Where

$n = p * q$ , ( $p$  and  $q$  are prime numbers which are selected and kept private).

$e$ , chosen such that  $e, (p - 1) * (q - 1)$  are co-primes

$d = (e^{-1}) \text{ mod } ((p - 1) * (q - 1))$

Encryption of plaintext  $M$ :

$$C = M^e \text{ mod } n \quad (1)$$

Decryption of cipher text  $C$ :

$$M = C^d \text{ mod } n \quad (2)$$

- Steganography is a method of hiding information in the ordinary plain text or image to avoid detection. LSB method is a steganography method in which the message to be hidden replaces the least significant bit of the pixel values.

The combination of these methods is to cipher the 128 bit key, by which the image is encrypted in AES encryption, using RSA algorithm and places the obtained ciphertext into the image with LSB method.

### 3.1.1 Encryption

The values of the image are traversed pixel by pixel and in doing so it encrypts each value and append them to the new data array (new\_data). The method uses ciphertext feedback (CFB) mode of operation. The key of length 16 bytes is produced randomly, and the initialization vector (IV) of size 16 bytes is to be used.

- As each pixel contains r, g, b values are converted into bytes for the encryption.

```
cfb_cipher = AES.new (key, AES.MODE_CFB, IV)
k = bytes (pixel [0], pixel [1], pixel [2])
enc_data = cfb_cipher.encrypt(k)
```

The encrypted value is in bytes. So it is converted back into rgb values and places in new\_data.

```
Val = rgb (enc_data)
data_b[ind] = Val
```

Finally, a new encrypted image is formed from new\_data.

- The key used is converted into hexadecimal and then ciphered and converted into bits for LSB method. The values of n, e, d are declared ahead of implementing the method.

```
num=key.hex()
X = (num[i] **e) %n; i = 0 to length of num.
Message = binary(X) (converting it into binary)
```

- A constant string is appended to the message to specify the end in decryption process. Each bit of the message is placed in the least significant bit of the pixel value of the encrypted image formed by AES method.

```
rgb = rgb[:-1]+ message[i] ,where i = 0 to length of message
```

, rgb is value of a pixel.

The success of LSB method does not show changes of the encrypted image, before and after the implementation of the process, to the naked eye.

### 3.1.2 Decryption

The decryption process is similar to encryption process except we retrieve the key from the encrypted image and use it to decrypt it to get original image. Same key and initialization vector are to be used to get the original image.

- The below step is repeated until the constant string is encountered.

```
message += rgb[-1]
```

- The message is converted into hexadecimal for deciphering with RSA. Converting Y into bytes gives the key used to encrypt the image.

```
message = hex(message)
Y += (message[i] **d) %n; i = 0 to length of message.
Key = bytes(Y)
```

- Each r, g, b value of the pixel is converted into bytes and then decrypted to get the original value at the pixel.

```
cfb_cipher = AES.new (key, AES.MODE_CFB, IV)
```

```
k = bytes (pixel [0], pixel [1], pixel [2])
enc_data = cfb_cipher.decrypt (k)
```

The encrypted value is in bytes. So it is converted back into rgb values and places in new\_data.

```
Val = rgb (enc_data)
data_b[ind]=Val
```

- Finally, the image formed by the new\_data is the decrypted image.

### 3.2 Hash Function (Chaotic Method)

Any function that converts an input of variable length (generally a large value) into an output of fixed small range value is called hash function. In this method of encryption, we use SHA-2 (a 1D hash) algorithm to create a 2D mask. The SHA-2 operation can be divided into two parts:

- Pre-processing:

It includes padding of the original message of length  $< 2^{128}$  to make the length a multiple of 1024 bits along with the length of original message. The padded message is then decomposed into blocks of 1024 bits each. Initial values that are used in hash generation are initialized.

- Hash computation:

The final message digest is formed after a number of rounds of hash value that is produced using the padded message from the pre-processing step. It uses functions, word logics, and other operations to get hash functions.

This method uses substitution-diffusion type hash-based image encryption. This has four stages, two of which involve both substitution and diffusion while the other two involves only the diffusion process. In the process substitution, each pixel value is interpreted as hexadecimal value and is converted using s-box of the AES encryption method.

In this algorithm, the image is divided into 4 subparts (Im.1, Im.2, Im.3, Im.4 each of  $128 \times 128$ ) for both encryption and decryption processes. The idea is to encrypt half on the image with the remaining half.

Description of the methods used in the algorithm:

- I. Substitution of subpart Im.X using S-box of AES algorithm( sbox(Im.X))
- II. XOR of columns of each row of subpart Im.X( XCR(Im.X)):
 

For a subpart, XOR all the  $r, g, b$  values with values corresponding to them in the column in row  $i$ , where  $i = 1, 2, \dots, 127, 128$ . The result is a matrix of size  $128 \times 1$ . The matrix is horizontally concatenated 128 times to create a matrix of  $128 \times 128$ .
- III. XOR of rows of each column of subpart Im.X( XRC (Im.X)):
 

For a subpart, XOR all the  $r, g, b$  values with values corresponding to them in the row in column  $i$ , where  $i = 1, 2, \dots, 127, 128$ . The result is a matrix

of size  $1 \times 128$ . The matrix is vertically concatenated 128 times to create a matrix of  $128 \times 128$ .

IV. Subparts hashing (Hash (Im.X, Im.Y)):

Subparts Im.X and Im.Y together create a matrix of size  $128 \times 256$ , say H. Each row of H is then divided to form four arrays of size 124 bytes. To these arrays, four keys of 1 byte each will be appended to form 128 byte array. For each row, the 4 sub-arrays and keys that are appended are shown below:

$$SA1 = H[1 : 124] + \text{Keys}(1 \text{ to } 4)$$

$$SA2 = H[51 : 174] + \text{Keys}(5 \text{ to } 8)$$

$$SA3 = H[100 : 223] + \text{Keys}(9 \text{ to } 12)$$

$$SA4 = H[132 : 255] + \text{Keys}(13 \text{ to } 16)$$

The resultant of hash function is 64 bytes for each sub-array. These results together, 256 bytes, are the output of hash function of each row. By the end of all the rows, a new matrix of size  $128 \times 256$  is created.

V. Substitution of subpart Im.X according to AES algorithm's inverse sbox (InvSbox (Im.X))

### 3.2.1 Encryption

The height and width of the image should be 256, else they are to be adjusted. In this algorithm, 16 keys each of 8 bits are used. The image is divided into 4 subparts,  $128 \times 128$  pixels each. In the method, lower half's information is used to encrypt the upper half and vice-versa.

**Remarks** In each step,  $\text{Im.X}_{\text{new}}$  is considered as subpart that is a result obtained by using  $\text{Im.X}_{\text{old}}$ . ( $\text{Im.X}_{\text{old}}$  is the subpart that is about to be used in present step, and  $\text{Im.X}_{\text{new}}$  is the result of the present step).

#### Step 1:

Choose the secret keys (total 16) to perform method IV. Substitute subparts Im.1 and Im.2 according to sbox of AES. Calculate XRC (Im.3), XCR (Im.3), XRC(Im.4), XCR (Im.4), and hash (Im.3 and Im.4). The results of the step are  $\text{Im.1}_{\text{new}}$ ,  $\text{Im.2}_{\text{new}}$ .

$$\text{Im.1}_{\text{new}} = \text{sbox}(\text{Im.1}_{\text{old}}) \oplus \text{XRC}(\text{Im.3}) \oplus \text{XCR}(\text{Im.3}) \oplus \text{Hash}(\text{Im.3}, \text{Im.4}). \quad (3)$$

$$\text{Im.2}_{\text{new}} = \text{sbox}(\text{Im.2}_{\text{old}}) \oplus \text{XRC}(\text{Im.4}) \oplus \text{XCR}(\text{Im.4}) \oplus \text{Hash}(\text{Im.3}, \text{Im.4}). \quad (4)$$

**Remarks** The result of the hash is  $128 \times 256$ , and the columns from 1 to 128 are used in step (1) and from 129 to 256 are used in step (2).



**Step 2:**

Substitute subparts Im.3 and Im.4 according to S-box of AES. Calculate XRC (Im.1), XCR (Im.1), XRC (Im.2), XCR (Im.2), and hash (Im.1 and Im.2). The results of the step are Im.3<sub>new</sub>, Im.4<sub>new</sub>.

$$\text{Im.3}_{\text{new}} = \text{sbox}(\text{Im.3}_{\text{old}}) \oplus \text{XRC}(\text{Im.1}) \oplus \text{XCR}(\text{Im.1}) \oplus \text{Hash}(\text{Im.1}, \text{Im.2}). \quad (5)$$

$$\text{Im.4}_{\text{new}} = \text{sbox}(\text{Im.4}_{\text{old}}) \oplus \text{XRC}(\text{Im.2}) \oplus \text{XCR}(\text{Im.2}) \oplus \text{Hash}(\text{Im.1}, \text{Im.2}). \quad (6)$$

**Remarks** The result of the hash is  $128 \times 256$ , and the columns from 1 to 128 are used in step (1) and from 129 to 256 are used in step (2).

**Step 3:**

Calculate XRC (Im.2), XCR (Im.2), XRC (Im.4), XCR (Im.4). The results of the step are Im.1<sub>new</sub>, Im.3<sub>new</sub>.

$$\text{Im.1}_{\text{new}} = \text{Im.1}_{\text{old}} \oplus \text{XRC}(\text{Im.2}) \oplus \text{XCR}(\text{Im.4}) \quad (7)$$

$$\text{Im.3}_{\text{new}} = \text{Im.3}_{\text{old}} \oplus \text{XRC}(\text{Im.4}) \oplus \text{XCR}(\text{Im.2}) \quad (8)$$

**Step 4:**

Calculate XRC (Im.1), XCR (Im.1), XRC (Im.3), XCR (Im.3). The results of the step are Im.2<sub>new</sub>, Im.4<sub>new</sub>.

$$\text{Im.2}_{\text{new}} = \text{Im.2}_{\text{old}} \oplus \text{XRC}(\text{Im.1}) \oplus \text{XCR}(\text{Im.3}) \quad (9)$$

$$\text{Im.4}_{\text{new}} = \text{Im.4}_{\text{old}} \oplus \text{XRC}(\text{Im.3}) \oplus \text{XCR}(\text{Im.1}) \quad (10)$$

All the 4 subparts are then appended accordingly to form a  $256 \times 256$  encrypted image.

**3.2.2 Decryption**

The decryption process is similar to encryption process except the order is reversed, and in place of S-box, we use InvSbox of AES algorithm. The encrypted image is divided into 4 subparts.

Step 1: This step is same as the step 4 of encryption process.

Step 2: This step is same as the step 3 of encryption process.

Step 3: Calculate XRC (Im.1), XCR (Im.1), XRC (Im.2), XCR (Im.2), and hash(Im.1 and Im.2). The result of the step is Im.3<sub>new</sub>, Im.4<sub>new</sub>.

$$\text{Im.3}_{\text{new}} = \text{InvSbox} (\text{Im.3}_{\text{old}} \oplus \text{XRC} (\text{Im.1}) \oplus \text{XCR} (\text{Im.1}) \oplus \text{Hash} (\text{Im.1}, \text{Im.2})) \quad (11)$$

$$\text{Im.4}_{\text{new}} = \text{InvSbox} (\text{Im.4}_{\text{old}} \oplus \text{XRC} (\text{Im.2}) \oplus \text{XCR} (\text{Im.2}) \oplus \text{Hash} (\text{Im.1}, \text{Im.2})) \quad (12)$$

**Remarks** The result of the hash is  $128 \times 256$ , and the columns from 1 to 128 are used in step (9) and from 129 to 256 are used in step (10).

#### Step 4:

Calculate XRC (Im.3), XCR (Im.3), XRC (Im.4), XCR (Im.4), and hash(Im.3 and Im.4). The result of the step is Im.1<sub>new</sub>, Im.2<sub>new</sub>.

$$\text{Im.1}_{\text{new}} = \text{InvSbox} (\text{Im.1}_{\text{old}} \oplus \text{XRC} (\text{Im.3}) \oplus \text{XCR} (\text{Im.3}) \oplus \text{Hash} (\text{Im.3}, \text{Im.4})) \quad (13)$$

$$\text{Im.2}_{\text{new}} = \text{InvSbox} (\text{Im.2}_{\text{old}} \oplus \text{XRC} (\text{Im.4}) \oplus \text{XCR} (\text{Im.4}) \oplus \text{Hash} (\text{Im.3}, \text{Im.4})) \quad (14)$$

**Remarks** The result of the hash is  $128 \times 256$ , and the columns from 1 to 128 are used in step (11) and from 129 to 256 are used in step (12).

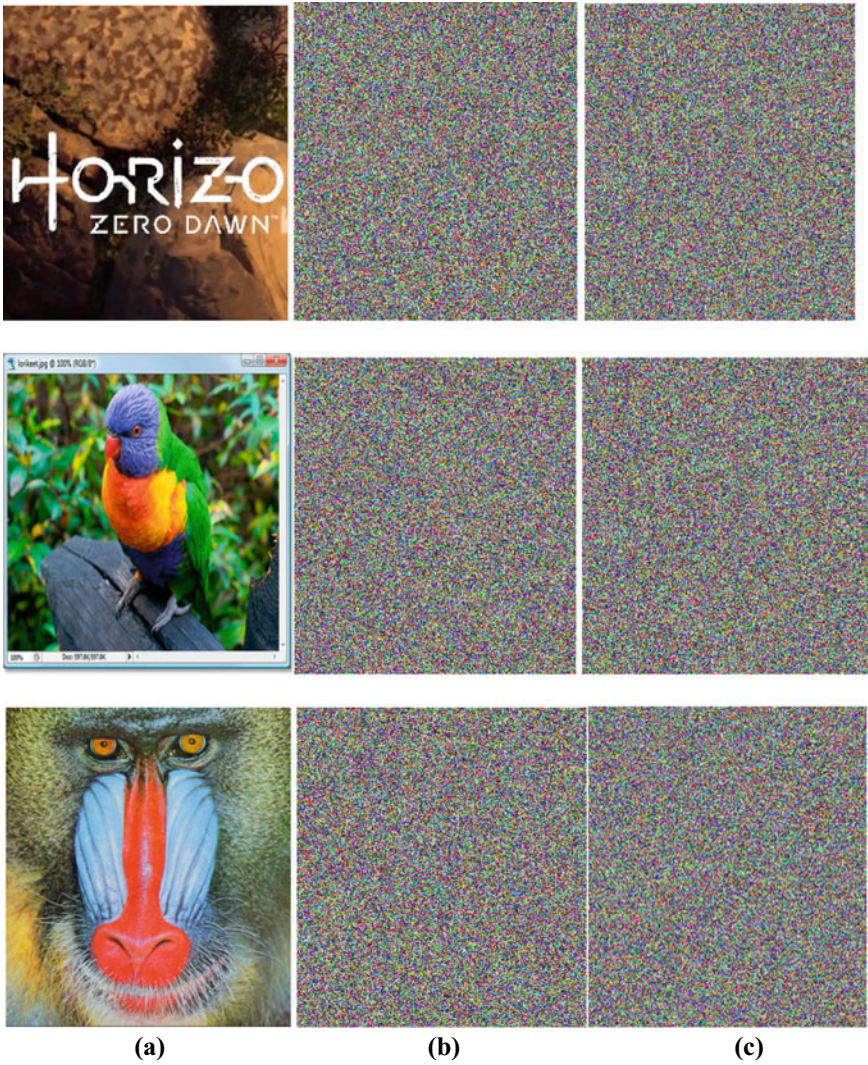
All the 4 subparts are then appended accordingly to form a  $256 \times 256$  decrypted image.

Below figures show the encrypted images using both algorithms (Fig. 1):

## 4 Security Test and Comparative Study

### 4.1 Statistical Analysis

Statistical analysis based on the histograms of original and respective cipher images, the correlation of adjacent pixels in the cipher images and the coefficient of correlation for original images and their respective cipher images are taken into consideration to examine the robustness of the image encryption algorithms.



**Fig. 1** a Plain original images b corresponding encrypted images using AES-RSA-LSB algorithm c corresponding hash-encrypted images

### 4.1.1 Histogram Analysis

An image-histogram is a histogram that acts as a graphical illustration of image pixel distribution in a digital image by indicating the no. of pixels at each tonal intensity level.

In Fig. 2, histogram of the original image 1(HorizonZero) of size  $250 \times 250$  demonstrates that the histogram of original image has particular pattern of r, g, and b components.

Figure 3 shows us the corresponding cipher image, encrypted using AES-RSA-LSB method, and its respective histogram.

Figure 4 shows us the hash-encrypted image of original image HorizonZero and its respective histogram.

Figure 5: The histogram of the original image 2 of size  $250 \times 250$  demonstrates that the histogram of original image has particular pattern of r, g, and b components.

Figure 6 shows us the corresponding cipher image, encrypted using AES-RSA-LSB method, and its respective histogram.

Figure 7 shows us the hash-encrypted image of original image 2 and its respective histogram.

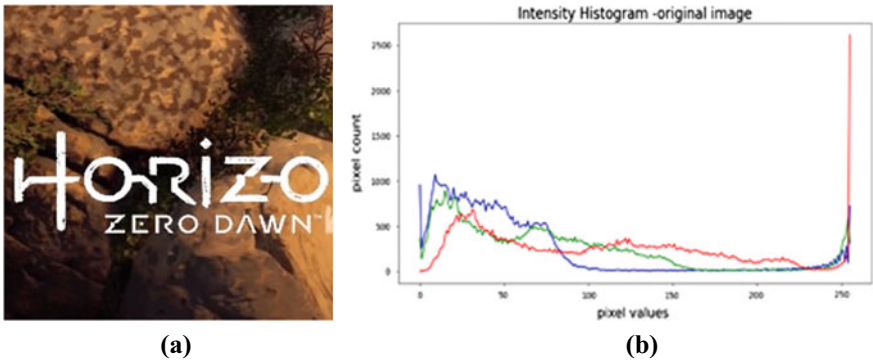


Fig. 2 a Original image HorizonZero, b histogram of original image HorizonZero

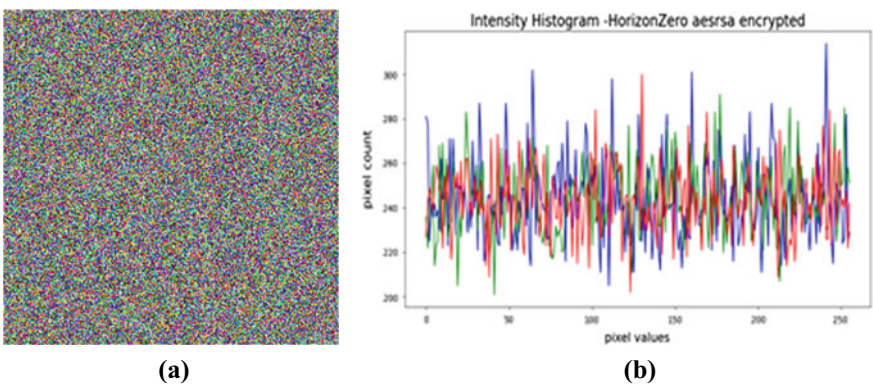


Fig. 3 a Encrypted image of original image HorizonZero using AES-RSA-LSB method, b histogram of (a)

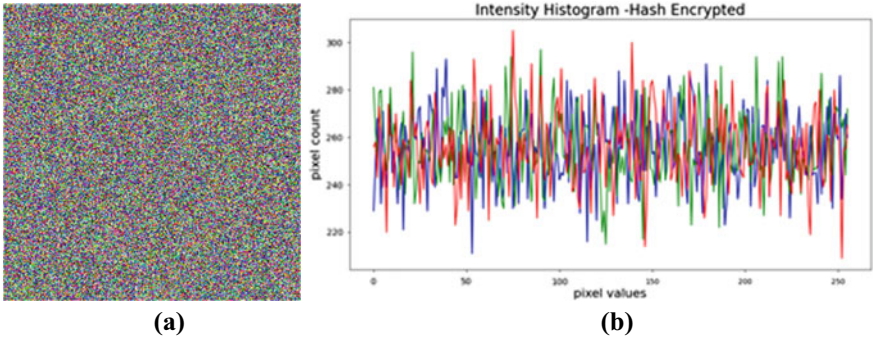


Fig. 4 a Hash-encrypted image of original image HorizonZero, b histogram of Fig. 4a

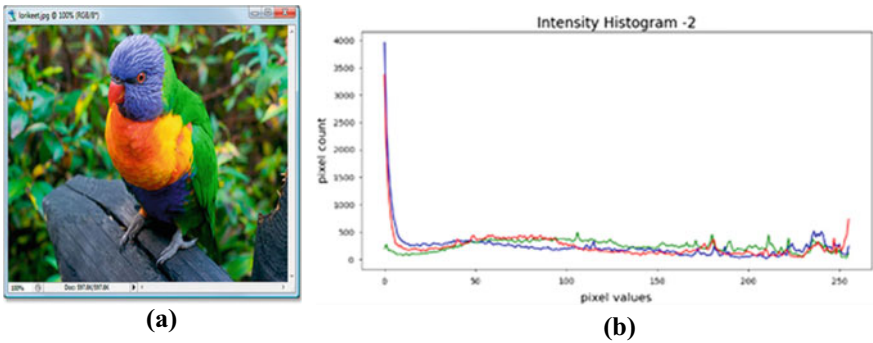


Fig. 5 a Original image 2, b histogram of original image 2

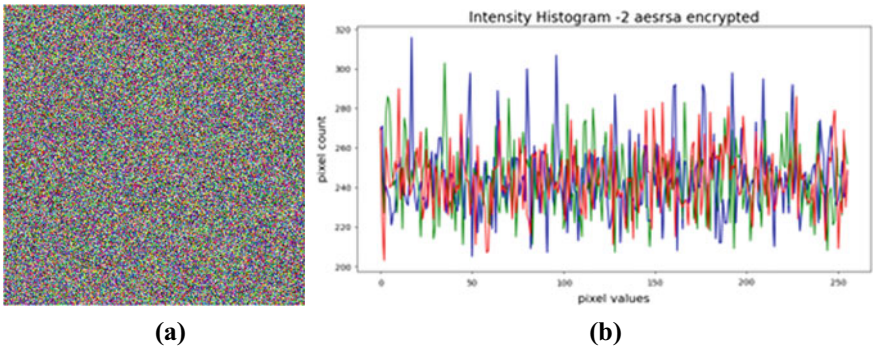


Fig. 6 a Encrypted image of original image 2 using AES-RSA-LSB method, b histogram of Fig. 6a

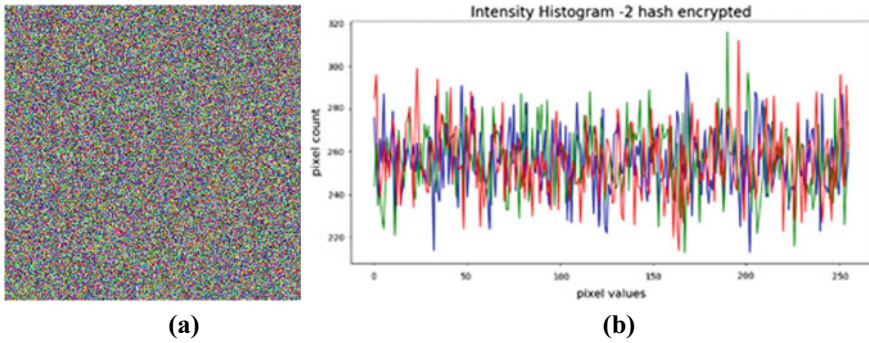


Fig. 7 a Hash-encrypted image of original image 2, b histogram of Fig. 7a

It should be noted that all pixels are uniformly distributed in the cipher image, therefore making cryptanalysis more difficult. This test is conducted for both algorithms using both images and more satisfactory results were obtained with the hash algorithm.

### 4.1.2 Correlation Coefficient Analysis

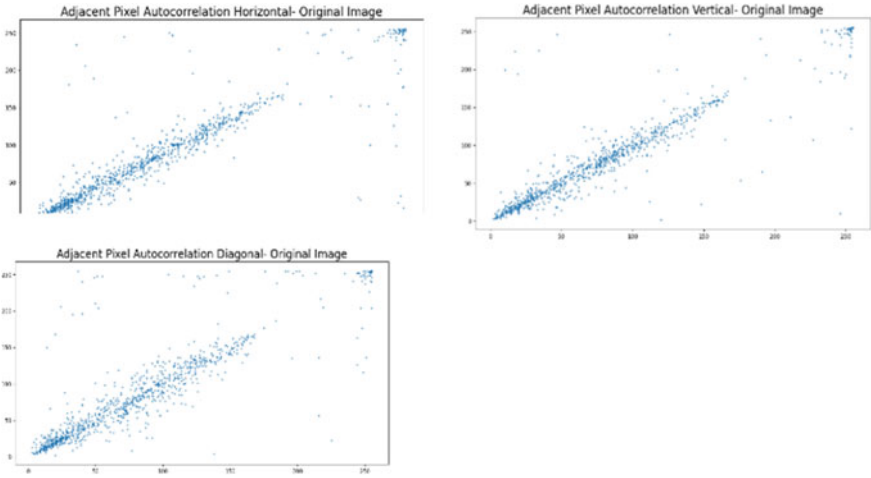
Correlation is a process of determining the probability that there exists a linear relation among two weighted values. Karl Pearson described the Pearson product-moment coefficient of correlation,  $r$ , in 1895. In image processing, pattern recognition, and statistical analysis, the Pearson’s coefficient of correlation,  $r$ , was widely used as a correlation measure. The Pearson’s coefficient of correlation for digital images is defined as

$$r_{XY} = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum (X_i - \bar{X})^2 \sum (Y_i - \bar{Y})^2}} \tag{13}$$

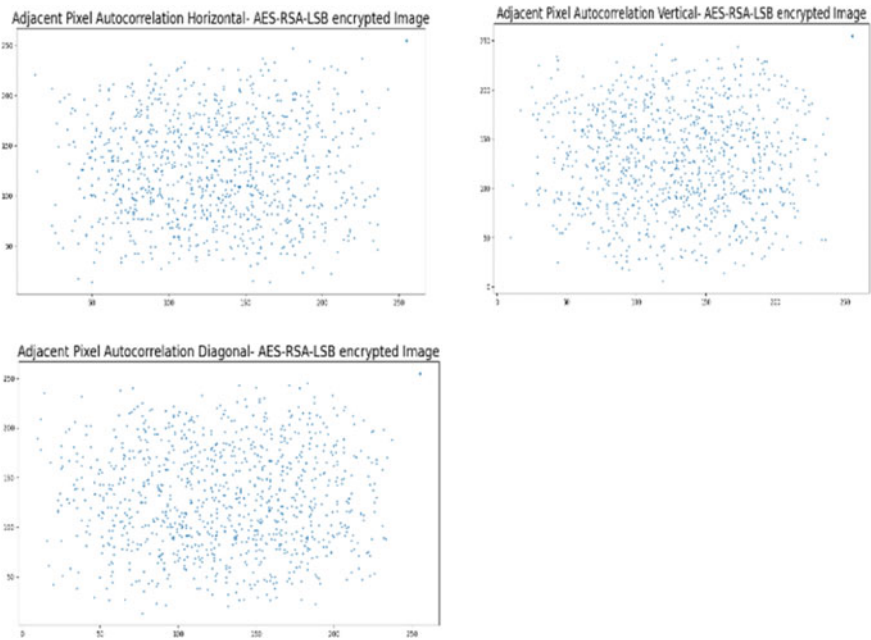
It calculates the relation between two adjacent pixel values. If they are absolutely identical, coefficient of correlation has the value  $r = 1$ . They are absolutely uncorrelated if  $r = 0$  and are absolutely anti-correlated if the value of  $r = -1$ . Horizontal correlation, vertical correlation, and diagonal correlation plotted between adjacent pixels of both the original image HorizonZero and images encrypted with encryption algorithm using AES-RSA-LSB, hash function are shown in the Fig. 8.

From Fig. 8a, it is clearly observable that there is a certain pattern in the correlation plot of adjacent pixels in original image HorizonZero, but Fig. 8b and c show us that there is uniform pixel distribution in encrypted images of both algorithms.

Horizontal correlation, vertical correlation, and diagonal correlation plotted between adjacent pixels of both the original image 2 and images encrypted with encryption algorithm using AES-RSA-LSB, hash function are shown in the Fig. 9.

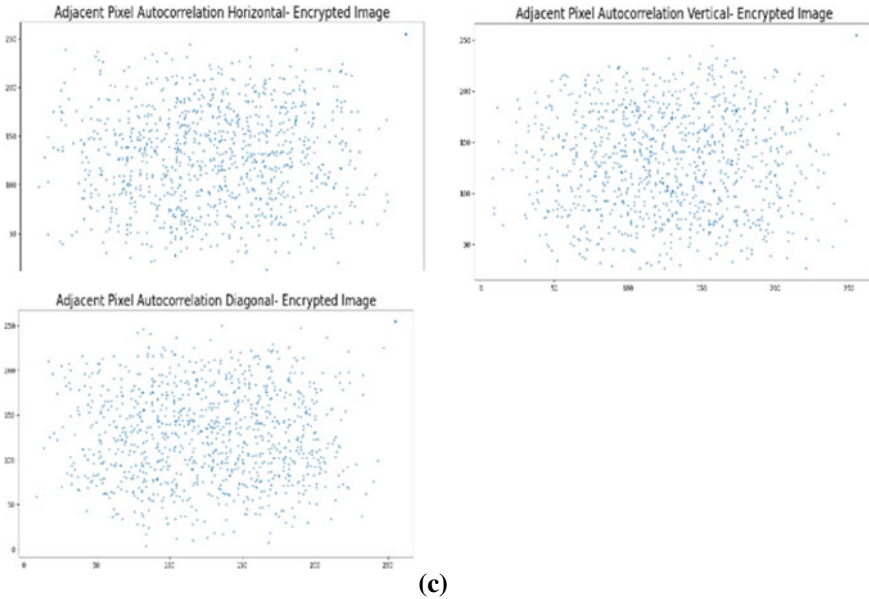


(a)



(b)

**Fig. 8** **a** Distribution of pixels in original image HorizonZero, **b** distribution of pixels in encrypted image using AES-RSA-LSB method, **c** distribution of pixels in hash-encrypted image



**Fig. 8** (continued)

From Fig. 9a, it is clearly observable that there is a certain pattern in the correlation plot of adjacent pixels in original image 2, but Fig. 9b and c show us that there is uniform pixel distribution in encrypted images of both algorithms.

Tables 1 and 2 show us that the values of correlation coefficient of encrypted images using both algorithms are closer to 0 than that of original images, but it is notable that correlation coefficients of hash-encrypted image are more precise to 0.

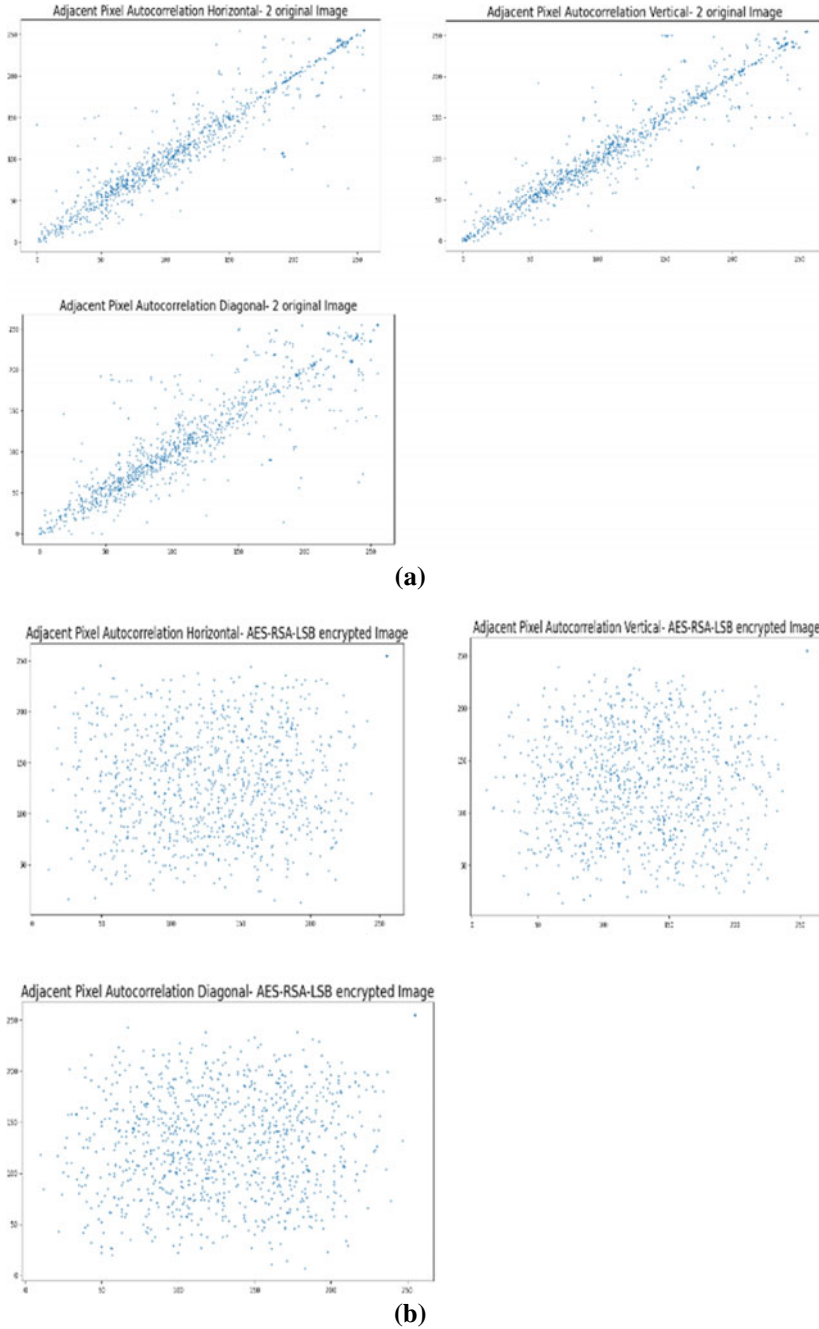
### 4.2 Information Entropy Analysis

Entropy is often used for the calculation of the amount of information within an image and the randomness that is indicated by the image’s texture and the amount of information within an image. The entropy  $H(s)$  is defined as

$$H(S) = - \sum_{i=1}^m p_i \log_2 p_i \tag{14}$$

where  $p_i$  stands for the probability of message. The ideal value of information entropy, when an image is encrypted, is considered to be 8, if the value is lower than this, then there remains a definite degree of certainty that its security might be compromised (Table 3).





**Fig. 9** **a** Distribution of pixels in original image 2, **b** distribution of pixels in encrypted image using AES-RSA-LSB method, **c** distribution of pixels in hash-encrypted image

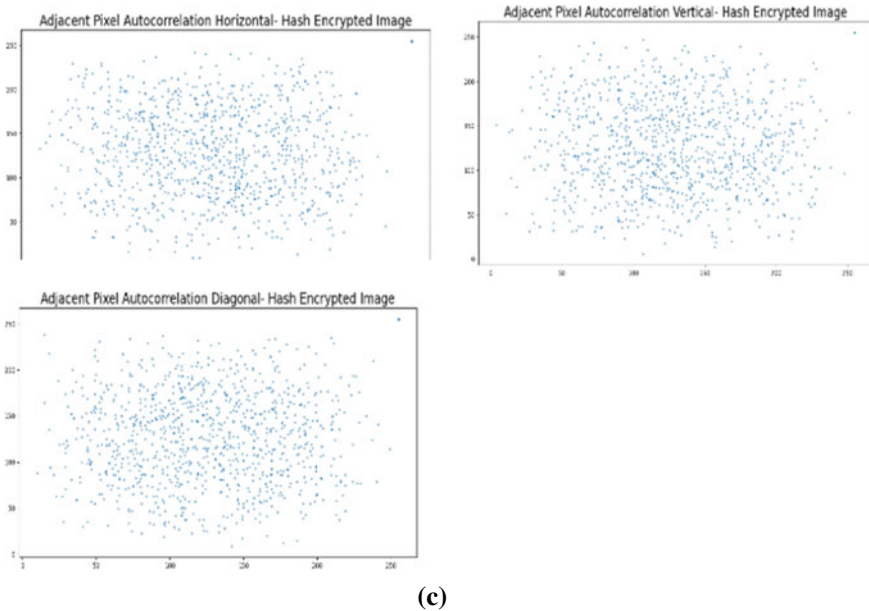


Fig. 9 (continued)

### 4.3 Encryption Quality–PSNR, NPCR, UACI Analysis

**PSNR Analysis:**

“Peak signal-to-noise ratio (PSNR)” can be defined as the ratio between the original image and its respective cipher image (Table 4).

$$PSNR = 10 \log_{10} \left( \frac{MAX_I^2}{MSE} \right) \tag{15}$$

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I(i, j) - K(i, j)]^2$$

$MAX_I$  = Maximum value of pixel in an original image

$m$  = Total number of rows in an original image

$n$  = Total number of columns in an original image.

The greater the value of PSNR, the more is the resemblance of cipher image and the original. Normally, a greater PSNR value should be able to correlate to a better quality image. Encryption scheme is considered good if the PSNR value is as low as possible. Without the learning of secret key, lower PSNR values can reason the challenge in obtaining the original image from the encrypted image.

**NPCR and UACI analysis:**

By analyzing both the NPCR and the UACI values for images encrypted with the AES-RS-LSB method and hash function, the evaluation of ability to show resistance against a chosen plain text attack (differential attack) was done. NPCR defines

**Table 1** Coefficient of correlation of adjacent pixels of original image HorizonZero and 2 and cipher images using AES-RSA-LSB method (horizontal, vertical, and diagonal)

Image name	Horizontal correlation		Vertical correlation		Horizontal correlation	
	Original image	Enciphered image	Original image	Enciphered image	Original image	Enciphered image
HorizonZero	0.9214918184319467	0.013909011716219577	0.9347604292095957	0.011044511955837444	0.8768273314184477	0.024230078775364844
2	0.9276754566366184	0.013302685540301992	0.8692175319999583	0.011286853264362157	0.9275945654901759	0.023489688715665452

**Table 2** Coefficient of correlation of adjacent pixels of original image HorizonZero and 2 and hash-encrypted images (horizontal, vertical, and diagonal)

Image name	Horizontal correlation		Vertical correlation		Diagonal correlation	
	Original image	Enciphered image	Original image	Enciphered image	Original image	Enciphered image
HorizonZero	0.9214918184319467	0.01155419667017823	0.9347604292095957	0.013058796261792245	0.8768273314184477	0.023844368140612762
2	0.9276754566366184	0.006025497419932511	0.9275945654901759	0.010301019107040054	0.8692175319999583	0.019655715741470316

**Table 3** Shows the entropy values for original image and cipher images

Image name	Original image	Entropy with algorithm 1	Entropy with algorithm 2
HorizonZero	7.4924423342698985	7.998747402442854	7.9990382671485065
2	6.599263778679899	7.998825871300561	6.786805859225379
Baboon	7.692590923186865	7.998818479848452	7.998972035956279

the variation rate of pixels between original image and cipher image, and UACI defines the average change in intensity of the original image and the cipher image. There will be better ability to prove resistance against a chosen plain text attack if the original image is highly modified when compared to the cipher image. Even a small modification created in original image results in observable change. Significant correspondence between original image and cipher image can be found using this method.

“Number of pixel change rate (NPCR)” is defined as

$$NPCR = \frac{\sum_{i,j} D(i, j)}{M \times N} \times 100\%. \tag{16}$$

“Unified average change in intensity (UACI)” is defined as

$$UACI = \frac{1}{M \times N} \left[ \sum_{i,j} \left[ \frac{|I(i, j) - K(i, j)|}{255} \right] \right] \times 100\% \tag{17}$$

$$D(i, j) = \begin{cases} 0 & \text{if } I(i, j) = K(i, j) \\ 1 & \text{otherwise} \end{cases} .$$

If the NPCR value is more than 99%, then the analysis is positive. The greater the values of NPCR and the UACI, the superior the proposed algorithm is (Table 5).

## 5 Conclusions

We introduced image encryption algorithms that are based on AES, RSA and LSB methods, and hash function (SHA-2(512)). First algorithm uses RSA algorithm to encrypt the generated keys, then conceals the encrypted key using LSB steganography in image encrypted using AES algorithm. The second algorithm uses SHA-2, a 1D hash algorithm, to create a 2D mask to encrypt the image. The method is a substitution-diffusion type hash-based image encryption. Results from the conducted experiments acknowledge that encrypted images using both the suggested methods generate histograms with uniform distribution in pixel. The pixels are more uniformly distributed in the hash-encrypted images. Coefficient of correlation of adjacent pixels

**Table 4** Shows the results of PSNR analysis

Image name	AES-RSA-LSB method		Hash encryption	
	PSNR	PSNR(db)	PSNR	PSNR(db)
HorizonZero	7.197269826170008	27.91007489488519	7.716409883513068	27.886869583915008
2	7.413889553591595	27.897974289640757	7.998875509097952	27.901030827344787
Baboon	8.9535938740058	27.90293351995701	9.775708777524772	27.898571899987957

**Table 5** Measurement of encryption quality

Image	Encryption with AES-RSA-LSB		Encryption with hash	
	NPCR	UACI	NPCR	UACI
HorizonZero	99.60906666666665	33.43224470588695	99.58826666666668	33.30574431373036
2	99.60640000000001	35.432288627456124	99.60106666666665	35.59527529412285
Baboon	99.6032	29.532580392163016	99.60053333333333	29.567472941182693

in enciphered images is decreased greatly when compared to original images. Moreover, values of correlation coefficient of encrypted images using both algorithms are closer to 0 than that of original images, but it is notable that the coefficients of correlation for hash-encrypted images are more precise to 0. The suggested algorithms have proven that the encrypted pictures have information entropy of precise to 8, which is ideal. Information entropy of hash-encrypted images HorizonZero and baboon is closer to 8. These results prove robustness of both algorithms, also proving better robustness of hash function. The PSNR values are less AES-RSA-LSB algorithm compared to hash function. The NPCR analysis shows both algorithms give NPCR values above 99%, although they are close AES-RSA-LSB algorithm gives a better edge in results and UACI analysis yield better results in both the algorithms, but hash function gives slightly better results, which conclude superior quality and security of both the algorithms.

## References

1. Moumen A, Sissaoui H (2017) Images encryption method using stenographic LSB method, AES and RSA algorithm. *Nonlinear Eng* 6(1):53–59
2. Seyedzade SM, Atani RE, Mirzakuchaki S (2010) A novel image encryption algorithm based on hash function. In: *Iranian conference on machine vision and image processing*. pp 5941167
3. Hafsa A, Sghaier A, Malek J, Machhout M (2021) Image encryption method based on improved ECC and modified AES algorithm. *Multimedia Tools Appl* 80:19769–19801
4. Arab A, Rostami MJ, Ghavami B (2019) An image encryption method based on chaos system and AES algorithm. *J Supercomput* 75(10):6663–6682
5. Alsaffar DM et al. (2020) Image encryption based on AES and RSA algorithms. In: *2020 3rd international conference on computer applications information security (ICCAIS)*, pp 1–5
6. Farah MB, Farah A, Farah T (2019) An image encryption scheme based on a new hybrid chaotic map and optimized substitution box. *Nonlinear Dyn* 8:1–24
7. Ping P, Xu F, Mao Y, Wang Z (2017) Designing permutation substitution image encryption networks with Henon map. *Neurocomputing*. 1–17
8. Shadangi V, Choudhary SK, Patro KAK, Acharya B (2017) Novel Arnold scrambling based CBC-AES image encryption. *Int J Control Theory Appl* 10:93–105
9. Shaktawat VAR, Rs S, Lakshmi N, Panwar A (2020) A hybrid technique of combining AES algorithm with block permutation for image encryption. *Rel Theory Appl* 15(1):15
10. Wang X, Zhu X, Wu X, Zhang Y (2018) Image encryption algorithm based on multiple mixed hash functions and cyclic shift. *Opt Lasers Eng* 107:370–379
11. Hua Z, Jin F, Xu B, Huang H (2018) 2D logistic-sine-coupling map for image encryption. *Signal Process* 149:148–161
12. Hua Z, Zhou Y, Huang H (2019) Cosine-transform-based chaotic system for image encryption. *Inf Sci* 480(1):403–419



# Speaker Identification Using Multiple Features and Models



A. Revathi, G. Gayathri, and C. Jeyalakshmi

## 1 Introduction

As we all know, speech is the best fastest way of interacting among humans. This reality created a spark among researchers to consider speech signal as an effective tool for interacting with computers. This creates the need for the machines to have a complete knowledge about the humans and a pathway to utilize speech for different accustoms. One such pathway is the speaker identification technology (SIT) where human's speech is used to unlock various technologies [1]. Speaker identification is a "one on many mapping" technique in which "the speaker can be identified by matching the unknown speaker's speech with templates of all speakers", called as speaker identification or in other words, an utterance from unknown speaker is analysed and compared with speech model of known speakers. Basically, speaker identification is a pattern recognition problem. Widely, text-independent systems are used for speaker identification process [2]. Text independent is nothing but the speech given at training and testing are different as the ultimate goal is to identify the person corresponding to the particular voice or speech and not the context of the speech. Text-independent system uses techniques such as acoustics and speech analysis. Acoustics is a branch of physics that deals with the oscillation of matter in solid, liquid and gases including terms such as vibration and sound. Speech analysis is the study of speech sounds, or in other words, it is the analysis of voice tones corresponding to the vocal folds. In speaker identification, predominantly speech analysis is used. The

---

A. Revathi · G. Gayathri (✉)

SASTRA Deemed to be University, Tirumalaisamudram, Thanjavur, Tamil Nadu, India

e-mail: [gayuganesan2000@gmail.com](mailto:gayuganesan2000@gmail.com)

A. Revathi

e-mail: [revathi@ece.sastra.edu](mailto:revathi@ece.sastra.edu)

C. Jeyalakshmi

K. Ramakrishnan College of Engineering, Samayapuram, Trichy, Tamil Nadu, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2022

T. Sengodan et al. (eds.), *Advances in Electrical and Computer Technologies*,

Lecture Notes in Electrical Engineering 881,

[https://doi.org/10.1007/978-981-19-1111-8\\_41](https://doi.org/10.1007/978-981-19-1111-8_41)

voice prints in this thesis is processed and stored using technologies such as vector quantization [3] and neural networks [4]. Here, I have used the k-means algorithm as classifier for VQ techniques and convolutional neural network for training the model in neural networks technique. Speaker identification technology (SIT) uses the power of voice in biometrics for authentication to recognize a speaker automatically and with high accuracy based on their voice [5]. This technology is used in surveillance for eavesdropping telephone conversations. Speaker identification is used in forensics department for backtracking suspect's voice during crimes. It is also used in online banking services, monitoring elderly people's health and helpful for people with dementia to identify who's speaking with them [6].

## 2 Speech an Overview

As we all know, speech is the best fastest way of interacting among humans. Speech is the output of the vocal tract system excited by the vibration of vocal cords due to quasi-periodic pulses of air for voiced speech or noise excitation for unvoiced speech. In other words, speech is the output of the vocal tract system excited by the vibration of vocal cords due to acoustic air pressure from the lungs. This acoustic air pressure is a function of time. The speech signal can be normally segmented into voiced segment and unvoiced segment. The source of excitation for the voiced segment is a quasi-periodic pulse of air, whereas for the unvoiced segment, it is a white noise signal. So, depending on how you shape your vocal tract different sounds can be produced such as fricatives, semivowels, consonants, diphthongs, vowels.

### 2.1 Formalizing the Speech

The vocal tract frequency response convolved with the glottal pulse results in the speech signal. This vocal tract frequency response acts as an impulse response. Let the speech signal be  $x(t)$ , glottal pulses represented as  $e(t)$ , excitation signal produced by the vocal tract system and  $h(t)$  be the vocal tract frequency response, i.e. impulse response.

$$\text{Speech, } s(t) = e(t) * h(t) \quad (1)$$

Normally, the log spectrum resembles the speech, spectral envelope resembles the vocal tract frequency response, and the spectral details resemble the glottal pulses. In the spectral envelope, the formants carry the identity of sound.

In terms of digital signal processing, vocal tract acts as a filter, by filtering the glottal pulses which consists of high-frequency noisy signals from the carrier information about the pitch. The speech signal  $s(n)$  is given by

$$s(n) = g(n) * h(n) \quad (2)$$

and

$g(n)$  = excitation signal

$v(n)$  = impulse response of vocal tract system.

Because of the digital processing of speech signals, before the extraction of features from the audio signals, let us get to know how the speech signal should be used for the analysis. First and foremost, the analogue speech signal has to be digitized and pre-emphasized before it is being used for analysis. Later, spectral analysis techniques are carried out to extract the features from the audio speech signal. It involves two steps. They are

1. Analogue speech signal from sound pressure wave is converted to digitized form.
2. Significant frequency components are emphasized, i.e. digital filtering.

The role of digitization is to produce high signal-to-noise ratio (SNR) on the sampled data of audio speech. After digitization comes the amplification process which is done by a pre-emphasis filter that boosts the spectrum of the audio signal up to 20 dB/decade since the natural voice signal has a negative slope of 20 dB/decade. In pre-emphasis, voice signals at high frequencies are amplified due to the effect of damping at higher frequencies. Here,

$$y(n) = x(n) - a * x(n - 1) \quad (3)$$

is our pre-emphasis filter, with filtering coefficient  $a = 0.95$ . In general, there are two major concerns in any speech communication system.

1. Preservation of the information contained in the speech signal.
2. Representation of the speech signal in an easier manner so that modifications can be done to the signal, without degrading the original information content present in the speech signal.

Hence, these two major concerns have to be dealt with before further processing of speech signals in any communication system.

## 2.2 Framing and Windowing

As the speech signal is not stationary for an infinite length, framing is done so that for a short period, it remains stationary since the glottal system cannot vary at once. So, framing is the splitting of the speech signal into smaller chunks. Before framing, the speech signal is filtered first using a pre-emphasis filter. Framing is done as machines cannot do computations with infinite data points, as the signal will be cut off at either ends leading to information loss. In framing, the speech signal is divided into frames of 8–16 ms length and shifted with overlapping of up to 50% so that the next frame

contains information about the previous frames. The frames of the speech signal are then multiplied with the window. Usually, windows enhance the performance of an FFT to extract spectral data. The window which we preferred for multiplication with frames of the speech signal is the Hamming window. The Hamming window is used to reduce the ripple caused in the signal so that we can get a clear idea of the original signal's frequency spectrum. Hamming window  $w(n)$  is given by

$$w(n) = 0.54 - 0.46 \cos(2\pi n / (N - 1)), \quad 0 \leq n \leq N - 1 \quad (4)$$

### 2.3 Understanding the Cepstrum

We know that the speech signal can be represented as follows:

$$\text{Speech, } s(t) = e(t) * h(t) \quad (5)$$

$$\text{Taking Fourier Transform, } s(w) = E(w) * H(w) \quad (6)$$

$$\text{Taking log on both sides, } \log(s(w)) = \log(E(w)) + \log(H(w)) \quad (7)$$

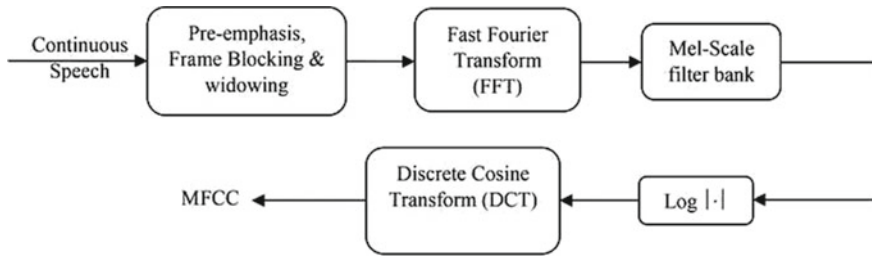
By using the log magnitude spectrum, we can separate the vocal tract information and the glottal pulse information which could not be done by normal spectrum. After which by taking IDFT of log magnitude spectrum, cepstrum is obtained. The physical separation of the information that is relative to the spectral envelope (vocal tract frequency response) is in the lower end of the quefrequency domain, and the information relative to spectral details (glottal pulse excitation) is in the higher end. Thus, the excitation  $e(t)$  can be removed after passing it through a low pass lifter.

Thus, after analysis of speech signals, it can be used to train the speaker models by extracting the feature vectors of each specified emotion. Later, the emotion model is constructed and test feature vectors are given to the classifier after which the emotional state of the test feature is being tracked.

## 3 Feature Extraction

### 3.1 Mel-frequency Cepstral Coefficients (MFCC) [2]

Mel-frequency cepstrum is a short-time power spectrum representation of a speech signal obtained after computation of discrete cosine transformation. The words



**Fig. 1** Block diagram of MFCC extraction

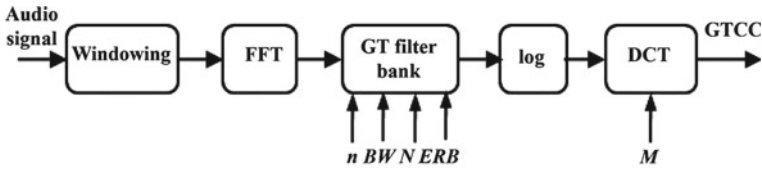
cepstrum, quefrency, liftering, and harmonic are just the wordplay of the word spectrum, frequency, filtering and harmonic, respectively. The former terms correspond to the frequency domain, and the later terms correspond to time-domain representation. MFCC is based on the characteristics of the hearing perception of humans. It is a fact that human ears use a nonlinear frequency unit to simulate the auditory system of humans. Usually, human ears perceive frequency logarithmically. This, in turn, necessitates an ideal audio feature that can be depicted in the time–frequency domain and has a perceptually relevant amplitude and frequency representation. One such audio feature is mel-frequency cepstral coefficients. Figure 1 describes the procedure for MFCC extraction.

The speech waveform after undergoing pre-emphasis, frame blocking and windowing, discrete Fourier transform of the signal is computed after which the log amplitude spectrum is obtained. Then, mel-scaling is performed. Mel-scale is a logarithmic scale. It is a perceptually relevant or perceptually informed scale for pitch. It is a fact that equidistant on the scale has the same “perceptual” distance. In the mel filter bank, the difference between the mel points is the same resulting in null weight, whereas the difference between frequency points is not the same on a scale of frequencies. After mel-scaling, the log amplitude spectrum undergoes discrete cosine transform, and the cepstrum of the speech signal is obtained. The advantages of MFCC are as follows:

1. MFCC describes the “large” structures of the spectrum, that is, it focuses on the phonemes.
2. It ignores the fine spectral structures like pitch.
3. Works well in speech and music processing.

### 3.2 Gammatone Cepstral Coefficients (GTCC) [7–10]

During the hearing, the gammatone cepstral coefficients capture the movement of the basilar membrane in the cochlea. GFCCs model the physical changes more closely and accurately that occur within the ear during the hearing and are therefore more representative of a human auditory system than mel-frequency cepstral coefficients. Unlike mel filter bank, here, gammatone filter bank is used as it models the human



**Fig. 2** Block diagram of GTCC extraction

auditory system and thereby uses ERB scale. This gammatone filter bank is often used in cochlea simulation's front end, thereby transforming complex sounds into multichannel activity as observed in the auditory nerve. The GFB is designed in such a way that the centre frequency of the filter is distributed in proportion to their bandwidth and is linearly spaced on the ERB scale. Generally, equivalent rectangular bandwidth (ERB) scale gives an approximation to the filter bandwidths in human hearing. That is, it models the filter either as rectangular bandpass or band-stop filters. The gammatone filter is given by

$$g(t) = at^{n-1}e^{-2\pi bt} \cos(2\pi f_c t + \phi) \quad (8)$$

where  $a$  is amplitude factor,  $t$  is time in seconds,  $n$  is filter order,  $f_c$  is the centre frequency,  $b$  is bandwidth and  $\phi$  is phase factor. Figure 2 indicates the modules used for GTCC extraction.

### 3.3 How Many MFCC and GTCC Coefficients?

Traditionally, the first 12–14 coefficients are computed for the MFCC analysis and 42 coefficients for the GTCC analysis, since they retain the information about formants and the spectral envelope (vocal tract frequency response) which is the requirement of the analysis, whereas the higher-order coefficients retain information about the fast-changing spectral details (glottal pulse excitation).

### 3.4 Why Discrete Cosine Transform Instead of IDFT?

- As it is a simplified version of Fourier transform.
- By using discrete cosine transform, we will be able to get only the real-valued coefficients, thereby neglecting the imaginary values, whereas in IDFT, both real and imaginary valued coefficients will be computed.
- Discrete cosine transform decorrelates energy in different mel bands.
- It reduces the dimension to represent the spectrum.

## 4 K-means Clustering [3]

K-means clustering belongs to the group of exclusive clustering where data points belong exclusively to one group. K-means clustering is a vector quantization method, specifically from signal processing, that aims to partition  $k$  clusters to  $n$  observations in which each observation belongs to the cluster that has the nearest mean (cluster centres or cluster centroid), serving as the preliminary model of the cluster. We use k-means clustering rather than hierarchical clustering because k-means works on actual observations creating a single level of clusters rather than the variation between every pair of observations leading to a multilevel hierarchy of clusters. The variable  $K$  indicates the number of clusters. The algorithm then runs iteratively to assign each data points to one of the  $k$  groups based on the features provided. K-means works by evaluating the Euclidean distance between the data points and the centre of the clusters, thereby assigning the data point to the nearest cluster. Based on the extracted mel-frequency cepstral coefficients and gammatone cepstral coefficients from the trained set of speakers, k-means clusters are formed. According to a k-means algorithm, it iterates over again and again, unless and until the data points within each cluster stop changing. At the end of each iteration, it keeps track of these clusters and their total variance and repeats the same steps from scratch but with a different starting point. The k-means algorithm can now compare the result and select the best variance out. For finding the value of  $K$ , you have to use the Hit and Trial method starting from  $K = 1$ . This  $K = 1$  is the worst-case scenario as the variations among the dataset are large. Each time when you increase the number of clusters, the variation decreases. When the number of clusters is equal to the number of data points, then in that case, variation will be zero. In this project, the optimal  $k$ -value is found to be four as the variation is reduced and a further increase in the number of clusters does not affect the variation much. That particular value of  $K$  is termed to be elbow point. We have preferred k-means as a classifier rather than other mentioned classifiers since it is very simple to carry out and also applies to large sets of data like the Berlin database. It has the speciality of generalization to clusters of any shape and size. Each cluster is assumed to be an emotion feature, and cluster centroid is formed using a k-means algorithm, and finally, emotion data is classified.

## 5 Convolutional Neural Network [4, 11]

### 5.1 Why not Fully Connected Neural Network?

We know that in computers, the images are read in the form of pixels in 3D plane, i.e. RGB plane. Say if an image has pixel values  $28*28*3$ , then the number of weights in the first hidden value of fully connected network will be 2352, whereas in real life, images have larger pixel values  $200*200*3$ , then in this case, the number of weights in the first hidden value of fully connected network will be 1,20,000. So, we need

to deal with huge number of parameters and thereby we require a greater number of neurons which can eventually lead to overfitting. That is why we do not use fully connected neural networks for image classification.

## 5.2 *Why Convolutional Neural Network?*

In CNN, a particular neuron is connected only to a small region of that layer unlike connecting to the complete layer as it was in fully connected networks. Thereby, CNN requires a smaller number of weights and a smaller number of neurons.

## 5.3 *What is Convolutional Neural Network?*

- Convolutional neural network is “a sort of feed-forward artificial neural network within which the property pattern between its neurons is inspired by the organization of animal visual area”.
- Visual cortex—“small regions of cells that are sensitive to specific regions of the visual field”. In other words, “some individual neuronal cells in the brain responds only in the edges of a certain orientation”. For example, some neurons respond on exposing to vertical edges and some responds on exposing to horizontal or diagonal edges.

## 5.4 *How CNN Works?*

CNN has four layers. They are:

- Convolution layer
- ReLU layer
- Pooling layer
- Fully connected layer

In this project, the spectrogram of 15 input speaker’s speech utterances is obtained. Normally, the signal strength or loudness of a signal can be represented visually over time at a range of frequencies called **spectrogram**. The spectrogram images of 15 input speakers are fed to the layers (Fig. 3).

**Convolution Layer:** In the convolution layer, the spectrogram images are piece by piece compared with the predefined filter patches to get the convolved output. Steps involved are:

- Feature and image patches are lined up
- Multiply each pixel of image with the corresponding feature pixel



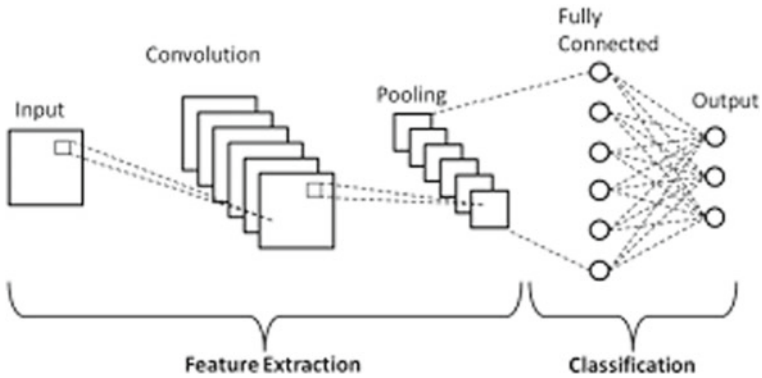


Fig. 3 Overview of convolutional neural network

- Add them
- Divide the feature by total number of pixels

**ReLU Layer**

- Every non-positive value from the filtered images is replaced with zeros.
- Done to prevent values from summing up to zero.
- ReLU stands for “rectified linear unit”. This transform function is used as a criterion.

$$F(x) = \begin{cases} n, & \text{for } n > 0 \\ 0, & \text{for } n < 0 \end{cases}$$

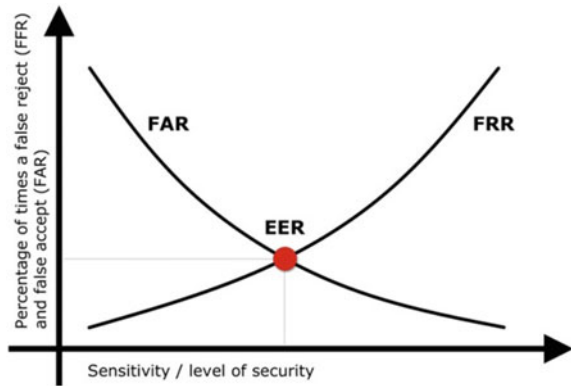
**Pooling Layer:** In this layer, the size of the image stack is reduced. Steps involved are:

- Choose window size (two or three)
- Choose stride = 2
- Window is moved over filtered images
- Maximum value is taken from each window

**Stacking up the Layers:** The above layers are stacked up one more time to minimize the size of the image stack.

**Fully Connected Layers**

- Actual classification takes place in this final layer.
- Here, our filtered and shrunk images are taken and lined up into a single sheet. That is, the values are listed in the form of the vector.
- Here, the number of fully connected layers taken is 15.

**Fig. 4** Equal error rate

## 5.5 Test Phase

Certain frames of spectrograms (nearly 80%) are given as input to the convolution layer which contains the predefined filters. Apart from the frames of spectrogram given to train the model, remaining frames of each speaker's spectrogram are given as a test input for classification. And finally, classification is done. During the classification, two rates have to be considered for accurate analysis. They are:

- False acceptance rate.
- False rejection rate.

When we **accept** a user whom we must actually have rejected, it is called false-acceptance rate. This issue is also called a **false** positive. False rejection rate occurs when we **reject** a user whom we should actually have accepted. As the false acceptances (FAR) decrease, the number of false rejections (FRR) increases and vice versa (Fig. 4).

The point at which the lines intersect is known as the equal error rate (EER). At this location, the percentage of false acceptances is equal to false rejections. Generally, FAR and FRR are configured in system by adjusting proper criteria, such that they are more or less strict.

## 6 Results and Discussions

### 6.1 Feature Extraction Results

A random utterance spoken by the speakers is tested against the trained cluster models to predict the particular speaker. Such that the minimum distance obtained from testing phase is used to determine the recognition rate, and further, we can

develop a confusion matrix or average recognition rate table. The results obtained from both the features are tabulated in Tables 1 and 2.

**Table 1** MFCC confusion matrix using clustering

	F1	F2	F3	F4	F5	F6	F7	F8	M1	M2	M3	M4	M5	M6	M7
F1	82.5	0	0	0	0	0	0	17.48	0	0	0	0	0	0	0
F2	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0
F3	13.29	0	86.70	0	0	0	0	0	0	0	0	0	0	0	0
F4	0	0	0	75.36	2.34	2.93	3.51	0	0	15.83	0	0	0	0	0
F5	0	0	0	0	100	0	0	0	0	0	0	0	0	0	0
F6	0	0	0	0	0	99.49	0.5	0	0	0	0	0	0	0	0
F7	13.58	0	0	0	1.44	2.31	24.27	15.60	2.89	0	39.88	0	0	0	0
F8	8.53	0	17.06	0	0	0	0	74.40	0	0	0	0	0	0	0
M1	0	0	0	0	0.34	0	0	0	80.20	0	19.45	0	0	0	0
M2	0	0	0	0	0	18.01	21.73	0	4.65	55.27	0	0	0.31	0	0
M3	2.99	0	0	0	0	0	0	0	0	0	97	0	0	0	0
M4	0	0	0	0	0	0	0	0	0	0	0	100	0	0	0
M5	0	0	0	0	6.72	0	0	0	0	0	0	0	93.27	0	0
M6	7.90	0	0	0	0	0	0	30.24	0	0	0	0	0	61.85	0
M7	0	63.36	0	0	0	0	0	0	0	0	0	0	0	0	36.63

F1–F8 female speakers/M1–M7 male speakers

**Table 2** GTCC confusion matrix using clustering

	F1	F2	F3	F4	F5	F6	F7	F8	M1	M2	M3	M4	M5	M6	M7
F1	82.5	0	0	0	0	0	0	17.48	0	0	0	0	0	0	0
F2	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0
F3	0	0	100	0	0	0	0	0	0	0	0	0	0	0	0
F4	0	0	0	96.18	0	0	3.81	0	0	0	0	0	0	0	0
F5	0	0	0	0	100	0	0	0	0	0	0	0	0	0	0
F6	0	0	0	0	0	100	0	0	0	0	0	0	0	0	0
F7	1.734	0	0	0	13.29	0	84.97	0	0	0	0	0	0	0	0
F8	0	0	9.95	0	0	0	0	90.04	0	0	0	0	0	0	0
M1	0	0	0	0	0.34	0	0	0	80.20	0	19.45	0	0	0	0
M2	0	0	0	0	0	0	0	0	0	95.96	0	0	4.03	0	0
M3	2.99	0	0	0	0	0	0	0	0	0	97	0	0	0	0
M4	0	0	0	0	0	0	0	0	0	0	0	100	0	0	0
M5	0	0	0	0	0	0	0	0	4.03	0	0	0	95.96	0	0
M6	0.27	0	0	0	0	0	0	20.70	0	0	0	0	0	79.01	0
M7	0	50	0	0	0	0	0	0	0	0	0	0	0	0	50

F1–F8 female speakers/M1–M7 male speakers

Based on the observation from the confusion matrix, it is evident that the overall speaker identification system's accuracy computed by means of gammatone cepstral coefficients (GTCC) is higher than that computed by means of mel-frequency cepstral coefficients (MFCC). It is observed that among the female and male speakers considered for evaluation of better classification system, the overall female speakers' accuracy is higher than that of overall male speakers' accuracy in both the feature extraction techniques (MFCC and GTCC). Though the overall female speakers' accuracy is higher, in the confusion matrix, it can be observed that some individual female speaker's accuracy is less than that of individual male speaker's accuracy. This can be dealt quantitatively in two ways.

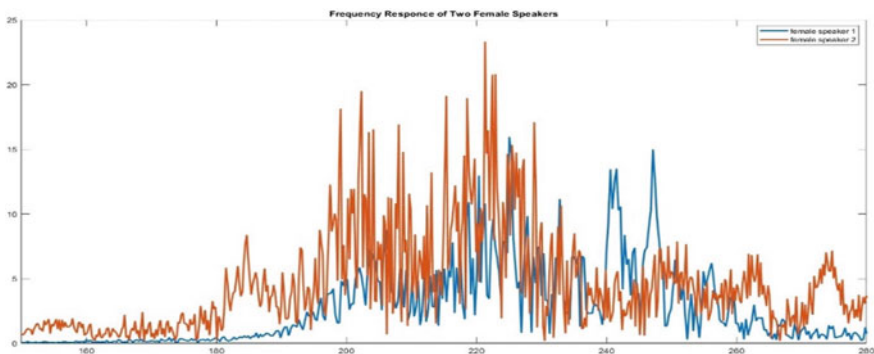
### 6.1.1 Quantitative Analysis

- Frequency Response Analysis

When the frequency response is plotted between two female speakers and two male speakers, it is observed that over a range of frequencies, two female speakers' frequencies tend to overlap each in most of the frequencies showing that there exists similarity between the female speakers' frequencies, whereas between two male speakers, there exists difference in frequency content (Figs. 5 and 6).

- Correlation Coefficient Analysis

When correlation analysis is carried out between two female speakers' speech utterances, between two male speakers' speech utterances and between a male speaker speech utterance and a female speaker speech utterance, it is observed that the correlation between two female speakers is higher than correlation between two male speakers. And when correlation is taken between a male speaker and a female speaker, it resulted in negative correlation coefficients which indicates that the overall performance of the system is good (Table 3).



**Fig. 5** Frequency response of two female speakers

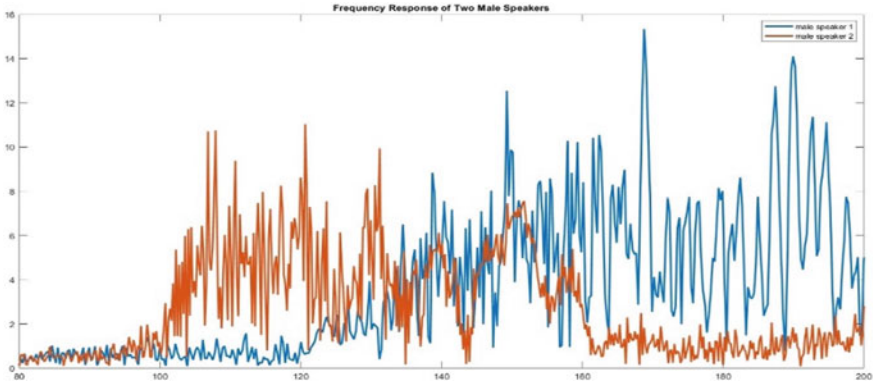


Fig. 6 Frequency response of two male speakers

Table 3 Correlation coefficient between speakers

(a) Female speakers		(b) Male speakers		(c) Male and Female	
1.0000	0.0264	1.000	0.0072	1.000	-0.0038
0.0264	1.0000	0.0072	1.000	-0.0038	1.000

## 6.2 Convolution Neural Network Results

The testing accuracy or average recognition rate obtained after classification is shown in Fig. 7.

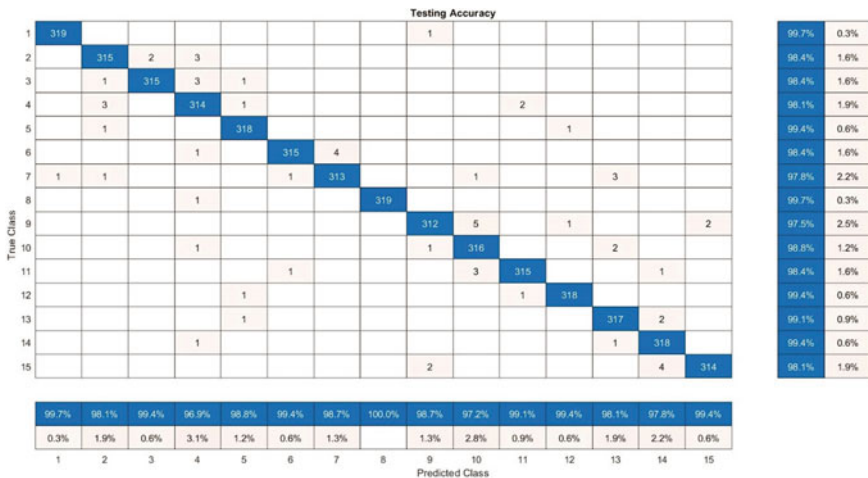


Fig. 7 CNN testing accuracy

So, in order to get accurate classification model, false acceptance rate (FAR) and false rejection rate (FRR) must be taken into consideration while modelling.

### 6.2.1 Effect of FAR and FRR

- If the value of FAR is set to the lowest possible value, then FRR tends to increase sharply. In other words, when the FAR is too low, the overall system will be more secure, and it will be less user-friendly such that it may falsely reject the correct users.
- If the value of FRR is set to the lowest possible value, then FAR tends to increase sharply. In other words, when the FRR is too low, the overall system will be more user-friendly and less secure, thereby enabling false users by mistake.

Therefore, it is our choice whether to give priority either to security or user convenience. If both do not want to be compromised, then we can set both the values to be the same, equal error rate.

## 7 Conclusions

In this paper, a speaker identification system is assessed by employing features such as mel-frequency cepstral coefficients (MFCC) and gammatone cepstral coefficients (GTCC), and VQ-based minimum distance classifier is used to classify the speakers from the speech utterances spoken by speakers. Additionally, convolutional neural network (CNN) was deployed to model a classification system for the speaker identification to enhance the accuracy furthermore. From the observations, it is evident that the GTCC feature tracks the particular speaker precisely as compared to the MFCC feature. Since the GTCC feature captures the movement of basilar membrane within the cochlea to mimic the human auditory system during the hearing, a speaker is accurately predicted compared to MFCC. The accuracy of the system using GFCC is improvised than MFCC. GTCC provides better overall accuracy of 90.12% as compared to MFCC whose overall accuracy is 77.79%. And, the accuracy obtained from the CNN model is 98.71%. A combination of features can be considered for improving the performance. This speaker identification system would find applications in the voice biometrics for authentication purpose, in surveillance for eavesdropping telephone conversations and in forensics department for backtracking suspect's voice during crimes. It is also used in Google's speech recognition system so as to unlock the gadgets with the speaker's voice that is used as a password for privacy protection.

## References

1. Jahangir R, Ali I (2017) Speaker identification through artificial intelligence techniques: a comprehensive review and research challenges. *Expert Syst Appl*
2. Leu F, Lin G (2017) An MFCC based speaker identification system. In: 2017 IEEE 31st international conference on advanced information networking and applications, pp 1055–1062
3. Escamilla E, Perez H, Martinez J, Suzuki MM (2012) Speaker recognition using MFCC and VQ techniques. *IEEE*
4. Albawi S, Al Zawi S, Mohammed TA (2017) Understanding of a convolutional neural network. In: 2017 International Conference on Engineering and Technology (ICET)
5. Furui S (2009) 40 years of progress in automatic speaker recognition. In: Nixon MS (ed) *Advances in Biometrics*, vol 5558. Springer, Berlin
6. Bansal PK, Sharma V (2013) A review on speaker recognition and challenges. *Int J Eng Technol* 2(5)
7. Fathima R, Raseena PE (2013) Gammatone cepstral coefficients for speakers identification. *Int J Adv Res Electr Electron Instrum Eng* 2
8. Valero X, Alias F (2012) Gammatone cepstral coefficients: biologically inspired features for non-speech audio classification. *IEEE Trans Multimedia* 14(6):1684–1689
9. Ayoub B, Jamal K, Arsalane Z (2016) Gammatone frequency cepstral coefficients for speaker identification over VoIP networks. In: 2016 International conference on Information Technology for Organizations Development (IT4OD), Fez, Morocco, pp 1–5
10. wang H, Zhang C (2019) The applications of gammatone frequency cepstral coefficients for forensic voice comparison under noisy conditions. *Aust Forensic Sci*
11. Chauhan R, Ghanshala KK, Joshi RC (2018) Convolutional Neural Network (CNN) for image detection and recognition. In: 2018 First International Conference on Secure Cyber Computing and Communication (ICSCCC)
12. Revathi AS and Jeyalakshmi C (2017) “Comparative analysis on the use of features and models for validating language identification system.” 2017 International Conference on Inventive Computing and Informatics (ICICI), 693-698.
13. Dharini D and Revathy A (2014) "Singer identification using clustering algorithm", *Communications and Signal Processing (ICC SP)*. 2014 International Conference on, pp. 1927-1931
14. Revathi AS, Jeyalakshmi C and Muruganantham T (2018) “Perceptual Features Based Rapid and Robust Language Identification System for Various Indian Classical Languages.”

# Performance Analysis on Flexible Modified Koch Fractal Patch Antenna for Wearable Healthcare Application



K. A. Malar and R. S. Ganesh

## 1 Introduction

For the past few decades, numerous progresses in the area of wearable devices and sensors which is useful in monitoring vital signs such as health care monitoring, military services, and sports. In recent years, smart watches have largely replaced smart phones, as have a variety of medical applications. Smart fabrics can collect vital signs from a human body, process them, and make decisions based on the data to some extent. Because antennas for wearable applications must operate in close proximity to people, their performance must be taken into account. The most important constraints to address are versatility, adherence, battery life, dimensions, and gravity. At the same time, such an adaptable wearable antenna must be simple and compact [1]. Also, the specific absorption rate (SAR) value for the body-worn antenna must be within the health and safety limits [2]. In the recent survey, different design configurations have been studied for the flexible wearable antennas including planar inverted F antenna (PIFA) [3], dipole, monopole, and patch antennas [4–6]. However, these antennas are not particularly small, have a very low impedance BW, and require flexible potting medium materials, which are critical for body-worn applications. Polymer-based ferroelectrics have made a strong impression in recent years for body-worn applications [7]. Considering the parameters like easy way of synthesis, miniaturized size, ease of interconnection with the nervous system, and good isolation between the human body and antenna, patch antennas are the best choice for wearable devices [8]. In this paper, a flexible miniaturized fractal patch antenna made of flexible textile material is presented for WBAN applications in the

---

K. A. Malar (✉) · R. S. Ganesh  
Noorul Islam Center for Higher Education, Kumaracoil, Tamil Nadu, India  
e-mail: [malarkrishnan.06@gmail.com](mailto:malarkrishnan.06@gmail.com)

R. S. Ganesh  
e-mail: [ganesh@niuniv.com](mailto:ganesh@niuniv.com)

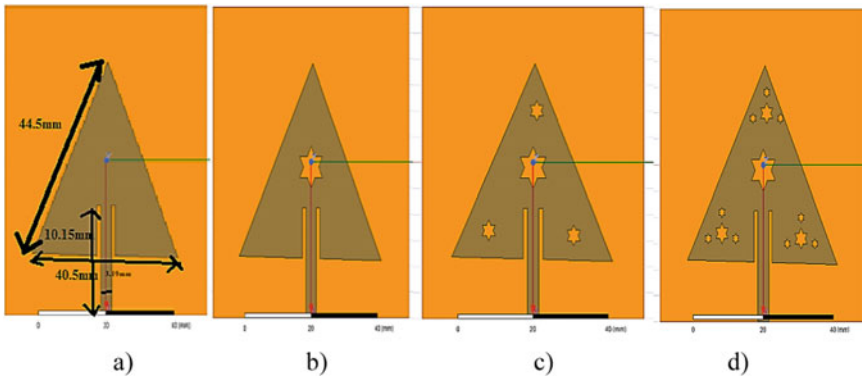


2.45 GHz ISM band. We use a Koch fractal with three iterations, which either allows us to raise the electrical length of the antenna without raising its size, resulting in a smaller antenna. The proposed antenna’s dimensions are  $44.5 \times 40 \times 1.27 \text{ mm}^3$ , and structural claim for damage polarity is demonstrated by measuring the antenna under various bending conditions. As a result of the numerical simulations, the proposed antenna appears to be the most suitable for wearable devices in terms of strength, reliability, and adaptivity.

## 2 Antenna Design

The preliminary antenna design is a conventional equilateral triangle with a full ground plane measuring  $60 \times 60 \text{ mm}^2$ . The preliminary antenna size is relatively comparable to other documented antennas for 2.45 GHz ISM band (Fig. 1a). On the patch, Koch fractal slot geometry is used to achieve miniaturization and high throughput (Table 1).

The modified Koch fractal slot antenna’s settings progress is shown in Fig. 1. The four steps to design the fractal antenna are as follows: (a) the traditional patch antenna



**Fig. 1** Fractal topology configuration, **a** basic patch, **b** 1st iteration, **c** 2nd iteration, **d** 3rd iteration

**Table 1** Optimized measurement for the antenna

Parameter	Value (mm)
Perpendicular height of triangle ( $L_p$ )	44
Base length of triangle ( $L_b$ )	40
Feed line length ( $L_f$ )	10.15
Feed line width ( $W_f$ )	3.19
Substrate height ( $h$ )	1.27
Substrate length ( $L_S$ )	60
Substrate width ( $W_S$ )	60

**Table 2** Dimensionality comparison of proposed antenna with existing antenna

References	Dimension( $L \times W \times H$ ) mm <sup>3</sup>	Substrate type
[5]	$39 \times 39 \times 0.508$	Flexible
[7]	$50 \times 50 \times 14.1$	Rigid
[9]	$60 \times 60 \times 3.25$	Rigid
[10]	$50 \times 50 \times 4$	Flexible
This work	$44.5 \times 40 \times 1.27$	Flexible textile

in triangular shape, (b) the Koch fractal slot in its first incarnation (shown in Fig. 1b), the fractal slot is obtained by replicating to second or third good enough condition of iteration with a scaling factor of roughly half (shown in Fig. 1c and d). Table 2 compares our proposed antenna's improved compressibility and effectiveness to relevant existing work.

The final fractal geometry achieves high gain to work in specified frequency. The top layer of the antenna is the flexible copper tape with equilateral triangle is modified in to a Koch fractal form and makes it as a slot in the center of the patch. This Koch fractal now acts as a basic unit; it is copied and scaled in to half and placed at the corners of the patch and then repeatedly placed at the corners with scaling. Cutting slots at the bottom of the patch increase capacitive coupling and modified triangle up to third iteration with an overall size of  $44.5 \times 40$  mm<sup>2</sup>. Additional iterations may be used, but it is subjected to only three to make it as simple, and main aim of this work to check flexibility, so only, bending analysis is preferred here. This fractal patch is connected to inset feed line with  $50 \Omega$  impedance. The copper is placed at the bottom layer in the antenna.

### 3 Results and Discussion

#### 3.1 Evaluation of Return Loss

The designed antenna is simulated by using Ansoft HFSS software. The return loss versus frequency plot of the designed antenna is shown in Fig. 2. From the diagram, the resonating frequency of the antenna is found to be 2.45 GHz, and it covers ISM band with good impedance match.

#### 3.2 Structural Conformability Evaluation (SCE)

The SCE of the designed antenna, which demonstrates its capacity to withstand some volume of SC, is a necessary prerequisite for wearable devices. The process used to evaluate structural conformability is bending analysis. Bending analysis is

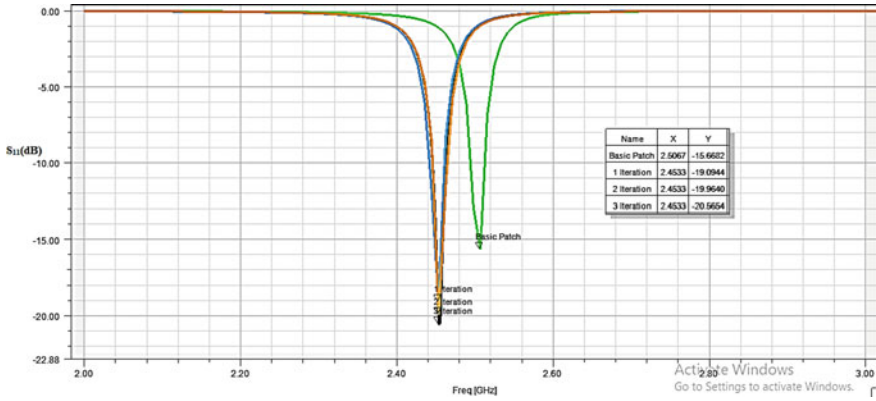


Fig. 2 Return loss versus frequency plot of the designed antenna

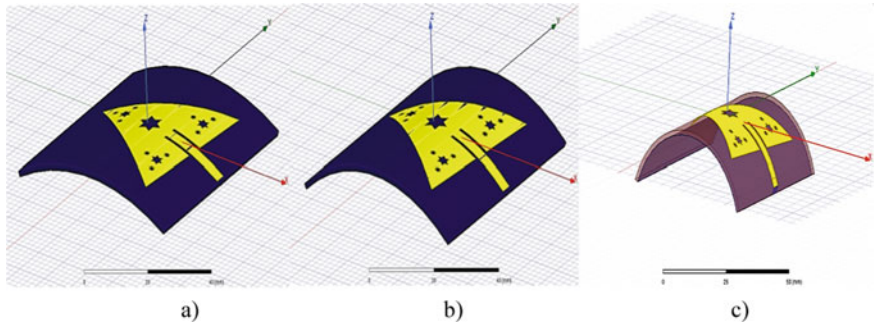


Fig. 3 Bending structure of the antenna, a  $R = 75$  mm, b  $R = 50$  mm, c  $R = 25$  mm

conducted and studied numerically in the HFSS simulation software with various radius of curvature along the x-axis (Fig. 3).

### 3.3 Far Field Radiation Pattern

The study of radiation in the far-field region has been done using HFSS software. The radiation pattern either with or without flexing is measured at the ISM band center frequency of 2.45 GHz, as shown in Figs. 4, 5, and 6. From the radiation pattern graph, maximum gain was found to be 6.5 dB, and the simulated radiation patterns are almost omni-directional. It can be seen that the radiation characteristics are almost same for all bending scenarios. Thus, the proposed antennas will work very well for wearable applications.

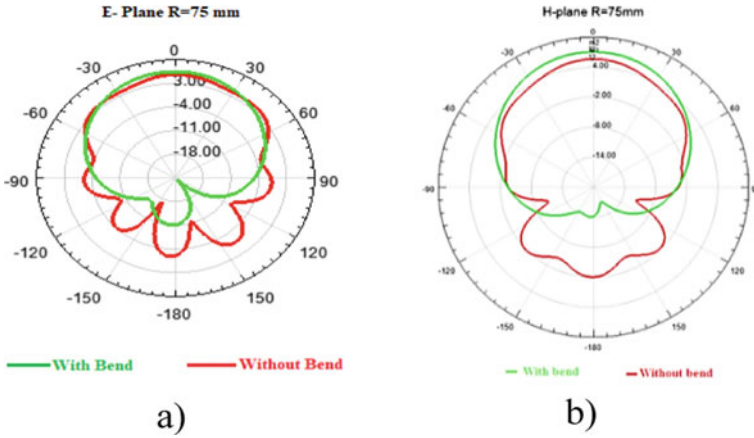


Fig. 4 Radiation pattern of antenna with and without the bend in, a E-plane and b H-plane at radius curvature  $R = 75$  mm

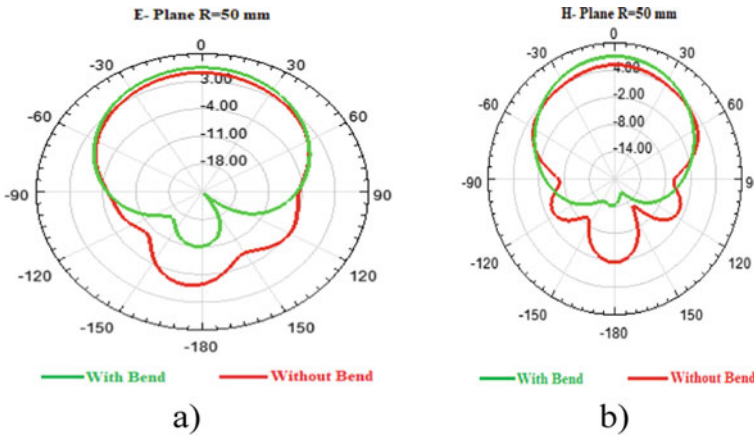
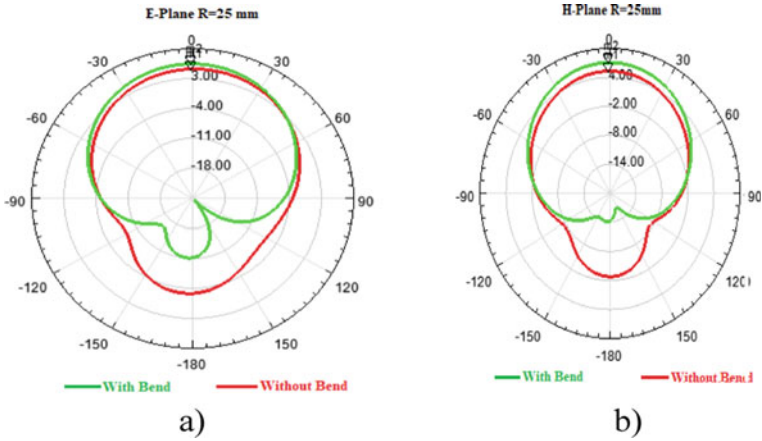


Fig. 5 Radiation pattern of antenna with and without the bend in, a E-plane and b H-plane at radius curvature  $R = 50$  mm

### 4 Analysis on SAR for Wearable Safety

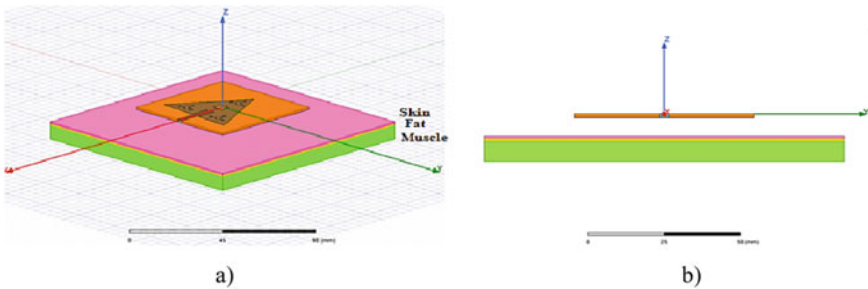
In this section, SAR of the 2.4 GHz antenna on a flat body phantom has been studied. Protection of human heart from the harmful radiation has been regularized by the International Commission on Non-Ionizing Radiation Protection (ICNIRP). The maximum SAR for ten gram of human phantom must not exceed 2.0 W/kg, according to ICNIRP regulations. Further, the Federal Communications Commission (FCC) also recommended that for one gram of tissue the value of SAR be no more than 1.6 W/kg [11]. To measure the effectiveness of a wearable antenna in an



**Fig. 6** Radiation pattern of antenna with and without the bend in, **a** *E*-plane and **b** *H*-plane at radius curvature  $R = 25$  mm

on-body case, a three-layer phantom modeled in HFSS v.19 with fat, muscle, and skin layers was used. To save computational complexity, the human tissue model is standardized with widths of  $120 \times 120 \times 42$  mm (skin layer = 1 mm, fat layer = 1 mm, and muscle layer = 40 mm) as shown in Fig. 7. Table 3 shows the electrical properties at 2.45 GHz and the depth of each tissue [12].

As shown in Fig. 7b, the antenna is placed 8 mm above the phantom to imitate the thickness of clothes in different layers. The designed antenna achieves good impedance matching and high gain in free space. However, due to the addition



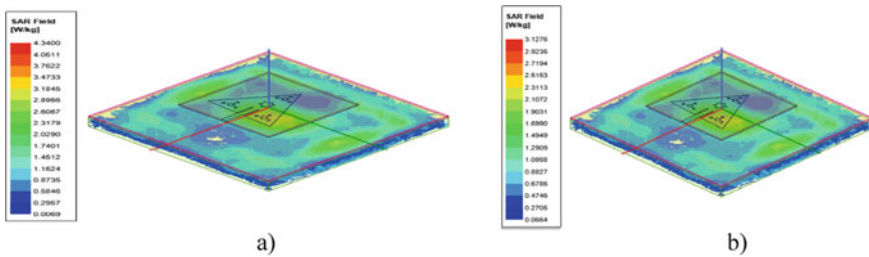
**Fig. 7** **a** Three-layer phantom model **b** Front view of three-layer phantom model

**Table 3** Dielectric properties of the tissue

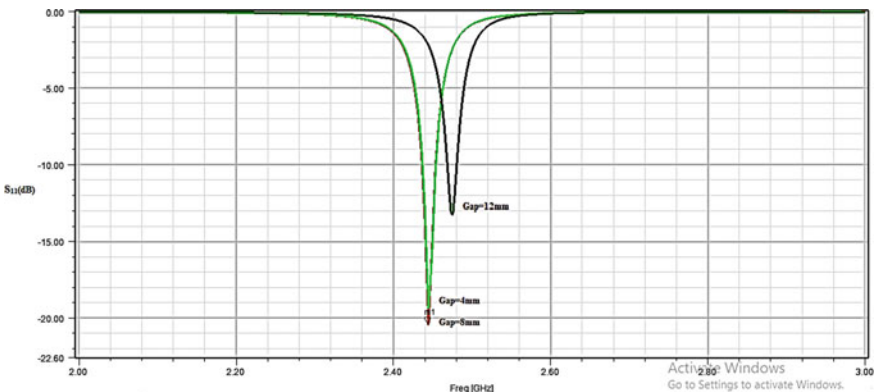
Tissue	$\epsilon_r$	$\sigma$ (S/m)
Skin	38.5	1.45
Fat	5.29	0.10
Muscle	52.7	1.73

of tissue, frequency detuning may occur after simply putting the antenna over the phantom model. As a result, the antenna’s effectiveness must be calculated using the phantom model with multiple tissue layers and distance.

The resonant frequency responsible for change whenever there is a gap has increased to 12 mm, as shown in Fig. 9, yet the band was covered properly. At a gap of 8–4 mm, good impedance matching is achieved. The calculated SAR values in one gram and ten gram of body tissue at 2.45 GHz are shown in Fig. 8. According to the IEEE/IEC 62704-1 standard [13, 14], the power input to an antenna was 0.5 W. The maximum SAR values for one gram and ten gram within tissues and organs at 2.45 GHz are shown in Table 4. 1 g body tissues absorbed significantly further electrical radiation than 10 g body tissues, as shown in the table. The strongest SAR for an average of 1 g body tissue is found to be 4.3400 W/kg. The simulation results did not exactly match the standard regulations once the antenna was placed near the human body. As a result, extra measures to reduce SAR must be chosen for safe operation standard level.



**Fig. 8** a Simulated SAR values for 1 g tissue in HFSS, b Simulated SAR values for 10 g tissue in HFSS



**Fig. 9** Simulated reflection coefficient on phantom model in HFSS

**Table 4** Antenna performance over three-layer phantom model

Distance (mm)	SAR (Unit: W/Kg)		Gain (dBi)	$\eta$ (%)
	1 g	10 g		
4	5.6758	4.21	6.2	27
8	4.3400	4.08	6.3	54
12	3.333	1.912	6.3	65

## 5 Conclusion

The flexible Koch-modified fractal antenna is designed and analyzed for wearable healthcare application. In this design, the effective gain improvement has been achieved by using Koch fractal slot with three iterations. Further, this compact design has a gain improvement of 6.5 dB and radiation efficiency of 89%. Also, the performance evaluation of the designed fractal antenna with its structural conformability is evaluated under three distinct curvature radii, and SAR is measured. The results clearly show this design is highly conformable for wearable applications.

## References

1. Fang G, Dutkiewicz E, Huq MA, Vesilo R, Yang Y (2011) Medical body area networks: opportunities, challenges and practices. In: International symposium on communications and information technologies, Hangzhou, China, pp 562–567
2. Zhang HH, Yu GG, Liu Y, Fang YX, Guangming S, Shang W (2020) Design of low-SAR mobile phone antenna: theory and applications. *IEEE Trans Antennas Propag* 69:698–707
3. Liu WC, Chen SH, Wu CM (2009) Bandwidth enhancement and size reduction of an implantable PIFA antenna for biotelemetry devices. *Micro Opt Tech Lett* 51:755–757
4. Scarpello ML, Kurup D, Rogier H, Ginsté DV, Axisa F, Vanfleteren J, Joseph W, Martens L, Vermeeren G (2011) Design of an implantable slot dipole conformal flexible antenna for biomedical applications. *IEEE Trans Antennas Propag* 59:3556–3564
5. Chien TF, Yang HC, Cheng CM, Luo CH (2010) Develop CPW-FED monopole broadband implantable antennas on the high dielectric constant ceramic substrates. *Micro Opt Tech Lett* 52(2010):2136–2139
6. Arif A, Zubair M, Ali M, Khan MU, Mehmood MQ (2019) A compact, low-profile fractal antenna for wearable on-body WBAN applications. *IEEE Antennas Wirel Propag Lett* 18:981–985
7. Raad K, Al-Rizzo HM, Issac A, Hammoodi AI (2016) A compact dual band polyimide based antenna for wearable and flexible telemedicine devices. *Prog Electromagn Res C* 63:153–161
8. Vladimir H (2013) Planar antenna in proximity of human body models. In: European conference on antennas propagation, Gothenburg, Sweden, pp 3309–3311
9. Tak J, Choi J (2013) Circular-ring patch antenna with higher order mode for on-body communications. *Micro Opt Tech Lett* 56:1543–1547
10. Paraskevopoulos A, de Sousa Fonseca D, Seager RD, Whittow WG, Vardaxoglou JC, Alexandridis A (2016) Higher-mode textile patch antenna with embroidered vias for on-body communication. *IET Microw Antennas Propag* 10:802–807
11. IEEE (1999) IEEE standard for safety levels with respect to human exposure to radio frequency electromagnetic fields, 3 kHz to 300 GHz, IEEE Standard C95

12. Gabriel C, Gabriel S, Corthout E (1996) The dielectric properties of biological tissues: I. literature survey. *Phys Med Biol* 41:2231–2249
13. Andreuccetti D, Fossi R, Petrucci C (1997) An internet resource for the calculation of the dielectric properties of body tissues in the frequency range 10 Hz–100 GHz. IFAC-CNR, Florence (Italy)
14. Zhou L, Fang S, Jia X (2020) Dual-band and dual-polarized circular patch textile antenna for on-/off-body WBAN applications. *IET Microw Antennas Propag* 14:643–648



# Intelligent ICH Detection Using K-Nearest Neighbourhood, Support Vector Machine, and a PCA Enhanced Convolutional Neural Network



Shanu Nizarudeen and Ganesh R. Shunmugavel

## 1 Introduction

Stroke remains a leading cause of fatality with intracranial haemorrhage (ICH) slightly more problematic than ischaemic [9]. Hypertension and trauma are the dominant causes of stroke [6], and the global lifetime risk of stroke has a higher probability of prevalence [7]. Based on the haemorrhage location, different types of ICH include epidural haemorrhage (EPH), intraparenchymal haemorrhage (IPH), intraventricular haemorrhage (IVH), subdural haemorrhage (SDH), and subarachnoid haemorrhage (SAH). Severe headache and lack of consciousness are the common symptoms, and misdiagnosis leads to mortality and morbidity [13]. Importantly, non-contrast computed tomography (NCCT) scans confirm haemorrhage, and it is observed that NCCT is sensitive even in the initial six hours of headache [20]. Further, additional burden on clinical management is imposed due to inadequate neuro-care amidst COVID-19 lockdown restrictions [23] and risk for ICH in COVID-19 patients [17]. Importantly, ICH is a crucial health problem with mortality or disability as its outcome, and clinical management of ICH involves earlier detection and aggressive treatment [24].

Intelligent detection algorithms applied various classifiers such as K-means [10], artificial neural network (ANN) [28], K-nearest neighbour (KNN) [1], support vector machine (SVM) [19], Bayesian, and decision tree [14]. The K-means algorithm classifies objects in the image as K-clusters. The distance from the cluster centroid determines the grouping criterion for a new object. Recently, artificial neural networks are successfully applied in image classification tasks. They are complex layers of

---

S. Nizarudeen (✉) · G. R. Shunmugavel  
Department of ECE, Noorul Islam Centre for Higher Education, Kumaracoil, Tamil Nadu, India  
e-mail: [shanu@ceconline.edu](mailto:shanu@ceconline.edu)

G. R. Shunmugavel  
e-mail: [ganesh@ninuniv.com](mailto:ganesh@ninuniv.com)

neurons with each layer receiving input from the previous layer. Another important algorithm KNN depends on the distance from K-nearest neighbours for classification. Recently, the SVM method provided accuracy in classification and regression tasks. The method depends on orienting a hyperplane used to separate two classes to increase the distance between the closest points in each class. Recently, an SVM-based automatic haemorrhage classification and segmentation algorithm achieved an accuracy of 92.46 and 94.13 for binary classification of IVH from normal. The paper presented a multiclass classification accuracy of 96.15, 95.96, and 94.87 for categorizing EDH, ICH, and SDH [27]. They applied feature extraction based on a hierarchical classifier. The primary classifier detects whether the pixels are IVH or normal, and the SVM-based second classifier identifies the subtype [27]. Further, Bayesian is an extremely scalable probability-based classifier that minimizes the classification error. The method is applied in muscle segmentation from clinical CT [8] and soft segmentation of brain data [15]. Another group implemented a decision tree-based stroke and haemorrhage detection algorithm and achieved premium performance [14]. However, the performance of these algorithms degrades in factors such as dependence of initialization for K-means, large execution time (for large data set) in SVM, black-box approach in ANN, a priori knowledge in Bayesian, and the unstable nature of decision tree. Studies depicted that CNN is a dominant method in image classification [2, 21, 25]. Moreover, studies demonstrated that CNN significantly improved automated detection of ICH [29]. Further, accuracy improved from 83.7 to 89.7%, and processing time reduced from 68 to 43 s using automated approaches [29].

CNN aids in the diagnosis of different ICH types, and image thresholding improves accuracy and performance [11]. Another group applied post-processing to deep CNN and improved specificity to 98% [18]. Ko et al. considered a CNN-LSTM network to assist radiologists in disease diagnosis [12]. Further, applying attention maps and predictive bias, the team explained how the method mimics the workflow of radiologists [16]. The algorithm achieved a specificity >92% and specificity 95% on two independent data sets [16]. A more recent work applied CNN with bidirectional LSTM for ICH detection and subtype classification. This method achieved performance comparable to a trained practitioner [3]. Thus, an increase in performance is noted in a hybrid CNN-based automated approach with an accuracy, sensitivity, and specificity of 97.5%, 98.3%, and 97.1%, respectively [4]. In this method, to select the highest ranking region, the classifier draws a set of bounding boxes of pre-defined shapes and sizes on the CT scan [4]. However, the method is complex and required dedicated hardware preventing widespread implementation. This paper outlines a simple and interesting CNN-based work applying feature selection using principal component analysis (PCA). We have implemented an intelligent ICH detection system using KNN and SVM. Further, we have provided a multi-channel visualization of layer-wise interpretation for a sample scan in the proposed model. We believe that this implementation can be applied as a stand-alone unit to assist radiologists especially in a clinical set-up with limited resources.

The methodology of the proposed algorithm is explained in Sect. 2. Here, we have defined the three different algorithms in our experiment. Result analysis and

discussions are presented in Sect. 3. Further, a brief explanation on the performance metrics applied to analyse the algorithms is given. It is followed by multichannel visualization of layer interpretations and finally conclusion in Sect. 4.

## 2 Methodology

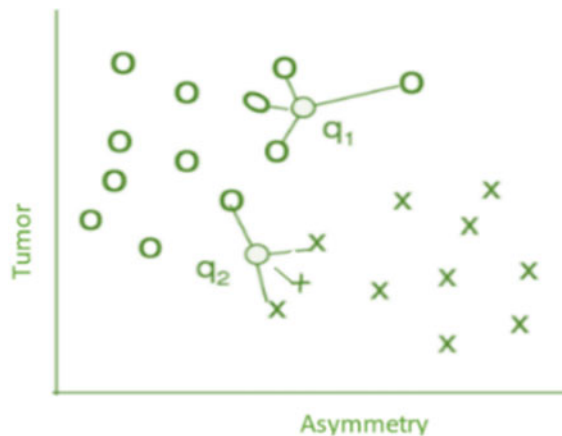
### 2.1 Data Set and Preprocessing

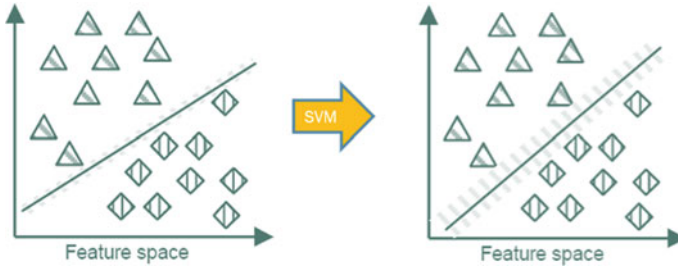
The data set comprises publicly available non-contrast CT images (NCCT) from <https://www.kaggle.com/>. The data set consists of 100 haemorrhage and 100 non-haemorrhage NCCTs. The preprocessing steps applied include data augmentation using intensity normalization, zoom (10%), height-width shift (10% each), and horizontal flip. For improving the performance of the classifier, we applied feature extraction using PCA. The PCA is an important dimensionality reduction tool that helps in improving classification accuracy [22]. Moreover, the elimination of highly correlated data using PCA reduces overfitting.

### 2.2 KNN-Based Classifier

The theme behind the KNN method is to classify samples based on their nearest neighbours [5]. Figure 1 shows a two-class problem based on nearest neighbours. The classification of  $q_1$  is simple as its four neighbours are in class O, while for  $q_2$ , three neighbours are from class X and one from class O. Assume that a new sample  $q$  is to be classified. At first, the distance between  $q$  and each sample  $s_i$  is computed

**Fig. 1** Two-class problem based on four nearest neighbours





**Fig. 2** Hyperplane (line) separates the two classes and SVM orients the hyperplane

using Eq. 1, where  $w_f$  is the weight of each feature  $f$  in the feature space  $F$ .

$$d(q, s_i) = \sum_{f \in F} w_f \delta(q_f, s_{if}) \quad (1)$$

Finally, K-nearest neighbours are selected using this distance metric, and a majority vote is applied to resolve the class of  $q$ . In our method, we applied the ball tree algorithm to compute the nearest neighbours using two different distance metrics: *Minkowski and Wasserstein*. We proposed to study the variations in performance using the two metrics.

### 2.3 SVM-Based Classifier

SVM is a popular approach in classification and regression. In this method, a hyperplane is inserted between the classes, and its orientation is altered to increase the separation between the nearest data points. Figure 2 shows the class separation using a hyperplane (line) [26]. The hyperplane depends on the selection of a kernel function, and based on the complexity of the problem, these functions are modified for decision-making. In our experiment, we analysed the classifier with three kernels: *SVM-linear, SVM-polynomial, and SVM-radial basis function (RBF)*.

### 2.4 CNN-Based Classifier

CNN models are employed in a variety of medical image classification tasks. Each model applies multiple convolutional and pooling layers. The convolutional layers slide a filter over the input image and learn better approximations of the image. In a deep CNN, earlier layers learn simple features such as edges and deeper layers learn more abstract features. Pooling layers provide robust translational invariance

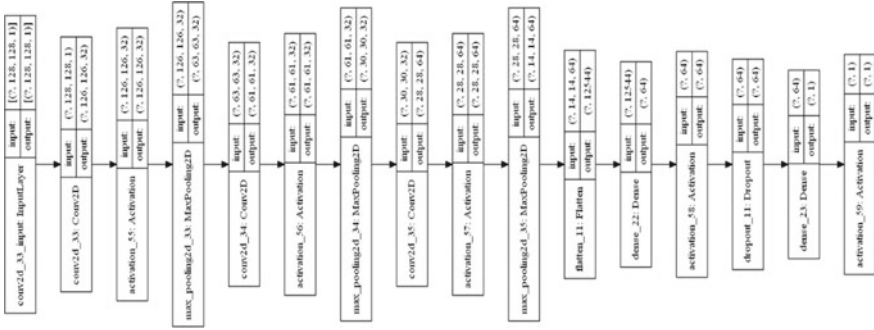


Fig. 3 Model architecture

and dropout layers prevent overfitting. The CNN architecture we implemented in our experiment is depicted in Fig. 3.

We analysed the model using 2D convolutions with (3, 3) kernels. To study the model’s performance on the input image size, we applied images of several sizes  $100 \times 100$ ,  $110 \times 110$ ,  $120 \times 120$ ,  $130 \times 130$ , and  $140 \times 140$  (Table 1).

Table 1 Type of layer, layer output, and the quantum of parameters in each layer of the proposed model

Serial #	Layer (type)	Output shape	Param #
1	conv2d_30 (Conv2D)	(None, 126, 126, 32)	320
2	activation_50 (Activation)	(None, 126, 126, 32)	0
3	max_pooling2d_30 (MaxPooling)	(None, 63, 63, 32)	0
4	conv2d_31 (Conv2D)	(None, 61, 61, 32)	9248
5	activation_51 (Activation)	(None, 61, 61, 32)	0
6	max_pooling2d_31 (MaxPooling)	(None, 30, 30, 32)	0
7	conv2d_32 (Conv2D)	(None, 28, 28, 64)	18,496
8	activation_52 (Activation)	(None, 28, 28, 64)	0
9	max_pooling2d_32 (MaxPooling)	(None, 14, 14, 64)	0
10	flatten_10 (Flatten)	(None, 12,544)	0
11	dense_20 (Dense)	(None, 64)	802,880
12	activation_53 (Activation)	(None, 64)	0
13	dropout_10 (Dropout)	(None, 64)	0
14	dense_21 (Dense)	(None, 1)	65
15	activation_54 (Activation)	(None, 1)	0

### 3 Results and Discussion

To evaluate our classifier, we first analyzed the average accuracy of KNN and SVM models for image size variation from  $100 \times 100$  to  $140 \times 140$  with an increment of ten pixels in either dimensions. In our method, KNN was benchmarked with two distance metrics and SVMs with three kernel functions. The KNN-Minkowski model outperformed the KNN-Wasserstein model with an average accuracy of 85% and 77.5%, respectively. In the SVM-based approach, both SVM-polynomial and SVM-RBF yielded an average accuracy of 87.5% (Table 2).

#### 3.1 Performance Metrics

We computed sensitivity ( $S$ ), specificity, and F1-score to estimate the classification accuracy of our algorithm. Sensitivity is the model's capability to recognize subjects without the disease. It is defined as

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (2)$$

where TN denotes true negatives and FP denotes false positives. Similarly, specificity is model's capability to recognize subjects with a disease.

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \quad (3)$$

where TP is the true positives and FN is the false negatives.

F1-score is the harmonic average of sensitivity and precision ( $P$ ) [11]. It is the measure of the algorithm's accuracy on a data set. The F1-score is defined as

$$\text{F1-score} = \frac{2}{\frac{1}{S} + \frac{1}{P}} \quad (4)$$

**Table 2** Benchmarking average accuracy in KNN and SVM models

Image size	KNN-Wasserstein	KNN-Minkowski	SVM-linear	SVM-polynomial	SVM-RBF
100	60	70	77.5	80	60
110	77.5	85	80	<b>87.5</b>	<b>87.5</b>
120	55	65	67.5	72.5	70
130	77.5	72.5	77.5	75	85
140	75	77.5	65	75	75

The maximum value obtained is highlighted in bold font

**Table 3** Performance metrics of CNN and PCA-CNN

	Sensitivity	Specificity	F1-score
CNN	1.0	0.91	0.95
PCA-CNN	1.0	1.0	1.0

Specificity and F1-score improved in PCA-CNN

Further, we noted that when PCA-extracted feature scans are applied to the CNN, there is a 9% improvement in specificity and a 5% improvement in F1-score. An F1-score of 1 shows absolute sensitivity and precision. The sensitivity was 100% in both cases as indicated in Table 3.

### 3.2 Multichannel Visualization of Layer Interpretations

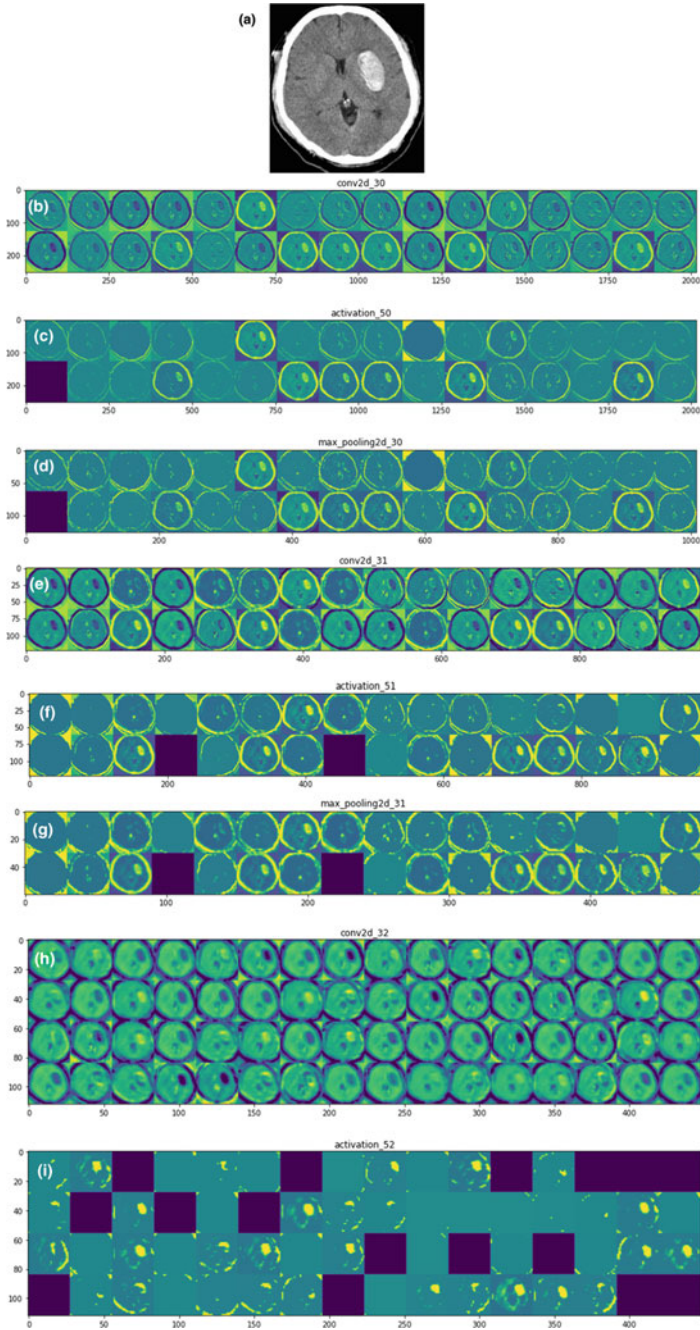
It is important to visualize and interpret representations learned by various layers in the model for each channel (width, height, and depth). We tested the model with a sample scan shown in Fig. 4a. The channel output of the first eight activation maps is then visualized to understand how the sample scan is decomposed into the convolutions of the network (Fig. 4b–i). We noted that the initial layers in Fig. 4b of the model “learn” edges, and later layers learn more abstract structures of the different tissues. The activations become sparse as we move from Fig. 4b to i.

### 3.3 Training and Validation Accuracy

To monitor the training accuracy of the model on unseen data, we have applied an 80:20 split. We further selected the validation data from the training set. We then trained the model for 100 epochs in mini-batches of ten samples and plotted training and validation accuracy in Fig. 5. Further, we plotted the training and validation loss in Fig. 6. The validation accuracy curve indicates that the algorithm provides excellent generalization. Also, training and validation loss decreased with every epoch as shown in Fig. 6

## 4 Conclusion

We have implemented an intelligent ICH detection system using KNN and SVM. We have analysed KNN with two distance metrics and SVM with three kernel functions. Further, we have developed a simple and interesting CNN model applying feature selection using principal component analysis. The sensitivity and F1-score improved by 9% and 5%, respectively. Furthermore, we have presented a visualization of layer



**Fig. 4** Multichannel visualization of layer interpretations. **a** Sample scan, **b–i** visual interpretations of layers 1–8



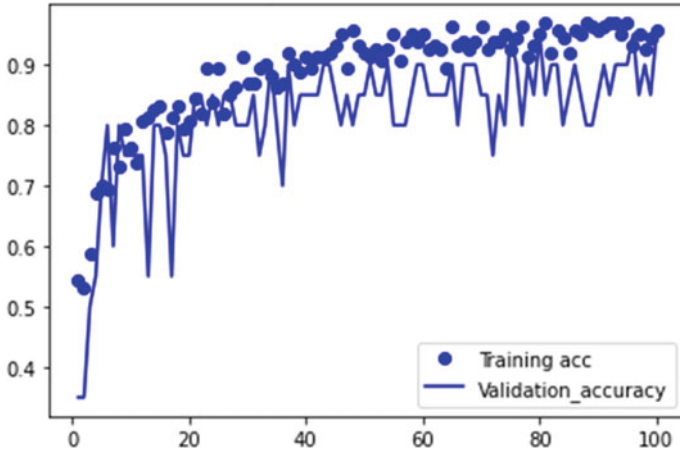


Fig. 5 Training and validation accuracy

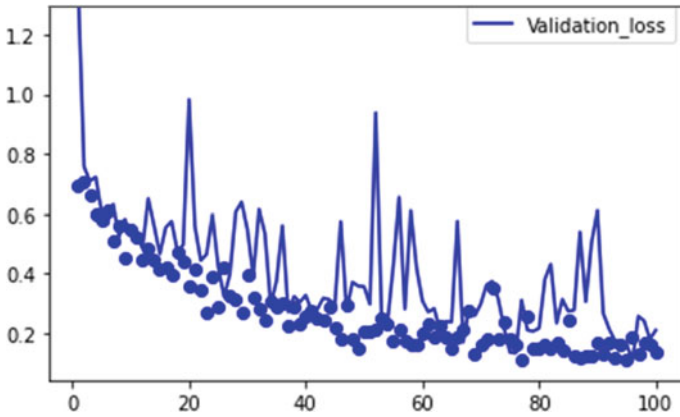


Fig. 6 Training and validation loss

representations for a sample scan to eliminate the black-box effect of the proposed CNN. Importantly, we believe that our model can be implemented as an assistive tool for the radiologist.

## References

1. Abu Alfeilat HA et al (2019) Effects of distance measure choice on K-Nearest neighbor classifier performance: a review. *Big Data* 221–248. <https://doi.org/10.1089/big.2018.0175>
2. Ben-Cohen A et al (2019) Improving CNN training using disentanglement for liver lesion classification in CT. In: *Proceedings of the annual international conference of the IEEE engineering*

- in medicine and biology society, EMBS, Institute of electrical and electronics engineers Inc., pp 886–889. <https://doi.org/10.1109/EMBC.2019.8857465>
3. Burduja M, Ionescu RT, Verga N (2020) Accurate and efficient intracranial hemorrhage detection and subtype classification in 3D CT Scans with convolutional and long short-term memory neural networks. *Sensors* 20(19):5611. <https://doi.org/10.3390/s20195611>
  4. Chang PD et al (2018) Hybrid 3D/2D convolutional neural network for hemorrhage evaluation on head CT. *Am J Neuroradiol* 39(9):1609–1616. <https://doi.org/10.3174/ajnr.A5742>
  5. Cunningham P, Delany SJ (2021) k-Nearest neighbour classifiers—a tutorial. *ACM Comput Surv* 54(6):1–25. <https://doi.org/10.1145/3459665>
  6. Dastur CK, Yu W (2017) Current management of spontaneous intracerebral haemorrhage. *Stroke Vasc Neurol* 21–29. <https://doi.org/10.1136/svn-2016-000047>
  7. Gorelick PB (2019) The global burden of stroke: persistent and disabling, *Lancet Neurol* 417–418. [https://doi.org/10.1016/S1474-4422\(19\)30030-4](https://doi.org/10.1016/S1474-4422(19)30030-4)
  8. Hiasa Y et al (2020) Automated muscle segmentation from clinical CT using Bayesian U-Net for personalized musculoskeletal modeling. *IEEE Trans Med Imaging* 39(4):1030–1040. <https://doi.org/10.1109/TMI.2019.2940555>
  9. Johnson CO et al (2019) Global, regional, and national burden of stroke, 1990–2016: a systematic analysis for the global burden of disease study 2016. *Lancet Neurol* 18(5):439–458. [https://doi.org/10.1016/S1474-4422\(19\)30034-1](https://doi.org/10.1016/S1474-4422(19)30034-1)
  10. Kamila UK, Bandyopadhyay O, Biswas A (2019) Detection of hemorrhagic region in brain MRI. In: *Lecture notes in networks and systems*, Springer, pp 383–391. [https://doi.org/10.1007/978-981-13-1217-5\\_38](https://doi.org/10.1007/978-981-13-1217-5_38)
  11. Ker J et al (2019) Image thresholding improves 3-dimensional convolutional neural network diagnosis of different acute brain hemorrhages on computed tomography scans. *Sensors (Switzerland)* 19(9). <https://doi.org/10.3390/s19092167>
  12. Ko H et al (2020) Feasible study on intracranial hemorrhage detection and classification using a CNN-LSTM network. In: *Proceedings of the annual international conference of the IEEE engineering in medicine and biology society, EMBS, July 2020*, pp 1290–1293. <https://doi.org/10.1109/EMBC44109.2020.9176162>
  13. Kowalski RG et al (2004) Initial misdiagnosis and outcome after subarachnoid hemorrhage. *J Am Med Assoc* 291(7):866–869. <https://doi.org/10.1001/jama.291.7.866>
  14. Kumar DV, Jaya Rama Krishniah VV (2016) An automated framework for stroke and hemorrhage detection using decision tree classifier. In: *Proceedings of the International Conference on Communication and Electronics Systems, ICCES 2016*, Institute of electrical and electronics engineers Inc. <https://doi.org/10.1109/CESYS.2016.7889861>
  15. Lauric A, Frisken S (2016) Soft segmentation of CT brain data soft segmentation of CT brain data (Jan 2007)
  16. Lee H et al (2019) An explainable deep-learning algorithm for the detection of acute intracranial haemorrhage from small datasets. *Nat Biomed Eng* 3(3):173–182. <https://doi.org/10.1038/s41551-018-0324-9>
  17. Loizou CP et al (2012) Video segmentation of the common carotid artery intima media complex. In: *2012 IEEE 12th International conference on Bioinformatics and Bioengineering (BIBE)*, IEEE, pp 500–505. <https://doi.org/10.1109/BIBE.2012.6399728>
  18. Majumdar A et al (2018) Detecting intracranial hemorrhage with deep learning. In: *Proceedings of the annual international conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, pp 583–587. <https://doi.org/10.1109/EMBC.2018.8512336>
  19. Milosevic M, Jovanovic Z, Jankovic D (2017) A comparison of methods for three-class mammograms classification. *Technol Health Care* 25(4):657–670. <https://doi.org/10.3233/THC-160805>
  20. Perry JJ et al (2011) Sensitivity of computed tomography performed within six hours of onset of headache for diagnosis of subarachnoid haemorrhage: prospective cohort study. *BMJ (Online)* 343(7817). <https://doi.org/10.1136/bmj.d4277>
  21. Polsinelli M, Cinque L, Placidi G (2020) A light CNN for detecting COVID-19 from CT scans of the chest. *Pattern Recogn Lett* 140:95–100. <https://doi.org/10.1016/j.patrec.2020.10.001>

22. Praveen K et al (2021) A simplified framework for the detection of intracranial hemorrhage in CT brain images using deep learning. *Curr Med Imaging Formerly: Curr Med Imaging Rev* 17. <https://doi.org/10.2174/1573405617666210218100641>
23. Pujol-Lereis VA et al (2021) COVID-19 lockdown effects on acute stroke care in Latin America. *J Stroke Cerebrovasc Dis: Offic J Nati Stroke Assoc* 30(9):105985. <https://doi.org/10.1016/j.jstrokecerebrovasdis.2021.105985>
24. Qureshi AI, Mendelow AD, Hanley DF (2009) Intracerebral haemorrhage. *Lancet* 1632–1644. [https://doi.org/10.1016/S0140-6736\(09\)60371-8](https://doi.org/10.1016/S0140-6736(09)60371-8)
25. Schwendicke F et al (2019) Convolutional neural networks for dental image diagnostics: a scoping review. *J Dent* <https://doi.org/10.1016/j.jdent.2019.103226>
26. Sewak M et al (2008) SVM approach to breast cancer classification. In: *Institute of Electrical and Electronics Engineers (IEEE)*, pp 32–37. <https://doi.org/10.1109/imscs.2007.46>
27. Shahangian B, Pourghassem H (2016) Automatic brain hemorrhage segmentation and classification algorithm based on weighted grayscale histogram feature in a hierarchical classification structure. *Biocybern Biomed Eng* 36(1):217–232. <https://doi.org/10.1016/j.bbe.2015.12.001>
28. Wang JL, Jin GL, Yuan ZG (2021) Artificial neural network predicts hemorrhagic contusions following decompressive craniotomy in traumatic brain injury. *J Neurosurg Sci* 65(1):69–74. <https://doi.org/10.23736/S0390-5616.17.04123-6>
29. Watanabe Y et al (2020) Improvement of the diagnostic accuracy for intracranial haemorrhage using deep learning-based computer-assisted detection. *Neuroradiology*. <https://doi.org/10.1007/s00234-020-02566-x>

# Feature Selection and Diagnosis Performance Evaluation of Breast Cancer



Anu Babu and S. Albert Jerome

## 1 Introduction

An umbrella that blends correlated diseases is called cancer. It will lead to the uncontrolled growth of cells. Uncharacteristic developments in the breast like lump, shape change, inverted nipple, fluid from the nipple, etc., can be symptoms of breast cancer. Extreme care must be given to prevent and cure this disease because WHO 2020 survey reveals that among different cancers, breast cancer occupies the first position [1].

It is not possible to avert any type of cancer. The early finding is the only measure to increase the persistence rate [2]. It is a very awkward process to make an accurate diagnosis into benign and malignant within a short period. This quandary can be resolved by including a machine learning model in the diagnosis process as a reviewer for physicians [3]. It is a model which identifies the pattern of disease from digital images.

The premier intention of this work is to develop a machine learning framework that will proficiently assort breast cancer-affected people and healthy ones. Supervised learning algorithms such as SVM,  $K$ -NN, NB, DT and RF are used for the recognition of affected ones. For classifier performance refinement, the Lasso feature selection method has been used. Performance of model with and without feature selection is evaluated by computing metrics like accuracy, misclassification, sensitivity, specificity, F1 score, MCC and  $J$ .

---

A. Babu (✉)

Department of Electronics and Communication Engineering, Noorul Islam Centre for Higher Education, Kumaracoil, Tamil Nadu, India  
e-mail: [anubabu123@gmail.com](mailto:anubabu123@gmail.com)

S. Albert Jerome

Department of Biomedical Engineering, Noorul Islam Centre for Higher Education, Kumaracoil, Tamil Nadu, India

## 2 Related Work

Various techniques have been used for the classification of breast cancer. [4] obtained 96% sensitivity, 93% specificity and 92% accuracy by employing grasshopper optimization algorithm over SVM classifier on Mammographic Image Analysis Society (MIAS) and Digital Database for Screening Mammography (DDSM). [5] introduced an integrated method for breast cancer detection employing minimal redundancy maximal relevance and chi-square algorithm with SVM, and an accuracy of 99.71% is achieved from minimal redundancy maximal relevance SVM. Ten cross-validations were proposed in [6] for the development of the nested ensemble technique for automated diagnosis. An accuracy of 98.07% was scored in both Naive Bayes and BayesNet algorithms.

## 3 Background

### 3.1 Data set

The experiments are carried out by retrieving fine needle aspiration of breast cancer images from the Breast Cancer Wisconsin Diagnostic data set [7]. The data set consists of 569 instances out of which 357 benign and 212 malignant cases. In this work, malignant is identified as a negative class with zero value, and benign is identified as a positive class with value one. For the prediction of cancer, the ten features which describe the characteristics of cell nuclei are being used.

### 3.2 Data Set Pre-processing

The pre-processing of the data set will aid in the precise diagnosis of cancer. The steps involved in this process are removal of missing value, standard scalar and correlation checking.

### 3.3 Feature Selection Algorithm: Least Absolute Shrinkage Selection Operator

Feature selection is a crucial step in classification which selects the most suitable features for predictive system. Lasso feature selection technique is used in this paper. The absolute coefficient feature values are amended for feature selection in Lasso. The selection process is performed by shrinking the coefficient values. The features

with zero coefficient values are removed from feature set. It will create a new subset of features which include only relevant features [8].

### ***3.4 Support Vector Machine***

SVM is a ML algorithm which is mainly used for classification. The finest classification function is computed by constructing a hyperplane which separates the positive and negative cases in feature space. The finest function is the one which inflates the margin between two classes. The shortest distance between the closest datapoint to the point on the hyperplane is called margin. The key benefit of SVM is the reservation of free space for classification of future process [9].

### ***3.5 K-Nearest Neighbour***

Predictive analysis is largely supported by  $K$ -NN classifier. It will make an evaluation between test data and train data that are closer. The training data which are quantified by  $n$  traits are stored in  $n$  dimensional metric space.  $K$ -NN searches for  $K$  training data which have similar traits of new data in metric space during the arrival of new data [9]. Minkowshi distance is used to compute the closeness of data.

### ***3.6 Naive Bayes***

Among all machine learning algorithms, simplest and easiest classifier is Naive Bayes which is based on Bayes theorem. Naive Bayes will generate a rule which can assign future objects to a class using given feature vectors of variables alone from a set of objects with known class and known feature vectors of variables. It is also called as idiot's Bayes since no skill is needed to perform this operation [9].

### ***3.7 Decision Tree***

Supervised learning algorithm that can be used to resolve both classification and regression issues is called DT. DT is a tree-like structure with root starting point, internal node of tree as test, branch as test outcome and class label as leaf. Primary function of algorithm is to make decisions, and best decision is the choice with maximum information or with minimum entropy [10].

### 3.8 *Random Forest*

RF is a machine learning technique that employs ensemble learning for solving complex issues. It will generate many decisions trees called forest which is trained by bagging. The ultimate verdict is based on average of output from various trees. The precision of outcome can be improved by having greater number of trees [10].

### 3.9 *Evaluation Measures*

Classifier performance has been evaluated by calculating performance metrics such as accuracy, misclassification, sensitivity, specificity, F1 score, Mathew's correlation coefficient (MCC) and Youden's index ( $J$ ). Performance measuring parameters used here are true positive (TP), true negative (TN), false positive (FP) and false negative (FN). Accuracy is the proportion of samples properly classified into total samples. Equation (1) is used to calculate accuracy of machine model.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (1)$$

Misclassification gives the amount of incorrectly classified samples. The lower the value, better the ML model. Equation (2) is applied to compute misclassification.

$$\text{Misclassification} = \frac{\text{FP} + \text{FN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (2)$$

Sensitivity is the probability that result of a ML model belongs to positive class given that sample presents cancer. Equation (3) gives sensitivity of a model.

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (3)$$

Specificity is the probability that result of a ML belongs to negative class given that samples not having cancer. Equation (4) is used to compute specificity.

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \quad (4)$$

Harmonic mean between sensitivity and specificity gives F1 score. The perfect value of F1 score is unity, and a value close to unity indicates a model with low FP and FN. Equation (4) is used to calculate F1 score.

$$\text{F1 Score} = \frac{2 * \text{Sensitivity} * \text{Specificity}}{\text{Sensitivity} + \text{Specificity}} \quad (5)$$

Mathew’s correlation coefficient is a quantity used to measure quality of a ML model. The range of MCC value is between  $-1$  and  $+1$ . MCC will produce a score near to unity only if prediction of all four categories (TP, TN, FP and FN) has good result. Equation (6) is used to calculate MCC.

$$MCC = \frac{(TP*TN) - (FP*FN)}{\sqrt{(TP + FN)(TP + FP)(TN + FP)(TN + FN)}} \tag{6}$$

The evaluation metric that summarizes the performance of a ML model is Youden’s index. Its value lies between zero and one. The larger the value, the better the performance. Equation (7) is used to compute  $J$ .

$$J = \text{Sensitivity} + \text{Specificity} - 1 \tag{7}$$

## 4 Proposed Methodology

The pre-processed data set has undergone the Lasso feature selection method for choosing significant features for precise breast cancer diagnosis. The features selected by Lasso for diagnosis are mean concavity, mean concave point, area error, compactness error, worst texture, worst area, worst smoothness, worst concave points and worst symmetry. Then performance checking of different classifiers is carried out on the full feature set and feature subset. Various performance assessment metrics were used for evaluation purposes.

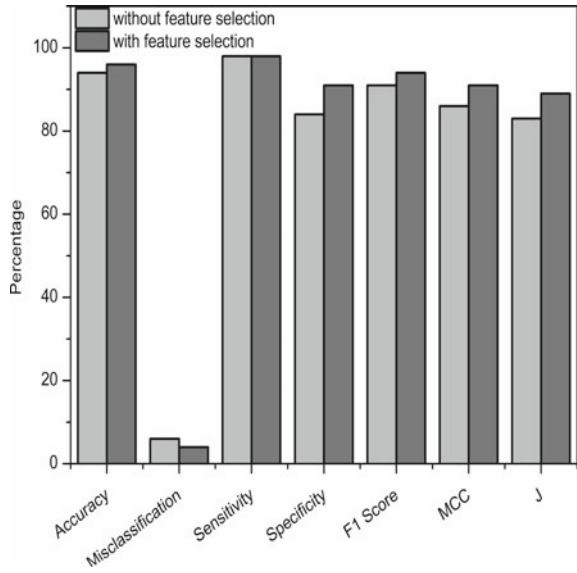
## 5 Results and Discussion

Classifiers performance on entire feature set and subset selected by Lasso algorithm is evaluated in the succeeding section. Primarily, the SVM classifier is considered, and its performance is evaluated in terms of accuracy, misclassification, sensitivity, specificity, F1 score, MCC and J on non-selected and selected feature set. The outcome of the evaluation process is shown in Fig. 1. The performance of the SVM classifier in terms of all evaluation metrics improved in a significant manner. The accuracy of the whole feature was 94% improved to 96% on selected features. Similarly, misclassification reduced from 6 to 4%. The value of detecting a healthy person remains the same as 98% which is a good value and declares 84 to 91% improvement in detection of a person with cancer. F1 score, MCC and J values also show increased value on feature subset.

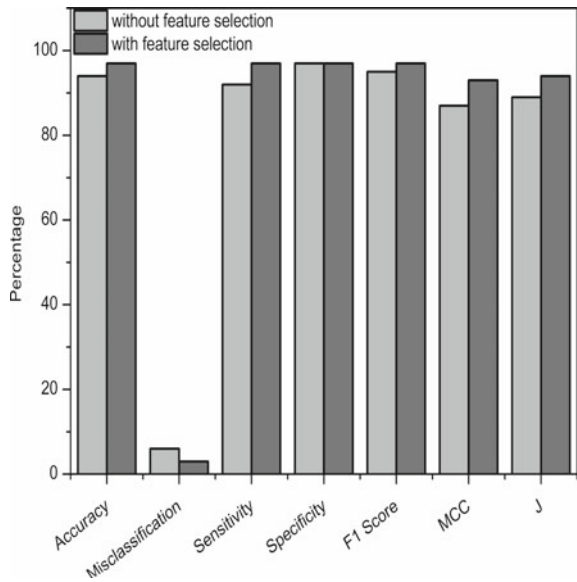
Features selected by Lasso were classified using  $K$ -NN, and the performance was evaluated by computing assessment parameters. According to the graphical result shown in Fig. 2,  $K$ -NN achieved very good performance after using Lasso feature



**Fig. 1** SVM classifier performance with and without Lasso feature selection

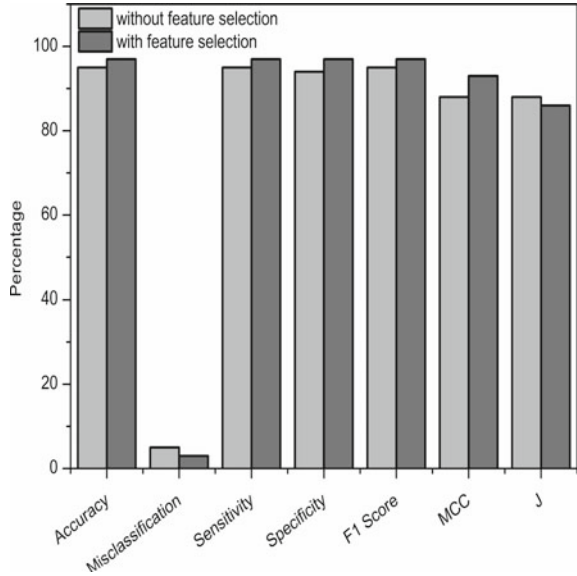


**Fig. 2** K-NN classifier performance with and without Lasso feature selection



selection. The values attained by classifier before applying feature selection are of 94% accuracy, 6% misclassification, 92% sensitivity, 97% specificity, F1 score 95%, 87% MCC and 89% J. Lasso feature selection results in increased value of 97% accuracy, 3% misclassification, 97% sensitivity, 97% specificity, F1 score 97%, 93% MCC and 94% J. The 97% sensitivity and specificity show that K-NN is the very

**Fig. 3** NB classifier performance with and without Lasso feature selection



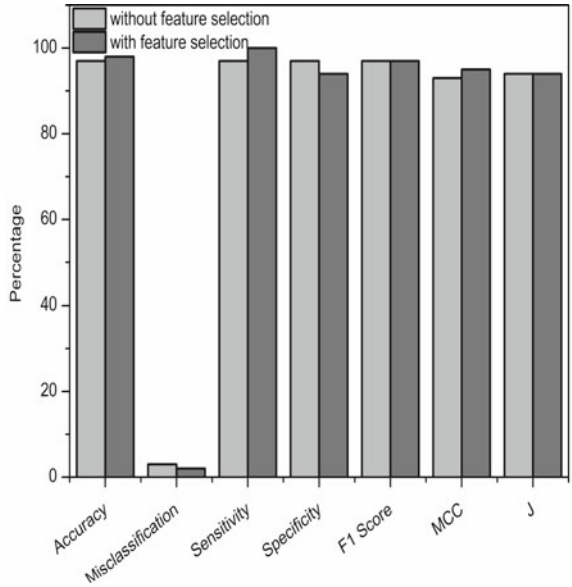
best algorithm for diagnosing healthy and cancer-affected people. This algorithm also can detect all categories of disease which is proved by 93% MCC and 94% J.

The pursuance of the NB classifier is assessed by computing various performance metrics for with and without feature selection. The results of the calculation are graphically demonstrated in Fig. 3. The results of NB with Lasso feature selection are high compared to that of without selection. The accuracy of NB on the full feature was 95% and is increased to 97% after Lasso feature selection. Misclassification reduced to 3 from 5%. This classifier has also shown improvement in the ability to diagnose healthy and cancerous patients with sensitivity and specificity values of 97% along with MCC 93% and F1 score 97%.

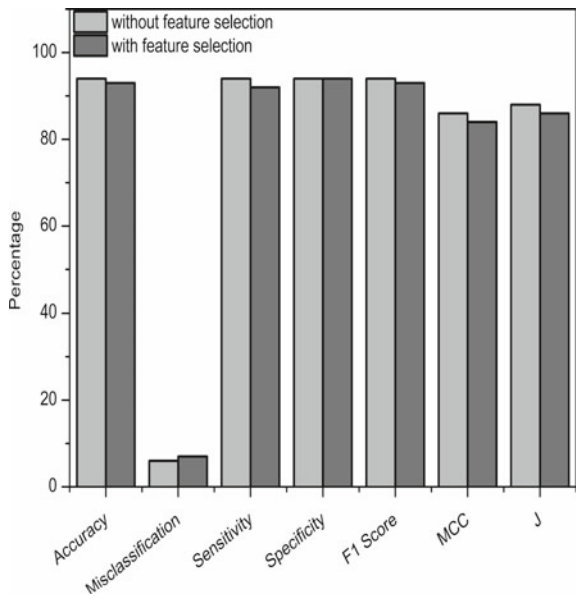
DT is the classifier that shows superior performance in identifying affected patients without feature selection with 97% accuracy, 3% misclassification, 97% sensitivity and specificity, 97% F1 score, 93% MCC and 94% J. Inclusion of the Lasso model results in even more enhanced detection rate with accuracy 98%, misclassification 2%, 100% sensitivity, 94% specificity, 97% F1 score, 95% MCC and 94% Youden’s index. The performance evaluation of the DT classifier with and without feature selection is shown in Fig. 4.

The graphical view of RF classifier outcome with and without Lasso feature selection is shown in Fig. 5. The results show that the application of the Lasso feature subset results in the degradation of classifier performance in all metrics. With full features, the RF classifier achieved 94% accuracy which reduced to 93%, misclassification increased from 6 to 7%, sensitivity and specificity decreased from 94 to 92%, and 94% F1 score became 93% along with MCC reduction from 86 to 84%. The entire performance factor J is also diminished from 88 to 86%. The

**Fig. 4** DT classifier performance with and without Lasso feature selection



**Fig. 5** RF classifier performance with and without Lasso feature selection



analysis shows that the RF classifier with Lasso feature selection is not suitable for breast cancer detection.

## 6 Conclusion

In this paper, a feature selection method that is appropriate for different machine learning classifiers is suggested. To motif, a breast cancer diagnostic model, classifiers like SVM,  $K$ -NN, NB, DT and RF have been used. The consummation of the classifier will be drastically affected by the redundant features. To overcome this, a feature selection method called least absolute shrinkage and selection operator is introduced. The SVM and  $K$ -NN classifiers with Lasso have shown the best performance in terms of all performance metrics which indicates the best breast cancer diagnostic. NB with Lasso also resulted in outstanding achievement in diagnosing healthy and cancerous patients. DT achieved 100% sensitivity with Lasso features along with very good improvement in other features. Lasso feature selection with RF classifier results proclaims that RF with Lasso features is not suitable for breast cancer diagnosis. The results culminate that Lasso feature selection with machine learning classifiers will aid in a superior breast cancer diagnosis system.

## References

1. WHO (2021). <https://www.who.int/news-room/fact-sheets/detail/cancer>. Accessed May 2021
2. Abd Elaziz M, Lu S, He S (2021) A multi-leader whale optimization algorithm for global optimization and image segmentation. *Exp Syst Appl* 175:114841
3. Shah D, Mangrulkar RS (2021) Machine learning techniques to classify breast cancer. In: Design of intelligent applications using machine learning and deep learning techniques (2021). Chapman and Hall/CRC, pp 245–256
4. Sha Z, Hu L, Rouyendegh BD (2020) Deep learning and optimization algorithms for automatic breast cancer detection. *Int J Imaging Syst Technol* 30(2):495–506
5. Ul Haq A, Li J, Memon MH, Khan J, Ud Din S (2020) A novel integrated diagnosis method for breast cancer detection. *J Intell Fuzzy Syst* 38(2):2383–2398
6. Abdar M, Zomorodi-Moghadam M, Zhou X, Gururajan R, Tao X, Barua PD, Gururajan R (2020) A new nested ensemble technique for automated diagnosis of breast cancer. *Pattern Recogn Lett* 132:123–131
7. Breast Cancer Wisconsin Dataset. [https://archive.ics.uci.edu/ml/datasets/breast+cancer+wisconsin+\(original\)](https://archive.ics.uci.edu/ml/datasets/breast+cancer+wisconsin+(original)). Accessed May 2021
8. Li JP, Haq AU, Din SU, Khan J, Khan A, Saboor A (2020) Heart disease identification method using machine learning classification in e-healthcare. *IEEE Access* 8:107562–107582
9. Subasi A, Khateeb K, Brahimi T, Sarirete A (2020) Human activity recognition using machine learning methods in a smart healthcare environment. In: Innovation in health informatics. Academic Press, pp 123–144
10. Cohen S (2021) The basics of machine learning: strategies and techniques. In: Artificial intelligence and deep learning in pathology. Elsevier, pp 13–40

# Grading and Classification of Retinal Images for Detecting Diabetic Retinopathy Using Convolutional Neural Network



Neetha Merin Thomas and S. Albert Jerome

## 1 Introduction

Diabetic retinopathy (DR) is a typical diabetes difficulty that happens when the retina's veins are harmed because of high glucose levels, bringing about expanding and leakage of blood vessels. It is a progressive disease which damages blood vessels and results in irreversible vision loss [1]. The early stage of diabetic eye disease is known as non-proliferative diabetic retinopathy, and most of the people with diabetics for more than 16 years have this stage. In this stage, tiny blood vessels in the retina get weaker and cause leaks and swelling in the retina. If new abnormal blood vessels are grown on the surface of the retina, then it is proliferative diabetic retinopathy stage. This is an advanced stage of retinopathy, which leads to permanent vision loss [2]. The studies show that by 2030 people with DR will increase to 191 million and among these 56.3 million will lose their vision due to DR [3]. DR has no initial symptoms and as the stages of DR gets advanced symptoms like blurred vision, poor night vision and dark spots will develop. Research shows that, with frequent and proper monitoring and treatment, at least 90% of the new cases might be reduced [4, 5].

The previous work in the detection of DR stages was based on feature extraction and classification of characteristics by the use of different machine learning techniques. However, high accuracy can be accomplished utilizing these techniques yet diagnosing retinopathy dependent on feature extraction is a complex procedure [6]. With the introduction of deep learning in medical image processing, the classification of images becomes an easy task. Convolutional neural network (CNN) is a division of deep learning algorithm used for pattern recognition and image classification [7].

---

N. M. Thomas (✉) · S. Albert Jerome  
Noorul Islam Centre for Higher Education, Kumaracoil, Tamil Nadu, India  
e-mail: [neetha.ec@stistvm.edu.in](mailto:neetha.ec@stistvm.edu.in)

In this work, a CNN model is developed to classify the retinal fundus image into different classes based on the severity of DR like mild, moderate, severe, and proliferative DR. Deep learning (DL) was used extensively in the detection and classification of DR recently. Even if the images collected are heterogeneous in nature, DL can learn the features of input images correctly.

In [8], a method for classification of retinal images into DR or non-DR using CNN proposed. Nearly 1000 good images were taken from Kaggle dataset and perform some translations on images to increase the size of the dataset and obtained an accuracy of 94.5%. The work [9] developed a weighted path CNN (WP-CNN) model for the detection of DR. Sixty thousand images were collected from different datasets then augmented and resized it and finally labeled it as referable or non-referable DR. The WP-CNN with 105 layers gave an accuracy of 94.23%. According to [10], a system is presented which extracts features of image and is classified using CNN model. The images taken from Kaggle dataset was preprocessed and extract features like blood vessels and microaneurysms. The extracted features were given to CNN pretrained VGG16 architecture for multi-label classification and got an accuracy of 95.41%. The work [11] used AlexNet, VggNet, GoogleNet, and ResNet for the classification of DR. The features are initially extracted and given to the pretrained network. The accuracy of VggNet is better in this work compared to other pretrained network. The accuracy of VggNet is 95.8%. In [12], an algorithm for detecting DR based on red lesion feature using two CNN models proposed. A custom-made CNN and a pretrained VGG16 model were used for training the images taken from different dataset to classify the lesions as red lesions or not and obtained a sensitivity of 0.94. The work [13] developed a CNN architecture for detecting DR and diabetic macular edema (DME). In this paper, two publicly available datasets Messidor-2 and eyepacs-1 were used to test the model and obtained a sensitivity of 96.1% and 97.5%, respectively, for multi-class classification. The work [14] proposed an automated model for diagnosing DR using CNN. In this study, DR is categorized into five different stages by developing a CNN model with 10 convolutional layers, 8 max-pooling layer, and 3 fully connected layer. The images from Kaggle dataset were used for evaluation and obtained an accuracy of 75%, sensitivity of 95%, and specificity of 30%. According to [15], the paper compared the performance of some CNN pertained architectures like VGG16, AlexNet, InceptionNet-V3 using retinal images from Kaggle dataset. Only 166 images are selected for comparing the performance and obtained an accuracy of 50.03%, 37.43%, and 63.23% for VGG16, AlexNet, and InceptionNet-V3, respectively. According to [16], an automated system for the detection of DR using CNN was developed. It consists of different stages: preprocessing, extract blood vessels using CNN, exudates with fuzzy C-means, and texture features of blood vessels, and exudates are extracted and finally classify the images using support vector machine (SVM). In [17], a computer-aided method for the classification of DR using two networks was presented. One is a coarse network which performs binary classification, and a fine network is employed to classify the DR into four stages based on the severity grade [18].

## 2 Methodology

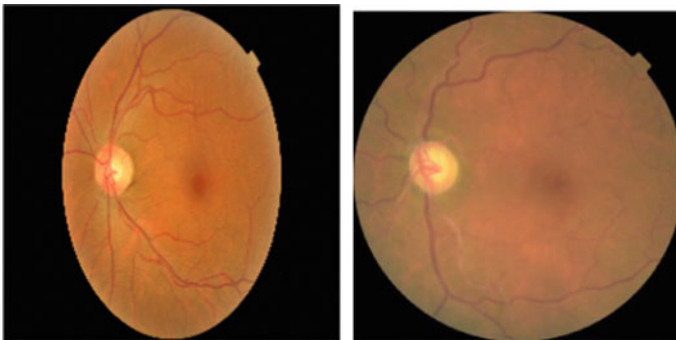
The images for this work were collected from Kaggle dataset. The images from the dataset are categorized into five classes such as No\_DR, mild, moderate, severe, and proliferative DR (PDR) based on lesions like microaneurysms, hemorrhages, and exudates.

### 2.1 Dataset

In this work, images are taken from APTOS 2019 dataset, it was a blindness, detection competition was conducted by Kaggle for DR detection. This dataset contains 3662 images, and these were separated into five different classes based on the severity such as No\_DR, mild, moderate, severe, proliferative DR.

### 2.2 Preprocessing

The images obtained from Kaggle dataset are heterogeneous in nature since it is taken with different fundus camera and has various illumination effects. In order to standardize the images, preprocessing is required. In this stage, the noise from the image is removed, resize the image into  $256 \times 256$ . Initially, the image is cropped to fit the image, the unwanted black broader is removed from the image. Gaussian filtering technique is applied to suppress the unwanted artifacts from the raw image, and these filtering techniques remove the noise as well as blur the image. Then the gaussian filtered image is resized into a smaller size as shown in Fig. 1 in order to reduce the computational complexity.



**Fig. 1** Original image and cropped resized image

### 2.3 Convolutional Neural Network Architecture (CNN)

Deep learning (DL) is a sub-division of artificial intelligence which imitates the behavior of human brain; it processes the data and gives out decision based on the information. Over the past few decades, deep learning method becomes more popular in many fields like object identification, translating languages, and medical image analysis. DL can distinguish precisely from the input information for characterization or division of classes and outperforms all conventional image analysis method [19].

A DL strategy does not have to separate the handcrafted highlights while it requires broad information for training the model. Conversely, AI methods require extraction of the handmade highlights; however, they do not require broad information for preparing the model [20]. CNN architecture mainly consists of three layers: a convolutional layer, a pooling layer, and a fully connected layer. Number of layers, the number of neurons in each layer, and number of filters required are determined based on the image to be classified [21].

The input image is given as whole to the CNN architecture. The proposed CNN consists mainly of three layers: first one is convolutional layer, where the features are extracted from image by convolving the images with different filters. The output of convolutional layer is known as feature map. Next these feature maps are given to pooling layer where the size of the feature map gets reduced as well as computational complexity. Then rectified linear unit (ReLU) activation followed by batch normalization is done on feature maps. Eight such layers are modeled in this proposed work in order to improve the accuracy of grading of the images into five classes. Finally, the extracted features are flattened before given to the fully connected layer.

The final set of layers comprises of one fully connected layer with 32 nodes, one dropout layer with 0.15 ratio, and finally another fully connected layer with five nodes with softmax activation, which provides the binary classification output (Fig. 2).

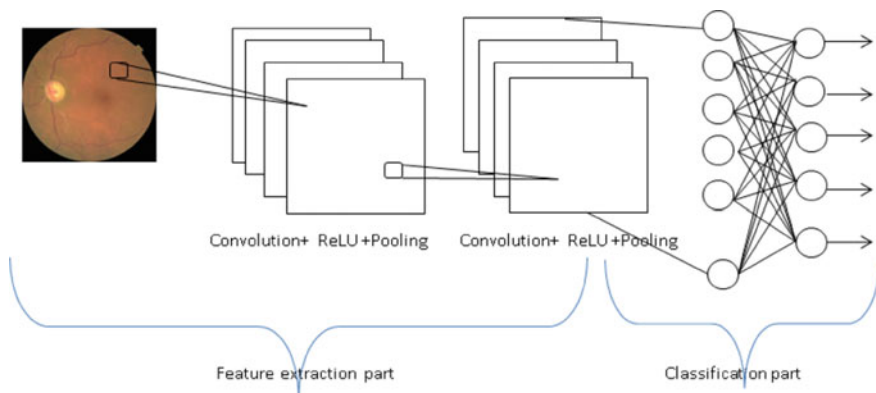


Fig. 2 Convolutional neural network architecture



### 3 Results and Discussion

The entire image set was divided into train, test, and validation set. 15% of the entire image is separated for testing purpose, 15% for validation purpose and the model is trained with the remaining 70% of images. The hyperparameters that were used in this model are given in Table 1. The model is first trained with 70% of the images that were divided into 32 batches. Adam optimizer is used in this work which helps to train the model efficiently. The loss function which defined here is binary cross-entropy which computes the cross-entropy loss between the true labels and the predicted labels. The loss function is calculated for each epoch.

The model is evaluated by finding the accuracy, precision, and recall using test dataset. For multi-class classification, the accuracy obtained is 85%. The accuracy is reduced because some of the mild DR images are misclassified as moderate DR. For a binary class classification, the accuracy of the model is 95.3%. If handcrafted features were included along with CNN classification model, the accuracy of multi-class classification can be improved. Table 2 shows the confusion matrix which explains the performance of the model with the validation dataset.

The accuracy, precision, and recall were calculated using Eqs. (1)–(3):

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

$$\text{Precision} = \frac{TP}{TP + FP} \tag{2}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{3}$$

**Table 1** Hyperparameters of the Proposed Model

Hyperparameters	Values
Optimizer	Adam
Loss function	Binary cross entropy
Batch size	32
Epoch	40
Learning rate	0.00001

**Table 2** Confusion matrix

Actual/predicted class	No DR	Mild	Moderate	Severe	PDR
No DR	268	3	0	0	0
Mild	4	28	24	0	0
Moderate	2	2	141	4	1
Severe	0	0	18	8	3
PDR	0	0	15	6	23

**Table 3** Performance of evaluation of multiclass classification

Grade	Precision	Recall
No DR	0.97	0.98
Mild	0.84	0.50
Moderate	0.71	0.94
Severe	0.44	0.27
PDR	0.85	0.52
Macro-average	0.76	0.64
Weighted average	0.84	0.85

**Table 4** Performance of evaluation of binary class classification

Grade	Precision	Recall
DR	0.94	0.95
No DR	0.95	0.94
Macro-average	0.94	0.94
Weighted average	0.94	0.94

Precision and recall as shown in Table 3 give a clear idea about how much our model correctly classifies the input images. Precision is defined as out of the examples that the learning algorithm marks as positive, how many are correctly positive. The precision for severe case is less compared to other cases, severe cases are not accurately identified by the model. Recall gives an idea about how many of the positive samples the learning algorithm retrieves as positive. Recall is less for mild and severe in this proposed work. The mild cases are wrongly identified as moderate DR, and thus the model fails to classify mild DR accurately.

The images are also classified into binary class such as DR or No DR using the same model and obtained an improved accuracy than the multiclass classification. The precision and recall for binary classification are shown in Table 4. The loss and accuracy curve for multiclass and binary class is shown in Figs 3 and 4, respectively.

## 4 Conclusion

Diabetic retinopathy can be detected early, which can greatly improve the chances of a successful recovery. However, clinical diagnosis is both expensive and time-consuming. The analysis of medical images in computer vision has a high processing speed and can provide improved prediction accuracy. In this research work, diabetic retinopathy images are taken from Kaggle dataset and are preprocessed to remove the artifacts and to improve the clarity by cropping, resizing, and filtering before classification purpose. Convolutional neural network is used for classification of images into five classes based on severity and obtained an accuracy of 85% as shown in Fig. 3 and obtained an accuracy of 95.3% for binary classification as shown in

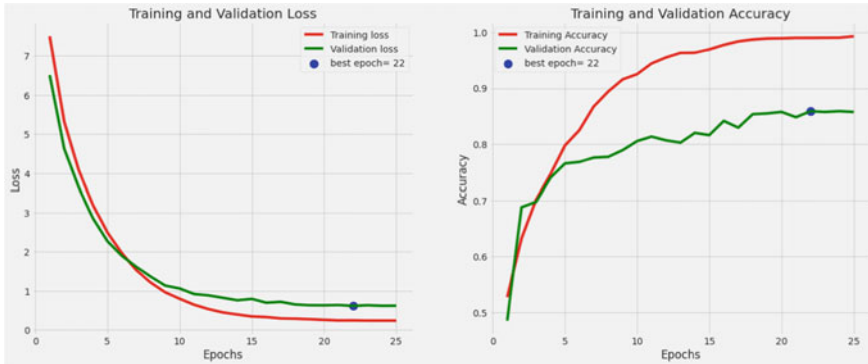


Fig. 3 Loss and accuracy curve of training and testing of the model for multiclass classification



Fig. 4 Loss and accuracy curve of training and testing of the model for binary class classification

Fig. 4. The accuracy of multi-class classification can be improved by incorporating handcrafted features along with CNN classification.

## References

1. Amin J, Sharif M, Yasmin M, Ali H, Lawrence S (2017) A method for the detection and classification of diabetic retinopathy using structural predictors of bright lesions. *J Comput Sci* 19:153–164
2. Pradhan A, Sarma B, Nath RK, Das A, Chakraborty A (2020) Diabetic retinopathy detection on retinal fundus images using convolutional neural network, vol 1240 CCIS. Springer, Singapore
3. Zheng Y, He M, Congdon N (2012) The worldwide epidemic of diabetic retinopathy. *Indian J Ophthalmol* 60(5):428–431
4. Tapp RJ et al (2003) The prevalence of and factors associated with diabetic retinopathy in the Australian population. *Diabetes Care* 26(6):1731–1737

5. Mansour RF (2018) Deep-learning-based automatic computer-aided diagnosis system for diabetic retinopathy. *Biomed Eng Lett* 8(1):41–57
6. Kar SS, Maity SP (2018) Automatic detection of retinal lesions for screening of diabetic retinopathy. *IEEE Trans Biomed Eng* 65(3):608–618
7. Abdelmaksoud E, El-Sappagh S, Barakat S, Abuhmed T, Elmogy M (2021) Automatic diabetic retinopathy grading system based on detecting multiple retinal lesions. *IEEE Access* 9(6):15939–15960
8. Xu K, Feng D, Mi H (2017) Deep convolutional neural network-based early automated detection of diabetic retinopathy using fundus image. *Molecules* 22(12)
9. Liu YP, Li Z, Xu C, Li J, Liang R (2019) Referable diabetic retinopathy identification from eye fundus images with weighted path for convolutional neural network. *Artif Intell Med* 99:101694
10. Raj MA, Al Mamun M, Faruk MF (2020) CNN based diabetic retinopathy status prediction using fundus images. In: 2020 IEEE Region 10 Symposium (TENSYP) 2020, pp 190–193
11. Wan S, Liang Y, Zhang Y (2018) Deep convolutional neural networks for diabetic retinopathy detection by image classification. *Comput Electr Eng* 72:274–282
12. Zago GT, Andreão RV, Dorizzi B, Salles EO (2020) Diabetic retinopathy detection using red lesion localization and convolutional neural networks. *Comput Biol Med* 116:2019
13. Gulshan V et al (2016) Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA J Am Med Assoc* 316(22):2402–2410
14. Pratt H, Coenen F, Broadbent DM, Harding SP, Zheng Y (2016) Convolutional neural networks for diabetic retinopathy. *Procedia Comput Sci* 90(July):200–205
15. Wang X, Lu Y, Wang Y, Chen WB (2018) Diabetic retinopathy stage classification using convolutional neural networks. In: Proceedings of 2018 IEEE 19th international conference on information reuse integration data science IRI 2018, pp 465–471
16. Raja C, Balaji L (2019) An automatic detection of blood vessel in retinal images using convolution neural network for diabetic retinopathy detection. *Pattern Recognit Image Anal* 29(3):533–545
17. Wu Z et al (2020) Coarse-to-fine classification for diabetic retinopathy grading using convolutional neural network. *Artif Intell Med* 108:101936
18. Alzubaidi L et al (2021) Review of deep learning: concepts, CNN architectures, challenges, applications, future directions, vol 8, no 1. Springer
19. Khan A, Sohail A, Zahoora U, Qureshi AS (2020) A survey of the recent architectures of deep convolutional neural networks. *Artif Intell Rev* 53(8):5455–5516
20. Alyoubi WL, Abulkhair MF, Shalash WM (2021) Diabetic retinopathy fundus image classification and lesions localization system using deep learning. *Sensors* 21(11):1–22
21. García G, Gallardo J, Mauricio A, López J, Del Carpio C (2017) Detection of diabetic retinopathy based on a convolutional neural network using retinal fundus images. *Lect Notes Comput Sci LNCS* 10614:635–642

# Ensured Configuration Security of FPGAs against CAD Attacks



Daliya A. John and M. Nirmala Devi

## 1 Introduction

A field programmable gate array (FPGA) is a system which is programmed after manufacturing to perform in specific applications. FPGAs are useful in high-performance computing, lower cost for prototyping, and shorter time to market. Since they have so many advantages, their application ranges from consumer products to military systems. There comes the need for FPGA security in applications. To create a design in FPGA, we can write HDL files, and HDL is synthesized to a bit file to configure the FPGA.

We aim to consider the attacks that can be injected from infected FPGA CAD tools, as these attacks are come from the infected software, that mount above the original design suite for SRAM FPGAs. Hardware Trojans can be inserted in the design flow at several stages without changing the actual design hardware description language (HDL) file.

- Proposed method is the implementation of an FPGA-oriented defense, by using the moving target defense principles that will create an unpredictable system for an attacker who is going to insert the hardware Trojans.
- FPGA-oriented defense consists of two defense lines. The first defense line is the method of changing the position of a part of design to a different area of FPGA. Second defense line is the random selection of one of the design replicas at runtime.

---

D. A. John (✉)

Department of Electronics and Communication Engineering, Amrita School of Engineering,  
Coimbatore, India  
e-mail: [daliyajohn96@gmail.com](mailto:daliyajohn96@gmail.com)

M. Nirmala Devi

Amrita Vishwa Vidyapeetham, Coimbatore, India  
e-mail: [m\\_nirmala@cb.amrita.edu](mailto:m_nirmala@cb.amrita.edu)

- We are analyzing the reduction in Trojan hit rates and the overheads in power, delay, and utilizations with FPGA.

## 2 Overview

There are several categories of papers are existing on protecting the FPGA. Among that one is the common FPGA mechanisms. In common FPGA security paper [1], the basic security features to all the threats and its applications are explained. The ways of achieving security features like information assurance and anti-tampering are well explained. A survey and an identification of relevant issues related to the security of FPGA are mentioned in [2]. Security of bitstream is explained in [3–5]. The IPs used in FPGA are coming from multiple vendors and are passed through several processes like design house, users, etc.

In [6], adding dummy logics to the unused resources will improve the security of the FPGA system and creating the IP for the FPGA. The detailed assessment of the power and delay overheads is included. In [7], the need for a secured hardware in IoT is discussed and mentioned that security of software alone is not enough in IoT technologies. Bogdanov et al. [8] proposed an authenticated encryption for protecting the bitstreams in FPGA, and it will prevent the attacks on bitstream in field. This also analyzed the protection against side channel attacks. In Mutarch paper [9], using physical and logical configuration keys making devices different in architecture. In [10], proposed a dynamic obfuscation method based on LUT to prevent the tampering in designs. A TRNG is designed in [11] which uses low power and generates random numbers using physical phenomenon. The paper [12] proposed a Trojan detection approach at the hardware level in decryption algorithm. Self-authentication [13] structure is implemented that will protect the bitstream from tampering attacks. Paper [14] implemented a security measure that focusses on the attacks from FPGA CAD tool. Data protection schemes in some of the wireless networks are also explained in some of the available papers [15].

In the existing papers, they are focusing only on some stages of the design flow. So, if we are not considering the other stages of design flow, the FPGA is still vulnerable to attacks. In our work, addressing the FPGA configuration flow.

## 3 Proposed Method

### 3.1 Threat Model

The non-trusted phase we are focusing is on the FPGA configuration flow which includes mapping, place and route, and bitstream generation stages as shown in Fig. 1. In this phase, we are considering two types of attacks. First:—The attacker does not have any idea about the design. According to his experience, placing the

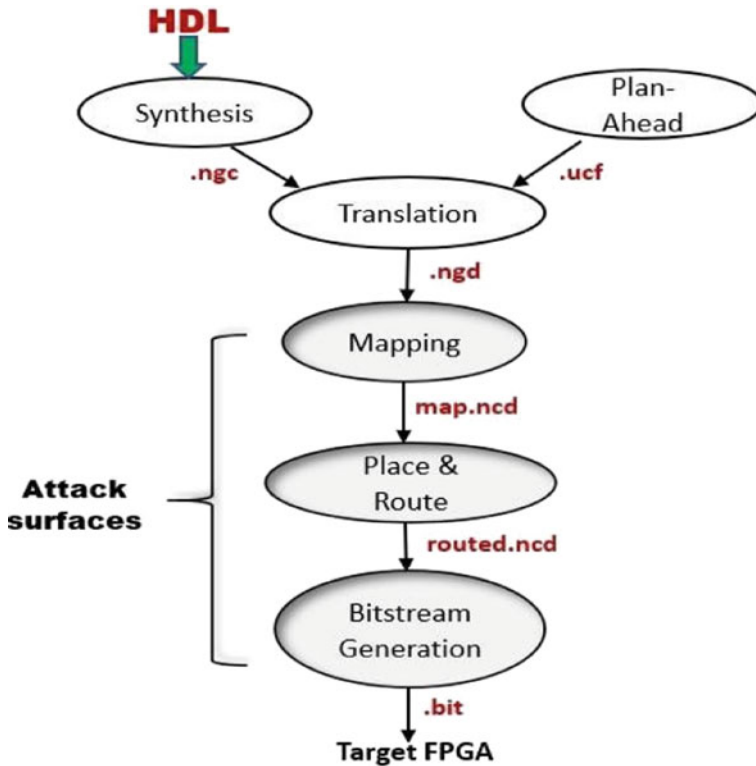


Fig. 1 Possible attack regions in the design flow of FPGA

Trojan in the most commonly using area in the FPGA die. Second:—The adversary can find the design position and function. We can inject the Trojans directly into the device for the analysis through the FPGA editor or by editing the .ncd file.

### 3.2 FPGA-oriented Defense

Moving target defense (MTD) is a method of making the adversary difficult to make an attack by varying the attack surface over time. By incorporating the MTD concept into the FPGA, we can develop an FPGA-oriented defense which create two types of uncertainties through two defense lines.

#### Defense Line 1: Selection of Slices

While using the default settings of FPGA for placement, that will make the design positions predictable for the attacker to insert the Trojans through the CAD tools. To prevent those situations, defense line 1 (DF1) is the changing of default settings in

the tool by editing the user constraints file (UCF). As shown in Fig. 2, mapping the design to a different location in the FPGA grid. We are assuming that the attacker does not have access to the UCF file.

### Defense Line 2: Selection of Replica

Defense line 2 (DFL2) is a method of replicating the design multiple times. Making one replica active and the remaining replicas inactive at that time as shown in Fig. 3. The selection of replica is done with the help of a pseudorandom selector. We can use any user-defined logic and some external inputs for pseudorandomly selecting the replica.

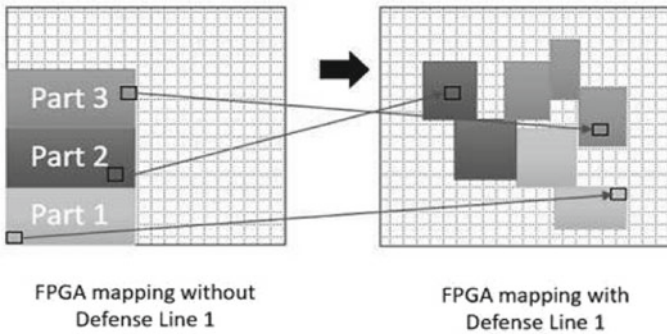


Fig. 2 Defense line 1 [14]

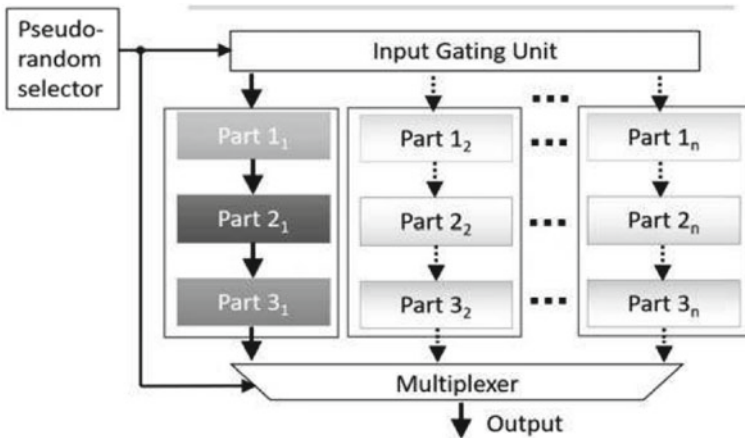


Fig. 3 Defense line 2 [14]



### Estimation of Rate of FPGA Trojan Hit

Reference design (baseline) is the original design before protection. Consider the baseline occupies  $\alpha$  slices and the entire FPGA contains  $\beta$  slices. Then, the Trojan hit rate  $T$  is the probability of a randomly selected slice is certainly one of the slices used in the design. For a blind attack, hardware Trojan hit rate,

$$T = \alpha / \beta \quad (1)$$

The Trojan hit rate analysis will differ in defense line 1 and defense line 2 based on the attack and the exploration space on the FPGA.

## 4 Experimental Analysis

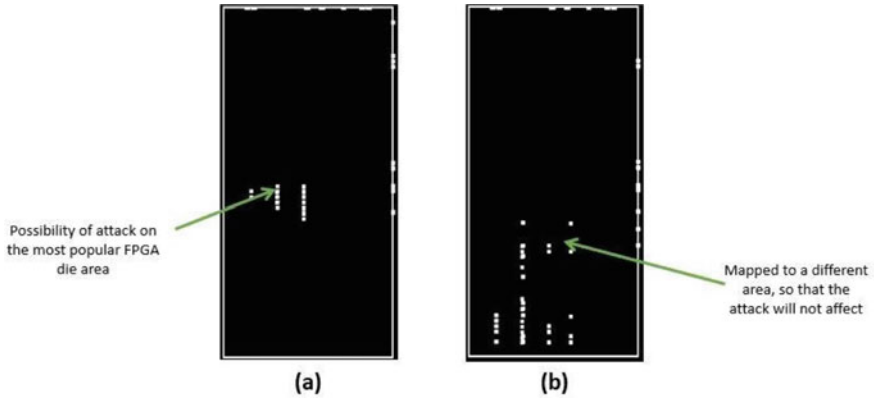
The design suite used in the project is Xilinx ISE 14.7 for synthesis, place, and route and generating the bitstreams. We used ISCAS 85 and ISCAS 89 benchmark circuits for analysis. These circuits are configured on a Spartan 6 FPGA board. The implemented design runs on an Intel 1.60 GHz quad core CPU with 8 GB RAM system. The attacks are created directly in the FPGA editor tool by altering the configuration of FPGA slice without disturbing the logic netlist. After implementing the defense lines, we can analyze the hardware Trojan hit rates and the utilization of slices and power and delay overheads.

### 4.1 Defense Line 1 Results

Here, we are mapping the design to a different location in FPGA by specifying slice positions with the help of UCF file. The output of DFL1 in ISCAS benchmark circuit c1355 is shown in Fig. 4. Figure 4a shows the design with default settings in the tool. Figure 4b shows the design after changing the default settings, that is, we specified the slice locations for LUT configurations. From Table 1, we can arrive at a conclusion that on an average of 50% of the LUT positions are placed to new positions on the FPGA die for each benchmark circuits.

### 4.2 Defense Line 2 Results

Here, we are duplicating our design and selecting one of the replicas during runtime. Unlike first type of attacks, second type of attacks know the exact slice positions occupied by the design. This DFL2 will activate one complete design replica with the help of pseudorandom selector which will improve unpredictability of the system.



**Fig. 4** Placement of design C432 by, **a** default settings from Xilinx ISE tool, **b** after the implementation of DFL1

**Table. 1** Change in LUT to new positions in DFL1 for some benchmark circuits

Circuit		LUTs	Slices	Rate of change in LUT positions
C432	Original	58	18	0.50
	Modified	58	42	
C1355	Original	62	30	0.51
	Modified	62	48	
C1908	Original	58	65	0.49
	Modified	58	80	
C6288	Original	530	185	0.50
	Modified	530	204	
S444	Original	33	13	0.48
	Modified	33	28	
S1488	Original	117	51	0.49
	Modified	117	75	
S13207	Original	180	81	0.43
	Modified	180	105	

This can further reduce hardware Trojan hit rates. Figure 5 shows the placement of design in FPGA editor after implementing DFL2.

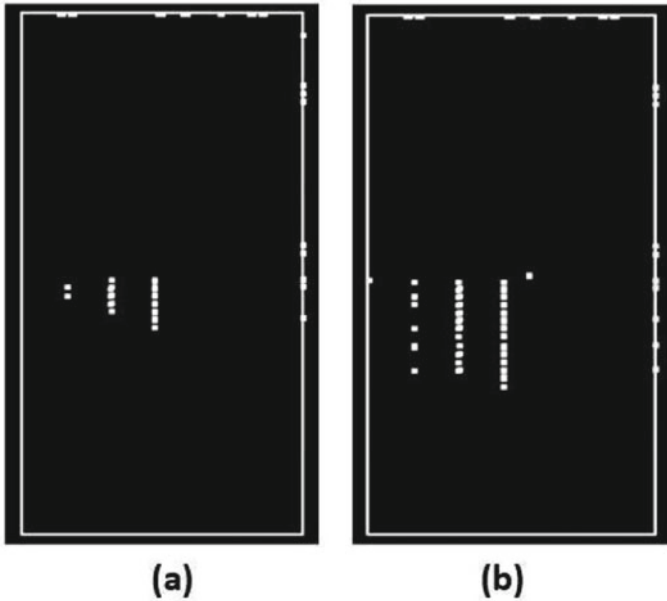


Fig.5 Routed NCD files of c432, **a** before and **b** after implementing DFL2

### 4.3 Trojan Hit Rate Assessment in DFL1 and DFL2

Here, we are implemented the FPGA-oriented defense mechanism in ISCAS 85 and 89 benchmark circuits. This complete assessment will give a detailed understanding of Trojan hit reduction by our method with the baseline as shown in Fig. 6. For

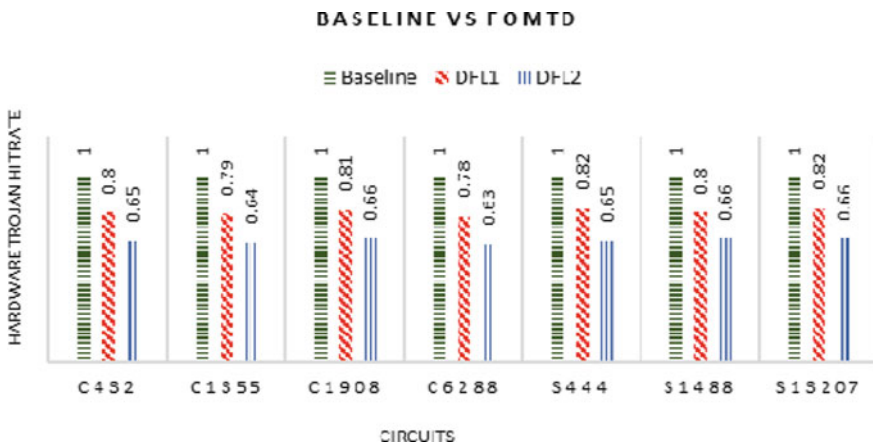


Fig.6 Trojan hit rate for baseline vs DFL1 vs DFL2

**Table. 2** Power consumption and worst-case delay by different methods

Circuits	Power(W)			Worst-case delay(ns)			Utilization (number of LUTs)		
	Baseline	DFL1	DFL2	Baseline	DFL1	DFL2	Baseline	DFL1	DFL2
C432	0.014	0.014	0.015	5.659	5.667	6.164	58	58	158
C1355	0.016	0.016	0.016	4.677	4.678	5.249	62	62	156
C1908	0.018	0.018	0.020	5.241	5.282	5.702	58	58	178
C6288	0.021	0.021	0.023	10.181	10.234	10.612	530	530	1123
S444	0.014	0.014	0.015	1.431	1.322	1.578	33	33	67
S1488	0.015	0.015	0.017	4.105	4.051	4.699	117	117	261
S13207	0.020	0.020	0.025	3.328	3.372	3.900	180	180	421

baseline design without any protection, we assume the Trojan hit rate as 1. A reduction of about 40% we can achieve through the FPGA-oriented defense.

#### ***4.4 Power, Delay, and Utilization Variations with Defense Method***

After synthesizing the Verilog codes of benchmark circuits in the Xilinx ISE tool, we will get power and delay details. There is a little variation in these metrics after implementing the FPGA-oriented defense. There is not much variation in power and delay for DFL1 implemented circuits. But these will show a variation in DFL2 implemented circuits as shown in the Table 2. DFL2 has a little increase in power, delay, and utilization of FPGA LUTs.

## **5 Conclusion**

Many papers exist that explained the threats like bitstream tampering, reverse engineering, etc., in the design and provided security for that separately. But there are limited papers are available that explains the attacks from CAD tools for FPGA configuration. We implemented the FPGA-oriented defense using the principles of moving target defense in Xilinx ISE 14.7 design suite. This method creates two types of unpredictability toward an attacker by the implementation of two defense lines. This is a generalized approach of preventing attacks, and this is suitable for SRAM-based FPGAs. The defense line 1 can change about 50% of the LUT to a new position. The two defense lines together can reduce the Trojan hit rate up to 40% when compared to the baseline design. Power, delay, and utilization of FPGA LUTs

are increasing in our method, and it will be a trade-off for achieving two types of unpredictability.

## References

1. Trimberger SM, Moore JJ (2014) FPGA security: Motivations, features, and applications. *Proc IEEE* 102(8):1248–1265
2. Majzoobi M, Koushanfar F, Potkonjak M (2012) FPGA-oriented security. In: *Introduction to hardware security and trust*, pp 1–38
3. Duncan A, Rahman F, Lukefahr A, Farahmandi F, Tehranipoor M (2019) FPGA Bitstream security: a day in the life. In: *Paper presented at the 2019 IEEE international test conference (ITC)*, Washington, DC, USA, pp 1–10
4. Horovitz K, Kainth M, Kenny R (2017) Protecting partial regions in FPGA bitstreams. In: *Paper presented at the 2017 IEEE 2nd international verification and security workshop (IVSW)*, Thessaloniki, pp 123–127
5. Sree Ranjani R, Nirmala Devi M (2018) Enhanced logical locking for a secured hardware IP against key-guessing attacks. In: *Paper presented at the VLSI design and test (VDATEC 2018)*, communications in computer and information science, Springer, Singapore, p 892
6. Hazari NA, Niamat M (2017) Enhancing FPGA security through Trojan resilient IP creation. In: *Paper presented at the 2017 IEEE National aerospace and electronics conference (NAECON)*, Dayton, OH, 2017, pp 362–365
7. Sidhu S, Mohd BJ, Hayajneh T (2019) Hardware security in IoT devices with emphasis on hardware Trojans. *J Sens Actuator Netw* 8(3):42
8. Bogdanov A, Moradi A, Yalcin T (2012) Efficient and side-channel resistant authenticated encryption of FPGA bitstreams. In: *Paper presented at the 2012 international conference on reconfigurable computing and FPGAs*, Cancun, pp 1–6
9. Karam R, Hoque T, Ray S, Tehranipoor M, Bhunia S (2017) MUTARCH: architectural diversity for FPGA device and IP security. In: *Paper presented at the 22nd Asia and South Pacific design automation conference (ASP-DAC)*, Chiba, pp 611–616
10. Baby J, Mohankumar N, Nirmala Devi M (2020) Reconfigurable LUT-based dynamic obfuscation for hardware security. In: *Advances in electrical and computer technologies, select proceedings of ICAECT 2019*, pp 963–973
11. Kavinkartik E, Suseendiran S, Mohankumar N (2020) Design and randomness evaluation of mixed-signal TRNG for low power applications. *Intell Comput Inf Control Syst* 105–113
12. Gayatri R, Gayatri Y, Mitra CP, Mekala S, Priyatharishini M (2020) System level hardware Trojan detection using side-channel power analysis and machine learning. In: *Paper presented at the 5th international conference on communication and electronics systems (ICCES)*, Coimbatore, pp 650–654
13. Zamanzadeh S, Jahanian A (2016) Self-authentication path insertion in FPGA-based design flow for tamper-resistant purpose. *ISC Int J of Inf Secur* 8(1):53–60
14. Zhang Z, Njilla L, Kamhoua CA, Yu Q (2019) Thwarting security threats from malicious FPGA tools with novel FPGA-oriented moving target defense. *IEEE Trans VLSI Syst* 27(3):665–678
15. Sudarsono A, Huda S, Fahmi N, Al-Rasyid MU, Kristalina P (2016) Secure data exchange in environmental health monitoring system through wireless sensor network. *Int J of Eng technol innov* 6(2):103–122

# Configuration Security of FPGA in IoT Using Logic Resource Protection



S. Meenakshi and M. Nirmala Devi

## 1 Introduction

Field-programmable gate array (FPGA) is widely used due to the flexible and scalable solutions it provides to interface with the different devices easily for Internet of Things (IoT) requirements. The main challenges in IoT are the security, unauthorized access, malicious control, and privacy, and hence, it is challenging to enhance the IoT security. The reconfigurable computing ability of FPGA has enabled more flexibility to modify some parts of the hardware; moreover, the increased logic in the FPGA has provided more complex designs and algorithms to be programmed on to the FPGA. To ensure the security of the hardware is not easy due to the increased cost and complexity in the designing, fabrication, and the deployment of the integrated circuit (IC), and hence, the outsourcing is necessary. The offshore foundries may introduce slight modifications in the design functionality that tends it to exhibit undesired behavior due to the malicious Hardware Trojans (HT).

The inserted Trojans can serve different purposes like reducing the reliability of the system or to function incorrectly. Re-fabrication at a different foundry would be time consuming and complicated, and this scenario is increasing due to the untrusted foundries. Hence, ensuring the security of the FPGA is an important concern at different phases of the manufacturing flow of the IC.

The Hardware Trojan is a malicious change made in a FPGA that is intended to change the original functionality or trustworthiness of a system. The Trojans target the design causing an undesirable behavior. A Trojan that is well designed may be not

---

S. Meenakshi (✉) · M. Nirmala Devi  
Department of Electronics and Communication Engineering, Amrita School of Engineering,  
Amrita Vishwa Vidyapeetham, Coimbatore, India  
e-mail: [meenakshiswapna2296@gmail.com](mailto:meenakshiswapna2296@gmail.com)

M. Nirmala Devi  
e-mail: [m\\_nirmala@cb.amrita.edu](mailto:m_nirmala@cb.amrita.edu)

visible directly as it can be inactive and inert within the design until some activation signal is received. This has been the reason for the need to prevent the Trojan insertion to make the FPGA reliable.

This work focusses on the prevention of the Trojans on the application space of the FPGA in order to reduce the chances of the abuse of the logic resources. In the proposed scheme, configuration security of FPGA is ensured by occupying the unutilized logic resources which is followed. Prevention of Trojan insertion is ensured by utilizing the unused spaces by shift registers to fill up the flip flops that are the sequential part and a chain of gates in order to fill up the lookup tables (LUT). The method is efficient in terms of protection of the FPGA as it does not leave any configurable logic resources. The scheme does not introduce any functional design alteration as the structure added to the functional design works independently.

## 2 Overview

There are different Hardware Trojan detection and prevention schemes for ensuring the security of FPGA. In [1], the influence on the security of FPGA due to different types of attacks is discussed. The attacks such as cloning, overbuilding, reverse engineering, and spoofing and the security measures to defend the threats are analyzed. The paper focusses on distinct types of FPGA and the FPGA design cycle highlighting on the unique aspects of FPGA security. It discusses two important security domains like information assurance and anti-tamper.

There are different schemes to protect the bitstreams like obfuscation methods used in [2–5] or by using adapted triple modular redundancy approach as in [6, 7]. In [2], in order to make less vulnerable to attacks like intellectual piracy and targeted malicious modification, an architecture that is liable to change for an FPGA is proposed. The architectural configuration is based on key, and the architecture will logically change over time. In [6], Trojan detection technique by a combination of side-channel analysis and logic testing called as adapted triple modular redundancy approach is attempted. The paper discusses the different Trojan attacks that is based on payload and activation characteristics. Compared to the conventional TMR approach, the proposed scheme introduces reduced power overhead without compromising on the security and performance. In order to reinforce the security of IoT, the Trojan infected design on the FPGA has to be identified. A random number generator using linear feedback shift register is used for creating randomness in cryptographic applications [7].

In [8], the ring oscillator-based free running Trojan is used to modify the configuration bitstream file. Some of the prevention/detection mechanisms are recommended like the unutilized pins grounding, filling up the unutilized spaces, and to check the CRC status, a dedicated hardware architecture can be used. In [9], the unutilized area in the ASIC is filled up with an authentication structure, and functional cells will spot out the malicious modification, if any. The proposed method does not cause the

functional design parameters to change, but there is no mechanism for the detection of Trojan.

In [10], the malicious hardware functionality is recognized using ring oscillators which fills a part of FPGA that detects HTHs when enabled near the ring oscillator. The proposed method does not provide a prevention method, instead it is only used for the detection of Trojan. In [11], the work utilizes all the usable empty spaces by filling using shift registers and gates in ASIC. The signals having more slack are used to identify the empty spaces which are critical. The original design parameters like power, area, and performance are not affected, and there is strong detection whether any malicious modification occurs.

In [12], the unutilized space is filled up using dummy logic for distinct resources by adding to the Xilinx Design Language (XDL) file, and the original design is not modified. This scheme does not detect whether an attack has occurred or not. In [13], the approach is based on FPGA bitstream protection that provides more security compared to the encryption with much lower overhead. The unutilized resources also called as FPGA dark silicon, within the configurable logic blocks, are exploited in this technique to enhance the security. Additional complexity in design verification is not introduced, and performance is not degraded.

### 3 Proposed Methodology

The following steps involved in proposed protection scheme are depicted in Fig. 1.

Step 1: Synthesize and implementation of the functional design.

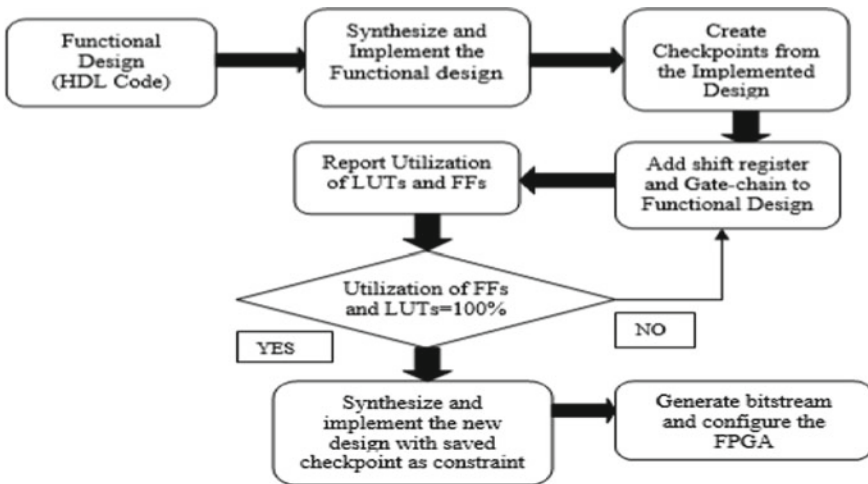
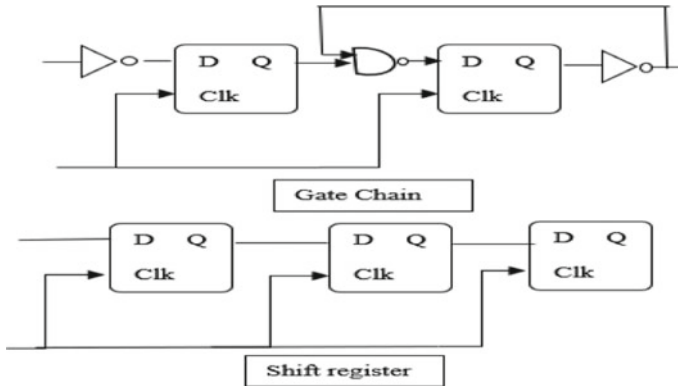


Fig. 1 Flow diagrams of the proposed protection scheme





**Fig. 2** Added chain structure [14]

Using the write-checkpoint instruction in the Tool Command Language (TCL) section of Vivado, the place and route of the current functional design are saved.

Step 2: Create checkpoints from the implemented design.

The critical path of the functional design and the specifications are saved by creating the checkpoints using the TCL instruction. A checkpoint file contains the current netlist, including the optimizations that is made during the implementation and the design constraints at any specific point of the design flow. After the functional design is implemented, the utilization report that shows the number of logic resources utilized is obtained and analyzed.

Step 3: Filling up the unused logic resources.

Once the implementation is complete and utilization rate is obtained, the remaining logic resources like the LUTs and FFs are filled up using shift registers and a chain of gates as shown in Fig. 2. The added structure is integrated to the functional design by adding hardware description language (HDL) code to the main design.

The flip flops of the device are filled up using shift registers, and the lookup tables are filled by using a structure of connected gates. The number of shift registers and chain of gates depends on the number of unused lookup tables and the flip flops. The filling up of LUTs is sufficient as the programmable part is implemented on the LUTs. To secure the FPGA against reverse engineering and to provide more security, we fill up the flip flops also using a shift register. The flip flops are used to implement the shift register logic, and a structure similar to the ring oscillator consisting of LUTs and FFs is used to build the added structure. The chain of gates can have different types of gates or only NOT gates such that it produces an output with feedbacks from the chain.

Step 4: Synthesize and implement the protected design.

The utilization of all the logic resources is followed, synthesizing and implementing the protected design saving the checkpoint as the constraint.

Step 5: Generate bitstream and configure the FPGA.

The bitstream file is generated from the final implemented design, and the FPGA is programmed with the new design to verify the protection scheme. The prevention method mentioned above does not introduce any changes in the functional design as the main design and added structure work independently. This is ensured by two modes in the proposed work. In the normal mode, the original design works, and in the test mode, the added chain structure works.

## 4 Experimental Analysis

### 4.1 Experimental Setup

The proposed scheme is implemented using Vivado 2017.4 design suite. The IWLS benchmarks [15] and ISCAS 89 benchmarks are used to implement the proposed scheme and fill up the unutilized resources such as the LUT and flip flops in order to validate that the security is enhanced. The proposed scheme is implemented on a Basys3 (xc7a35tcp236-1) FPGA board, and the results are validated. The experiments were run on a PC as shown in Fig. 3 with Intel Core i5 3.10 GHz Processor and 8 GB RAM operating on Windows 10.

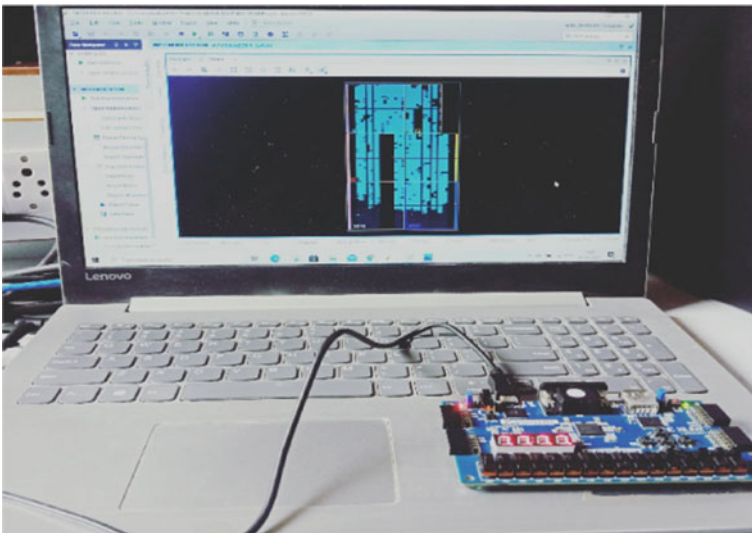
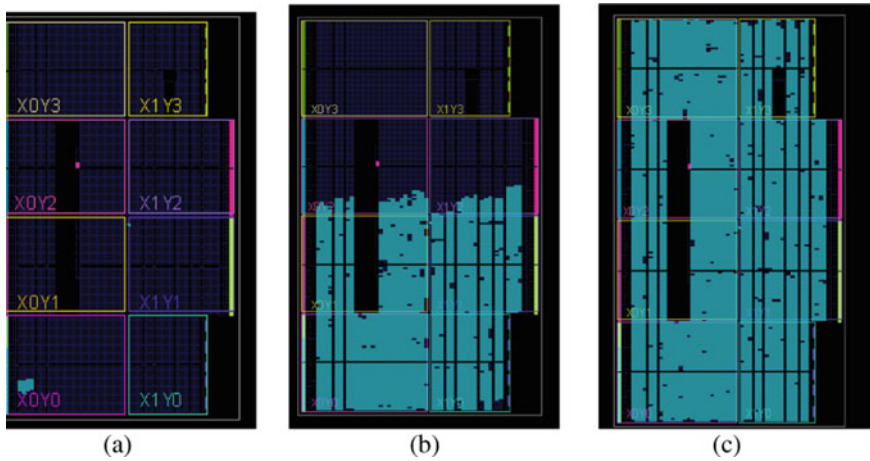


Fig. 3 Experimental setups for logic resource protection



**Fig. 4** **a** Implemented design of I2C before protection, **b** implemented design of I2C partial filling, **c** implemented design of I2C after protection

## 4.2 Results and Analysis

Figure 4a shows the implemented design of inter-integrated circuit (I2C) design before protection, and Fig. 4b, c shows the rate of resource utilization by partially filling and completely filling the unutilized resources for protection. The experimental results proved the FPGA protection before and after filling the unused resources by comparing the resource utilization report. The checkpoint is saved as a constraint, and hence, adding the shift register and the chain with the gates does not alter the power consumed and the critical path.

The utilization rate of the LUTs and FFs after the protection is obtained as 100% as given in Table 1. The power consumption of the original design is not affected due to the two modes of operation.

Different time constraints are used so that the functional design and the extra added structure do not work simultaneously. The power in the normal mode is same as the power of the functional design as given in Table 2. It shows that our scheme does not impose any performance overhead to the functional design. The critical path delay before and after protection for the design remains the same.

The FPGA static power is a constant irrespective of the fact that whether the resource is utilized or not. The dynamic power in the test mode is about 20 times the power in the normal mode due to the added structure, and the static power remains the same in both the modes.

**Table 1** Comparison of utilization of resources of proposed scheme with unprotected design

Benchmark	Unprotected design			Protected design	
	I/O utilization (in %)	LUT utilization (in %)	Used flip flop (in %)	LUT utilization (in %)	Used flip flop (in %)
I2C	11.00	0.22	0.09	100	100
SPI	30.00	0.75	0.18	100	100
TV80	15.33	1.89	0.28	100	100
Mem-ctrl	89.00	1.76	0.78	100	100
s298	3.67	0.02	0.01	100	100
s13207	44.00	0.24	0.19	100	100
s15850	33.33	0.15	0.10	100	100
s38417	45.33	1.97	1.22	100	100

**Table 2** Comparison of power and performance of protected and unprotected design

Benchmark	Unprotected design		Protected design		
	Power (in W)	Critical path (in ns)	Power in normal mode (in W)	Power in test mode (in W)	Critical path (in ns)
I2C	0.081	8.415	0.081	0.457	8.415
SPI	0.167	9.867	0.167	0.713	9.867
TV80	0.108	15.76	0.108	0.672	15.76
Mem_ctrl	0.134	12.75	0.134	0.863	12.75
s298	0.081	4.228	0.081	0.317	4.228
s13207	0.117	3.743	0.117	0.709	3.743
s15850	0.137	6.812	0.137	0.915	6.812
s38417	0.124	7.140	0.124	0.727	7.140

## 5 Conclusion

The field-programmable gate arrays (FPGA) used in most of the applications is at risk due to the Hardware Trojan Horse (HTH) insertion. The reconfiguration property, however, makes it vulnerable to design alteration as compared with application specific integrated circuits (ASIC) even after implementation of the design. The FPGA used to meet the IoT requirements is to be protected against the Hardware Trojan attacks in the current scenario. Until the hardware for the IoT device, which is the FPGA not protected, it is not possible to reinforce the IoT security. The FPGA is vulnerable to various threats like cloning, reverse engineering, overbuilding, etc. Hence, the security of the FPGA that depends on the hardware and software and the data that is handled is of importance. The proposed scheme aims at providing configuration security of FPGA by occupying the unutilized logic resources. The

prevention of Trojan insertion is ensured by utilizing the unused spaces by shift registers to fill up the flip flops that are the sequential part and a chain of gates in order to fill up the lookup tables (LUT). The prevention method utilizes all the logic resources that are at risk of application space attacks. The method is efficient in terms of protection of the FPGA as it does not leave any configurable logic resources. The scheme does not introduce any functional design alteration as the structure added to the functional design works independently. Hence, the proposed scheme does not increase the power overhead of the functional design and ensures 100% utilization of the logic resources.

## References

1. Trimberger SM, Moore JJ (2015) FPGA security: motivations, features, and applications. *Proc IEEE* 102(8):1248–1265
2. Karam R, Hoque T, Ray S, Tehranipoor M, Bhunia S (2017) MUTARCH: architectural diversity for FPGA device and IP security. In: Paper presented at design automation conference (ASP-DAC), 22nd Asia and South Pacific, pp 611–616
3. Baby J, Mohankumar N, Nirmala Devi M (2019) Reconfigurable LUT-based dynamic obfuscation for hardware security. In: *Advances in electrical, computing, communications and sustainable technologies (IEEE ICAECT)*, 2019. Lecture notes in electrical engineering, vol 672. Springer, Singapore
4. Kumar AVP, Bharathi S, Meghana C, Anusha K, Priyatharishini M (2019) Toggle count based logic obfuscation. In: Paper presented at 3rd international conference on electronics, communication and aerospace technology (ICECA), Coimbatore, India, pp 809–814
5. Sree Ranjani R, Nirmala Devi M (2018) Enhanced logical locking for a secured hardware IP against key-guessing attacks. In: Paper presented at 22nd international symposium on VLSI Design and Test, VDAT
6. Mal-Sarkar S, Karam R, Narasimhan S, Ghosh A, Krishna A, Bhunia S (2016) Design and validation for FPGA trust under hardware Trojan attacks. *IEEE Trans Multi-Scale Comput Syst* 2(3):186–198
7. Durga RS, Rashmika CK, Madhumitha ONV, Suvetha DG, Tanmai B, Mohankumar N (2020) Design and synthesis of LFSR based random number generator. In: Paper presented at 2020 3rd international conference on smart systems and inventive technology (ICSSIT), Tirunelveli, India, pp 438–442
8. Chakraborty RS, Saha I, Palchaudhuri A, Naik GK (2013) Hardware Trojan insertion by direct modification of FPGA configuration Bitstream. *IEEE Design Test* 30(2):45–54
9. Xiao K, Tehranipoor M (2013) BISA: built-in self-authentication for preventing hardware Trojan insertion. In: Paper presented at hardware-oriented security and trust (HOST), 2013 IEEE international symposium, pp 45–50
10. Kitsos P, Voyiatzis AG (2014) FPGA Trojan detection using length-optimized ring oscillators. In: Paper presented at digital system design (DSD), 2014 17th Euromicro conference, 2014, pp 675–678
11. Ba PS, Palanichamy M, Dupuis S, Flottes ML, Di Natale G, Rouzeyre B (2015) Hardware Trojan prevention using layout-level design approach. In: Paper presented at circuit theory and design (ECCTD), 2015 European conference, pp 1–4
12. Khaleghi B, Ahari A, Asadi H, BayatSarmadi S (2015) FPGA-based protection scheme against hardware Trojan Horse insertion using dummy logic. *IEEE Embed Syst Lett* 7(2):46–50
13. Duncan A et al (2019) FLATS: filling logic and testing spatially for FPGA authentication and tamper detection. In: Paper presented at 2019 IEEE international symposium on hardware oriented security and trust (HOST), McLean, VA, USA, pp 81–90

14. Labafniya M, Saeidi R (2019) Secure FPGA design by filling unused spaces. *ISC Int J Inf Secur* 11(1):47–55
15. Albrecht C (2005) IWLS 2005 benchmarks. In: International workshop for logic synthesis (IWLS): <http://www.iwls.org>

# A Novel Construction of Multiband Compact Microstrip Patch Antenna and Its Applications in WiMAX, HiPERLAN



Pritam Singha Roy and Moumita Guha

## 1 Introduction

In many wireless broad band communication systems, application of the multiband frequencies are very commonly used. Hence, the multiband antennas have more requirement in telecommunication system. The most common and suitable antenna for wireless communication and broad band system is microstrip patch antenna and very flexible sophisticated devices due to its low profile structure, it is a most suitable structure for wireless broad band communication. The microstrip antenna is the major choice to operate in wireless communication because the size of communication device depends on this antenna. Therefore, one of the best methods to minimize the dimension of the communication system is by reducing the antenna. The most common way of them is by proper adjustment of ground plane and patch dimension [1–10]. The suitable method of patch adjustment is cutting the slots in proper sizes. [1, 2] obtained the minimized size antenna by 50% and obtained 400 MHz resonance frequency range. [3] reported that designed structure has been minimized by 84%. Multi-range frequencies are used in parallel polarization planes and by reduction of antenna size. The performance of the antenna depends on the thickness of the textile substrate material [4]. In [5, 6], the miniaturized antenna was fabricated with resonant frequency of 3.5 GHz for WiMAX application. [7] reported rectangular slits at two sides of the patch. Compared with the conventional rectangular patch antenna, this antenna can improve reduction in patch size up to 71%. The proposed antenna designed by us also represents the compact multiband frequency of antenna is obtained by cutting the double rectangular slots with different dimensions on radiating surface edges of the microstrip patch material. We designed a compact microstrip patch antenna, and it operates in multiband frequency range by properly

---

P. S. Roy (✉) · M. Guha

Government College of Engineering and Textile Technology, Berhampore, India

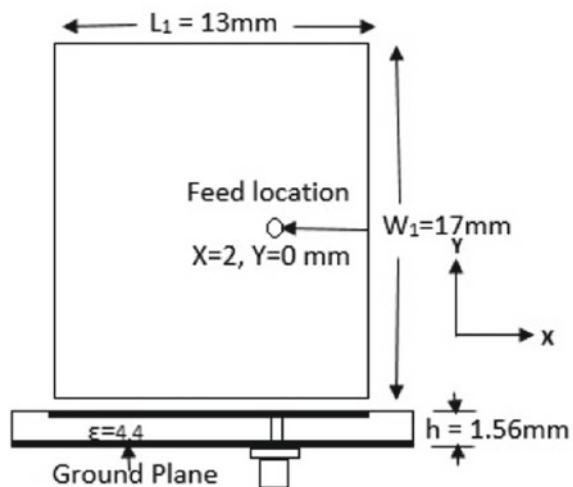
e-mail: [pritam.singha.roy@gcetbt.ac.in](mailto:pritam.singha.roy@gcetbt.ac.in)

adjusting the radiating slots and ground plan. Due to the presence of slots in the radiating patch, the multiple band of resonance frequencies found and dimension of the microstrip patch antenna has been minimized. The proposed designed compact antenna is better than rectangular plane or ground plane structure [1, 2]. The unequal rectangular-shaped compact multi-frequency microstrip patch antenna proposed by us should be applied in many telecommunication systems because of its compact profile structure and overall excellent working features.

## 2 Basic Structure Geometry

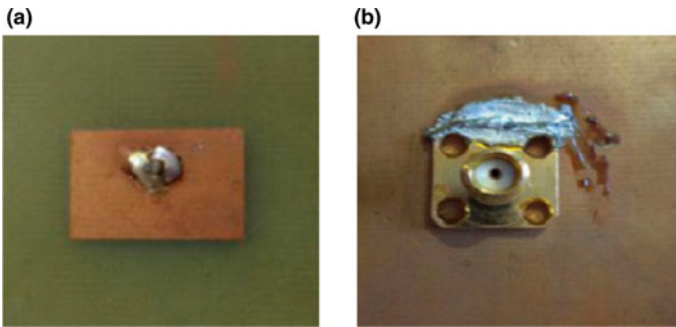
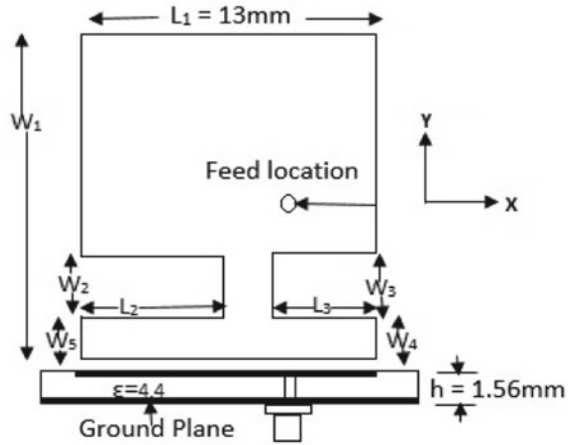
Basic parameters of conventional patch antenna are calculated by the conventional mathematical equations [8–10]. The dimension of the antenna is found by using generalized equations and found as  $L = 16$  mm and  $W = 12$  mm at the center operating resonance frequency at 5.6 GHz. Here, the conductive height is 1.567 mm, and selected dielectric material FR4 is 4.4. The conventional unslotted patch antenna geometry structure shown in Fig. 1 having  $L_1 = 13$  mm and  $W_1 = 17$  mm using Coaxial probe feed with probe radius is 0.5 mm has been positioned at a location ( $X = 2$  mm,  $Y = 0$  mm) w.r.t the origin of the unslotted conventional antenna. The structure of proposed slotted rectangular microstrip antenna is shown in Fig. 2. Here, one pair of unequal rectangular slots are placed at the both opposite sides of the patch antenna. Thus, minimized the size of the antenna and increases the band of frequencies from the designed antenna. Similarly, coaxial probe feeding used at the location  $X = 1.6$  mm in conventional unslotted rectangular antenna. The optimized data of the proposed antenna dimension are:  $W_1 = 17$  mm,  $L_1 = 13$  mm,  $W_2 = 2.5$  mm,  $W_3 = 2.5$  mm,  $L_3 = 6.5$  mm,  $L_2 = 5.5$  mm,  $W_4 = 1.5$  mm,  $W_5 = 1.5$  mm,

**Fig. 1** Structure of unslotted conventional antenna





**Fig. 2** Proposed slotted patch antenna



**Fig. 3** Fabricated unslotted rectangular conventional antenna, **a** top view and, **b** bottom view

and  $X = 1.6$  mm. The fabricated designed conventional antenna and proposed slotted antenna are shown in Figs. 3 and Fig. 4, respectively.

### 3 Parameter Analysis of Proposed Compact Antenna

The variation of antenna parameters with the variation of different values of dimensions of the slotted microstrip antenna have been measured by MoM-based simulator [11, 12]. For a single rectangular-slotted antenna given in Fig. 5 exhibits two different band of resonance frequencies are 6.2 GHz and 7.51 GHz with optimum value of slot dimension as  $W_2 = 3.5$  mm and  $L_1 = 5$  mm having below  $-10$  dB level of reflection coefficient. For proposed bi-rectangular slotted microstrip antenna with dimension  $W_3 = 3$  mm and  $L_3 = 6$  mm shown in Fig. 6. Because of the dual slot, the resonant frequencies are shifted and obtained at 5.42, 6.72, and 8.36 GHz below of  $-10$  dB

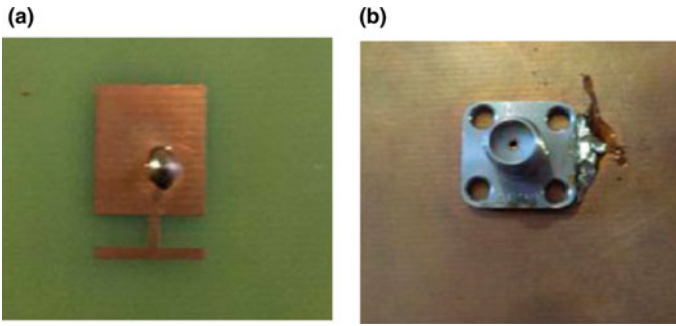


Fig. 4 a Top and, b bottom fabricated structure of designed antenna

Fig. 5 Reflection coefficient of single-slotted antenna

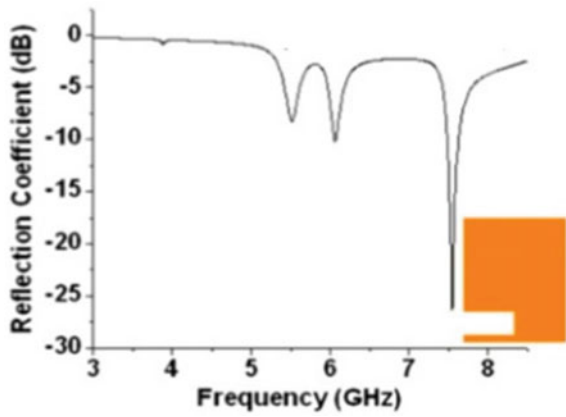
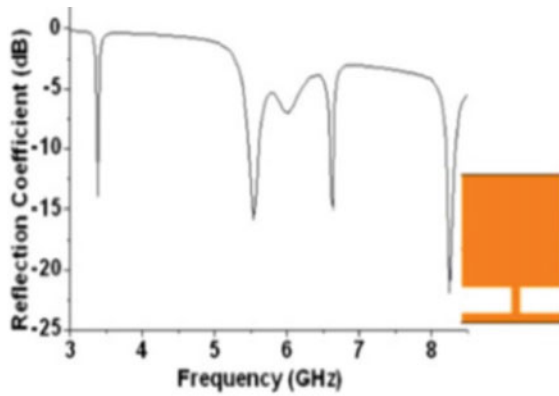
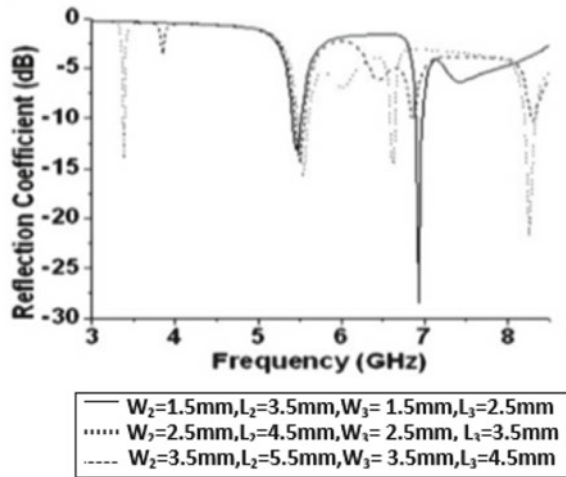


Fig. 6 View of reflection coefficient and frequency characteristics



**Fig. 7** Variation of ref. coefficient with dimension



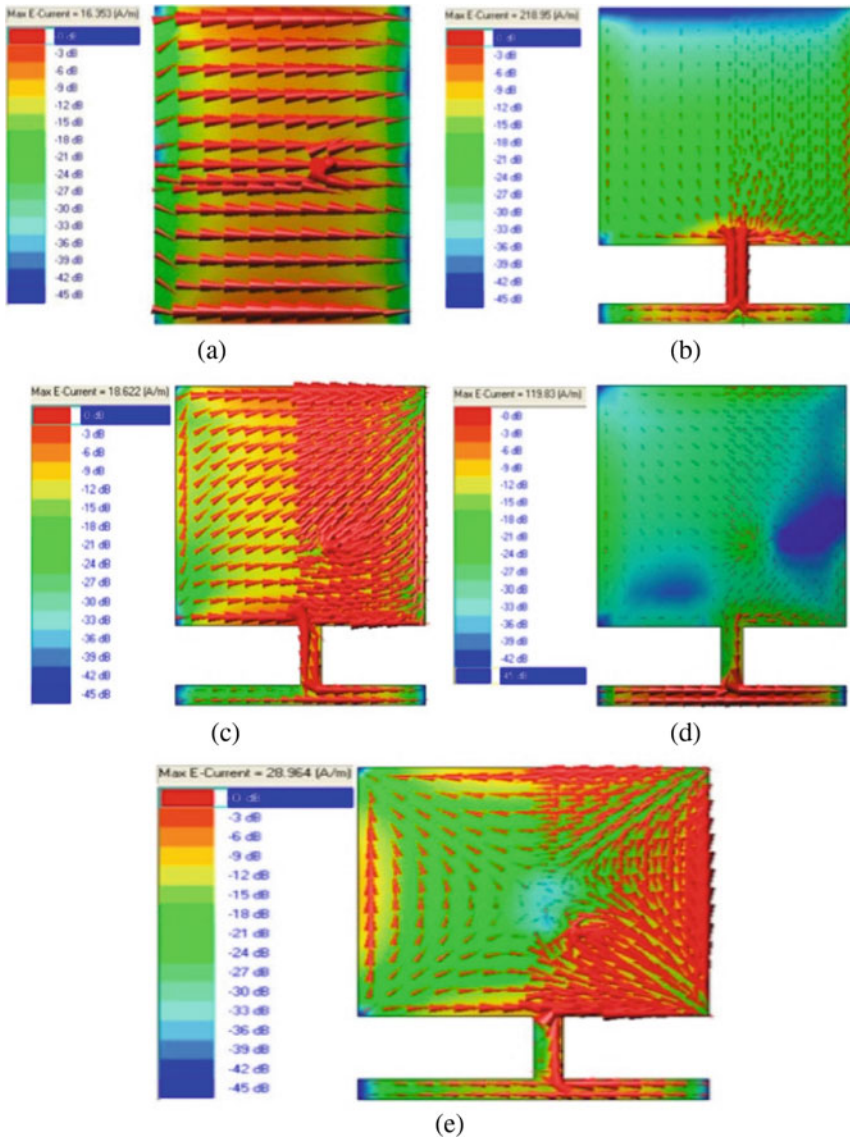
**Table 1** Variation of simulation data with different dimension

Slot dimensions (mm)	Resonance frequency (GHz)	Reflection coefficient (dB)
$W_2 = 1.5, L_2 = 3.5, W_3 = 1.5, L_3 = 2.5$	5.55, 6.84	-14, -27.6
$W_2 = 2.5, L_2 = 4.5, W_3 = 2.5, L_3 = 3.5$	5.40, 6.74, 8.31	-15, -12, -11
$W_2 = 3.5, L_2 = 5.5, W_3 = 3.5, L_3 = 4.5$	3.42, 5.42, 6.72, 8.36	-13.86, -14.85, -14.30, -22.80

(ref. coefficient). The variation characteristics of ref. coefficients for different slot dimensions for a fixed interval of deviation is shown in Fig. 7 referred to the Table 1. From the simulated results, it is found that for any variations of dimension of slotted antenna, the resonance frequency may be varied. Thus, the antenna will not fulfill the properties of compactness.

### 4 Conventional Versus Design Proposed Antenna

Here, we discussed the variation of current distribution of conventional and proposed slotted antenna. With the use of surface current distribution, the working principle of proposed and conventional antenna is analyzed. For different resonant frequencies, current distribution of the unslotted conventional antenna and proposed antenna is shown in Fig. 8a–e. Here, current density has been increased with increases the cutting slots Fig. 8b–e. Figure 8a–e shows that the current density characteristics are different for both conventional and proposed antenna because of their double cutting slots. Fig. 8b–e shows the maximum surface current density than conventional antenna in



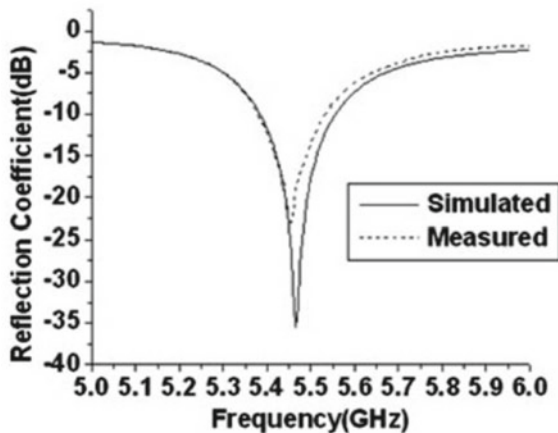
**Fig. 8** Current distribution of conventional antenna at, **a** 5.55 GHz and proposed slotted antenna at, **b** 3.42 GHz, **c** 5.42 GHz, **d** 6.72 GHz, **e** 8.36 GHz

Fig. 8a. The number of resonance frequencies can be increased by introducing the open-ended slot at the patch surface edges without any disturbance of the surface current and smoothly distributed the electric and magnetic fields.

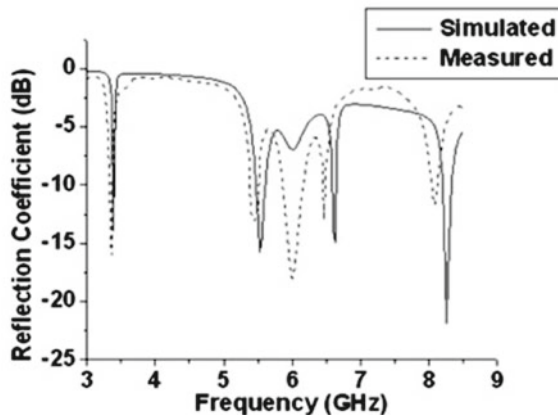
### 5 Antenna Results and Discussion

At first, the simulated results of reflection coefficients of slotted compact antenna and conventional antenna were measured and studied using IE3D simulation software. Most common analyzer Agilent E5071B is used to measure the value of reflection coefficient of proposed compact antenna. The simulated and measured results of conventional antenna and proposed slotted antenna are shown in Figs. 9 and 10, respectively. In both cases, the obtained frequency is 5.55 GHz with reflection coefficient of about  $-36$  and  $-24$  dB, respectively. Due to the introduction of the cutting slots in proposed antenna, the multi-resonance frequencies found at 3.42 GHz, 5.42 GHz, 6.72 GHz, and 8.36 GHz with reflection coefficients are about  $-19.85$ ,  $-14.86$ ,  $-14.30$ , and  $-22.80$  dB, respectively. Here, the simulated result is verified according to the change of the measured value and resonates at

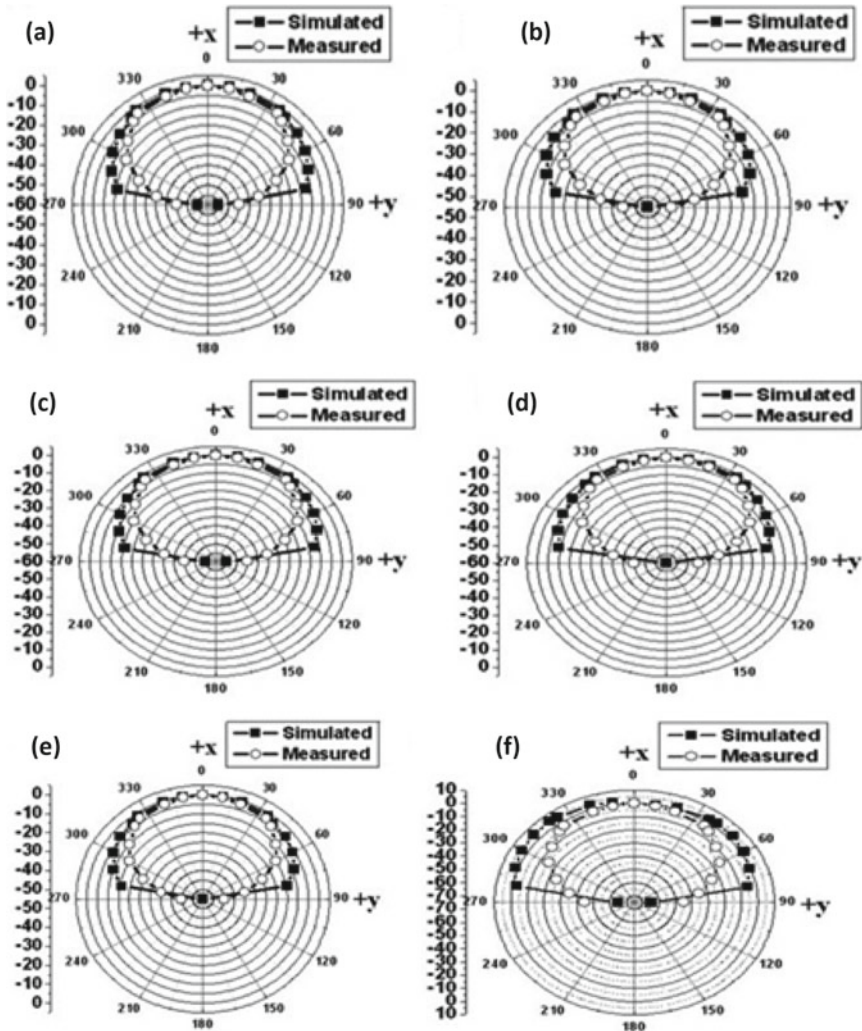
**Fig. 9** Ref. coefficient versus frequency of conventional antenna



**Fig. 10** Ref. coefficient versus frequency of design antenna

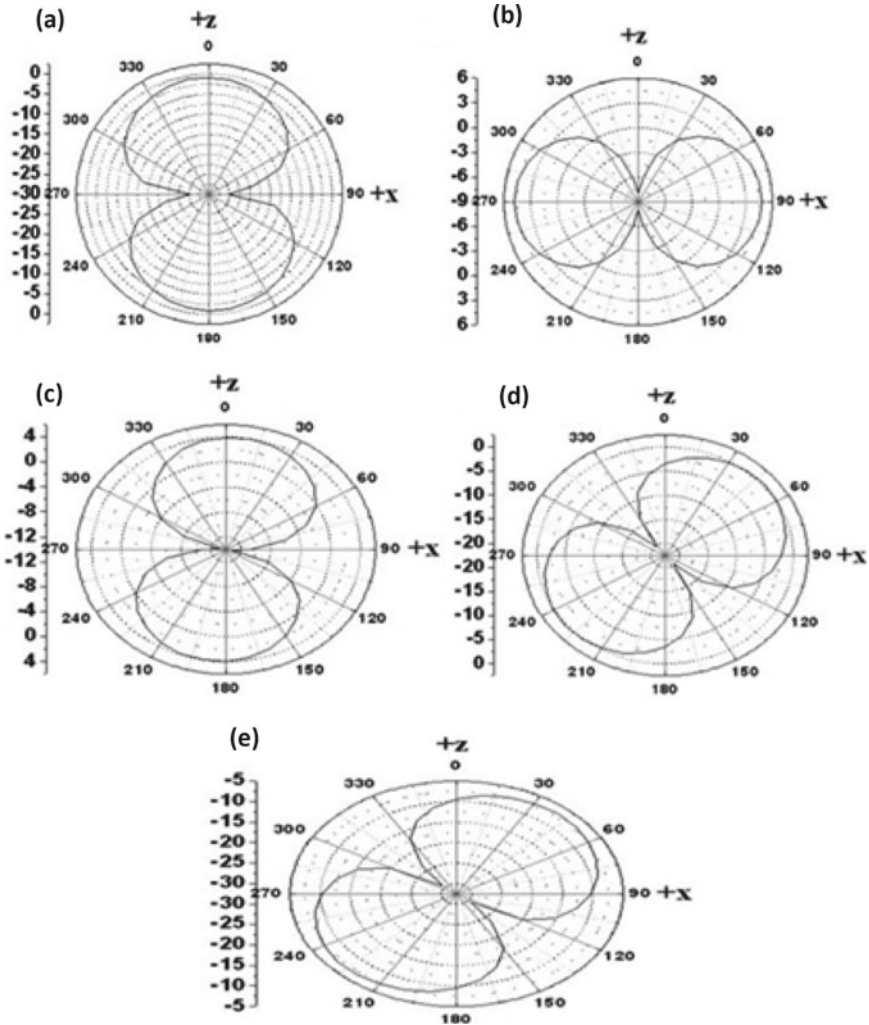


3.42 GHz, 5.42 GHz, 6.1 GHz, 6.72 GHz, and 8.36 GHz having reflection coefficient of about  $-17$ ,  $-14$ ,  $-18$ ,  $-14$ , and  $-12$  dB, respectively. There is a little variation of response due to the improper soldering of SMA connector to the patch. Figures 11a–f show that the E-field radiation pattern for both conventional and proposed designed antenna. Here, conventional antenna shows the radiation pattern at 5.55 GHz, whereas proposed antenna shows the different resonant frequencies at (3.42, 5.42, 6.1, 6.72, and 8.36 GHz). Electric field radiation pattern shows in Figs. 11a–f, unidirectional pattern for all different frequencies. The maximum radiation is obtained along  $0^\circ$ .



**Fig. 11** E plane radiation pattern of conventional antenna at, a 5.55 GHz and proposed antenna at, b 3.42 GHz, c 5.42 GHz, d 6.1 GHz, e 6.72 GHz, f 8.36 GHz at  $90^\circ$

Also, found a stable response throughout different operating frequencies. Figs. 12a–e shows that the H-field radiation for both conventional and proposed antenna in azimuth direction  $90^\circ$ . The both side radiation pattern occurs for proposed compact slotted antenna due the different dimension of slots. It can be observed that the maximum and minimum radiations found along  $0^\circ$  and  $180^\circ$  ( $3.42\text{ GHz}$ ), respectively. In proposed antenna, the radiation likes figure “eight” at  $5.55\text{ GHz}$  toward  $90^\circ$  and  $270^\circ$ . For reduction of antenna size (due to introducing slots) by shifting its resonant frequency, it can be applied as DGS.

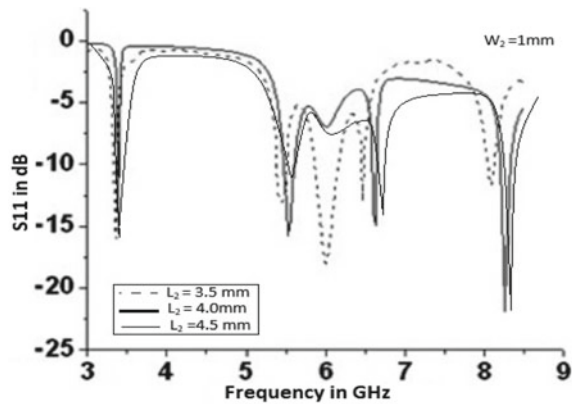


**Fig. 12** H plane radiation pattern of designed antenna at, **a**  $3.42\text{ GHz}$ , **b**  $5.42\text{ GHz}$ , **c**  $6.1\text{ GHz}$ , **d**  $6.72\text{ GHz}$ , **e**  $8.36\text{ GHz}$   $90^\circ$

**Table 2** Measured results of the conventional and proposed slotted compact antenna

Type of antenna	Resonance Freq (GHz)	Reflection coefficient (dB)	10 dB BW (MHz)	Frequency ratio	Gain (dBi)	3 dB beam width (°)
Conventional antenna	5.55	-24	152		5.3	67
Proposed slotted compact antenna	$f_1 = 3.42$	-17	62	$f_2/f_1 = 1.58$	4.2	66
	$f_2 = 5.42$	-14	122	$f_3/f_1 = 1.78$	5.3	67
	$f_3 = 6.22$	-17	264	$f_4/f_1 = 1.86$	5.2	73
	$f_4 = 6.62$	-14	52	$f_5/f_1 = 2.44$	4.7	71
	$f_5 = 8.36$	-12	92		3.0	69

**Fig. 13** Variation of reflection coefficient with  $L_2$  (Proposed antenna)



Another things that bidirectional radiation patterns of proposed slotted antenna are obtained at 6.1 GHz, 6.72, and 8.36 Hz but not concentrated perfectly along  $0^\circ$  and  $180^\circ$ . Whereas, the radiation patterns at 6.72 and 8.36 GHz are exactly identical. The measured results of both conventional and proposed slotted antenna are given in Table 2. Figure 13 shows the equivalent circuit of slotted antenna. It consists of four series RLC circuit connected in parallel with different values of  $R$ ,  $L$ , and  $C$ .

Due to the introduction of cutting slots, the resonance frequency has been shifted toward lower side of the antenna. The variation of length for fixed with the resonance frequency has been shifted shows in Fig. 13. This method of antenna design is very useful for modern communications system.

## 6 Conclusion

Our proposed compact slotted antenna with multiple slots at the sides edges on the patch antenna; the size has been minimized, and more than 68% has been introduced.



This proposed compact antenna exhibits more suitable radiation with proper gain and bandwidth. In radiation pattern of the proposed compact antenna, the 3dB beam width varies in within  $67^\circ$  to  $73^\circ$  which is very practical achievement of patch antenna. Experimental results show that the antenna resonates at 3.42, 5.42, 6.1, 6.72, and 8.36 GHz covering WiMAX (802.16d) bands. The frequencies commonly used are 3.5 and 5.8 GHz for 802.16d and 2.3, 2.5, and 3.5 GHz for 802.16e which depends upon the use of different countries. HiPERLAN is a European family of standards High Performance European Radio LAN communication in the 5.15–5.3 GHz range. Finally, the proposed slotted compact patch antenna should be promising for various advanced communication system standards like WiMAX, WLAN, and HiPERLAN frequency band.

## References

1. Malek NA, Mokhtar NH, Ali K, Isa FN, Rahman FD (2018) Design of small antennas for 400 MHz applications. pp 99–103. <https://doi.org/10.1109/ICCCE.2018.8539246>
2. Roy PS, Chakraborty S (2016) Design of h-slotted microstrip patch antenna with enhanced bandwidth for C-band application. *Int J Innov Technol Explor Eng* 6(3):43–46. <http://www.ijitee.org/wp-content/uploads/papers/v6i3/C2348086316.pdf>
3. Deng L, Li SF, Lau KL, Xue Q (2012) Vertical meandering approach for antenna size reduction. *Int J Antenn Propag* 2012. Article ID 980252. <https://doi.org/10.1155/2012/980252>
4. Roy PS, Guha M, Roy CS (2020) Wide band rectangular wearable microstrip ring antenna with textile substrate and its performance after washing. *J Phys Conf Ser* 1706: 012072. <https://doi.org/10.1088/1742-6596/1706/1/012072>
5. Ajay VG, Mathew T (2017) Size reduction of microstrip patch antenna through metamaterial approach for WiMAX application. In: 2017 international conference on wireless communications, signal processing and networking (WiSPNET), March 2017, pp 22–24. <https://doi.org/10.1109/WiSPNET.2017.8299782>
6. Agarwal VK, Shaw AK, Das MK, Mukherjee J, Mandal K (2012) A novel compact dual frequency microstrip antenna. In: 2nd international conference on computer, communication, control and information technology (C3IT-2012) on February 25–26, 2012. <https://doi.org/10.1016/j.protcy.2012.05.067>
7. Roy CS, Roy PS (2015) Design of slotted microstrip patch Antenna with enhanced bandwidth in C-band applications. In: 2015 IEEE international conference on electrical, computer and communication technologies (ICECCT), 2015, pp 1–4. <https://doi.org/10.1109/ICECCT.2015.7226186>
8. Balanis CA (1997) *Antenna theory analysis and design*, 2nd edn. Wiley, New York
9. Roy PS, Chakraborty S (2016) Analysis of microstrip patch antenna with enhanced bandwidth for wireless application. *Int J Electron Comput Eng* 7(4):220–223. ISSN No: 2249-071X
10. Tarbouch M, El Amri A, Terchoune H, Barrou O (2018) A compact microstrip patch antenna based on fractal geometry on the ground plane. In: 2018 international conference on advanced communication technologies and networking (CommNet), Marrakech, 2018, pp 1–8. <https://doi.org/10.1109/COMMNET.2018.8360245>
11. Roy PS, Chakraborty S (2012) Design of C-slotted microstrip antenna using artificial neural network model. *Int J Res Sci Adv Technol* 2
12. Roy PS, Chakraborty S (2017) Bandwidth analysis and optimization of  $\Psi$ -shaped microstrip patch antenna with artificial neural network. *Int J Innov Sci Res* 2(2):29–306. ISSN No: 2351-8014. <http://www.ijisr.issr-journals.org/abstract.php?article=IJISR-17-057-0>

# Integrated Smart Alert System for Industrial Applications using Transceiver Module Analysis



Vaithiyathan Dhandapani , Joel Jacob Thomas, and Y. Durga Sraavanthi

## 1 Introduction

In our day-to-day lives, we always solicit to get notified of anything and everything, be it an alarm to wake up, a reminder to accomplish tasks or to know where our spectacles or keys are lying. In such a world of hustle-bustle, we are in desperate need of an alerting system that can also monitor every detail of our lifestyle keenly. We are talking small scale, so it is imaginable how important it is for industries.

In a more practical scenario, every year, thousands of medical samples are prone to get discarded from the laboratories due to a lack of maintenance of the required atmospheric factors. In the case of vaccines (say, for COVID-19), high heat may cause degradation in the quality, and freezing may end up losing its ability. One of the recent incidents that shocked the nation was when five frontline workers fell victim to a major fire in Serum Institute of India (SII) in Pune, an institute dedicated to the world's vaccination run against COVID-19, due to an unnoticed anomaly [1]. Even in other industries, we have seen accidents causing tremendous loss due to the hike in temperature of the storage facility, or increased exposure to light for plantations, and this insists for an integrated solution in keen monitoring that can alert us of any such anomaly possibility and analyze the situation well beforehand.

Hence, we study the development of a secure, continuously integrated, and smart alert system with minimal hardware, maximum efficiency, user-friendliness, and portability using considerable attributes. We enhance the system with a prediction model to get a broader picture of how and when the system is probable to fall witness to a deviation in the attributes from the standard range, using polynomial regression in machine learning. One key risk with transforming the earlier heavy hardware to

---

V. Dhandapani (✉) · J. J. Thomas · Y. D. Sraavanthi  
Department of Electronics and Communication Engineering, National Institute of Technology,  
Delhi, New Delhi 110036, India  
e-mail: [dvaithiyathan@nitdelhi.ac.in](mailto:dvaithiyathan@nitdelhi.ac.in)

nanotech with the use of Internet is that the device always is connected to some or the other cloud. Eavesdropping into one of the devices or its environment can grant access to the entire service, and this is a high-potential risk [2]. Vulnerability of IoT devices at the end of gateway is high. This gives rise to the research on how conventional methods to enhance security can be automated for each time the system is accessed [3]. The Mirai attack of 2006 that enslaved the poorly secured routers and devices into self-spreading in US is a perfect reason to block bots from accessing any system [4]. Therefore, to validate secure access, a double-level authentication is implemented to grant access only to authorized personnel. Given the recent advances in technology, it is plausible to integrate IR with a system that senses the environment in real-time for overall diagnosis [5]; along with securing the software implementation of the system, the hardware implementation is also secured using motion sensors to analyze any intrusion in the premises of observation.

## 2 Literature Review

The past few years have experienced a driven expansion and advancement in technology that constantly improve our lifestyle now. The demand for more has only been on the rise ever since. An annual growth of 42% until 2023 is expected for the global sensor market, and it is not just for healthcare, according to [6]. With such a pivotal role, it is justified to claim that COVID-19 has been a catalyst for IoT. The key to sustained utilization is efficient and result-driven developments, so that the providers as well as consumers are advancing in harmony in the years to come.

Several pieces of work have been demonstrated in the domain, which includes sensing as well as alerting broadly. Although these efforts have made a change to the thought process of the industry, a lot of shortcomings have pulled them back with time, and these inabilities have not let them deliver a positive and powerful impact. A research from Hubei University on temperature and humidity systems in hospitals makes use of a proportional-integral-derivative controller for observing key temperature and depicting humidity from it using a consequent algorithm, which concludes that the factors under observation can be correlated [7]. An experimental study on real-time monitoring of industrial machines with SMS has turned out successful and deploys an alert for every change in peak traditionally [8]. This introduces the alerting capabilities of a machine, although the heavy hardware usage could cause a security breach. Enhancing this with a seamless prediction is an event triggered by this study. A group of six contributed the attribute-based control with authorized search in cloud storage, which provides a trap door for outlying attributes and changes its access using its hybrid algorithms [9]. This became beneficial for the defense sectors as data are protected in real-time. The data manipulation possibilities had to be overcome as it became prey to decryption. There arises the need for the safe storage of

data. A study from computer science majors has established the use of probability-based methods and ANN techniques for data fusion [10]. This provides a structured and mathematical implementation of pour-and-store data. Another study in artificial intelligence for cloud-assisted smart factories suggests AI for network layers of application, cloud, and device levels, giving an efficient technique for handling data in individual layers [11]. It can be ideally expanded to cross-disciplinary data on further study, which is what this research has caught upon.

The challenge is to develop an integrated system to alert a user based on the continuously collected data of factors affecting the observational environment (like temperature, light) and its variations in an atmosphere and also can perform various data analytics techniques to get insights into the collected data. An aggregation of sensors with an infrastructure-less connection is what forms a wireless sensor network (WSN), enabling components to build a social network of things in itself, like Fig. 1.

A WSN is good as long as the information it puts on table is good [12]. Thus, securing the system is integral. Table 1 acts as a guide to the various abbreviations and acronyms used in the paper. Goals aimed through the research are to automate processes, to increase efficiency, to reduce costs, better storage, better collaboration, and eventually reduce the ecological footprint, to join hands in sustainability.

**Fig. 1** Social network of things



**Table 1** Table defining the nomenclature used in the paper

S. No	Short form	Definition
1	IC	Integrated circuit
2	ADC	Analog to digital converter
3	LDR	Light dependent resistor
4	PIR	Passive infrared sensor
5	GPIO	General purpose input/output pin
6	CPU	Central processing unit
7	LED	Light emitting diode
8	OS	Operating system
9	VM	Virtual machine
10	API	Application programming interface
11	ML	Machine learning
12	UI	User interface (front-end design)
13	GUI	Graphical user interface
14	Bot	short for “robot”
15	2FA	Two-factor authentication
16	P2P	Point to point
17	IoT	Internet of things
18	ANN	Artificial neural network

### 3 System Overview

#### 3.1 System Components

The heavily integrated and smart system are introduced to the user in a secure and consistent environment, be it small-scale or large-scale. The base model for integration is adapted from Azure as in Fig. 2. Users shall be able to access a hosted data server to extract the latest sensor values out of the temperature sensor IC, or the LDR light sensor, from respective terminals. There is an optional field provisioned to set up a threshold. Then, we analyze the data extracted from the server of the ambiance under observation. Different data visualization techniques are a prerequisite to study the sensed values in this system. Moreover, a predicted set of values is also displayed to the user to plan his/her next move concerning the asset. By his/her, we intend humans alone, and to decline access to any hacker/cracker/bot, the development of a strictly smart login feature with two-factor authentication comes into play. In addition to this, to ensure the system hardware and its ambiance are safe, the passive IR motion sensor is integrated to attract any movement within the premises of the system.

ESP8266 is the transceiver module that can enable itself as a Wi-Fi adapter to any microcontroller with any communication interface and possesses a multitude of pins,

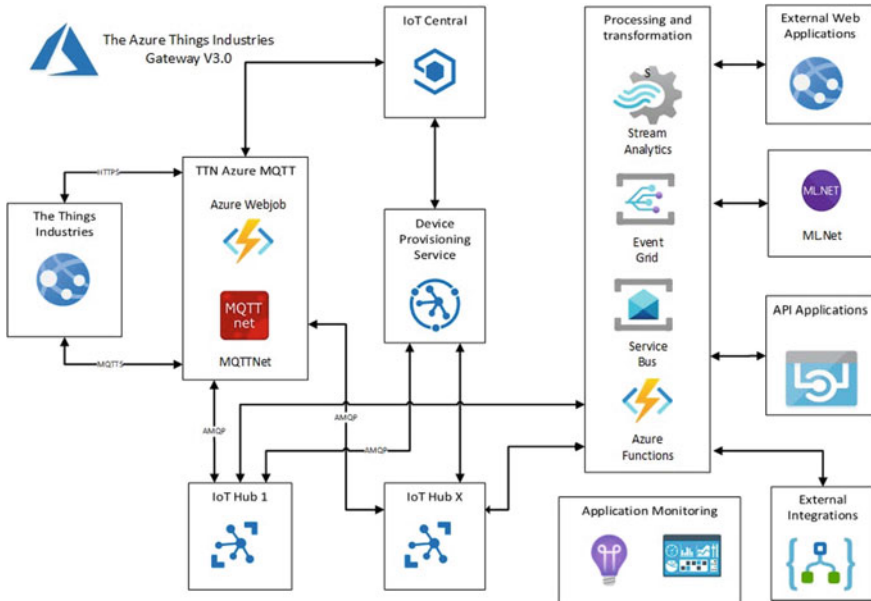


Fig. 2 Diagram showing the base IoT integration model the system adapts

shown in Fig. 3. With the capabilities of a microcontroller and Wi-Fi networking, it is a standalone device with high compatibility.

LM35 is a precision temperature IC that gives an easy interface with any microcontroller that has an ADC function. Every increase in degree temperature is accounted for an increase of 10 mV, and the input voltage is received as temperature as below:

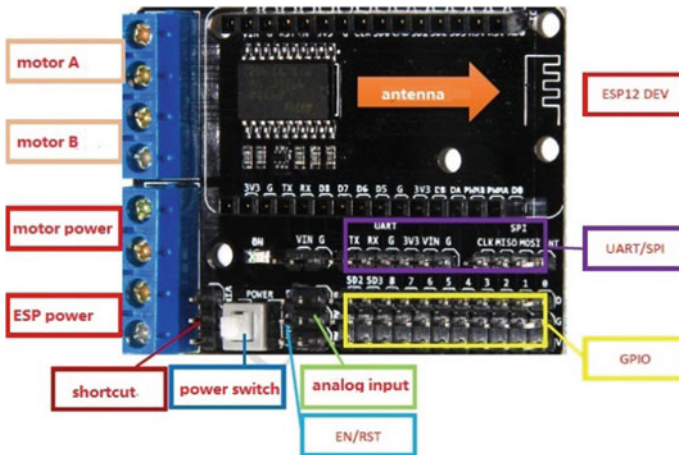


Fig. 3 Schematic of the Wi-Fi module ESP8266-12S IC

$$V_{out} = 10 \text{ mV}/^{\circ}\text{C} \times T \tag{1}$$

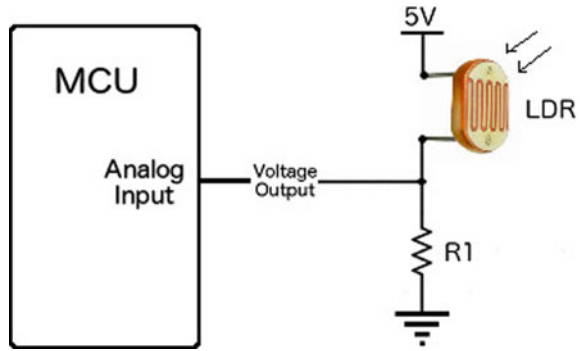
where  $V_{out}$  is the output voltage, and  $T$  is the temperature in  $^{\circ}\text{C}$ .

Light dependent resistor (LDR), otherwise known as photo-resistor, enables photoconductivity. DC multi-meter measures the voltage across LDR. The connection of LDR with the module is as depicted in Fig. 4. Light closer to the resistor decreases the resistance value of LDR and results in a proportional decrease in voltage drop.

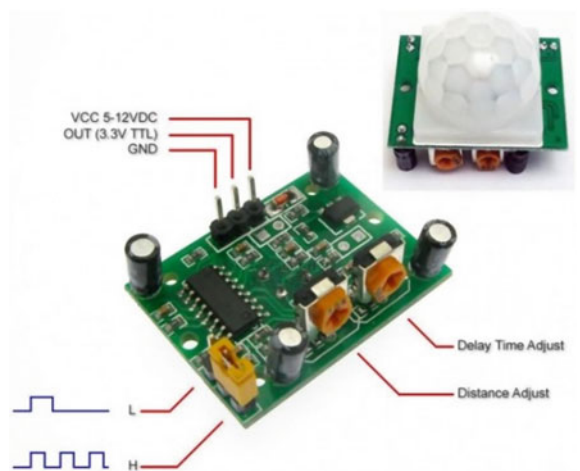
PIR HC-SR501 is a passive infrared sensor with two operating modes (H and L) to detect the presence/motion of beings and has pins as shown in Fig. 5. Typically used at 5 V voltage (extendable), it needs to be calibrated well before direct use due to its high sensitivity (for around 2 min). It has promising use in extracting information on human motion, abstracting other irrelevant information [13].

The tactile switch is another passive component that is used for tactile input received from the user and allows easy interfacing with the board. The pressed

**Fig. 4** LDR connection diagram with terminals



**Fig. 5** Schematic of the PIR HC-SR501 sensor (upper and lower side views)



button makes the switch ON, and the released button makes the switch OFF. The tactile response is consequently defined by the perception of touch on this switch.

**Hard Indications:** Certain LEDs are used to debug the system board and diagnose: Blue LED tells us about the status of Wi-Fi and green LED about the cloud connectivity. The blue LED has 4 states. If both the LEDs are stable, the module is ready for further hardware connections. Orange LED is set up with intensities in the range (0–1024) to alert the user of real-time deviations. Piezoelectric buzzer rings with set intensities in the range (0–255), also to alert the user.

**Soft Indications:** Seamless integration with real-time SMS services through Twilio, personal email services through Mailgun API, and collateral alerts through Telegram channel using BotFather to create our bot are the focused alerts in our scope.

### 3.2 Software Initialization

Before we begin to run the system, we need to initialize it with a set of prerequisites. They include installing Python and certain site packages required, setting up a Web control app using a VPL [14], defining workflow and train model on Azure machine learning, and linking APIs to communicate, to integrate the software and hardware.

**Python Installation:** Python is a high-level programming language first developed by Rossum. Its simplicity allows developers to implement several machine learning algorithms, and it has a multitude of inbuilt packages. The smart system makes use of PyCharm IDE for Python up to v3.10. Pip (Python packaging index) is a package manager which installs and handles software packages that are written in Python. Modules used: json, requests, math, statistics, time, configuration.

Run in console: `Python -m pip install--upgrade pip`.

**Microsoft Azure:** The cloud services API is linked to the system, and activation is done through base model of Azure IoT. Azure ML, meanwhile, helps us to stream and save data, preprocess data, enhance predictive modeling, and publish the predictive model. Azure portal enables all services for cloud training, to set:

- Compute Instance: VM type: CPU, VM size: Standard\_DS11\_v2.
- Compute Cluster: VM type: CPU; VM size: Standard\_DS11\_v2; VM priority: Dedicated; min, max number of nodes: 0, 2; idle seconds prior to minimization: 120.

Training pipelines for the respective observed datasets can be generated in Azure ML Studio for predictive modeling of the system.

**Control App (Web):** Bubble.io is a robust VPL platform supporting no-code architecture that enables users to create innovative Web applications and enables to build of powerful GUI to control output devices connected. Refer Fig. 6. Steps to create control app: Log in and scroll down to “create new” → Name it and start with a



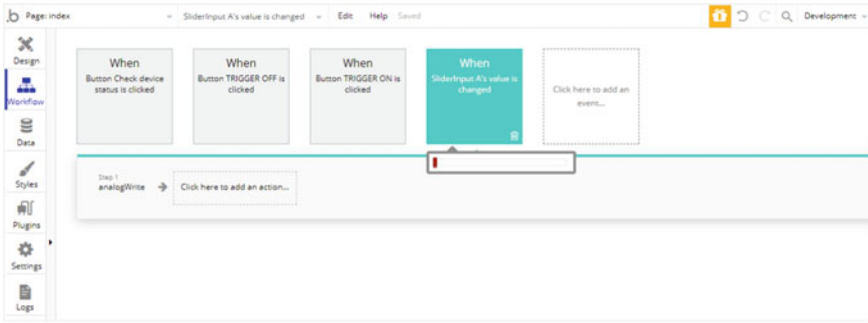


Fig. 6 Example of workflow devised in Bubble.io for controlling

blank page → Add the IoT plug-in and install it → Open the page and design the UI → Assign action to each element in workflow → Preview and publish the app as live. This enables the authorized user to control the system remotely, basic as in Fig. 7.

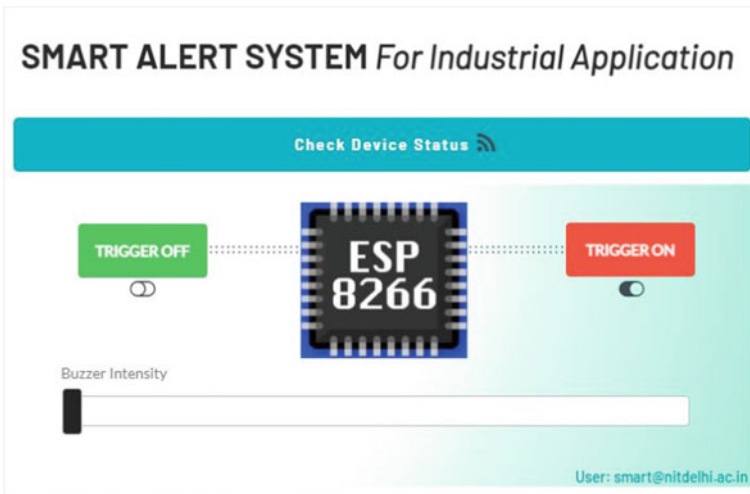


Fig. 7 Control app–base UI for our system

## 4 System Flow

### 4.1 Login Analysis

Although a rule of thumb for passwords is that it should not be common and known only to the user, humans tend to set the same passwords for multiple platforms and moreover have weaker memory. This paves way for the risks of malicious logins. Security properties to be kept in mind while developing include confidentiality, integrity, authentication, authorization, availability, and non-repudiation [15]. 2FA accounts itself as a second-level login authentication, wherein the next move/activity of the user gets clearance. Primarily, we focus on the sensing layer as it is where an attack initiates and that is where this tactile switch 2FA comes into play. Steps to interface system with tactile switch and its response:

- Two factors the login authorizes on: login credentials (configurable) and tactile switch response.
- If credentials are invalid: second chance to login pops up.
- If the given credentials are valid, then a function is called which takes input from the tactile switch (2nd authentication). If the user presses the tactile switch: circuit completes, the value becomes zero and login will be successful.
- If the login is unsuccessful after 2 attempts, (suspicious login): there the buzzer buzzes with a high volume to alert admin and triggers software alerts, whichever assigned. It will be accessible for user to login after a certain time which will be assigned (10 s, here).

### 4.2 Ambience Analysis

This system breaks down the problem and integrates the alarming capability with enhanced prediction on parameters such as temperature and light for now (as these are the main factors on which the quality of any specimen in an industry depends on; extendable on large scale). As per [16], by using this analysis in capturing information from even humans of their internal attribute values, we can speed up the COVID-19 detection process due to the instant temperature samples we can take and so on.

**Temperature Sensing:** Configuring this function requires consideration of factors like probe, thermal buffer, constant measurement, data storage, and alarming capacity. LM35 establishes this requirement in our system.

**Light Sensing:** The change in resistance of sensor which is then converted to change in voltage is factored upon intensity of light falling. LDR establishes this smoothly and outputs analog voltage proportional to the light receptive of it.

**Z-Score:** In our system, this analysis lets us study the deviation from sensed values, and when input to our algorithm gives us a better accurate prediction of probable

data. This is called feature scaling. The standard score for z-score is calculated as below:

$$z = \frac{X - \mu}{\sigma} \tag{2}$$

where  $\mu$  is the mean,  $X$  is the score, and  $\sigma$  is the standard deviation.

**Regression:** Regression analysis helps develop a connection between the target and predictor variables using independent variables. Under-fitting and over-fitting can be a running problem when taking all data points into consideration, wherein the training dataset may fit the algorithm developed but not the test dataset or vice versa, respectively. We ought to avoid this and observe considerable number of data points to study. Polynomial regression is used in a linear model with non-linear dataset, hence deriving the equation from a linear regression model as below:

$$Y = b_0 + b_1x + b_2x^2 + \dots + b_nx^n \tag{3}$$

where  $Y$  is the dependent output,  $b_0 \dots b_n$  are the regression coefficients, and  $x$  is the independent input. Polynomial regression is performed on the data values implied through the GPIO inputs of the sensors which is trained by following Fig. 8, and prediction is made based on its z-score.

- Let the sensors calibrate with the ambience under observation for a few minutes
- As and when the system is up, the sensors record all deviations
- The collected data are analyzed and further divided into two factions: trigger and predict. The trigger is alarmed when the deviation is observed, and the prediction is a result of all the values taken into consideration for the given ambience.

The extracted excerpt from the big observed data of the experimented setup with the base sensors of temperature and light, enhanced with 2FA, is as shown in Table 2. This data give the notion that accesses to the sensors and its values for further use can be granted only for authenticated and successful logins, and this explains how correlated the attributes are with each other.

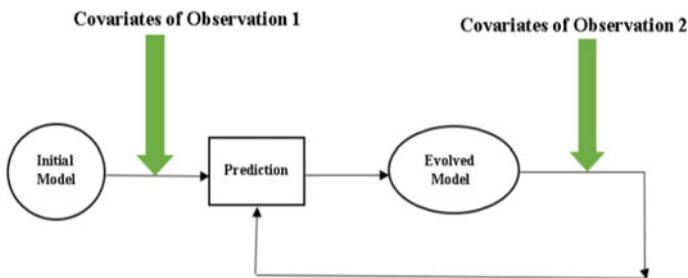


Fig. 8 Predictive model training methodology

**Table 2** Observed training data excerpt taken at a point of analysis

Time_stamp	Temp value	Light value	Login success (1/0)
18/02/2021 10:46	105	465	1
18/02/2021 10:51	100	421	1
18/02/2021 10:54	98	216	1
18/02/2021 10:55	97	204	1
18/02/2021 10:57	96	221	1
18/02/2021 10:58	96	211	1
18/02/2021 11:03	4	192	1
18/02/2021 11:03	3	195	1
18/02/2021 11:04	354	1024	1
18/02/2021 11:06	358	1024	1
18/02/2021 11:07	358	1024	1
18/02/2021 11:12	208	1024	1
18/02/2021 11:14	182	521	1
18/02/2021 11:22	146	173	1
18/02/2021 11:27	327	1024	1
18/02/2021 11:27	332	1024	1
18/02/2021 11:31	178	301	1
18/02/2021 11:33	164	322	1
18/02/2021 11:33	161	385	1

### 4.3 Intrusion Analysis

“With great power, comes great responsibility.” Although we have developed a user-specific continuous system, securing the system within the premises is the next vital task as discussed. Steps to initiate the analysis:

- Let the sensor takes a minute or two to stabilize (settling time 20–120 s)
- Keep the IR sensor white hemisphere facing up
- Due to very high sensitivity, the sensor demands motionless atmosphere to correctly detect and report any movement in premises
- If the login is unsuccessful after two attempts (suspicious login): buzzer buzzes with high volume for admin and triggers software alerts, whichever assigned.
- Hence, leave the sensor connected and run the code, the respective results are output and based on those, alert is triggered.

## 5 Results

### 5.1 Prototype

As the integrated prototype is setup and deployed, the live feed is continuously streamed in cloud and controlled using our control app. The login analysis is undergone through a 2FA. Given below is an instance of the same (Figs. 9 and 10), with the authentic credentials as username = “smart@nitdelhi.ac.in” and password = 1212. This is configurable in the backend of any system.

Successful login is defined by matched validation of credentials entered by the user as well as positive tactile response from the tactile switch at user’s end.

Unsuccessful login is defined by unmatched validation of credentials entered by the user, or the negative tactile response from the tactile switch at user’s end. This triggers all indications and holds the user back for 10 s straight until authorized person is notified of the attempt.

Given below is a demonstrative instance of the hard indications when it is placed under low-temperature and low-light conditions (Figs. 11 and 12). When the temperature is raised a little, the orange LED and buzzer trigger and simultaneously notify the user with various alerts described above, even when the lighting condition stays the same. This is how smart the system can be to notify the user of all kinds of deviation in any of the observational parameters anywhere across the globe.

```
C:\Users\Hp\PycharmProjects\venv\Scripts\python.exe
C:/Users/Hp/PycharmProjects/venv/Lib/site-packages/loginalert.py

Enter username:smart@nitdelhi.ac.in

Enter password:1212
Press and hold button for 1 sec
{'value': '1', 'success': 1}
{'value': '1', 'success': 1}
{'value': '1', 'success': 1}
{'value': '1', 'success': 1}
{'value': '1', 'success': 1}
Timeout. Please login again

Enter username:smart@nitdelhi.ac.in

Enter password:1212
Press and hold button for 1 sec
{'value': '1', 'success': 1}
{'value': '1', 'success': 1}
{'value': '1', 'success': 1}
{'value': '0', 'success': 1}
Successfully logged in!

Process finished with exit code 0
```

**Fig. 9** Code output of the 2FA for success by second attempt

```
C:\Users\Hp\PycharmProjects\venv\Scripts\python.exe
C:/Users/Hp/PycharmProjects/venv/Lib/site-packages/loginalert.py

Enter username:smart@nitdelhi.ac.in

Enter password:1212
Press and hold button for 1 sec
{'value': '1', 'success': 1}
{'value': '1', 'success': 1}
{'value': '1', 'success': 1}
{'value': '1', 'success': 1}
{'value': '1', 'success': 1}
Timeout. Please login again

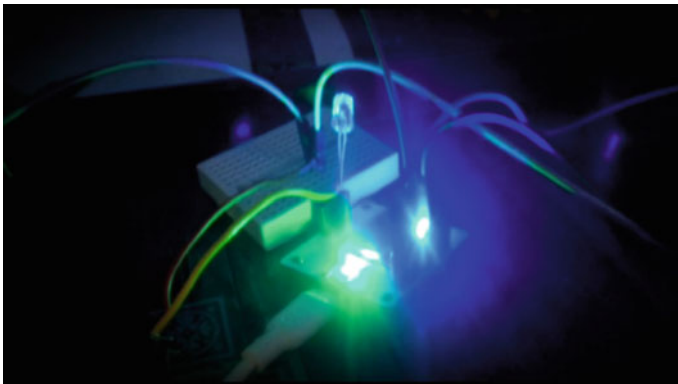
Enter username:smart@nitdelhi.ac.in

Enter password:1212
Press and hold button for 1 sec
{'value': '1', 'success': 1}
{'value': '1', 'success': 1}
{'value': '1', 'success': 1}
{'value': '1', 'success': 1}
{'value': '1', 'success': 1}
Timeout. Please login again

Maximum tries reached. Please wait for another 10 seconds

Process finished with exit code -1
```

**Fig. 10** Code output of the 2FA for unsuccessful login even by second attempt



**Fig. 11** Ambience analysis resulting in safe indication

Figures 13 and 14 show the intrusion analysis resulting in an unsafe alert through soft yet instant indications like Email and Telegram, where the automated messages are timely sent P2P for user-specific reception.

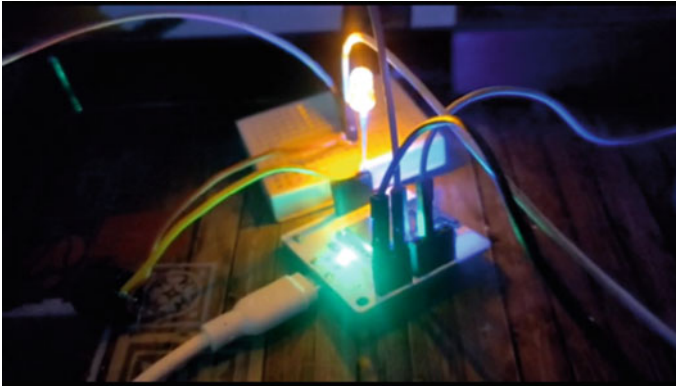


Fig. 12 Ambience analysis resulting in unsafe, alert indication



Fig. 13 Mailgun API alerting the user through email upon intrusion

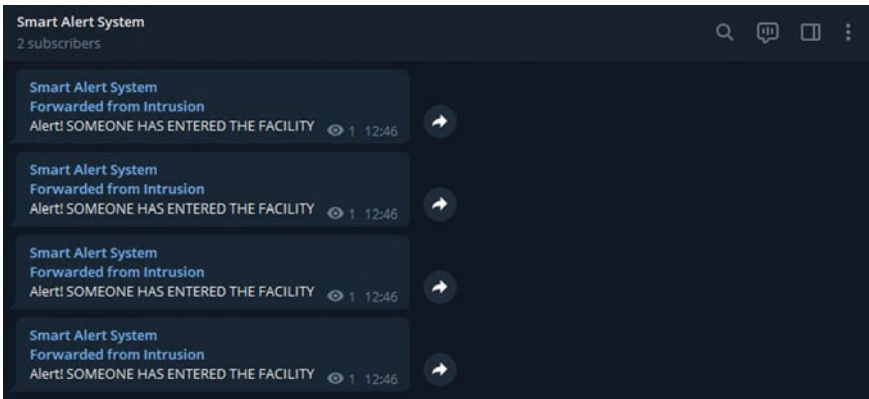


Fig. 14 Bot administered for Telegram channel alerting the user upon intrusion

### 5.2 Visualization

The streamed data can be visualized using various techniques. Although there are stepped graphs and multiple graphs, they need multiple parameters because they are

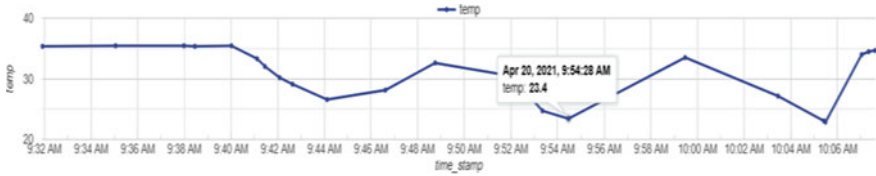


Fig. 15 Line visualization of obtained data from LM35. Blue line = Initial observed data

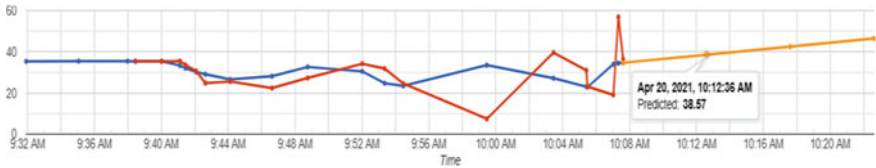


Fig. 16 Line visualization of predicted data from LM35. Blue line = Initial observed data. Red line = Prediction possibility history. Orange line = Final predicted data

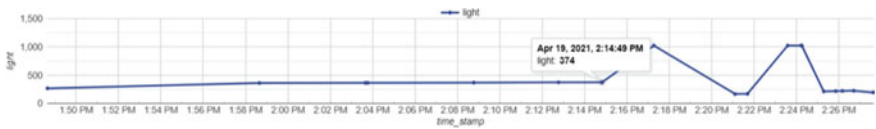


Fig. 17 Line visualization of obtained data from LDR. Blue line = Initial observed data

basically graphs used for comparison; line graph is best used for sensed values. Given in Figs. 15, 16, 17, and 18 are the line visualizations of output from the temperature and light sensors, respectively, at a given period of analysis. The prediction considering a data frame of 9 data points is obtained from the recorded data.

Furthermore, the intrusion analysis cannot necessarily have a relation among its parameters by nature, but the experimental intrusion that leads to an alert can be analyzed using Table 3, wherein the stability time delineates the calibration it takes.

The experimental results prove that the integrated network of sensors can do wonders in safe and real-time data extraction from the environment. An efficiency of 92% (as measured by an instance of predicted to actual value of temperature) is achieved. This establishes the accuracy that the data points are trained and tested with.

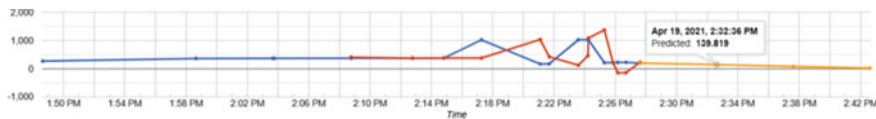


Fig. 18 Line visualization of predicted data from LDR. Blue line = Initial observed data. Red line = Prediction possibility history. Orange line = Final predicted data



**Table 3** PIR observation excerpt for three instances

Time_stamp	Stability time (s)	Distance (m)	Trigger
20/04/2021 11:56	94	± 20	No
20/04/2021 12:45	55	± 5	Yes
21/04/2021 09:32	46	± 10	Yes

## 6 Conclusion

To summarize, a smart, highly integrated, predictive, and automated alert system have been devised and studied, which proposes the analysis of modern WSN components, to meet all kinds of communication needs tailored to all industries. The proposed system has taken into consideration all related constraints and improved the existent shortcomings. The analysis technique model has resulted in efficient data handling that is upfront for effective decisions on a large scale.

The use of this system provides automated data capture, which can save organizations considerable time and labor costs. Effective sensing can lead to effective prediction and thus save top industries the headache of spontaneous maintenance. Alerting media can be extended, with a few lines of code to ensure their connection throughout the globe. This research paves way for Industry 5.0, the next revolution to smarter handling of data, as it is a promising hand to ensure the interaction and collaboration between mankind and machine.

## References

1. BBC India News page, Serum Institute: fire at world's largest vaccine producer kills five. <https://www.bbc.com/news/world-asia-india-55753586>. Last Accessed 14 May 2021
2. Fremantle P, Aziz B (2018) Cloud-based federated identity for the internet of Things. *Annal Telecommun* 73:415–427
3. Kponyo JJ, Agyemang JO, Klogo GS, Boateng JO (2020) Lightweight and host-based denial of service (DoS) detection and defense mechanism for resource-constrained IoT devices. *ScienceDirect Internet of Things* 12
4. Das S, Amritha PP, Praveen K (2021) Detection and prevention of Mirai attack. In: Reddy V, Prasad VK, Wang J, Reddy K (eds) 3rd international conference on soft computing and signal processing (ICSCSP 2020), vol 1325, Springer (2021), pp 79–88
5. Llaguno-Munitxa M, Bou-Zeid E (2021) Sensing the environmental neighborhoods: mobile urban sensing technologies (MUST) for high spatial resolution urban environmental mapping. In: Yuan PF, Yao J, Yan C, Wang X, Leach N (eds) Springer proceedings of the 2020 Digital FUTURES—the 2nd international conference on computational design and robotic fabrication (CDRF 2020), Shanghai (2021), pp 124–133
6. Fierce Electronics—Electronics page. <https://www.fierceelectronics.com/electronics/role-iot-sensors-covid-19>. Last Accessed 21 July 2021
7. Guo B, Wang X, Zhang X, Yang J, Wang Z (2016) Research on the temperature and humidity monitoring system in the key areas of the hospital based on the Internet of Things. *Int J Smart Home* 10(7):205–216
8. Joshi SS, Kulkarni VM (2019) Real time monitoring of industrial machines using IOT with SMS alert system. *Int Adv Res J Sci Eng Technol (IARJSET)* 6(3)

9. Hao J, Liu J, Wang H, Liu L, Xian M, Shen X (2019) Efficient attribute-based access control with authorized search in cloud storage. *Spec Sect Secur Priv Cloud IoT IEEE Access* 7:182772–182783
10. Alam F, Mehmood R, Katib I, Albogami NN, Albeshri A (2017) Data fusion and IoT for smart ubiquitous environments. *IEEE Access* 5:9533–9554
11. Wan J, Yang J, Wang Z, Hua Q (2018) Artificial intelligence for cloud-assisted smart factory. *Spec Sect Collab IoT IEEE Access* 6:55419–55430
12. Olariu S, Wadaa A, Wilson L, Xu I Q, Eltoweissy M, Jones K (2004) Providing holistic security in sensor networks. In: *AICT proceedings of IFIP advances in information and communication technology*, vol 169. Springer (2004), pp 123–134
13. Luo X, Liu T, Liu J, Guo X, Wang G (2012) Design and implementation of a distributed fall detection system based on wireless sensor networks, Springer *EURASIP J Wirel Commun Netw* 118
14. Digital Technologies Hub—Visual Programming page. <https://www.digitaltechnologieshub.edu.au>. Last Accessed 21 July 2021
15. Al-Garadi MA, Mohamed A, Al-Ali AK, Du X, Ali I, Guizani M (2020) A survey of machine and deep learning methods for Internet of Things security. *IEEE Commun Surv Tutor* 22(3):1646–1685 (2020)
16. Nasajpour M, Pouriyyeh S, Parizi RM, Dorodchi M, Valero M, Arabnia HR (2020) Internet of Things for current COVID-19 and future pandemics: an exploratory study. *J Healthc Inf Res* 325–364

# Analyzing Beamforming for Secure Transmission of MIMO-NOMA-Based CRN using Power Allocation Based on Second-Order Perturbation



L. Thulasimani and A. Hylis Sharon Magdalene

## 1 Introduction to MIMO-NOMA Model

Consider MIMO-NOMA uplink-downlink communicating [1, 2] premises in which the BS transmits information to the many users simultaneously. The BS is provided with  $M$  antennas, and an individual user is specified with  $N$  antennas, respectively. Also consider the assumption  $N > \frac{M}{2}$  in order that the conception of signal alignment is implemented, a scenario which is more common than the implementation suggested by Peng et al. [3]. This scenario is relevant to several communication assumptions, for instance limited cells in 5G cloud radio access networks (C-RANs) (China Mobile Research Institute 2011) and heterogeneous networks [3], in which the low-cost BS is distributed with larger density. The users are consistently distributed in a disc  $D$ , which contains  $D_1$  and  $D_2$ , respectively. Therefore, the cell is restricted by the BS. The inner part of the disc is  $D_1$ , and the radius is  $r_i$ . Then,  $D_2$  is the outer part of the disc, and the BS is fixed at the middle of the disc, and  $r_o$  is the radius of the disc (Fig. 1).

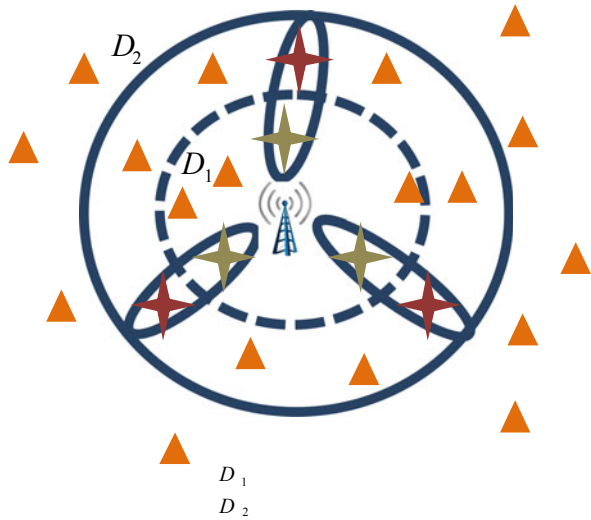
Many existing examinations on NOMA [4, 5] are suggested to decrease the system load by pairing of two users in NOMA implementation [6] and have established that it is idealistic to match users who have the distinct channel conditions [7]. Depending upon the view, the disc is classified into two parts and is represented in Fig. 1.12. The first part having minor disc and is referred as  $D_1$ , with radius  $r_i$  ( $r_i < r_o$ ) and the BS settle at its source. The second part is like a ring and is represented as  $D_2$ , making from  $D$  by get rid of  $D_1$ . Consider that the  $U$  pairs of users are chosen, which may

---

L. Thulasimani (✉) · A. Hylis Sharon Magdalene  
Department of Electronics and Communication Engineering, PSG College of Technology,  
Coimbatore, Tamil Nadu, India  
e-mail: [ltm.ece@psgtech.ac.in](mailto:ltm.ece@psgtech.ac.in)

A. Hylis Sharon Magdalene  
e-mail: [hsm.ece@psgtech.ac.in](mailto:hsm.ece@psgtech.ac.in)

**Fig. 1** Representative of the considered system model



be one SU, that are randomly situated in  $D_1$  and are paired with another PU, which is also randomly situated in  $D_2$ . This approach manipulates the data of the channel users, and so it is applicable nevertheless the distance from the SU to the BS is larger than or equal to that of the PU to the BS. The user which has the more sophisticated user pairing system can promote better performance in the MIMO-NOMA systems.

Consider a complex channel model with large scale path loss and semi-static Rayleigh fading. Specifically, the channel matrix from the BS to the user  $u$  is  $H_u = \frac{G_u}{\sqrt{L(d_u)}}$ , where  $G_u$  represents a Rayleigh fading channel gains with  $N \times M$  matrix, and the distance from the BS to the user is indicated as  $d_u$ . The consequential path loss is described as follows:

$$L(d_u) = \begin{cases} d_u^\nu, & \text{if } d_u > r_0 \\ r_0^\nu, & \text{otherwise} \end{cases} \tag{1}$$

The path loss exponent is indicated as  $\nu$  and  $r_0$  annuls a uniqueness, where the distance is limited. Assume that,  $r_1 \geq r_0$  and  $r_1$  are applied to modify the analytic results. For notational simplicity, the channel matrix from the BS to the user  $u$  is constituted as  $H_u^H$ . This global CSI is considered to be usable in the BS and the users. The downlink and uplink MIMO-NOMA systems are depicted in the adopting two subdivisions, respectively [8].

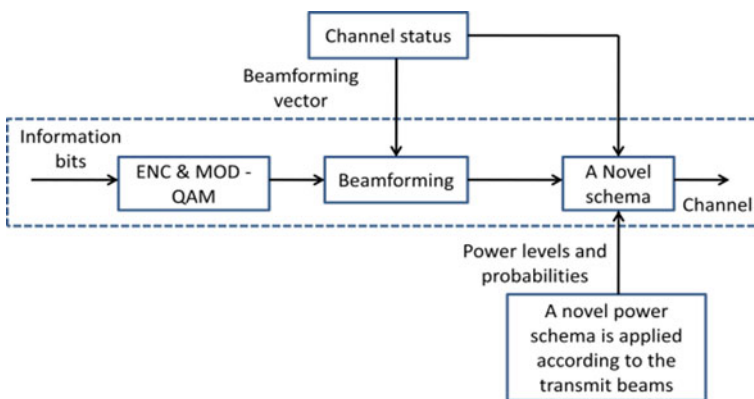
In this paper, the secrecy rate has been calculated according to the one that must recognize the eavesdropper's CSI. Such information is not preferable to be available in many assumptions, especially those involving solely passive eavesdroppers. Assumption of the transmitter has more number of antennas than the intended recipient so that the consistent channel has a non-trivial null space. One of the approaches is to broadcast an artificial interference in this null space, and such interference has

no impact on the receiver. So the interference model was developed. The high-SNR performance of the instance in the technique was expressed to be almost optimal, and the optimal power distribution among the data and the interference is analyzed. As a result, a different approach has been taken in which the transmitter minimizes the transmit power required to ensure the identified QoS at the desired receiver.

When the only distribution of the eavesdropper’s channel is recognized, then focusing on the situation where the transmitter has access the eavesdropper’s instantaneous CSI. An algorithm is developed that optimally use the information in the case where the eavesdropper’s CSI is recognized. The key consideration is the accuracy of the available CSI. Thus, the impact of the imperfect CSI on the secrecy rate of the single-antenna wiretap channel has been investigated. So the technique that depends on the knowledge of the eavesdropper’s channel in the multiple antenna is very sensitive to little perturbations in the CSI. If CSI is not accounted for the primary channel then the cause of interference to the desired recipient will be more. This results significant degradation in the desired user’s performance.

The block diagram of applying a novel second-order perturbation beamforming scheme in the beamforming system is shown in Fig. 2.

The robust schemes are developed that are restricted to the CSI errors, and transmission of a single data stream is performed by using a beamforming scheme [9]. Then, the novel second-order perturbation technique is proposed to minimize the transmit power that is required for the desired receiver to achieve the target QoS in the presence of the CSI errors. In turn, this maximizes the power availability to transmit a jamming signal that disrupts the ability of the eavesdroppers to recover the desired signal. The robust the novel second-order perturbation technique is proposed based on the knowledge of the statistics of the CSI errors. Hence, the algorithms provide the following gain: (1) minimize the impact of the jamming interference at the desired receiver when the CSI errors are presented, (2) it requires less transmit power to achieve the error free CSI, (3) then maximizes the power usability for



**Fig. 2** Block diagram of initializing a novel second-order perturbation technique in the beamforming system

degrading the eavesdropper channel. The simulation results demonstrate that the resulting secrecy capacity is significantly improved.

## 2 Methodology

### 2.1 A Novel Second-Order Perturbation Approach for the Secure Beamforming in MIMO-NOMA-Based CRN

The cascaded transmission methods used in the MIMO-NOMA framework are applied for the secure communication in the multi-cell model. The intruders will have an attempt to wiretap the channel information. The BS does not allow the intruders to restrict the data by applying a receiving beamforming technique. A novel second-order perturbation beamforming technique is implemented at the BS for the better finding of a data outflow between the cells. The received signal of the  $l$ th cluster in cell  $A$  for the PU and the SU is expressed as [10],

$$R_{su}^{l,A} = G_{su}^{l,A} \sum_{m=1}^L s_r^{l,A} + w_{su}^{l,A} \tag{2}$$

$$R_{pu}^{l,A} = G_{pu}^{l,A} \sum_{m=1}^L s_r^{l,A} + H_{pu}^{l,A} \sum_{m=1}^L s_r^{l,B} + w_{pu}^{l,A} \tag{3}$$

Let the  $M \times 1$  superposition signal vector be  $v^{l,A}$  for the  $l$ th cluster of cell  $A$  that holds the novel second-order perturbation technique of vector,  $s_r^{l,A} \in \mathbb{C}^{M \times 1}$  is implemented for the novel second-order perturbation schemes and the AWGN process be  $w_{pu}^{l,A}$  and  $w_{su}^{l,A}$ . It is found in the PU and the SU in the receiver position of the  $l$ th cluster of cell  $A$ . Then,  $l \in \underline{\underline{L}}\{1, 2, \dots, L\}$ .

Then, the secrecy rate of the SU and a PU is calculated as

$$Sr_{su}^{l,A} \underline{\underline{\text{def}}} \left[ \log_2 \left( 1 + \psi |G_{su}^{l,A} s_r^{l,A}|^2 \right) - \max_{m,m \neq l} \log_2 \left( 1 + \psi |G_{su}^{m,A} s_r^{m,A}|^2 \right) \right]^+ \tag{4}$$

$$Sr_{pu}^{l,A} \underline{\underline{\text{def}}} \left[ \log_2 \left( 1 + \psi |G_{pu}^{l,A} s_r^{l,A}|^2 \right) - X_9 - X_{10} \right]^+ \tag{5}$$

where the SNR transmission at the BS is  $\psi$ . The estimation  $[e]^+$  is set as  $\max(0, e)$ . Then,  $X_9 \underline{\underline{\text{def}}} \max_{m,m \neq l} \log_2 \left( 1 + \psi |G_{pu}^{m,A} s_r^{m,A}|^2 \right)$  and

$X_{10} \stackrel{\text{def}}{=} \max_m \log_2 \left( 1 + \psi \left| H_{\text{pu}}^{m,A} s_r^{m,B} \right|^2 \right)$ . The information leakage is assumed in the second term RHS of the above two equations in the same cell. For the neighboring cell, the data outflow is found in the third term RHS in the [11] equation. The information leakage will be prevented by employing a novel second-order perturbation technique in the BS, in which the BS has a CSI for all the users. The CSI is assessed at the receiving end, and renders response to the BS by the feedback channel.

The novel second-order perturbation technique is applied in the BS which depends upon the power allotment, and an adaptation-based rule to meet the SU security needs. Then make sure that the essential information range requisites for the PUs are satisfied. In particular, the power allocation constants are designed throughout the multiple access phase and the channel approximations in such a manner which will enlarge the operation of the legal users and reduce the functioning of an eavesdroppers [4]. The total difficulties in the interference are reduced through the novel second-order perturbation technique which executes effectively than all other methods.

The detailed steps of the novel second-order perturbation technique for the secure communication in MIMO-NOMA-based CRN system are given in Algorithm 1.

**Algorithm 1** A novel second-order perturbation beamforming scheme for mitigating the interference in the system model.

Step I: Initialize the beamforming input  $S_u, \forall u \in \{1, 2, \dots, M\}$ , for all the clusters, and initialize the PA coefficients  $\alpha_{u,v} \forall u \in \{\text{PU}, \text{SU}\}$  with the data symbols  $v$  of PU, SU data bits while keeping the MIMO-NOMA PA constraints.

Step II: The PA coefficient established for all the users is given in Eqs. (13) and (15).

Step III: A novel second-order perturbation beamforming scheme is applied for the computation of the interference covariance matrix in the cluster  $m$  of a cell  $A$  is expressed as

$$S_r^{l,A} = \text{span} \left\{ \omega_u^{c,A} t_n^H \bar{G}^H \left( \bar{G} P_r \bar{G}^H + \sigma_w^2 I \right)^{-1} \bar{G} t_n \right\} \tag{6}$$

A novel second-order perturbation scheme has a transmit covariance  $P_r$  and a noise covariance  $\sigma_w^2$ , and the transmitting beamformer is denoted as  $t$  of  $N \times 1$  dimensions. The transmit covariance of the beamformer  $P_r$  is calculated using Eq. (13).

Step IV: The power constraints matrix is  $T$  of  $N$  beams is calculated in [4] as

$$T = \begin{bmatrix} \sqrt{t_1} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \sqrt{t_n} \end{bmatrix} \tag{7}$$

A novel second-order perturbation power allocation coefficient for the  $N$  beams is calculated using Eq. (15). The transmit covariance of the beamformer restricts the

power usage using Eq. (13), and the constraints take place according to the  $N$  beams formed during the transmissions in  $T$ .

Step V: The interference covariance matrix calculated for the neighboring cell  $Y_r^A$  using the novel second-order perturbation scheme in accordant with the Eq. (11) that minimizes the interference leakage as possible.

Step VI: Compute the beamforming matrices using step 4 for the cell  $A$  and cell  $B$ , where the power constraints have been even more limited in the users using the novel second-order perturbation technique.

Step VII: Then, initializing the interference covariance matrices by the values obtained in the step 5 and replace the beamforming matrices as the error free matrices in the systems.

Step VIII: Distribute the transceiver filters using the novel second-order perturbation scheme with limited power constraints for both the users in the data communications.

## 2.2 Secure the Downlink MIMO-NOMA Framework by using a Novel Second-Order Perturbation Technique

Let the  $M \times 1$  superposition signal vector be  $s^{m,A}$  for the  $l$ th cluster of cell  $A$  that has a novel second-order perturbation vector of,  $s_r^{l,A} \in \mathbb{C}^{M \times 1}$ . A novel second-order perturbation technique is applied for the secure communication among the cluster users in the same and the nearer cells and is used at the BS. A novel second-order perturbation scheme  $s_r^{l,A}$  in [1] has non-trivial solutions. The cell  $A$  of a  $c$ th cluster in the novel second-order perturbation technique is indicated as  $s_r^{l,A} = Y^A S_r^{l,A}$  and  $S_r^{c,A} \in \mathbb{C}^{V \times 1}$  where  $Y^A \in \mathbb{C}^{M \times V}$  is the alignment matrix. Then, the alignment matrix  $Y^A$  is separated into  $Y_5^A Y_r^A$ , where the transmitting novel second-order perturbation beamforming vector is  $Y_r^A$  for the neighboring cell to prevent the outflow, and the nearby cell alignment to cancel the interference in the channel matrix is  $Y_5^A$ . A novel second-order perturbation vector  $Y_r^A$  for the nearby cell is derived as follows:

$$\left[ G_{\text{pu}}^{1,A} G_{\text{su}}^{1,A} \dots G_{\text{pu}}^{l-1,A} G_{\text{su}}^{l-1,A} G_{\text{pu}}^{l+1,A} G_{\text{su}}^{l+1,A} \dots G_{\text{pu}}^{L,A} G_{\text{su}}^{L,A} \right]^H s_r^{l,A} = 0 \quad (8)$$

$$\left[ G_{\text{pu}}^{1,A} Y^A G_{\text{su}}^{1,A} Y^A \dots G_{\text{pu}}^{l-1,A} Y^A G_{\text{su}}^{l-1,A} Y^A G_{\text{pu}}^{l+1,A} Y^A G_{\text{su}}^{l+1,A} Y^A \dots G_{\text{pu}}^{L,A} Y^A G_{\text{su}}^{L,A} Y^A \right]^G s_r^{l,A} = 0 \quad (9)$$

$$\left[ H_{\text{pu}}^{1,B} Y_5^A H_{\text{pu}}^{2,B} Y_5^A \dots H_{\text{pu}}^{L,B} Y_5^A \right]^T Y_r^A = 0 \quad (10)$$



where  $\mathbf{H}_{\text{pu}}^{L,B} \mathbf{Y}^A \mathbf{S}_r^{l,A} \in \text{span}(\mathbf{H}_{\text{pu}}^{L,B}) \forall \mathbf{Y}^A \mathbf{S}_r^{l,A}$ . A novel second-order perturbation scheme is expressed as

$$\mathbf{Y}_r^A = \text{span} \left\{ \omega_u^{c,A} \mathbf{Q} \mathbf{t}_n^H \overline{\mathbf{H}}^H \left( \overline{\mathbf{H}} \mathbf{P}_r \overline{\mathbf{H}}^H + \sigma_w^2 \mathbf{I} \right)^{-1} \overline{\mathbf{H}} \mathbf{t}_n \right\} \tag{11}$$

A non-trivial solution is assessed as follows  $\mathbf{S}_r^{l,A}$  in (7.5). Then, the solution for the novel second-order perturbation scheme is determined as

$$\mathbf{S}_r^{l,A} = \text{span} \left\{ \omega_u^{c,A} \mathbf{t}_n^H \overline{\mathbf{G}}^H \left( \overline{\mathbf{G}} \mathbf{P}_r \overline{\mathbf{G}}^H + \sigma_w^2 \mathbf{I} \right)^{-1} \overline{\mathbf{G}} \mathbf{t}_n \right\} \tag{12}$$

A novel second-order perturbation beamforming scheme has a transmit covariance  $\mathbf{P}_r$  and a noise covariance  $\sigma_w^2$  is found, and the transmitting beamformer is denoted by  $\mathbf{t}$ , which is used for the information signal of  $N \times 1$  dimensions. The transmit covariance of the beamformer  $\mathbf{P}_r$  is derived as

$$\mathbf{P}_r = \frac{(1 - \omega_u^{l,A}) \mathbf{Q}}{N - 1} \mathbf{T} \mathbf{T}^H \tag{13}$$

where the power allocation coefficient is  $\omega$ , the power constraints matrix is  $\mathbf{T}$ , and the total power transmitted to  $N$  beams is determined as

$$\mathbf{T} = \begin{bmatrix} \sqrt{t_1} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \sqrt{t_n} \end{bmatrix} \tag{14}$$

where the power transmitted to the  $j$ th beam is  $\sum_{j=1}^N t_j = Q$ , and the novel second-order perturbation power allocation coefficient for the  $N$  beams is calculated as

$$\omega_{\text{pu,su}}^{l,A} = \min \left( \omega_{\text{pu}}^{c,A}, \frac{\sigma_w^2 \psi}{\mathbf{t}_n^H \overline{\mathbf{G}}^H \mathbf{t}_n \overline{\mathbf{G}}} \right) \tag{15}$$

The transmit covariance of the beamformer already restricts the power usage in (13), and the constraints take place according to the  $N$  beams formed during the transmissions in (14).

A novel second-order perturbation power constraints in (15) restricts the power usage of the PU and the SU, and hence, the interference will not be caused in both the users transmission [9]. The power level will be restricted according to the users transmission view point. Thus, the interference user has different data rate and power level to be transmitted, and hence, the real transmission will take place according to the power level in (15) which is oriented to the users resources. Then, the signal

vector  $s_r^{l,A}$  has the solution that is applied at the BS for the better validation of the technique.

### 2.3 Performance Metrics and Analysis

#### 2.3.1 Secrecy Rate of the System and Its Analysis

The secrecy rate [12] of the novel second-order perturbation scheme is determined as

$$Sr_{su}^{l,A} \underline{\text{def}} \log_2 \left( 1 + \psi \omega_{su}^{l,A} |G_{su}^{c,A} s_r^{l,A}|^2 \right) \tag{16}$$

$$Sr_{pu}^{l,A} \underline{\text{def}} \log_2 \left( 1 + \psi \omega_{pu}^{l,A} |G_{pu}^{c,A} s_r^{l,A}|^2 \right) \tag{17}$$

For a cluster, the total secrecy rate is given by

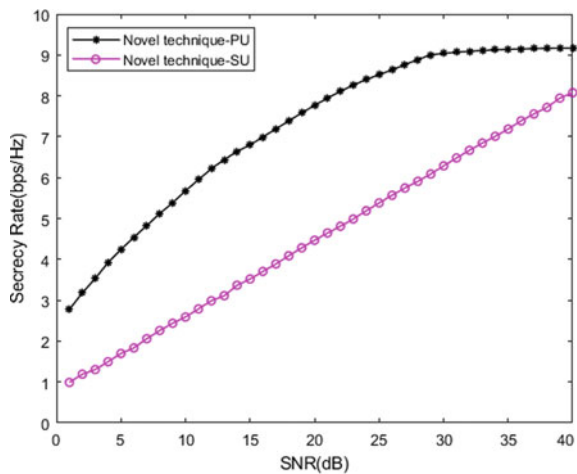
$$Sr_{total}^{l,A} = Sr_{pu}^{l,A} + Sr_{su}^{l,A} \tag{18}$$

Equations (16) and (17) convert the total secrecy rate of the cluster to the total throughput of the cluster.

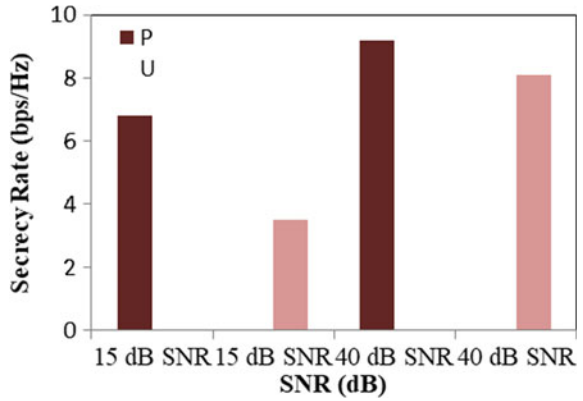
The secrecy rate of the system through the novel second-order perturbation scheme is represented in Fig. 3.

From Fig. 4 it can be observed that the secrecy rate achieved by the novel second-order perturbation technique for the PU and the SU for the 15 dB SNR is 6.8024 and

**Fig. 3** Secrecy rate of the system through the novel second-order perturbation method



**Fig. 4** Secrecy rate analysis of the system through the novel second-order perturbation technique for the PU and the SU for 15 dB and 40 dB SNR



3.5214, respectively. Then, for the values for the 40 dB SNR is 9.1669 and 8.078, respectively.

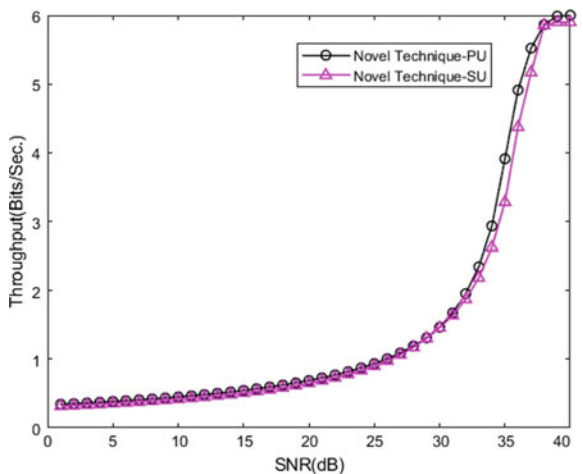
### 2.3.2 Throughput of the System and Its Analysis

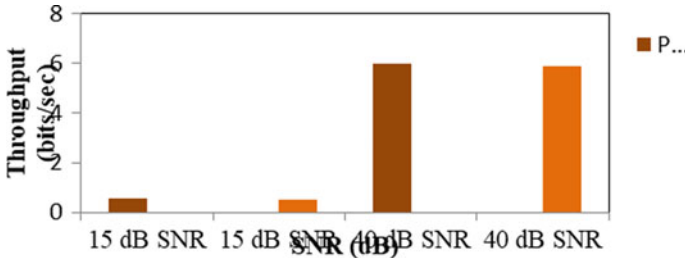
The throughput of the PU and the SU of a channel is calculated as

$$T t_{u(\text{pu,su})} = \psi \omega_u^{l,A} [ |G_u^{l,A} s_r^{l,A}| ]^2 \tag{19}$$

Then, the throughput of the system for the novel second-order perturbation beamforming strategy is represented in Fig. 5.

**Fig. 5** Throughput of the system through the novel second-order perturbation method





**Fig. 6** Throughput analysis of the system through the novel second-order perturbation technique for the PU and the SU for 15 dB and 40 dB SNR

The throughput of the system is increased though the novel second-order perturbation scheme. It can be observed for the 15 and 40 dB SNR and is shown in Fig. 6. The throughput achieved for the 15 dB SNR for the PU and the SU is 0.5468 and 0.5102 and for the 40 dB SNR for the PU and the SU is 6.0 and 5.9.

**2.3.3 BER of the System and Its Analysis**

The BER calculation for the PU and the SU is expressed as

$$BER_{u(pu,su)} = \frac{1}{2} \psi \omega_u^{l,A} Q \left( \frac{2 * E[|s_r^{l,A}|]^2}{w_u^{l,A}} \right) \tag{20}$$

And the Gaussian  $Q$ -function is  $Q(\cdot)$  and  $er = \frac{2 * E[|s_r^{l,A}|]^2}{w_u^{l,A}}$  which is defined as  $Q(er) = \frac{1}{2\pi} \int_{er}^{\infty} e^{-\frac{z^2}{2}} dz$ . The BER of the system through the novel second-order perturbation scheme is shown in Fig. 7.

In Fig. 8, it can be observed that the BER is very much reduced in 15 dB SNR and has the value of 0.001 for the PU and 0.0013 for the SU. For the 40 dB SNR, the value for the PU is 0.0004 and for the SU is 0.0006.

**2.3.4 Capacity of the System and Its Analysis**

The channel capacity of the PU and the SU is increased by this method and is expressed as follows

$$C_{u(pu,su)} = ChB * \psi \omega_u^{l,A} \log_2 \left( 1 + \frac{s_r^{l,A}}{w_u^{l,A} ChB} \right) \tag{21}$$

The channel bandwidth ChB is defined as 1 MHz. The capacity of the system through a novel second-order perturbation scheme is represented in Fig. 9.

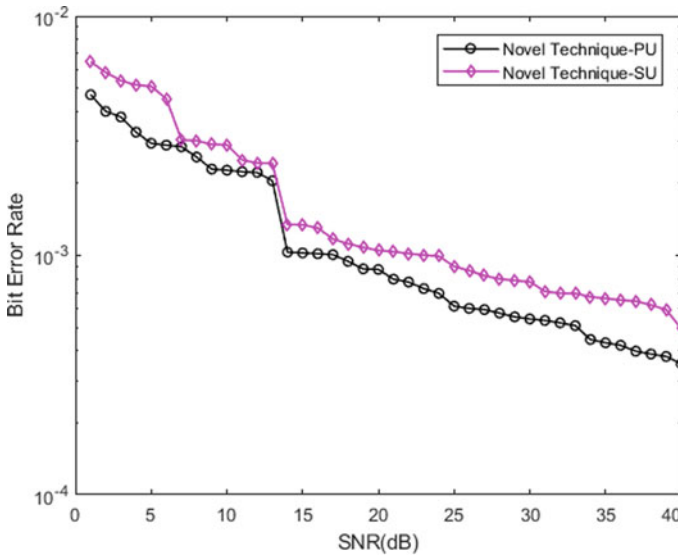
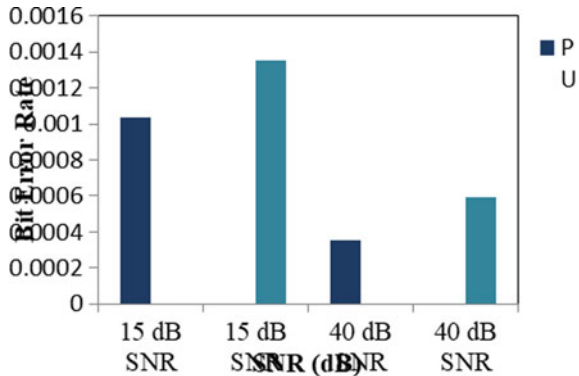


Fig. 7 BER of the system through the novel second-order perturbation technique

Fig. 8 BER analysis of the system through the novel second-order perturbation technique for the PU and the SU for 15 dB and 40 dB SNR

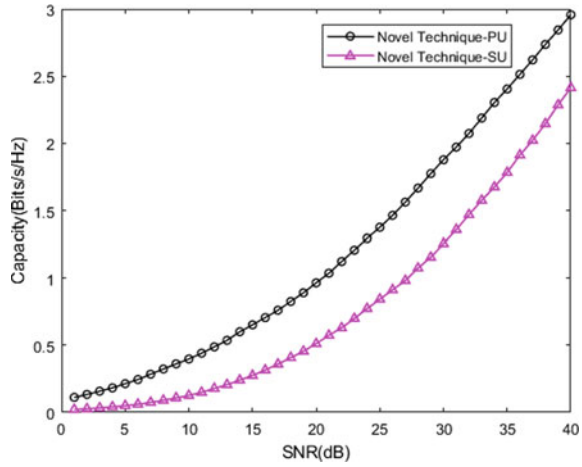


The capacity of the system is increased though the novel second-order perturbation scheme. It can be observed for the 15 dB and 40 dB SNR and is shown in Fig. 10. The capacity achieved for the 15 dB SNR for the PU and the SU is 0.6524 and 0.2753, and the capacity achieved for the 40 dB SNR for the PU and the SU is 2.9578 and 2.4132.

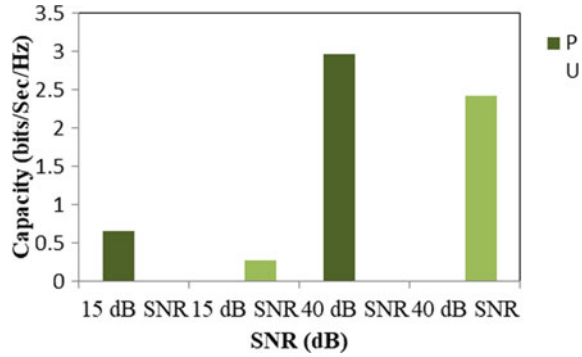
### 2.3.5 SOP of the System and Its Analysis

The SOP for the PU and the SU is expressed as follows:

**Fig. 9** Capacity of the system through the novel second-order perturbation technique



**Fig. 10** Capacity analysis of the system through the novel second-order perturbation technique for the PU and the SU for 15 and 40 dB SNR



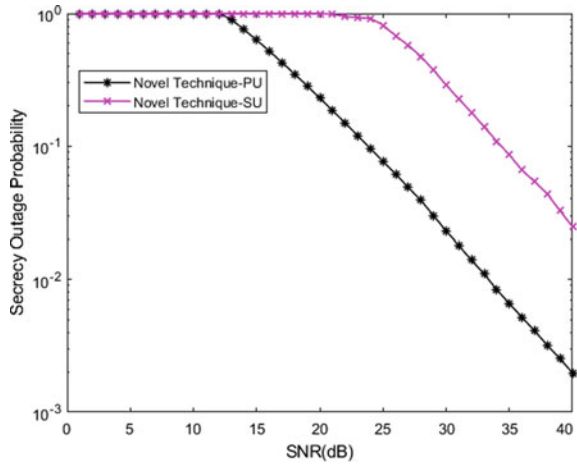
$$SOP_{pu}^{l,A} \stackrel{\text{def}}{=} \Pr \left[ \log_2 \left( 1 + \psi |G_{pu}^{l,A} s_r^{l,A}|^2 \right) \right] \tag{22}$$

$$SOP_{su}^{l,A} \stackrel{\text{def}}{=} \Pr \left[ \log_2 \left( 1 + \psi |G_{su}^{l,A} s_r^{l,A}|^2 \right) \right] \tag{23}$$

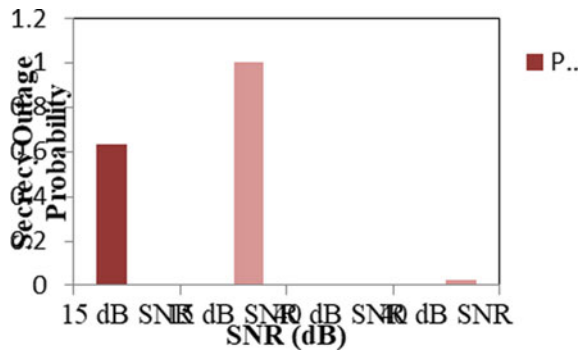
Then, the SOP of the system through the novel second-order perturbation scheme is shown in Fig. 11 respectively.

In Fig. 12 it is seen that the SOP is greatly reduced in PU in 15 dB and 40 dB SNR through the novel second-order perturbation scheme. The decreased values of the PU in 15 dB SNR are 0.6303, and the decreased values of the PU and the SU in 40 dB SNR are 0.002 and 0.0247, respectively.

**Fig. 11** SOP of the system through the novel second-order perturbation technique



**Fig. 12** SOP analysis of the system through the novel second-order perturbation technique for the PU and the SU for 15 and 40 dB SNR



**2.3.6 Ergodic Secrecy Capacity of the System and Its Analysis by Applying the Novel Second-Order Perturbation Technique**

The ergodic secrecy capacity is determined by the mean of the secrecy rate for the fading effect in the channel. This is defined as the assumption in which the precise CSI is not found at the transmitters. The power of every element in  $H_{pu}^{l,c}$  is recognized and follows exponential distribution of the mean  $\sigma_{pu}^{l,c}$ . The beamforming vector  $s^{m,A}$  has a novel second-order perturbation scheme vector as  $s_r^{l,A}$  in [1] and can be written as

$$s_r^{l,A} = I - \frac{\left(G_{pu}^{l,A}\right)^H Y_5^A G_{pu}^{l,A} Y_t^A}{\left|G_{pu}^{l,A}\right|^2} \tag{24}$$

Let the unary matrix be  $H$  and is separated into  $H\psi H$  and the diagonal matrix be  $\psi$ . A non-zero element of  $N(V - 1)$  in diagonal matrix is 1. The distribution of  $G_{\text{pu}}^H$  is same as the  $G_{\text{pu}}$ . In accordance with the projection matrix impotency,

$$\left( I - \frac{\left( G_{\text{pu}}^{l,A} \right)^H Y_5^A G_{\text{pu}}^{l,A} Y_t^A}{\left| G_{\text{pu}}^{l,A} \right|^2} \right)^2 = I - \frac{\left( G_{\text{pu}}^{l,A} \right)^H Y_5^A G_{\text{pu}}^{l,A} Y_t^A}{\left| G_{\text{pu}}^{l,A} \right|^2} \tag{25}$$

Let  $G_{\text{pu}}^{l,A} (s^{m,A})^H s_t^{m,A} \left( G_{\text{pu}}^{l,A} \right)^H = G_{\text{pu}}^{l,A} \psi \left( G_{\text{pu}}^{l,A} \right)^H = z_9$  which follows  $\text{Gamma}\left(N(V - 1), \sigma_{\text{pu}}^{l,c}\right)$ , i.e.,  $\left| G_{\text{pu}}^{l,A} s_t^{l,A} \right|^2 \sim \text{Gamma}\left(N(V - 1), \sigma_{\text{pu}}^{l,c}\right)$ . The value of  $\alpha$  is similar to the number of antennas at the PU and the SU. Then, the ergodic secrecy capacity of the PU and the SU is calculated as follows

$$\begin{aligned} \overline{\text{Er}}_{\text{pu}}^{L,c} &= \frac{1}{\sigma_{\text{pu}}^{l,c}} \int_0^\infty \log_2(1 + \psi \omega z_9) z_9^{\alpha-1} e^{-\frac{z_9}{\sigma_{\text{pu}}^{l,c}}} dz_9 \\ &= \log_2\left(1 + \psi \omega \sigma_{\text{pu}}^{L,c}\right) + (\log_2 e) \left(\frac{\alpha}{\alpha}\right) \end{aligned} \tag{26}$$

$$\begin{aligned} \overline{\text{Er}}_{\text{su}}^{L,c} &= \frac{1}{\sigma_{\text{su}}^{l,c} \alpha} \int_0^\infty \log_2(1 + \psi \omega z_{10}) z_{10}^{\alpha-1} e^{-\frac{z_{10}}{\sigma_{\text{su}}^{l,c}}} dz_{10} \\ &= \log_2(1 + \psi \omega \sigma_{\text{su}}^{l,c}) + (\log_2 e) \left(\frac{\alpha}{\alpha}\right) \end{aligned} \tag{27}$$

where  $z_9 \sim \text{Gamma}\left(\alpha, \sigma_{\text{pu}}^{l,c}\right)$ ,  $z_{10} \sim \text{Gamma}\left(\alpha, \sigma_{\text{su}}^{l,c}\right)$  and  $\alpha = N(S - 1)$ .

$$\alpha = \int_0^\infty z_9^{\alpha-1} e^{-z_9} dz_9 \tag{28}$$

And  $'\alpha = \frac{d\alpha}{d\alpha}$ , where  $'\alpha$  converges uniformly as  $1 < \alpha < \infty$ . Let  $\alpha$  be the justification of differentiation and the derivation of the total ergodic secrecy capacity  $\overline{C}_{\text{CRN}}^{L,c}$  is expressed as

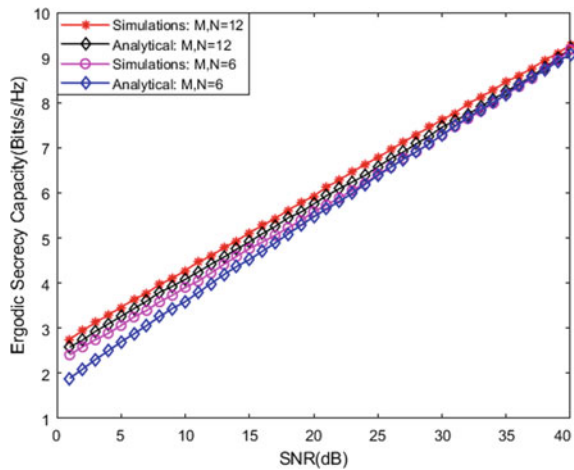
$$\begin{aligned} \overline{\text{Er}}_{\text{CRN}}^{L,c} &= \lambda \int_0^\infty \int_0^\infty \log_2(1 + \psi \omega z_9 + \psi \omega z_{10} + \psi \omega z_9 z_{10}) \\ &\quad \times (z_9 z_{10})^{\alpha-1} e^{-\left(\frac{z_9}{\sigma_{\text{pu}}^{l,c}} + \frac{z_{10}}{\sigma_{\text{su}}^{l,c}}\right)} dz_9 dz_{10} = \log_2\left(1 + \psi \omega \sigma_{\text{pu}}^{l,c}\right) \\ &\quad + \log_2(1 + \psi \omega \sigma_{\text{su}}^{l,c}) + (2 \log_2 e) \left(\frac{\alpha}{\alpha}\right) \end{aligned} \tag{29}$$



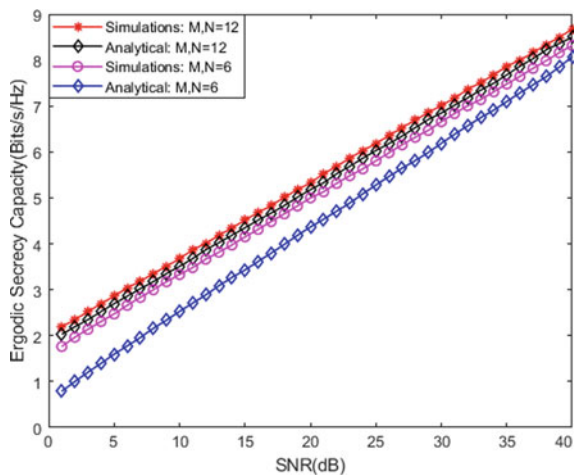
where  $\lambda = \frac{1}{\sigma_{pu}^{I_c} \sigma_{su}^{I_c}} (\alpha)^2$ . The analytical calculations are examined as  $z_9, z_{10} \gg 1$  which involves  $\log_2(1 + z_9) \approx \log_2 z_9$  and  $\log_2(1 + z_{10}) \approx \log_2 z_{10}$ . Then, the simulation and the analytical evaluations of the ergodic secrecy capacity of the system for the PU and the SU through the novel second-order perturbation technique are shown in Figs. 13 and 14, respectively.

In Fig. 15 the simulation and analytical values of the 15 dB SNR for the ergodic secrecy capacity with the 12 user and BS antennas for the PU are 5.1184 and 4.9145 and for the SU is 4.5184 and 4.3514. Then, the simulation and analytical values of the ergodic secrecy capacity for the 6 BS and user antennas with the 15 dB SNR for the PU are 4.7664 and 4.5214. For the SU, the value has 4.1528 and 3.4314. It can be

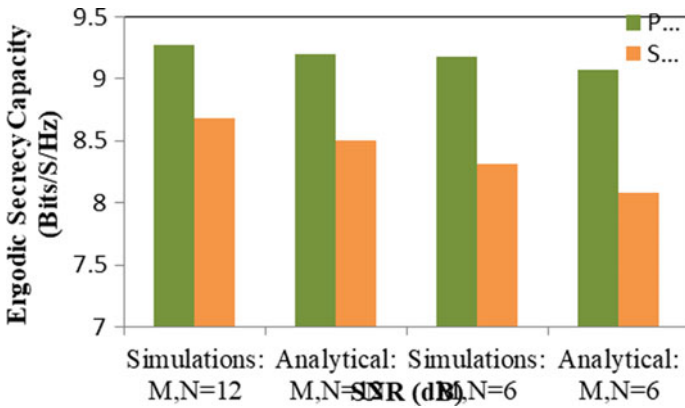
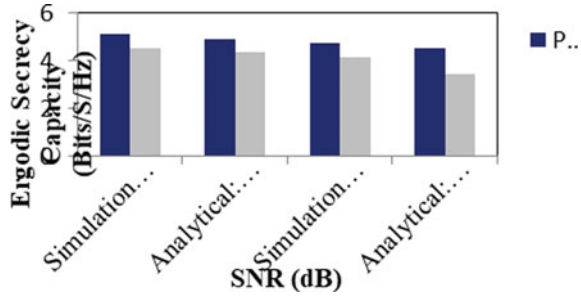
**Fig. 13** Ergodic secrecy capacity of the system for the PU (Primary User) through the novel second-order perturbation technique



**Fig. 14** Ergodic secrecy capacity of the system for the SU (Secondary User) through the novel second-order perturbation technique



**Fig. 15** Ergodic secrecy capacity analysis of the system for the PU and the SU with 15 dB SNR through the novel second-order perturbation technique



**Fig. 16** Ergodic secrecy capacity analysis of the system for the PU and the SU with 40 dB SNR through the novel second-order perturbation technique

seen that the ergodic secrecy capacity of the PU in 15 dB and 40 dB SNR is greater than the ergodic secrecy capacity of the SU.

In Fig. 16, the simulation and analytical values of the 40 dB SNR for the ergodic secrecy capacity with the 12 user and BS antennas for the PU are 9.2782 and 9.2001 and for the SU is 8.6782 and 8.5027. Then, the simulation and analytical values of the ergodic secrecy capacity of the system for the 6 BS and user antennas with the 40 dB SNR for the PU are 9.178 and 9.0734. For the SU, the value has 8.3164 and 8.078. It can be observed that the ergodic secrecy capacity of the PU in both 15 and 40 dB SNR is greater than the ergodic secrecy capacity of the SU.

### 3 Results and Discussion

The simulation parameters of the novel second-order perturbation scheme were found in figures shown above. The performance analysis of the secrecy rate, SOP, BER, throughput, and the channel capacity of the system is obtained by using the simulation

parameters for different SNR rate were analyzed. The performance analysis of the ergodic secrecy capacity is simulated for different SNR rate.

The resulting values of the novel second-order perturbation scheme for the evaluation of the secrecy rate, secrecy outage probability, throughput, capacity, and the BER for the 15 dB SNR and the 40 dB SNR of the PU and the SUs are represented in figures, respectively.

The resulting values of the novel second-order perturbation scheme for the evaluation of the ergodic secrecy capacity for the 15 dB and the 40 dB SNR of the PU and the SUs are shown in figures.

## 4 Conclusion

A novel second-order perturbation-based beamforming approaches are presented to improve the secure transmission of data in the wireless communications with multi-antenna nodes. The proposed technique allocates a novel second-order perturbation transmit power technique to the channel in order to achieve better SNR for the desired user. The remaining power is available as broadcast artificial noise to interrupt the interception of the signal by a passive eavesdropper. The novel second-order perturbation approaches rely on the availability of accurate CSI, and their performance can be quite sensitive to imprecise channel estimates. The simulations results were presented to demonstrate the validity of the analysis and to illustrate the sensitivity of the algorithms that depend on the precise CSI. To reduce the impact of the CSI errors, the robust beamforming schemes are novel second-order perturbation to efficiently recover from the channel estimation errors. The proposed techniques performed well for moderate CSI errors, and the channel mismatch can be eliminated by applying this technique.

## References

1. Higuchi K, Kishiyama Y (2013) Non-orthogonal access with random beamforming and intra-beam SIC for cellular MIMO downlink. In: Proceedings of the IEEE 78th vehicular technology conference, pp 1–5
2. Jiang M, Li Y, Zhang Q, Li Q, Qin J (2017) Secure beamforming in downlink MIMO nonorthogonal multiple access networks. *IEEE Signal Process Lett* 24(12):1852–1856
3. Peng M, Wang C, Li J, Xiang H, Lau V (2015) (Second quarter), Recent advances in underlay heterogeneous networks: Interference control, resource allocation, and self-organization. *IEEE Commun Surv Tutor* 17(2):700–729
4. Sun Q, Han S, Chin-Lin I, Pan Z (2015) On the ergodic capacity of MIMO NOMA systems. *IEEE Wirel Commun Lett* 4(4):405–408
5. Tian M, Zhang Q, Zhao S, Li Q, Qin J (2017) Secrecy sum rate optimization for downlink MIMO nonorthogonal multiple access systems. *IEEE Signal Process Lett* 24(8):1113–1117

6. Saito Y, Benjebbour A, Kishiyama Y, Nakamura T (2013b) System-level performance evaluation of downlink non-orthogonal multiple access (NOMA). In: Proceedings of the IEEE 24th annual symposium on personal, indoor and mobile radio communications, pp 611–615
7. Haijun Z, Ning Y, Keping L, Miao P, George KK, Victor CML (2018) Secure communications in NOMA system: subcarrier assignment and power allocation. *IEEE J Sel Areas Commun* 36(7):1441–1452
8. Shahira HA, Ahmed HM, Yasmin AF, Samy SS (2019) Maximizing the number of users in clustered MIMO NOMA systems under rate constraints. *Mob Netw Appl* 24(9):618–629
9. Chen J, Yang L, Alouini MS (2018) Physical layer security for cooperative NOMA systems. *IEEE Trans Veh Technol* 67(5):4645–4649
10. Ding Z, Poor HV (2016) Design of massive-MIMO-NOMA with limited feedback. *IEEE Signal Process Lett* 23(5):629–633
11. Haykin S (2005) Cognitive radio: brain-empowered wireless communications. *IEEE J Sel Areas Commun* 23(2):201–220
12. Zhang Y, Wang HM, Yang Q, Ding Z (2016) Secrecy sum rate maximization in non-orthogonal multiple access. *IEEE Commun Lett* 20(5):930–933

# A High Merit Factor Preamble Sequence for Better Synchronization in WiMAX



Suma Sekhar, Sakuntala S. Pillai, and S. Santhoshkumar

## 1 Introduction

As the impact of data-intensive mobile applications and services is increasing, the requirements for reliable high-rate communication are also shooting up. Physical layer specifications of WiMAX rely on orthogonal frequency division multiplexing (OFDM) considering its ability for high-bandwidth data transfer without being much affected by frequency selective fading [1, 2]. To improve the communication rate and quality of data, OFDM is combined with multiple input multiple output (MIMO) systems, where multiple antennas are employed at both the transmitter and receiver [3–5]. Mobile WiMAX uses MIMO OFDM to avail reliable high-rate data transfer with enhanced spectral efficiency [6]. But when propagated through an impaired channel with severe distortions, the advantages of OFDM can be maintained only if the OFDM subcarriers are orthogonal. The synchronization errors in the form of carrier frequency offset (CFO) and time offset (TO) may pull down the orthogonality between the subcarriers and introduce intersymbol interference (ISI) and intercarrier interference (ICI) [7–9]. As the number of antennas increases, the deleterious effects produced by synchronization errors become more severe in MIMO OFDM. So, the accurate estimation and correction of the frequency and time offsets are imperative

---

S. Sekhar (✉)

Department of Electronics and Communication Engineering, LBS Institute of Technology for Women, Poojappura, Thiruvananthapuram, Kerala, India  
e-mail: [sumasekhar@lbsitw.ac.in](mailto:sumasekhar@lbsitw.ac.in)

S. S. Pillai

Department of Electronics and Communication Engineering, Mar Baselios College of Engineering and Technology, Thiruvananthapuram, Kerala, India

S. Santhoshkumar

Department of Electronics and Communication Engineering, College of Engineering, Trivandrum, Thiruvananthapuram, Kerala, India

in MIMO OFDM-based communication systems. Like any other MIMO OFDM-based communication system, the realization of precise synchronization algorithms is one of the major challenges in WiMAX. As the OFDM system can put up with offsets of the order of subcarrier spacing only, the precise acquisition of symbol timing and CFO is crucial to avoid performance degradation [10, 11]. Even though different algorithms are available, data-aided synchronization method using a standard preamble sequence embedded in the OFDM header is followed in WiMAX standards to track down the synchronization errors. Preamble-based synchronization facilitates faster and more reliable estimation of frequency and timing offsets compared to blind synchronization methods.

The effectiveness of data-aided synchronization schemes in MIMO OFDM depends on the quality of sequences embedded in the preamble structure of the OFDM frame. Simulation results reveal that the proposed new preamble based on the constant amplitude zero autocorrelation (CAZAC) sequence attains improved accuracy in the time estimation procedure.

The rest of the paper is arranged as follows. In Sect. 2, the mathematical models of transmitted and received baseband OFDM signals are presented. Section 3 introduces the system model used for estimating frequency and time offsets from the preamble sequence. In Sect. 4, the parameters that are highly desired for the preamble sequence are discussed. Section 5 is devoted to a new preamble based on the CAZAC sequence. To demonstrate the superior performance of the proposed preamble, corresponding results and analysis are provided in Sect. 6, and finally, the paper is concluded in Sect. 7.

## 2 Signal Model for Transmitted and Received OFDM

In the WiMAX transmitter, the input data bits are initially randomized to improve coding performance, channel encoded, reordered by an interleaver, and mapped to constellation points through digital modulation schemes like quadrature amplitude modulation or quadrature phase-shift keying. The complex data symbols are further modulated on  $N$  parallel and orthogonal subcarriers using inverse fast Fourier transform (IFFT) operation, and the samples of the resulting OFDM symbol can be mathematically modeled as

$$d_l[n] = \frac{1}{N} \sum_{k=0}^{N-1} D_l(k) e^{j2\pi kn/N} \quad (1)$$

Here,  $D_l(k)$ ,  $l$ ,  $k$ , and  $N$  represent the complex data symbol, OFDM symbol index, subcarrier index, and the number of IFFT points.

At the receiver, OFDM demodulation is easily implemented through a fast Fourier transform (FFT) processor. A received baseband signal in the presence of carrier frequency offset  $\beta$  and timing offset  $\lambda$  can be expressed as

$$r_l[n] = \frac{1}{N} \sum_{k=0}^{N-1} H_l(k) D_l(k) e^{j2\pi \frac{(k+\beta)(n+\lambda)}{N}} \tag{2}$$

The timing offset causes the FFT window to be improperly placed as delayed or before time to the actual zero index, resulting in incorrect demodulation of the OFDM signal.

### 3 System Model for Estimating Frequency and Time Offsets

A preamble incorporated with selected sequences in a well-designed structure is prefixed to the data frame before transmission. The frequency and timing offsets corrupting the transmitted signal can be measured by making use of the preamble sequences extracted from the OFDM frame as shown in Fig. 1. The fundamental processing performed in the synchronization unit of the receiver is the correlation operation between either two identical repeated sequences or received preamble sequence and its reference copy set aside the synchronization unit.

The OFDM signal collected by the receiver is modeled by Eq. 2. The signal corrupted by time and frequency offsets is processed by the receiver analog front end, analog to digital converter, and offset error compensator before extracting the transmitted data. The preamble part excerpted from the OFDM frame is processed through correlation operation to estimate offset errors.

The standard structure of the WiMAX preamble [12] can be mathematically modeled as

$$P_W = [C \ P_{WS} \ P_{WS} \ P_{WS} \ P_{WS} \ C \ P_{WL} \ P_{WL}] \tag{3}$$

In Eq. 3,  $C$  represents cyclic prefix and  $P_{WS}$  and  $P_{WL}$  represent a short sequence with 64 samples and a long sequence with 128 samples that can be generated from

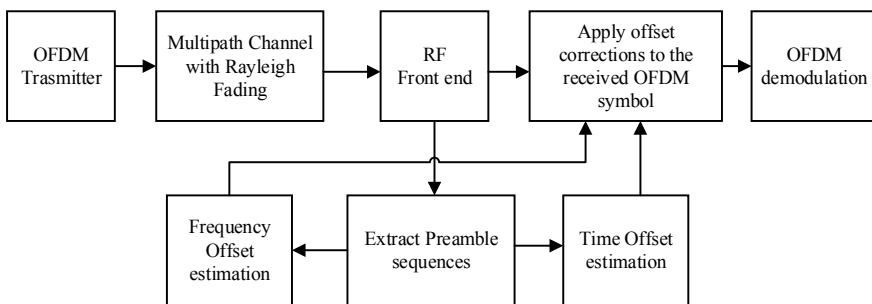


Fig. 1 System model for frequency and time offset estimation

the sequence  $P_W$  as shown in Eqs. 4, 5, and 6.

$$P_W[i] = \begin{cases} \pm 1 & \pm j, i = -100, \dots - 1, 1, \dots + 100 \\ 0, & \text{otherwise} \end{cases} \quad (4)$$

$$P_{WS}[i] = \begin{cases} 2P_W^*(i), & i \bmod 4 = 0 \\ 0, & i \bmod 4 \neq 0 \end{cases} \quad (5)$$

$$P_{WL}[i] = \begin{cases} \sqrt{2}P_W^*(i), & i \bmod 2 = 0 \\ 0, & i \bmod 2 \neq 0 \end{cases} \quad (6)$$

The timing synchronization errors can be quantitatively measured by correlating the preamble sequence extracted from the received OFDM frame with the replica copy stock up on the receiver. The number of samples by which the correlation peak is shifted from zero index provides an indication of time offset which can be directly computed by identifying the index corresponding to shifted peak.

## 4 Desired Parameters for Preamble Sequences

The accurate detection of offset errors demands sequences with excellent correlation properties and low peak to average power ratio (PAPR). If the sequence used in a preamble provides peaks at zero shift of autocorrelation and zero values for the off-peak autocorrelations, then the preamble is said to have excellent correlation properties. So, the quality of autocorrelation property of a sequence can be evaluated using the parameters [13] peak sidelobe level (PSL), integrated sidelobe level (ISL), and merit factor (MF).

Aperiodic autocorrelation function of the sequence  $b = [b_1, b_2 \dots b_N]^T$  can be represented by Eq. 7 as

$$A[k] = \sum_{n=1}^{N-k} b_n b_{n+k}^* \quad (7)$$

The parameters PSL, ISL, and MF are defined by Eqs. 8, 9, and 10 as

$$\text{PSL} = \max\{A[k]\}_{k=1}^{N-1} \quad (8)$$

$$\text{ISL} = \sum_{k=1}^{M-1} |A[k]|^2 \quad (9)$$



$$\text{MF} = \frac{|A[0]|^2}{\sum_{\substack{k = -(N-1) \\ k \neq 0}}^{N-1} |A[k]|^2} \quad (10)$$

PAPR of the sequence  $b[n]$  is defined by Eq. 11 as

$$\text{PAPR} = \frac{\max[b_n]^2}{E[b_n]^2} \quad (11)$$

Sequences with low PAPR (ideally unity or zero dB) help to get rid of nonlinear distortions from power amplifiers, zero autocorrelation property enhances correlation properties, and diminished sidelobes improve estimation accuracy [14, 15]. Careful selection and design of mathematical sequence inserted in the preamble structure are essential in the design of synchronization circuits in the receiver block.

## 5 CAZAC-Based Preamble

A CAZAC sequence is a complex sequence with excellent correlation properties and constant amplitude. A CAZAC sequence of odd length  $M$  can be defined by Eq. 12 as

$$\text{CZ}(m) = e^{-\frac{j\pi m(m+1)}{M}}, \quad m = 1, \dots, M \quad (12)$$

Its defining property of providing zero autocorrelation with its shifted versions makes it attractive in time localization applications and, hence, in time synchronization. The constant amplitude of the CAZAC sequence assures transmission efficiency. As the autocorrelation sidelobes are considerably low in the CAZAC sequence, it holds the desirable properties of low ISL value and high MF compared to other commonly used sequences for synchronization. Considering the attractive values for the desired parameters, a new CAZAC-based preamble is proposed for WiMAX by redefining Eq. 4 by Eq. 12. After replacing  $P_w[i]$  by the CAZAC sequence, the short and long CAZAC sequences  $P_{CS}$  and  $P_{LS}$  are generated as per Eqs. 5 and 6.

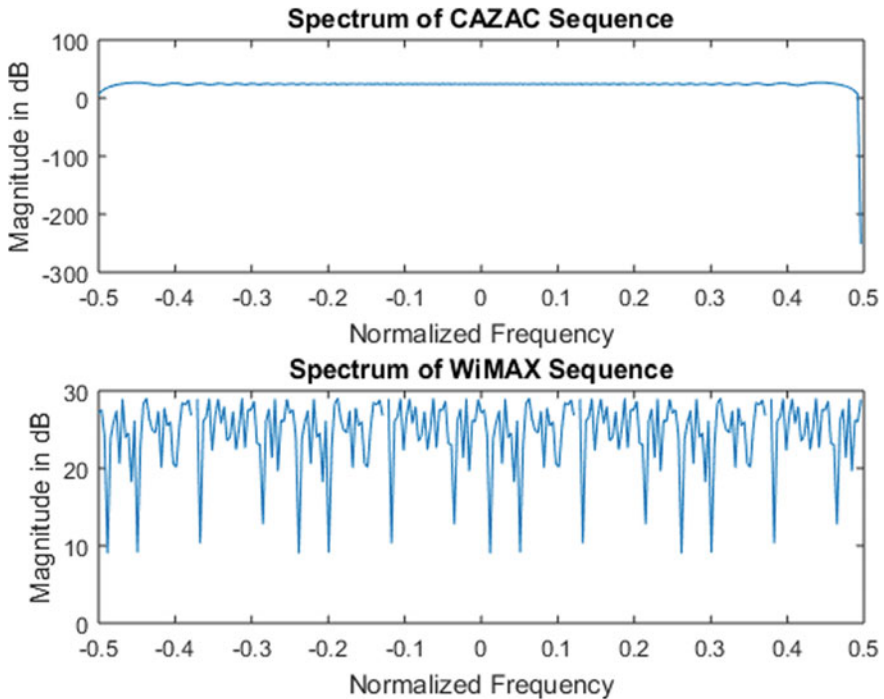
## 6 Results

As the sequences with reduced autocorrelation sidelobes improve synchronization accuracy, the values of the parameters PSL, ISL, MF, and PAPR of different sequences of length 256 are compared through simulations and summarized in Table 1. As the CAZAC sequence gives better values for the parameters PSL, ISL, PAPR, and MF, it is further compared with the standard WiMAX sequence and is verified that it

**Table 1** Comparison of CAZAC sequence with other synchronization sequences based on the parameters PAPR, PSL, ISL, and MF

Sequence	Parameters			
	PAPR	PSL	ISL	MF
PN sequence	2.44	142	12.31e+05	0.0532
Gold	4.13	87	22.94e+04	0.2857
Kasami	3.55	85	38.52e+04	0.1701
Golay	0	33	21.7e+03	3.01
CAZAC	9.64e-16	7.69	25.97e+02	25.24
WiMAX sequence	7.09	74	8.53e+05	0.3072

possesses superior parameter values than WiMAX preamble. The frequency spectrum of both the WiMAX and CAZAC preambles is plotted and given in Fig. 2. The results show that CAZAC sequence is giving a spectrum with high flatness than WiMAX preamble. The spectrum with constant amplitude enhances the transmission efficiency.

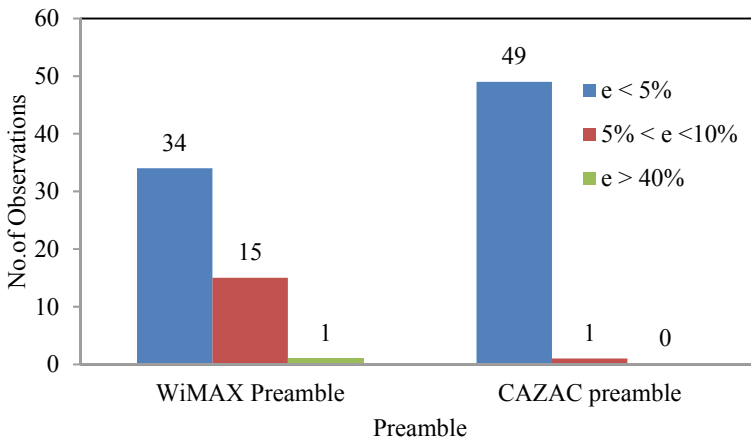


**Fig. 2** Comparison of spectrums of WiMAX and CAZAC preambles

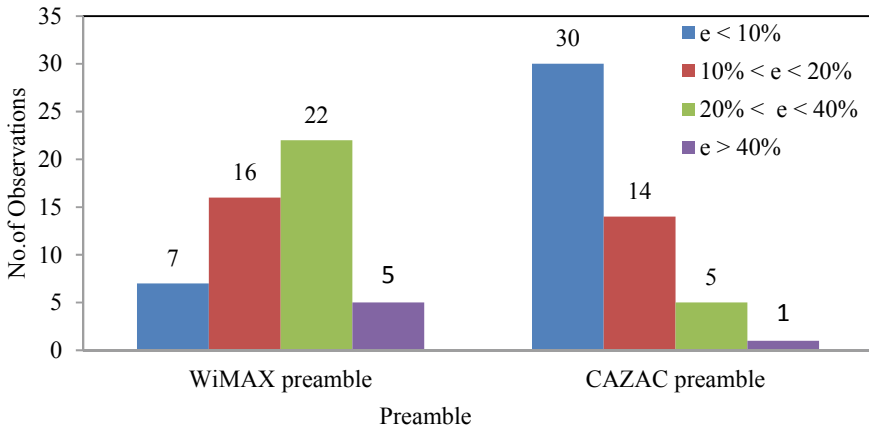
To avoid time synchronization errors, the FFT window must be correctly placed over the received signal. The time offset error is mathematically modeled as the number of samples by which the received signal samples are shifted. The meticulous detection and compensation of this delay before demodulating the signal is very important. To compare the performance in delay estimation, two different preambles composed of CAZAC and standard WiMAX sequences are simulated and transmitted over a multipath Rayleigh faded channel with three paths under both the medium- and severe-distorted conditions. The synchronization unit in the receiver computes the correlation between the preamble extracted from the received OFDM frame and a local copy of the same sequence stored at the receiver. The time offset is obtained from the sample delay which is estimated from the index of the sample with largest correlation magnitude.

When the delays are estimated from a medium-distorted channel using a standard WiMAX preamble, the error percentage was less than 5% for 34 observations, between 5 and 10% for 15 observations, and greater than 40% for one observation as shown in Fig. 3. When a preamble based on the CAZAC sequence was used, the error percentage was less than 5% for 49 observations and between 5 and 10% for only one observation out of 50 trials as shown in Fig. 3. The experiment was repeated with a severely distorted channel for 50 observations using both sequences and comparison shown in Fig. 4.

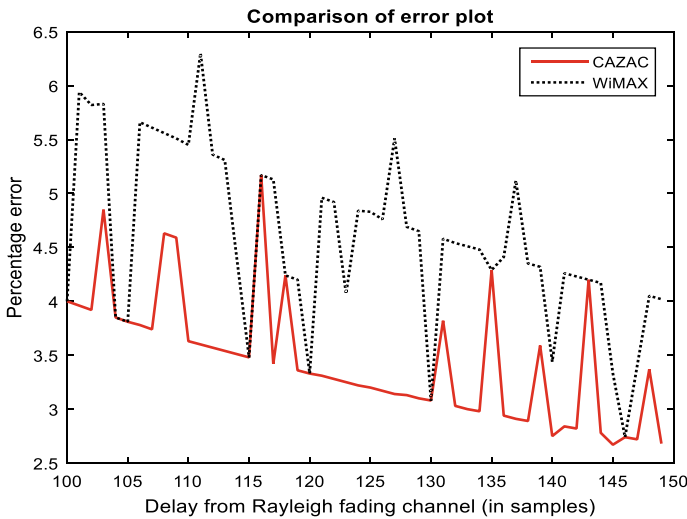
It was noted that corresponding to a WiMAX sequence, the estimation error was less than 10% for 7 observations, between 10 and 20% for 16 observations, between 20 and 40% for 22 observations, and greater than 40% for 5 observations. When a CAZAC sequence is used for delay estimation, the error percentage was less than 10% for 30 observations, between 10 and 20% for 14 observations, between 20 and 40% for 5 observations, and greater than 40% for 1 observation. The percentage error is plotted corresponding to delays of 100–150 samples in Fig. 5 for some medium-



**Fig. 3** Comparison of error percentage between WiMAX and CAZAC preambles in a medium-distorted channel



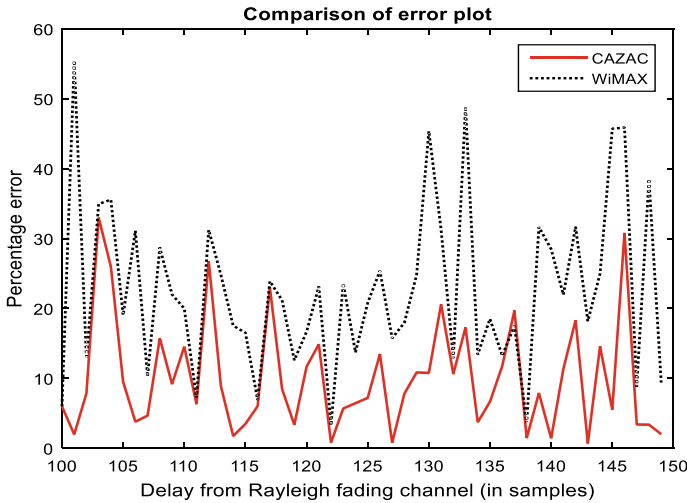
**Fig. 4** Comparison of error percentage between WiMAX and CAZAC preambles in a severe-distorted channel



**Fig. 5** Comparison of percentage error corresponding to delays of 100–150 samples in a medium-distorted channel

distorted channels and in Fig. 6 for a severely distorted channel. It is clear that error is less as well as error percentage is more stable when a CAZAC-based preamble is used.

The frequency offset is then estimated by computing the angle of correlation between two identical long preamble sequences. Frequency offsets estimated using WiMAX and CAZAC sequences are mathematically modeled by Eqs. 13 and 14.  $L_{WL}$  and  $L_{CL}$  represent the length of WiMAX and CAZAC-based preambles which are taken as 128.



**Fig. 6** Comparison of percentage error corresponding to delays of 100–150 samples in a severe-distorted channel

$$\widehat{\beta}_W = \angle \left\{ \frac{\sum_{i=0}^{L_{WL}-1} P_{WL}(i) P_{WL}(i + L_{WL})}{2\pi} \right\} \tag{13}$$

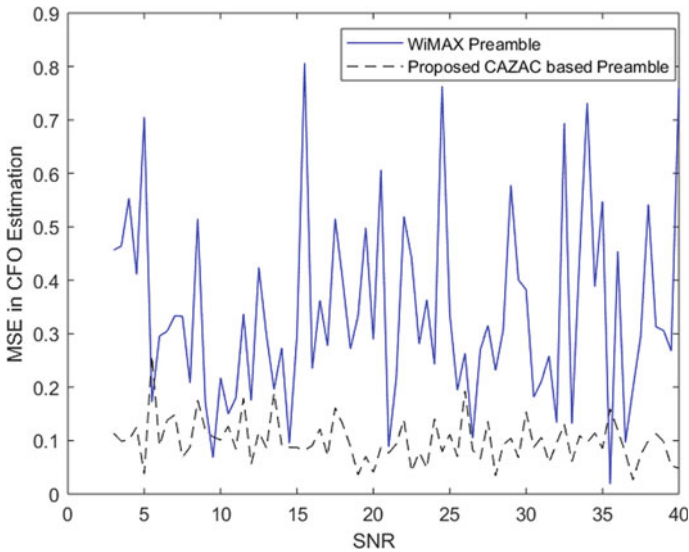
$$\widehat{\beta}_C = \angle \left\{ \frac{\sum_{i=0}^{L_{CL}-1} P_{CL}(i) P_{CL}(i + L_{CL})}{2\pi} \right\} \tag{14}$$

The mean square error (MSE) in FO estimation using both preambles is plotted in Fig. 7.

As the results show the reduced errors for CAZAC-based preamble in frequency offset estimation, it can be considered as a better option for improving synchronization in WiMAX.

## 7 Conclusion

WiMAX adopted the OFDM-based physical layer considering its high-data rate transmission efficiency in a frequency selective fading channel. But the time synchronization errors result in the sampling of the OFDM symbols at incorrect instants and frequency synchronization errors introduce phase distortions in the received signal. Unless detected meticulously and compensated properly, these errors give rise to ISI, ICI, and degradation in system performance. A new CAZAC-based preamble with reduced integrated sidelobe levels and improved MF is proposed. The results of simulated experiments show that both the time and frequency synchronization errors



**Fig. 7** Comparison of MSE in frequency offset estimation using both the WiMAX and CAZAC sequences

can be more accurately detected using the new preamble when compared with the existing standard WiMAX preambles.

## References

1. Chang RW (1966) Synthesis of band-limited orthogonal signals for multichannel data transmission. *Bell Syst Tech J* 45:1775–1796
2. Saltzberg BR (1967) Performance of an efficient parallel data transmission system. *IEEE Trans Commun Technol COM-15*:805–811
3. Foschini GJ, Gans MJ (1998) On limits of wireless communications in a fading environment when using multiple antennas. *Wireless Pers Commun* 6
4. Chizhik D, Foschini G, Gans M, Valenzuela R (2002) Keyholes, correlations and capacities of multi element transmit and receive antennas. *IEEE Trans Wireless Commun* 1:361–368
5. Goldsmith A, Jafar SA, Jindal N, Vishwanath S (2003) Capacity limits of MIMO channels. *IEEE J Sel Areas Commun* 21(5):684–702
6. Li YG, Winters JH, Sollenberger NR (2002) MIMO-OFDM for wireless communications: signal detection with enhanced channel estimation. *IEEE Trans Commun Technol* 50:1471–1477
7. Cimini L Jr (1985) Analysis and simulation of a digital mobile channel using orthogonal frequency division multiplexing. *IEEE Trans Commun Technol* 33:400–411
8. Pollet T, Van Bladel M, Moeneclaey M (1995) BER sensitivity of OFDM systems to carrier frequency offset and Wiener phase noise. *IEEE Trans Commun* 43:191–193
9. Speth M, Fechtel SA, Fock G, Meyr H (1999) Optimum receiver design for wireless broad-band systems using OFDM—part I. *IEEE Trans Commun* 47(11):1668–1677
10. Awoseyila AB, Kasparis C, Evans BG (2009) Robust time-domain timing and frequency synchronization for OFDM systems. *IEEE Trans Consum Electron* 55(2):391–399

11. Mostofi Y, Cox DC (2006) Mathematical analysis of the impact of timing synchronization errors on the performance of an OFDM system. *IEEE Trans Commun* 54(2):226–230
12. Granado J, Torralba A, Chávez J, Oria C, Baena-Lecuyer V (2009) Frequency offset estimation in WIMAX receivers. *IEEE Trans Consum Electron* 55(3)
13. Soltanian M, Stoica P (2012) Computational design of sequences with good correlation properties. *IEEE Trans Signal Process* 60:2180–2193
14. Chung CD, Chen WC (2018) Preamble sequence design for spectral compactness and initial synchronization in OFDM. *IEEE Trans Veh Technol* 67(2)
15. Tang L, Zhu Y, Fu Q (2018) Fast algorithm for designing periodic/apperiodic sequences with good correlation and stop band properties. *EURASIP J Adv Signal Process* 1:1–13

# Deep Learning Technique-Based Pulmonary Embolism (PE) Diagnosis



S. Vijayachitra, K. Prabhu, M. Abarana, A. Deepa, and L. Loga Priya

## 1 Introduction

Pulmonary embolism (PE) can be considered as a health issue that influences nearly 700,000 patients per annum. It is actually an impasse in one of the pulmonary arteries in lungs. In many persons, it is identified as a blood clot that travels to the lungs by deep veins in various parts of human body. Diagnosis of PE can be complicated to do, and for which, more number of imaging techniques have been done to assist in this process. PE can be identified in 2% of contrast-enhanced CT scans taken for various reasons other than for the detection of PE.

In most of the cases, the root cause for the pulmonary embolism is due to artery blockage happened in the lungs. The main reason for such artery blockage is blood clot/lump that developed in a deep vein in legs, and it further travels to the lungs, and lastly, there will be blockage in a smaller lung artery. All the blood clots are the reason for pulmonary embolism, which are further introduced in the veins of deep leg.

A pulmonary embolism will be the reason for a lack of blood flow that will further leads to lung tissue damage. It can provide low blood oxygen levels that can be harmful for the other organs in the body also. As per the medical practitioner's statement after analysis of the causes and reasons, PE can be a blood clot that will induce very serious life-threatening problems, and even death can be the final.

The major symptoms of pulmonary embolism include the following:

- Severe chest pain
- Pain while breathing
- Heartbeat increase

---

S. Vijayachitra (✉) · K. Prabhu · M. Abarana · A. Deepa · L. Loga Priya  
Department of Electronics and Instrumentation Engineering, Kongu Engineering College,  
Perundurai, Tamil Nadu 638060, India  
e-mail: [dr.svijayachitra@gmail.com](mailto:dr.svijayachitra@gmail.com)



- Fast and rapid breathing and continuous cough.

To perform the diagnosis and management of PE, it is relied on the imaging techniques. Currently, computed tomography pulmonary angiography is the preferred system to provide an accurate diagnosis quickly. It is well known that CT can also provide certain information on causes of acute chest pain felt by the person [1]. Chest radiograph (CR) is a technique with limited utility which will provide various findings about the PE. Another methodology called ventilation-perfusion (VP) scan provides ventilation-perfusion mismatches in the patients and demonstrates different classification categories ranging from normal to high levels. CT scanning is more preferable method which will provide X-rays to develop cross-sectional representations of testing person, and in particular, CTPA, also said to be a study of pulmonary embolism, will develop three-dimensional pictures that can show various defects represented as PE inside the deep arteries in lungs.

## 2 Deep Learning Technique

Deep learning technique is considered as a branch of an artificial intelligence function that mimics the behavior of brain in terms of handing out the data to utilize for detecting different matters, languages translation, speech recognition, and also for decision-making purposes. With the proposed deep learning, it is convenient to identify/learn without the need of any supervision by human, representing from information that is of any kind like unstructured or unlabeled type.

In deep learning technique, a convolutional neural network is preferably employed for analyzing any image. These networks can take in any type of input image; based on the importance, it is assigned the learnable weights and biases in the image. It is possible to distinguish different images based on the weights/biases added [2, 3].

Convolutional neural network is said to be deep network since it is based on the number of layers it contains. In general, in any artificial neural network (ANN), various signals travel between nodes, and weights have been assigned to the corresponding nodes. A heavier weighted node will have more effect on the next layer of nodes than the nodes which have less weight. The last layer of the architecture will provide an output by compiling the weighted input nodes.

Requirement of a suitable hardware for a deep learning system is essential because it needs to process large amounts of data which will further involve various mathematical calculations. The structure of the network is somewhat similar to that of human brain, particularly the way in which the neurons have connected in it. In our brain system, every neuron can react to stimuli only in certain specific field of reception, and a collection of that fields can overlies in order to envelop the overall area of visualization.

### 3 Pulmonary Embolism Detection

A pulmonary embolism will be the reason for a lack of blood flow that will further leads to lung tissue damage [4]. It can provide low blood oxygen levels that can be harmful for the other organs in the body also. The symptoms of pulmonary embolism will be different based on involvement of lungs, previous heart/lung disease, and the size of the clots diagnosed from the patients.

CT scan images obtained from the patients have to be trained with the deep learning neural network, and as per the desired conditions given to the network, various decisions such as LEFT-sided PE, RIGHT-sided PE, and normal (without any PE) will be made. Depending on the number of images given for training the neural network, the perfect decisions can be obtained [5].

Here, for obtaining the images from patients, CT pulmonary angiography is preferred. CT scanning is more preferable method which will provide X-rays to develop cross-sectional representations of testing person, and in particular, CTPA, also said to be a study of pulmonary embolism, will develop three-dimensional pictures that can show various defects represented as PE inside the deep arteries in lungs.

From which, it is observed that the data have been combined from several X-rays to produce a detailed image of structures inside the body. Normally, CT scans can able to produce 2D images of various sections of the body, and the data can also be used to construct 3D images as a result. CT scans consist of hundreds of images in the format of DCM (DICOM images) that needs a detailed review to identify the position of clots within the pulmonary arteries.

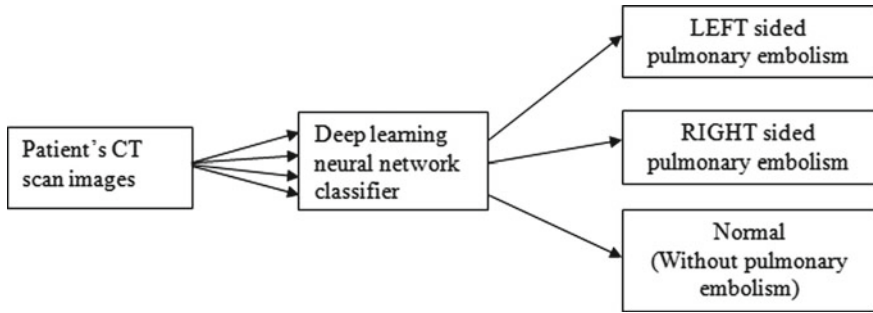
In CTA images, the dark spots have been identified as regions of PE among the bright regions of blood arteries. The radiologist will do injection of the contrast material, and then, the developed images can be recorded in with respect to time period. During this time, both the PE regions and veins will have same kind of gray levels in the obtained image, and then, it will be differentiated from the vein with respect to its high contrast. Hence, a computer-aided system is essential to do effective diagnosis of pulmonary embolism.

The complete block diagram for the purpose of diagnosis of PE using deep learning is shown in Fig. 1.

After gaining the advantage of using this technology, there is progression in the utilization of this kind of assistance for PE diagnosis.

Based on the number of images, the convolutional neural network will be effectively trained well. The DICOM medical file of a single patient consists of multiple numbers of images, which are of high resolution due to which the file size can be quite large. These files further need large storage for processing of images. To simplify the view of any file of DICOM type, it is easy for the user to know the entire details of the image. When DICOM files are read and processed by a user-generated program wise, it is compulsory to know various parts of the standard.

As per the classification stated in Fig. 1, the various categories identified from the testing images given to the deep neural network, the decisions for different kinds are



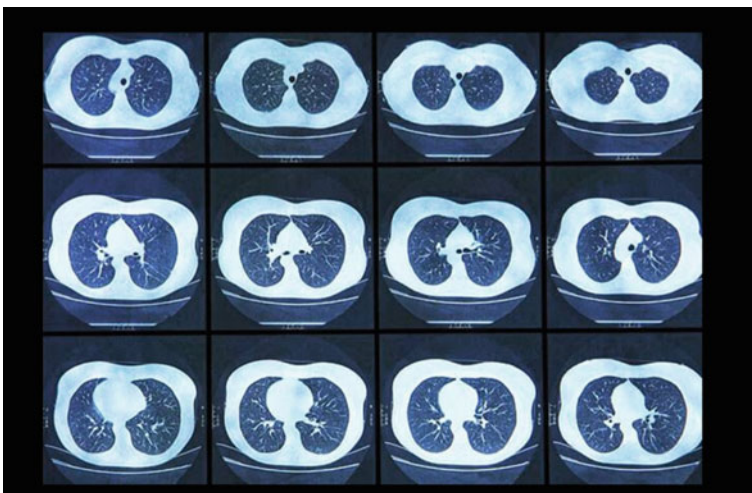
**Fig. 1** Block diagram of the diagnosis of PE using deep learning technique

made [6]. Figures 2, 3 and 4 show the samples obtained from various patients which are further used for training the neural network.

If the person is not affected with pulmonary embolism, it refers to the normal (NO pulmonary embolism) (Fig. 2).

If the left side of the heart is unable to pump sufficient amount of blood, the blood pressure of that person will be dropped. It is termed as LEFT-sided PE. This will further lead to death in a short period after the pulmonary embolism occurs (Fig. 3).

The blockage found in the right side of the lungs or heart that has moved from deep veins in the legs to other parts of the body through the bloodstream. It is then called as RIGHT-sided PE (Fig. 4).



**Fig. 2** CT images obtained from Patient-1 and diagnosed as normal

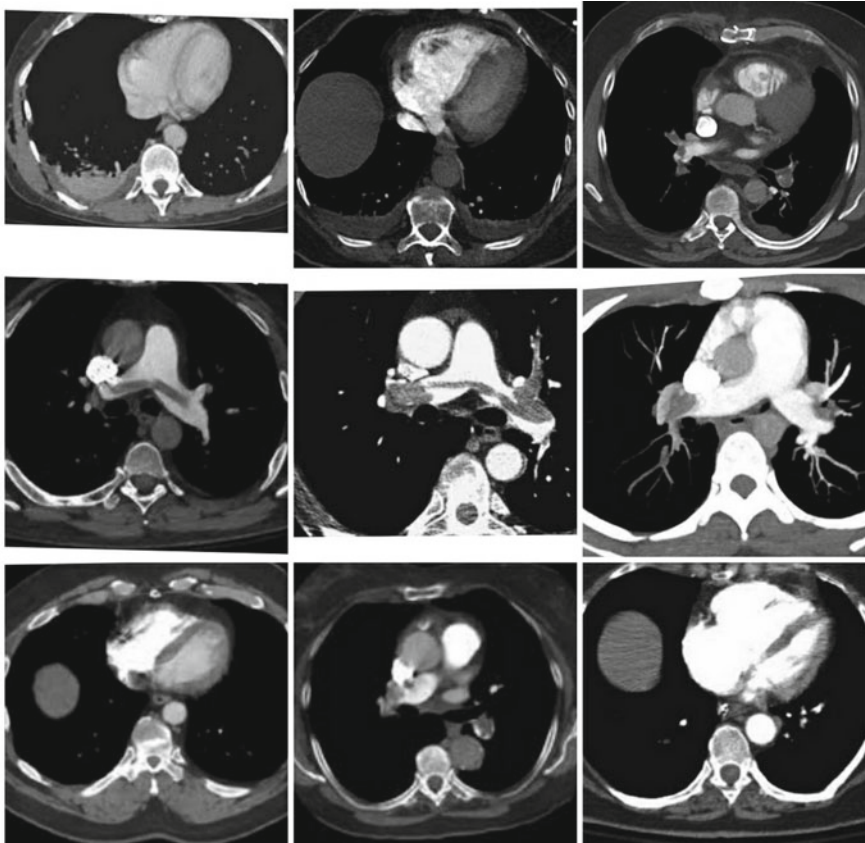
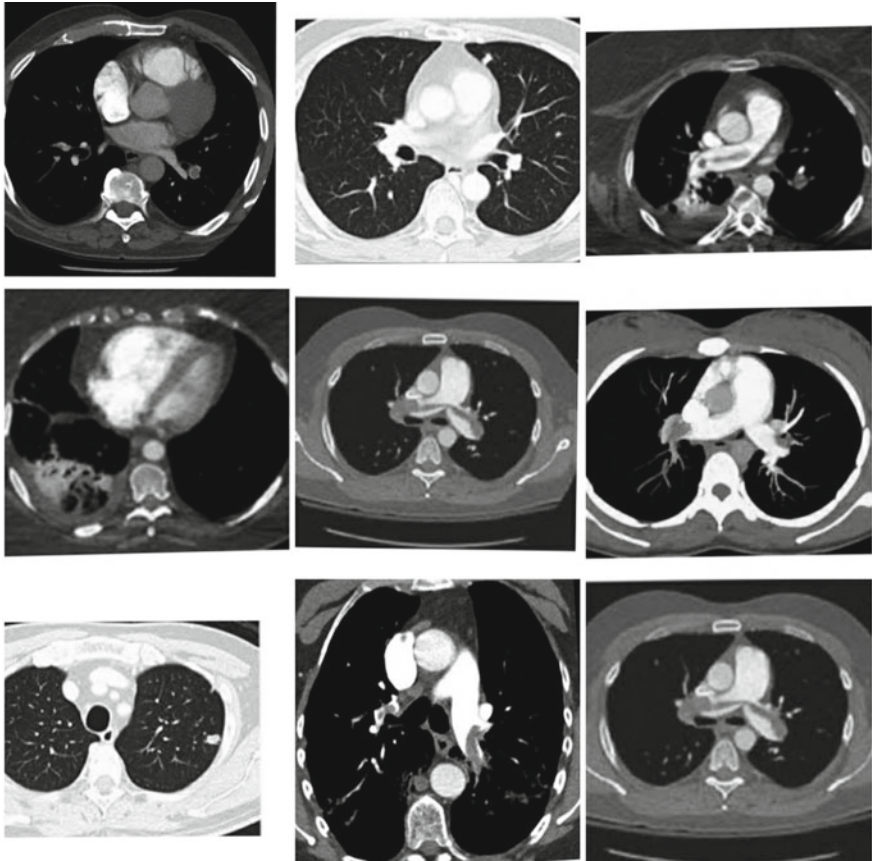


Fig. 3 CT images obtained from Patient-2 and diagnosed as LEFT-sided PE

#### 4 Results of Pulmonary Embolism Diagnosis

This model gets trained by feeding both input image files as DICOM images and output as its specified category as normal, LEFT-sided PE, and RIGHT-sided PE. By the application of deep learning neural network and training the network with more number of CT scan images obtained from more number of persons and fed as different slices, the neural network can provide best learning [7]. From this advantage, it is further tested the neural network with testing slices obtained from various persons, the appropriate classification decision was obtained.

If it is given the DICOM images to the neural network model, the network will process the input and gives the output in the format for better human understanding. As a sample of results, Fig. 5 shows the decision made as LEFT-sided PE, and Fig. 6 shows the decision made as RIGHT-sided PE.



**Fig. 4** CT images obtained from Patient-3 and diagnosed as RIGHT-sided PE

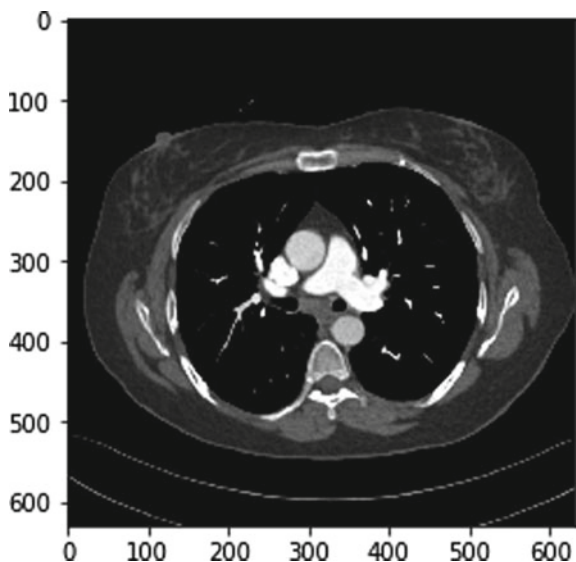
Table 1 shows the results of classification of PE diagnosis in terms of normal, LEFT-sided PE, and RIGHT-sided PE.

By this technique, it can be assured that the medical practitioner, or even the lab technicians, can easily identify the results of PE and from which the patients can get their treatment faster. It reduces the time delay for diagnosing the disease and gives effective classification.

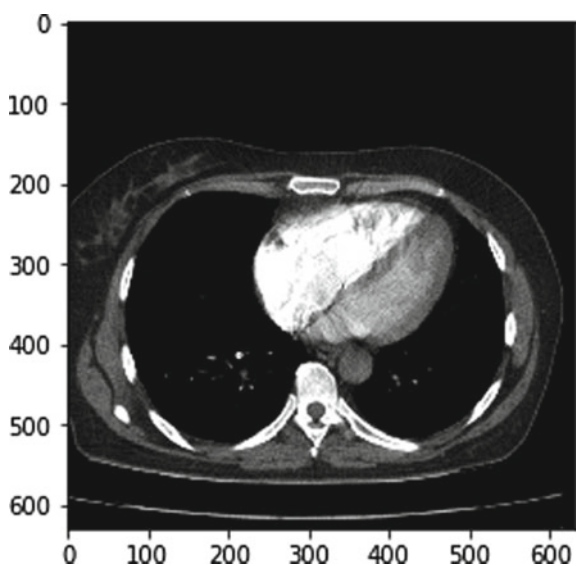
## 5 Conclusion

There is a deep learning neural network-based PE diagnosis was observed effectively. By the application of this method, it is classified the pulmonary embolism as normal, LEFT-sided PE, and RIGHT-sided PE from various input test images fed to the

**Fig. 5** LEFT-sided PE



**Fig. 6** RIGHT-sided PE



**Table 1** Classification results of PE diagnosis

Patient Id	No. of patients' images trained	No. of normal (NO PE)	No. of LEFT-sided images classified	No. of RIGHT-sided images classified
75d23269adbd	15	0	12	3
291c07d4a7c0	15	1	10	4
396cdd37b9bf	18	1	14	3

neural network. From the results, it is assured that by implementing this technique, the medical practitioner, even the lab technicians, can obtain faster and accurate results from pulmonary embolism CT scans. It further reduces the human errors and time consumption for diagnosing the PE. Faster diagnosis of PE will further helps to start the treatment faster and reduces the death rate happening due to delay in diagnosing.

## References

1. Kuriakose J, Patel S (2019) Acute pulmonary embolism. *Radiol Clin North Am* 48(1):31–50
2. Yang X (2019) A two-stage convolutional neural network for pulmonary embolism detection from CTPA images. *IEEE Access*
3. Huang S-C, Kothari T, Banerjee I (2020) PENet—a scalable deep-learning model for automated diagnosis of pulmonary embolism using volumetric CT imaging. *npj Digit Med*
4. Liu W, Liu M, Guo XJ (2020) Evaluation of acute pulmonary embolism and clot burden on CTPA with deep learning. *Eur Radiol* 30:3567–3575
5. Tao Q, Yan W, Wang Y (2019) Deep learning-based method for fully automatic quantification of left ventricle function from cine MR images, a multivendor, multicenter study. *Radiology* 290:81–88
6. Al Aref SJ, Anchouche K, Singh G (2019) Clinical applications of machine learning in cardiovascular disease and its relevance to cardiac imaging. *Eur Heart J* 40(24):1975–1986
7. Banerjee I, Sofela M, Yang J (2019) Development and performance of the pulmonary embolism result forecast model (PERFORM) for computed tomography clinical decision support

# Analysis of Daubechies 2 Wavelet in WPM System for Adhoc Network



M. B. Chakole and S. S. Dorle

## 1 Introduction

The introduction of wavelet is not very old. It has gained the popularity within the short duration of time and wavelet theory developed by leaps and bound amongst the researchers. Researchers are attracted to carry the research in wavelets looking towards its improved features and improved characteristics as compared to other multi-carrier modulation (MCM) techniques [1]. Nowadays, a large number of researchers are using the wavelet theory concept in the field of wireless communication to get the improved performance in the system. Also, several books and research publications related to wavelet theory are easily available for researchers for carrying out their research work. Major feature of the wavelet transform is its working capability in time domain and frequency domain and provides both the time–frequency information of the signal simultaneously [2]. In MCM, data stream is divided to obtain the bit streams. Bit streams in MCM can be modulated with the use of multiple narrow bandwidth carriers. The orthogonality of the signal is maintained in MCM by adjusting the signal pulses in the frequency and time plane. For any of the application, there are chances of losing the orthogonality of the signal during the travel through the wireless channel. This is due to intersymbol (ISI) and interchannel (ICI) interferences which disturb the signal at receiver [3].

Wavelet packet modulation (WPM) is a multi-carrier modulation technique based on wavelets. The greatest motivation to follow WPM system is the flexibility

---

M. B. Chakole (✉)

Department of Electronics and Telecommunication Engineering, Yeshwantrao Chavan College of Engineering, Nagpur, India  
e-mail: [mdole@gmail.com](mailto:mdole@gmail.com)

S. S. Dorle

Department of Electronics Engineering, G H Raisoni College of Engineering, Nagpur, India  
e-mail: [sanjay.dorle@raisoni.net](mailto:sanjay.dorle@raisoni.net)



and modulation functionality. The WPM system consists of compactly supported wavelets which is an advantage for designing decomposition and reconstruction filters used for the implementation of filter banks. During the simulation of wavelet packet decomposition phase, the given signal is divided into two components, i.e. the low-frequency component and high-frequency components. In software design for simulation by fixing the frequency resolution at desired level packet decomposition can be done. By using Daubechies (dB2) wavelet, design for two band filter banks implementation is achieved. By making use of filter banks, adhoc network can be established for smooth communication [4].

Adhoc network is consisting of the mobile and immobile nodes that are generally referred as the self-organizing networks. They are called as self-organizing networks because the topology of the adhoc network is dynamic in nature [5, 6]. Self-organizing networks also have the feature of self-recovering in many of the applications. Any of the adhoc network application requires the fast communication of data amongst the nodes in an autonomous mode. Autonomous nodes of the adhoc networks may have the feature of flexibility within the network so as to implement in large number of applications like defence activity and big farms/gardens. At the same time, threats may arise due to the flexibility. Resolving the threat due to flexibility secure routing needs to implement within an adhoc network.

In an adhoc network, routing is the emergent issue. It deals with the problems like mobility of the nodes and limited bandwidth. In this research work, an attempt is successfully made to improve the WPM system for efficient routing amongst nodes of the adhoc network. Further out of large number of available wavelets, suitable wavelet is identified for the WPM system. The proposed WPM system is experimented under various channels (AWGN, Rayleigh and Rician) using different modulation schemes (BPSK, QPSK and 16-PSK) related to the required adhoc network parameters like energy consumed, BER and throughput.

## 2 Related Work

Wavelet packet modulation (WPM) with the advantage of flexibility and high bandwidth efficiency is a perfect multi-carrier modulation technique. This advantage is useful in the wavelets for the next generation wireless communication systems. Gao, Li et al. in their research work explained the WIMAX system-based WPM system for optimizing the bit error rate (BER). BER performance is improved by applying the WPM technique in WIMAX system [7]. But the performance in BER results due to the consumption of high bandwidth.

In one of the research works, authors compared the performance of WPSE/WPM systems with systems having fast Fourier transform (FFT) and orthogonal frequency division multiplexing (OFDM). The results of their research indicate improved performance of WPSE/WPM system as compared to traditionally used FFT/OFDM system for the bit-error-rate (BER) [8] parameter.

Researchers also studied the distortions induced in the channel during the MCM. For the same, researchers proposed a time-domain equalizer for suppressing the distortions generated due to the MCM technique. Time domain equalizers are designed with the help of wavelet filters, and disadvantage is that for getting the improved performance of equalizer within the system, large number of equalizer taps to be implemented which ultimately results in more number of filters [9].

Researchers proposed a new algorithm used in fast wavelet packet transform to save computational period. The designed algorithm improves the demodulation and modulation of signal by increasing the used frequency band. As a result, the ISI available within the channel gets reduced and the performance of the system is improved. Here the system performance is slightly improved by increasing the bandwidth [10].

### 3 Wavelet Packet Modulation

WPM having orthogonal wavelet packet bases is derived from a multi-resolution analysis. Multiresolution analysis is the technique of transmitting the data within the channel to generate parallel data sub-streams by the division of the input signal. Within the channel, the signal is modulated and multiplexed at different carrier frequencies. At the beginning of the process, a pair of quadrature mirror filter (QMF) is considered to derive the orthogonal bases. QMF consists of a half-band high-pass and low-pass filter (impulse responses  $p[m]$  and  $q[m]$ , respectively) of length  $l$ . These filters share a tight relationship given as [11]:

$$q[l - 1 - n] = (-1)^n p[n] \quad (1)$$

The adjoints or duals of Eq. 1 are the complex conjugate time variants given by Eq. 2: These complex conjugate time variants are reversed in nature.

$$\begin{aligned} p'[n] &= p * [-n] \\ q'[n] &= q * [-n] \end{aligned} \quad (2)$$

The pair  $\{p'[n], q'[n]\}$  is the analysis filter pair for generating the modulation of data at the transmitter side through wavelet packet carriers. The opposite side of transmitter is the receiver side where the reverse procedure is followed as compared to transmitter side. In receiver side, the synthesis filter pair derives the wavelet packet carriers exact dual for demodulation process.

$$P * (w + \pi)P'(w) + Q * (w + \pi)Q'(w) = 0 \quad (3)$$

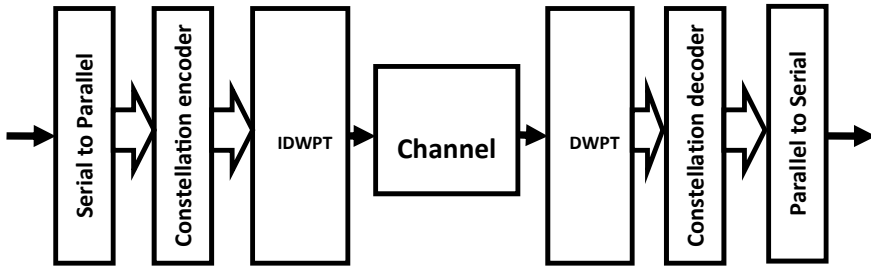


Fig. 1 WPM system functional block diagram

where  $P(w)$ ,  $Q(w)$ ,  $P'(w)$  and  $Q'(w)$  denote the magnitude of the four filters mentioned in the above equations in the frequency domain. represents the perfect reconstruction condition.

The quadrature modulated frequency filters as shown in Eqs. 1 and 2 develop wavelet packet bases  $\{\gamma_l^k\}$ . The final version of WPM modulation signal  $x[n]$  is generated due to the combination of the wavelet packet bases  $\{\gamma_l^k\}$  weighted with complex data symbols  $a_{u,k}$  of various parallel streams  $k$  and data index  $l$  is represented by Eq. 4:

$$x(n) = \sum_u \sum_{k=0}^{M-1} a_{u,k} \gamma_l^k(n - uM) \tag{4}$$

The synthesis filter pair  $\{p[n], q[n]\}$  at the receiver demodulates the data through a process mentioned above equations. Figure 1 shows the functional block diagram of WPM system.

The inverse discrete wavelet packet transform (IDWPT) is available on the transmitter side. It is designed with the synthesis tree by dividing the single signal into various parallel streams. The discrete wavelet packet transform (DWPT) available at receiver side divides the signal received from the transmitter with the help of wavelet packet analysis tree.

### 4 Simulation Results

Adhoc network is designed in the MATLAB platform for the proposed WPM system. Table 1 shows the proposed simulation parameters of the system. During the simulation, the distance amongst the adjacent neighbouring nodes of adhoc network is assumed as 100 m. For evaluating, the major performance parameters like energy consumption, BER and throughput of the adhoc network Daubechies (db2) wavelet are used. To measure the performance of the system, most widely used modulation

**Table 1** Proposed parameters for WPM system

Modulation	BPSK, QPSK, 16-PSK
Sub-carriers	64 bits
Code length	1000 bits
Number of levels	4
Distance amongst neighbouring nodes	100 m
Channels	AWGN, Rayleigh, Rician
Wavelet	Daubechies (db2)
SNR range	20–40 dB

schemes (BPSK, QPSK and 16-PSK) were considered over various channel conditions (AWGN, Rayleigh and Rician). Sub-carriers 64 bits have been considered for the modulation scheme, and the adhoc network was iterated for the code length of 1000 bits.

### **Working of Daubechies (db2) Wavelet for Adhoc Network**

The proposed WPM system for adhoc network is efficient if Daubechies (db2) wavelet is used under different channel conditions. The working of db2 wavelet for each of the modulation scheme was verified.

#### ***4.1 BPSK Modulation Using Daubechies (Db2) Wavelet***

The working of Daubechies (db2) wavelet within BPSK modulation scheme for an adhoc network parameters BER, throughput and energy consumed in three channel conditions (AWGN, Rayleigh and Rician) is shown, respectively, in Figs. 2, 3 and 4.

From Figs. 2, 3 and 4, it is clear that, as compared to Rayleigh and Rician channel condition, under AWGN channel condition the working of Daubechies (db2) wavelet in BPSK modulation scheme is good.

#### ***4.2 Working of Daubechies (Db2) Wavelet for QPSK Modulation***

The working of Daubechies (db2) wavelet within QPSK modulation scheme for an adhoc network parameters BER, throughput and energy consumed in three channel conditions (AWGN, Rayleigh and Rician) is shown, respectively, in Figs. 5, 6 and 7.

From Figs. 5, 6 and 7, it is clear that, as compared to Rayleigh and Rician channel condition, under AWGN channel condition the working of Daubechies (db2) wavelet in QPSK modulation scheme is good.

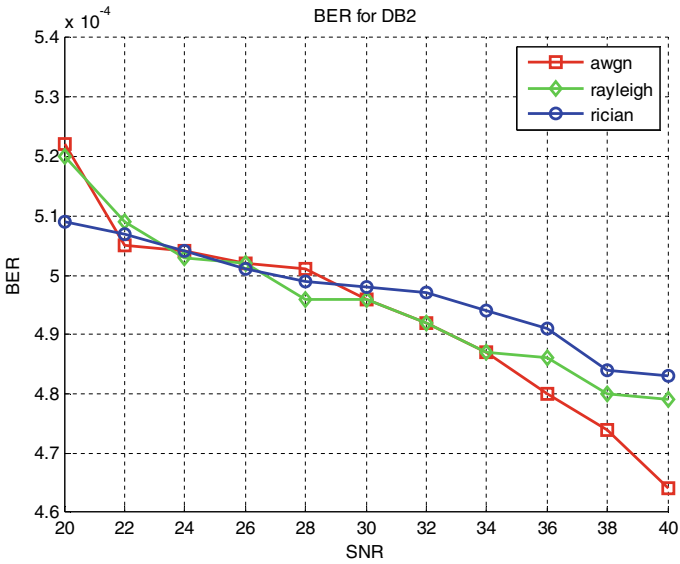


Fig. 2 Plot of db2 for BER in BPSK modulation scheme

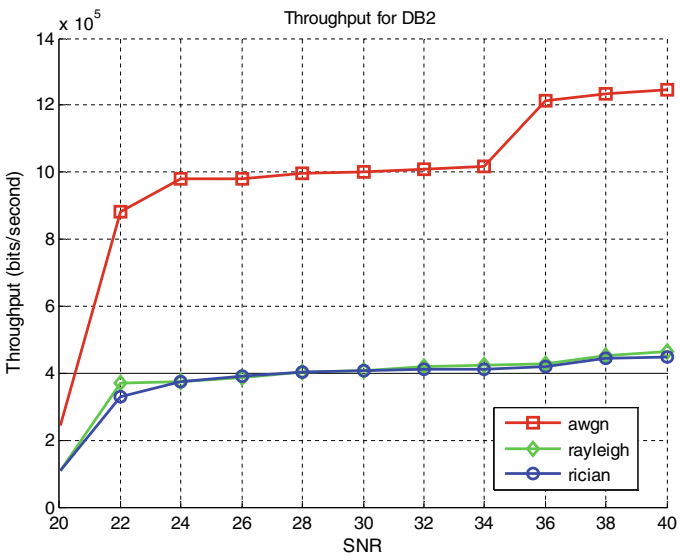


Fig. 3 Plot of db2 for throughput in BPSK modulation scheme

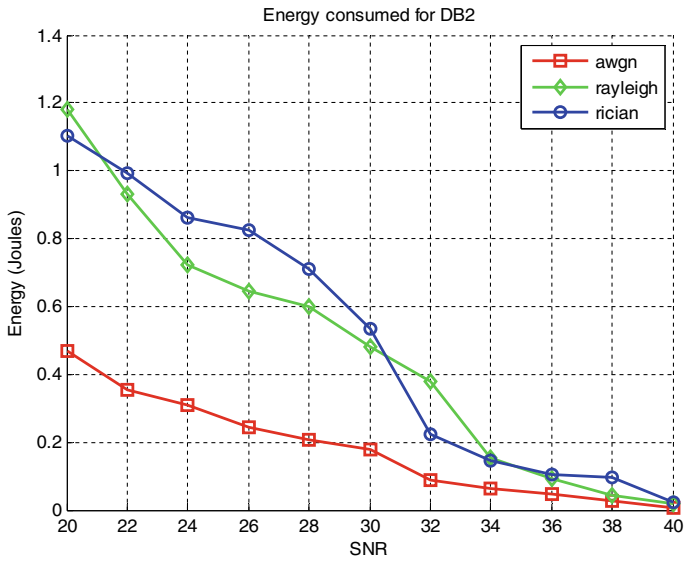


Fig. 4 Plot of db2 for energy consumed in BPSK modulation scheme

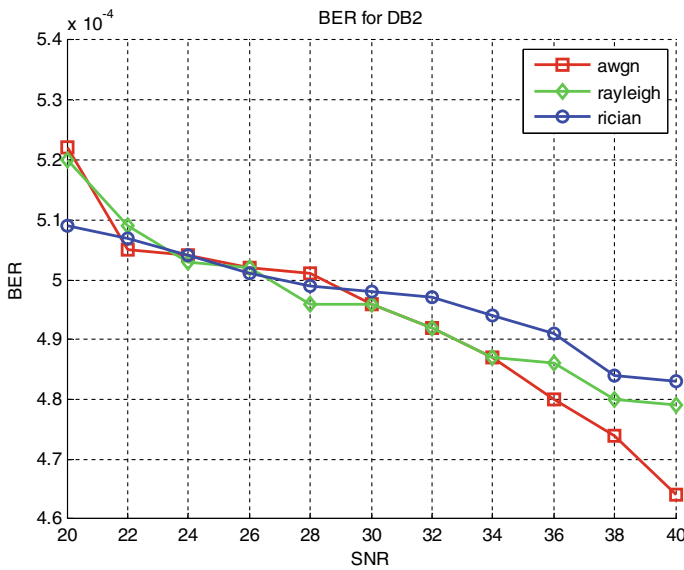


Fig. 5 Plot of db2 for BER in QPSK modulation scheme

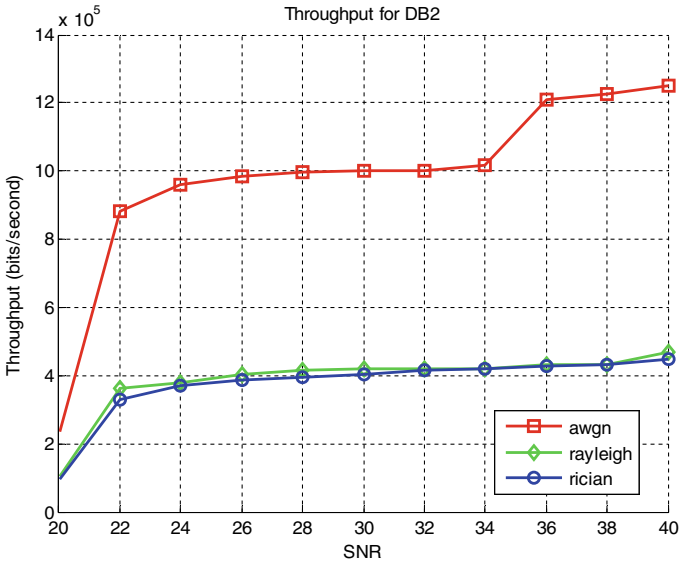


Fig. 6 Plot of db2 for throughput in QPSK modulation scheme

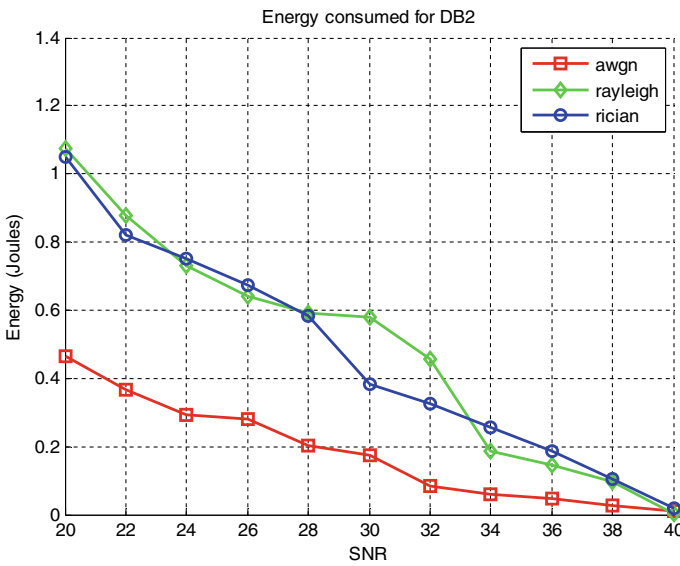


Fig. 7 Plot of db2 for energy consumed in QPSK modulation scheme

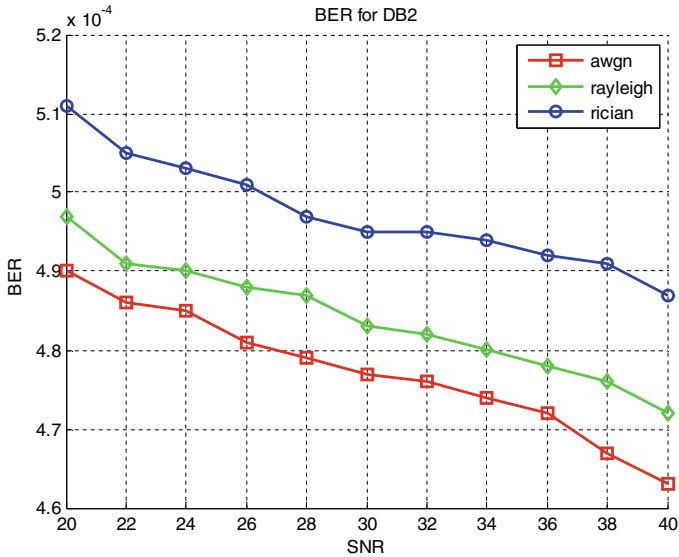


Fig. 8 Plot of db2 for BER in 16-PSK modulation scheme

### 4.3 Working of Daubechies (Db2) Wavelet for 16-PSK Modulation

The working of Daubechies (db2) wavelet within 16-PSK modulation scheme for an adhoc network parameters BER, throughput and energy consumed in three channel conditions (AWGN, Rayleigh and Rician) is shown, respectively, in Figs. 8, 9 and 10.

From Figs. 8, 9 and 10 it is clear that, as compared to Rayleigh and Rician channel condition, under AWGN channel condition the working of Daubechies (db2) wavelet in BPSK modulation scheme is good.

From Figs. 2, 3, 4, 5, 6, 7, 8, 9 and 10, the discussions clearly indicates the working of Daubechies (db2) wavelet is satisfactory. Amongst three channel conditions the working of db2 wavelet is found good in AWGN channel (Tables 2 and 3).

After comparison of all three channel conditions (Rayleigh, AWGN and Rician), under AWGN channel condition, it was observed that db2 almost work equally suitable for all the modulation schemes. But the average performance of db2 under AWGN channel condition for QPSK modulation scheme is good as compared to the BPSK and 16-PSK modulation schemes.



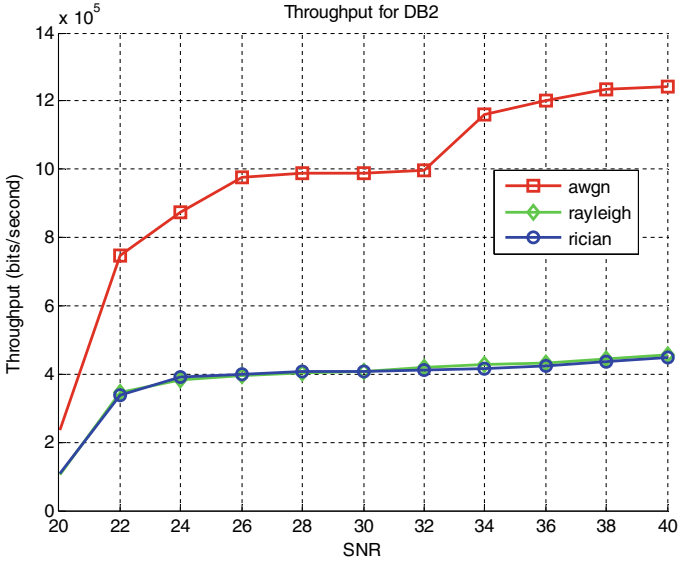


Fig. 9 Plot of db2 for throughput in 16-PSK modulation scheme

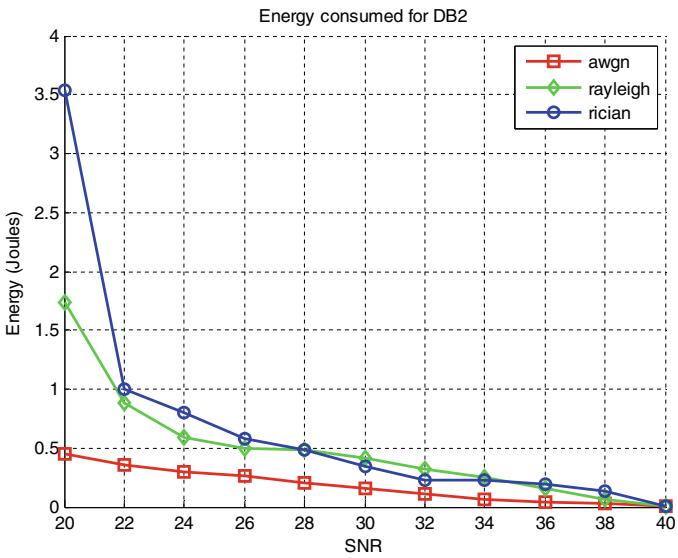


Fig. 10 Plot of db2 for energy consumed in 16-PSK modulation scheme

**Table 2** Analysis of db2 wavelet in WPM system using three different modulation schemes for adhoc network

Modulation schemes	Channels	Parameters		
		BER	Throughput (bits/s)	Energy (J)
BPSK	AWGN	0.000493	1,074,230	0.18433
	Rayleigh	0.000496	418,138	0.4771
	Rician	0.000497	411,777	0.5104
QPSK	AWGN	0.000493	1,075,506	0.18232
	Rayleigh	0.000496	420,750	0.48932
	Rician	0.000505	410,838	0.46855
16-PSK	AWGN	0.000495	1,073,633	0.18524
	Rayleigh	0.000503	418,279	0.49121
	Rician	0.000504	414,639	0.58649

**Table 3** Evaluation of db2 wavelet in WPM system using three different modulation schemes over AWGN channel for adhoc network

Parameters	SNR range	Average value BPSK	Average value QPSK	Average value 16-PSK
BER	20–40 dB	0.000493	0.000493	0.000495
Throughput (bits/s)		1,074,230	1,075,506	1,073,633
Energy (J)		0.18433	0.18232	0.18524

## 5 Conclusion and Future Scope

The WPM system is tested for the use in adhoc network. For adhoc network, three important parameters of the network like BER, throughput and energy consumed were analysed using WPM system. By using Daubechies (db2) wavelet with different channel conditions like Rayleigh, AWGN and Rician channels separately in separate three different modulation schemes BPSK, QPSK and 16-PSK, efficient WPM system for adhoc network will be obtained.

Finally, it is concluded that the WPM system by using Daubechies (db2) wavelet with QPSK modulation over AWGN channel condition gives a better performance of BER, throughput and energy consumption within an adhoc network.

The research carried out in this work leaves an ample scope for extension and applications of wavelet packet modulation system for wireless communication like mobile adhoc network (MANET), wireless fidelity (Wi-Fi).

It can be concluded that the work under investigation of the design of WPM system for adhoc network (communication purposes) can be extended towards multifold applications in the area of transportation systems. The practical development of improved WPM system is one of the major potential research directions in future.

## References

1. Chakole MB, Dorle SS (2020) Design of WPM system using QPSK modulation for different wavelets. *Int J Innov Technol Explor Eng (IJITEE)* 9(3). ISSN: 2278-3075
2. Van Bouwel C, Potemans J, Schepers S, Nauwelaers B, Van de Capelle A (2000) Wavelet packet based multicarrier modulation. In: *IEEE Benelux chapter on vehicular technology and communications. Symposium on communications and vehicular technology, SCVT-2000. Proceedings (Cat. No. 00EX465)*, pp 131–138. <https://doi.org/10.1109/SCVT.2000.923351>
3. Nawaz T, Baig S (2011) Wavelet OFDM—a solution for reliable communication in a frequency selective Rayleigh fading channel. In: *IEEE IBCAST*, pp 413–417
4. Chakole M, Dorle S (2017) Evaluation of WPM-BPSK based peer-to-peer network using wavelets for different channels. In: *The 8th IEEE international conference on computing communication and networking technology. IIT Delhi, India*
5. Lindgren A, Schelén O (2002) Infrastructured ad hoc networks. In: *Proceedings of the 2002 international conference on parallel processing workshops (international workshop on ad hoc networking)*
6. Meshram SL, Dorge PD (2017) Design and performance analysis of mobile ad hoc network with reactive routing protocols. In: *IEEE international conference on communication and signal processing, ICCSP*, pp 443–447
7. Ghaith A, Hatoum R (2013) Performance analysis of the wavelet-OFDM new scheme in AWGN channel. In: *ICCIT-2013 wireless communication and signal processing. IEEE*, pp 225–229
8. Ariananda DD, Lakshmanan MK, Nikookar H (2011) A wavelet packet transceiver for spectral analysis and dynamic spectrum access. *IEEE*
9. Li N, Gao Y, Du C, Hao X (2012) Optimizing performance of WIMAX system based on WPM. In: *The 7th international conference on computer science & education (ICCSE 2012)*, Melbourne, pp 372–375
10. Bajpai A, Lakshmanan MK, Nikookar H (2011) Channel equalization in wavelet packet modulation by minimization of peak distortion. In: *IEEE 22nd international symposium on personal, indoor and mobile radio communications*, pp 152–156
11. Effendi MR, Mengko TLR, Gunawan AH, Munir A (2019) Performance evaluation of wavelet packet modulation for wireless digital communications. In: *International symposium on networks, computers and communications (ISNCC)*, pp 1–4. <https://doi.org/10.1109/ISNCC.2019.8909167>

# Adiabatic Multiplexer and Delay Flip-Flop



Y. Syamala  and K. Srilakshmi 

## 1 Introduction

In recent years, as the transistor density is increasing on a single silicon chip, which improves the performance in most of the electronic circuits by significant orders of magnitude. Unfortunately, the increased dissipation of power and energy of the systems also depends on enhancement of the performance. In such cases, the reduction in reliability and increase in cost due to expensive packaging and cooling. In present scenario, minimization of power is one of the main considerations of the system due to so many reasons like improving the battery life of smart devices, reliability of the system and packaging [1]. Researchers are concentrating as the requirements of low power in micro-electronics systems at different levels like technology, circuit, and gate level for the reduction of energy dissipation by an electronic circuit. In VLSI circuits using CMOS, the most dominant source of power is the switching power due to charging and discharging of  $C_L$ . In order to minimize such power, an alternative method is introduced as an adiabatic switching [2]. In this technique, the load capacitance charging and discharging carried out as the small amount of energy loosed and recovery of energy stored in capacitors is achieved. Adiabatic logic is one technology which reduces the power dissipation. Power clock signal generator and digital core are the two main parts in an adiabatic system. Adiabatic families used here are positive feedback adiabatic logic (PFAL) and 2N-2N2P [3]. Both logics operate with a four-phase power clock supply and a main section of

---

Y. Syamala (✉) · K. Srilakshmi

Department of Electronics and Communication Engineering, Seshadri Rao Gudlavalleru Engineering College, Gudlavalleru, Andhra Pradesh 521356, India

e-mail: [Syamala@gecgudlavallerumic.in](mailto:Syamala@gecgudlavallerumic.in)

K. Srilakshmi

e-mail: [srilakshmi@gecgudlavallerumic.in](mailto:srilakshmi@gecgudlavallerumic.in)

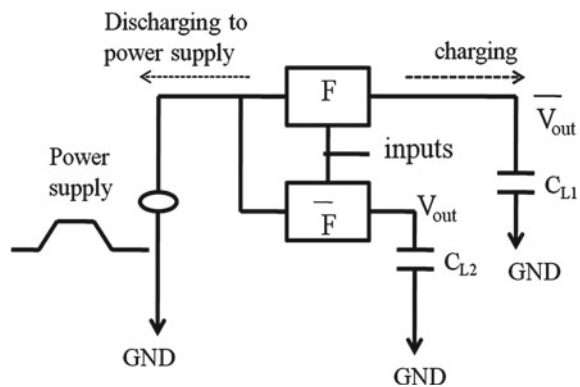
cross-coupled inverters. Logic blocks in PFAL are connected from clock to output nodes and from output to ground in 2N-2N2P [4–6].

In this paper, different sections are organized as basic information about adiabatic logic is given in Sect. 2. In Sect. 3, design of multiplexer using CMOS and adiabatic logics is presented. In Sect. 4, the D flip-flop circuit using both logics is depicted along with the results of simulation waveforms. Analysis and comparison of various performance parameters of designed circuits using both logics are explained in Sect. 5. Finally, conclusion is given in Sect. 6.

## 2 Adiabatic Logic

Due to switching activity of input patterns, power dissipation that takes place from circuit is converted into heat which is ultimately released to the environment like global warming. To minimize dissipation of power circuit, designers either can reduce supply voltage, node capacitance, or lowering number of switching events. To achieve energy dissipation below this lower limit of  $C_L V_{DD}^2$ , a novel technique is used which is termed as an adiabatic logic [7]. The researchers introduced a circuit-level approach called adiabatic which is possible to realize ultra-low-power computing applications without reducing  $V_{DD}$ . An adiabatic logic refers to the thermodynamic processes that exchange no heat with the environment. This leads to a minimization of power dissipation at the cost of slower performance of operation. In the literature, adiabatic switching has been identified as well suited for ultra-low-power applications. An adiabatic logic is used in ultra-low-power applications where conventional energy is minimal and performance is not an issue such as bio-medical, robotics, IoT, space, and deep sea [8–11]. To achieve energy efficiency, the adiabatic gate has to follow two basic rules. (i) Transistor will never get completely ON where the voltage is between drain and source. (ii) Transistor is never completely turned off where current is present in it. The basic topology of an adiabatic gate is given in Fig. 1.

**Fig. 1** Topology of an adiabatic gate



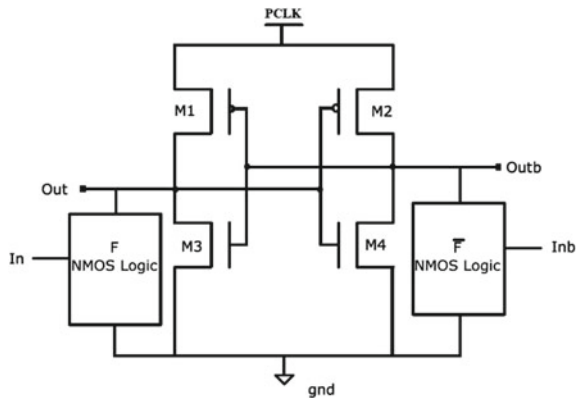
**2N2N-2P Logic**

2N2N-2P logic mainly derived from 2N2P and efficient charge recovery logic (ECRL) to reduce coupling effect which belongs to the partially adiabatic logic family shown in Fig. 2. Its main problem over 2N2P is cross-coupled nature of N-MOSFET switches which results into non-floating outputs for the recovery phase. The latest version of 2N2P is 2N2N-2P with only difference that it has extra two NMOS transistors which are connected in series with PMOSFETs where cross-coupled PMOS transistors common to both logics [12]. This logic structure is similar to static RAM which has a cross-coupled inverters.

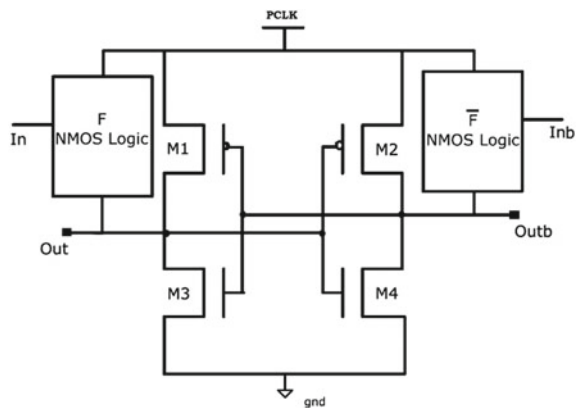
**Positive Feedback Adiabatic Logic (PFAL)**

The topology of a PFAL gate is given in Fig. 3. To realize the logic functions, two NMOS networks are used to generate both true and complementary outputs. The input NMOS network is connected in parallel to the PMOS transistors. In a PFAL gate, as the output is not floating therefore all outputs have full logic swing. Finally,

**Fig. 2** Structure of 2N2N-2P logic



**Fig. 3** Structure of PFAL gate



PFAL shows the best performance in terms of energy consumption, useful frequency range, and robustness against technology variations.

### 3 Design of 4 × 1 Multiplexer

A multiplexer is an electronic component which is a type of digital switch, or it can also be termed as data selector. This circuit has “*n*” selection lines,  $2^n$  input data lines, and a single data output line.

#### 4 \* 1 Multiplexer Using CMOS Logic

Based on the functionality of logic expression, circuit is designed using CMOS logic.

In which, the pull-up network consists of four sets of three PMOS transistors each in series and connects them in parallel, and the pull-down network has a similar inversely connected circuit; the output is then connected to an inverter to obtain the final output is presented in Fig. 4. The general logic diagram and expressions are given in [12]. The functionality has been verified using CAD tool, and simulation

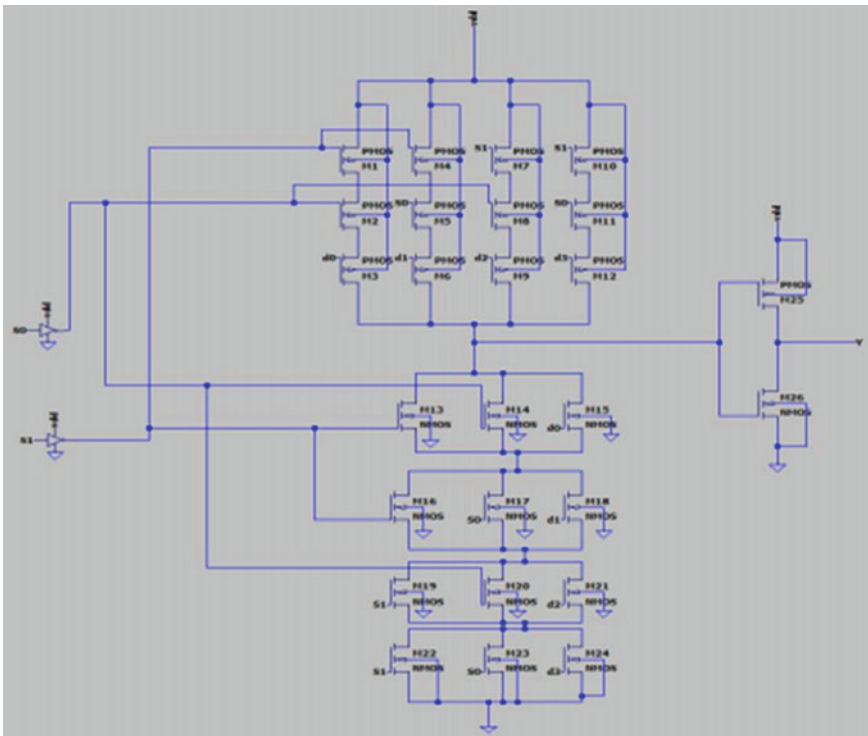
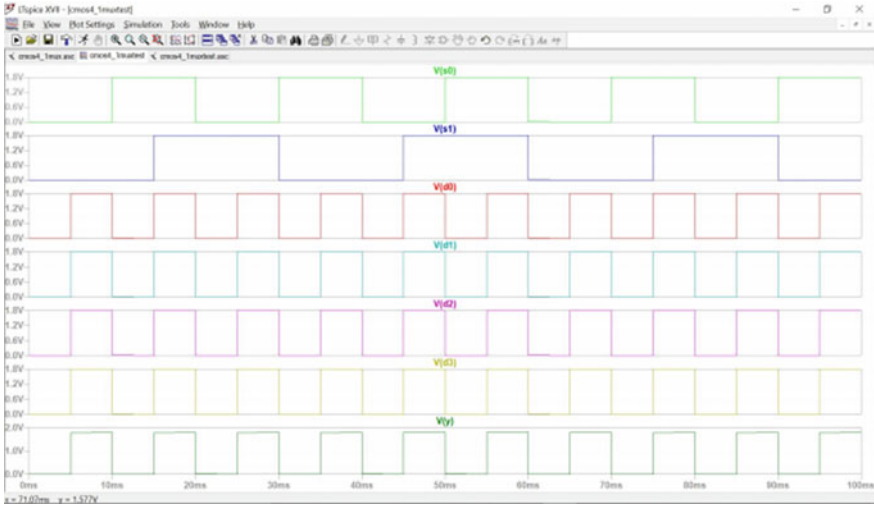


Fig. 4 CMOS transistor level diagram of 4 \* 1 multiplexer



**Fig. 5** Functional verification of 4 \* 1 data selector using CMOS

waveform is shown in Fig. 5. For S0S1 = “11,” output produces the data which are available at D3 to the output Y.

**4 \* 1 Multiplexer Using Adiabatic Logic**

The adiabatic technique incorporated in the design of 4 \* 1 multiplexer here is positive feedback adiabatic logic (PFAL) shown in Fig. 6. This is a form of dual rail circuit with partial recovery logic. The circuit mainly consists of a pair of cross-coupled inverters that avoids logic-level degradation on the output; on either sides, a couple of NMOS transistors are connected which forms n-tree sections which are left-side n-tree consists of three NMOS transistors connected in series, and four similar sets are connected together in parallel. Right-side n-tree consists of three NMOS transistors connected in parallel, and four such sections are connected in series. The coupled inverter and both n-trees are connected via a power clock to provide power supply; both true and complementary outputs are generated by this topology.

The operation performed by 4 \* 1 multiplexer with four selection lines s0 and s1 and their inverse, so it is to be used to select between input lines d0, d1, d2, d3, and provide the same via a common output lines out and out\_bar shown in Fig. 7.

**4 Delay Flip-Flop**

The delay flip-flop is an edge-triggered device which transfers input data to Q on clock rising or falling edge. Data latches are level-sensitive devices such as data and transparent latch.



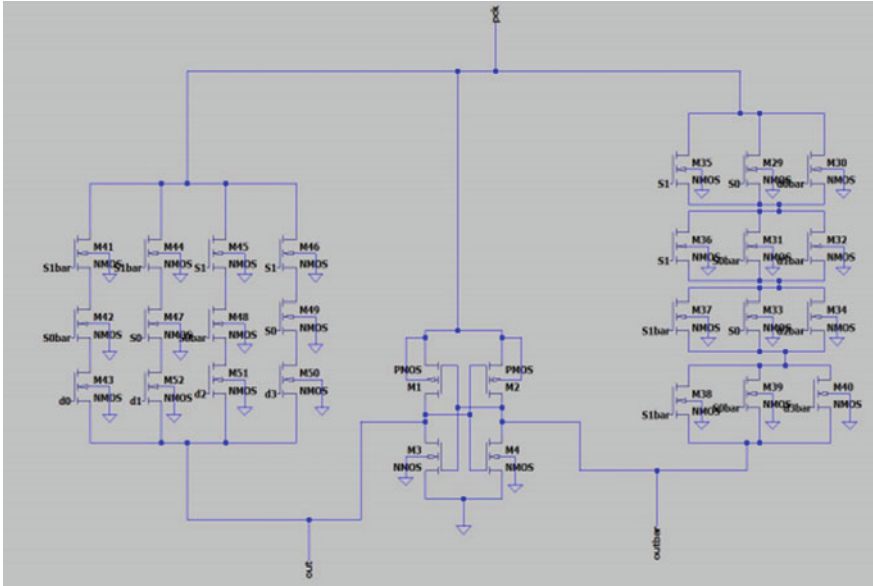


Fig. 6 Schematic diagram for 4 \* 1 multiplexer using adiabatic logic

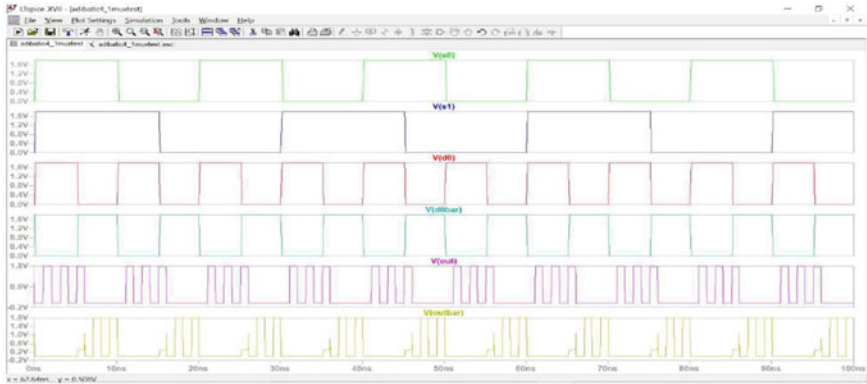


Fig. 7 Simulation waveform for 4 \* 1 multiplexer using adiabatic CMOS logic

### D Flip-Flop Using CMOS Logic

The D flip-flop circuit can be explained by dividing the complete design into three sections connected with a common power supply  $V_{DD}$  and ground. The first section consists of a PMOS transistor and two NMOS transistors connected in series with PMOS transistor at the top followed by NMOS transistors; a common input  $D$  is given to PMOS and bottom NMOS, and a clock input is given to the center NMOS transistor.

Second section consists of a PMOS transistor and NMOS transistor connected in series with PMOS transistor receiving its input from the combination of PMOS and middle NMOS from first section, and the NMOS receives its input from the combination from of two NMOS circuits from first section; this produces an output  $Q$  as shown in Fig. 8. Third section is an inverter circuit with input  $Q$ , so it produces the final output  $Q$ .

The D flip-flop will store and output, whatever logic level is applied to its data terminal, so long as the clock, input is HIGH. Once clock input goes LOW, the set and reset inputs of the flip-flop are both held at logic level “1” is given in Fig. 9.

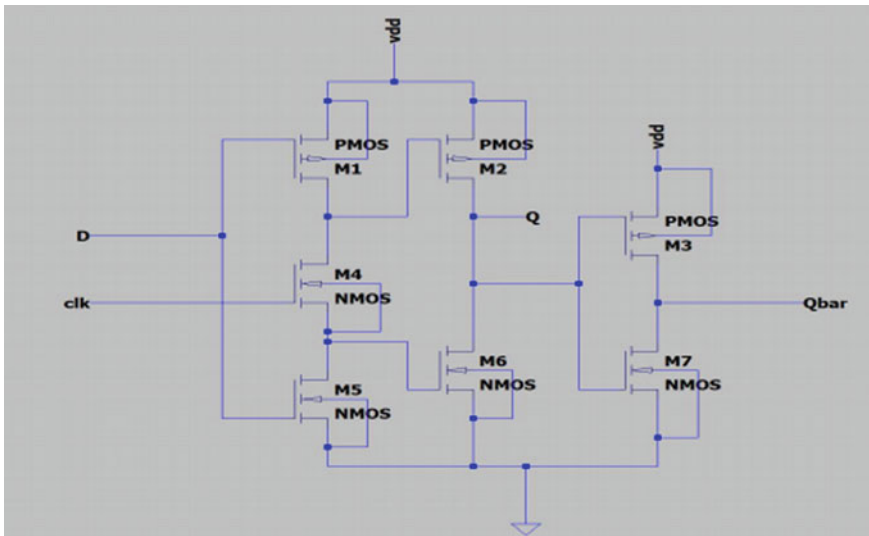


Fig. 8 Schematic diagram for D flip-flop using CMOS logic

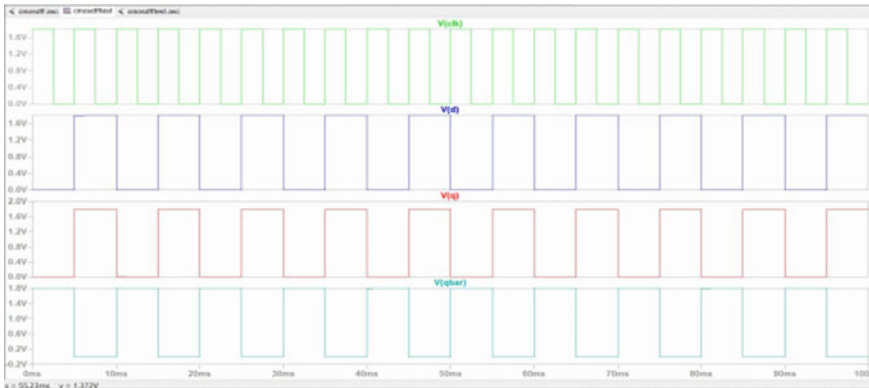


Fig. 9 Simulation waveform for D flip-flop using CMOS logic

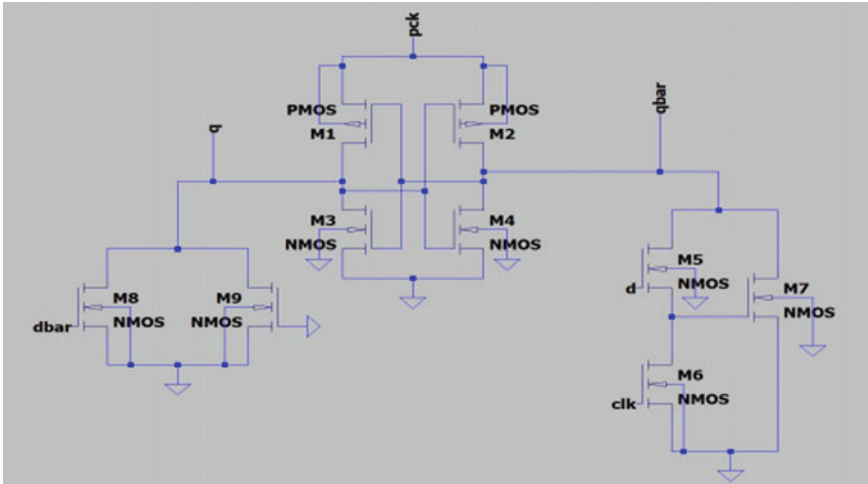


Fig. 10 Schematic diagram for D flip-flop using adiabatic logic

### D Flip-Flop Using Adiabatic Logic

The adiabatic technique incorporated in the design of D flip-flop here is 2N-2N2P. This technique reduces the coupling effects. The circuit mainly uses a cross-coupled latch of two PMOSFETs and two NMOSFETs with a power supply, and on either sides, a couple of NMOS transistors are connected which forms n-tree sections. The left-side n-tree is called functional block and consists of two NMOS transistor coupled and provided with its respective inputs. The right-side n-tree is called functional bar block and consists of two NMOS transistors connected in series, and a NMOS transistor is connected in parallel to the two NMOS set; the output from the two NMOS set is provided as input to the third NMOS in series, and by providing its respective inputs, the circuit is designed. The schematic using CMOS and functional verification of D flip-flop is given in Figs. 10 and 11, respectively.

An adiabatic D flip-flop has two inputs  $d$ ,  $dbar$  and a clock as another input, finally provides output via  $q$  and  $qbar$ .

## 5 Power Consumption Analysis and Comparison

The functional verification of designed adiabatic multiplexer and D flip-flop has been done with different supply voltages of 1.8, 3.3, and 5 V. The power dissipation of the designed circuits is determined at 180 nm technology node with  $W = 2 \mu\text{m}$  and  $L = 0.18 \mu\text{m}$ . The area in terms of number of devices is also calculated and compared with conventional CMOS logic.

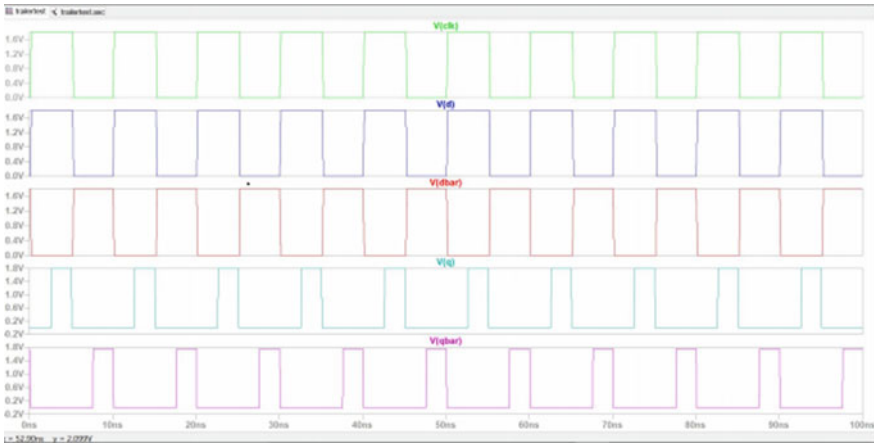


Fig. 11 Simulation waveform for D flip-flop using adiabatic CMOS logic

### 4:1 Multiplexer

The variation of power dissipation for the designed PFAL and CMOS multiplexers with different supply voltages is shown in Fig. 12. From the simulation results, it was observed that an average percentage of power dissipation improvement in adiabatic multiplexer is 33.08% as compared with conventional CMOS design.

### D Flip-Flop

Similarly, for the D flip-flop, the power dissipation curves are shown in Fig. 13. Average percentage of power dissipation improvement in 2N2N-2P D flip-flop in comparison with conventional CMOS is 78.48% observed from the results.

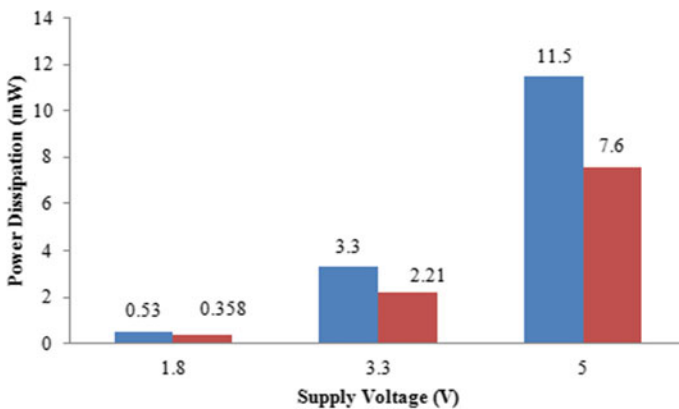
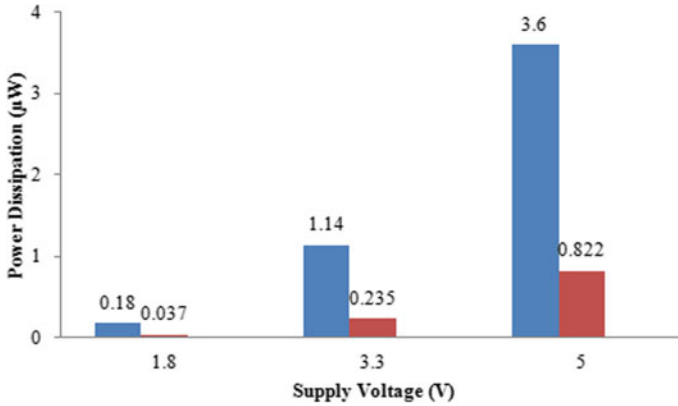


Fig. 12 Comparison plot for designed multiplexer circuits



**Fig. 13** Power plot comparison for delay flip-flop

**Table 1** Performance parameters of designed circuits

Performance parameters	4 × 1 multiplexer		Delay flip-flop	
	CMOS	Adiabatic	CMOS	Adiabatic
Transistor count	26	28	07	09
Area (µm <sup>2</sup> )	9.36	10.08	2.52	3.24
Total power dissipation (µW) at a supply voltage of 1.8 V	530	358.8	180	37.8

The performance comparison of multiplexer and D flip-flop using CMOS and adiabatic logic is presented in Table 1.

## 6 Conclusion

In this work, mainly concentrated on minimization of the power dissipation and APP. The digital logic circuits of multiplexer and D flip-flop are designed and implemented using CMOS and adiabatic logic. The circuits are synthesized using CAD tools, and the results observed that the adiabatic designed circuits have lesser power dissipation than CMOS logic. Adiabatic logic circuits were designed using 180 nm technology at three different DC supply voltages which are 1.8, 3.3, and 5 V on LTspice tool. On comparison with conventional CMOS designs, an average of 33.08% energy saving in the design of PFAL adiabatic multiplexer and an average of 78.48% energy saving in the design of D flip-flop are observed which designed using 2N-2N2P adiabatic logic was. The proposed adiabatic designs may have certain applications in the design of high-end circuits and will be felt useful in area which require better performance with less power dissipation.

## References

1. Chandrakasan SS, Brodersen R (1992) Low-power CMOS digital design. *IEEE J Solid State Circuits* 27(4):473–484
2. Athas WC, Svensson L, Koller JG, Tzartzanis N, Chou EY (1994) Low-power digital systems based on adiabatic-switching principles. *IEEE Trans VLSI Syst* 2(4):398–407
3. Chaudhuri A, Saha M, Bhowmik M, Pradhan SN, Das S (2015) Implementation of circuit in different adiabatic logic. *IEEE J Low Power Electron*. <https://doi.org/10.1109/ECS.2015.7124923>
4. Vetuli A, Di Pascoli S, Reyneri LM (1996) Positive feedback in adiabatic logic. *Electron Lett* 32(20):1867–1869
5. Kramer A, Denker JS, Flower B, Moroney J (1995) 2nd order adiabatic computation with 2N-2P and 2N-2N2P logic circuits. In: *Proceedings of the international symposium on low power design (ISLPED)*, pp 191–196
6. Amirante E, Bargagli-Sto A, Fischer J, Iannaccone G, Schmitt-Landsiedel D (2001) Variations of the power dissipation in adiabatic logic gates. In: *International workshop on power and timing modeling, optimization and simulation (PATMOS)*, pp 9.1.1–9.1.10
7. Alioto M, Palumbo G (2001) Power estimation in adiabatic circuit: a simple and accurate model. *IEEE Trans VLSI Syst* 9(5):608–615
8. Konwar S, Singha TB, Roy S, Vanlalchaka RH (2014) Adiabatic logic based low power multiplexer and demultiplexer. In: *International conference on computer communication and informatics*. <https://doi.org/10.1109/ICCCI.2014.6921808>
9. Srilakshmi K, Tilak AVN, Srinivasa Rao K (2018) Ultralow-power and secure S-box circuit using FinFET based ECRL adiabatic logic. *J Sci Technol* 10(3):10–17
10. Srilakshmi K, Tilak AVN, Srinivasa Rao K, Syamala Y (2018) Energy efficient 64-bit asynchronous adder. In: *Lecture notes in electrical engineering (LNEEE)*, vol 476. Springer, pp 499–508
11. Srilakshmi K, Tilak AVN, Srinivasa Rao K (2018) Ultralow-power multiplier for DSP applications. In: *Proceedings of seventh international conference on advances in computer science and application*
12. Konwar S, Singha TB, Roy S, Vanlalchaka RH (2014) Adiabatic logic based low power multiplexer and demultiplexer. In: *International conference on computer communication and informatics*

# A MobileNet-V2 COVID-19: Multi-class Classification of the COVID-19 by Using CT/CXR Images



N. Mahendran and S. Kavitha

## 1 Introduction

COVID-19 is a spherical or pleomorphic, and it has the envelop particles consisting of mono-stranded RNA which are incorporated with a nucleoprotein. The capsid is located inside the nucleoprotein which is comprised of matrices of proteins [1]. COVID-19 is one of the most dangerous virus and has emerged from China and spread to the worldwide in short period of time. The World Health Organization has announced this COVID as worldwide pandemic. Many countries have been introduced the lock down to reduce the further spread of COVID. Among the world, many of the countries followed this lock down to control the pandemic. In India, the government and people have more issues those are related to the economic, agricultural, social, educational, political, psychological, etc.

In the May month of 2020, WHO have registered positive COVID cases of 3,759,967 and the amount of death had increased to 259,474 throughout the world. Due to this severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection, there are totally 212 countries have affected. In India, the first COVID case had been registered in the state of Kerala on 30th Jan, 2020. Drastically, the number of cases are increased. On 8th May of 2020, the Indian Council of Medical Research (ICMR) collects 1,437,788 samples to the National Institute of Virology (NIV), Pune. As of Aug 12th, 2021, totally 32,077,706 positive cases registered in the India among that 429,669 deaths are occurred. Hence, the death rate is increased to 1.34% compared to 2020. The clinical symptoms have totally varied from minor infection in lungs to major level infections like MERS and SARS-CoV, COVID-19 symptoms are fatigue, cough, fever and shortness of breath. Among these, diarrhea also one

---

N. Mahendran · S. Kavitha (✉)

Department of Electronics and Communication Engineering, M. Kumarasamy College of Engineering, Karur, TN, India

e-mail: [skavitha495@gmail.com](mailto:skavitha495@gmail.com)

of the symptoms in COVID-19 infection and approximately 20–25% are intestinal symptoms [2, 3].

The evaluation of the radiological treatments and CT scans is widely utilized in the part of COVID diagnosis [4, 5]. On the other hand, the slice of chest CT scans have vital proof for appropriate COVID-19 findings [6]. The 100% prospective and confidence findings of the COVID-19 disease in Thoracic-CT (T-CT) scans are acquired with the help of air bronchograms, amalgamation, perilobular pattern, inverse halo and ground-glass opacity (GGO)  $\pm$  crazy-paving [7]. Hence, the coronavirus have changed their protein layer rapidly and available in multiple versions such as delta and delta+. The main goal of the paper is to deal with the deep learning approaches to carry over each level processing of the COVID-19 detection.

## 2 Literature Survey

In the twenty-first century, COVID-19 is the most dangerous virus stated by [8] which is diagnosed with the help of variety of testing approaches such as CXR images, RT-PCR testing, and CT scans. The present-day COVID database has not the ability to build such triage system because COVID cases tended toward more severe as well as heterogeneous in nature. Hybridized deep learning architecture was developed by [9] in which the unsupervised anomaly detection has been clearly solved for multiple spatiotemporal data. Dong [10] reviewed the computing and characteristic imaging models that had been processed in the COVID-19 management. CT scans, RT-PCR, MRI, and lung ultrasound images are used for detect, treat, and follow-up of the COVID-19. These computing models are explored with the aid of AI techniques. Horry [11] demonstrated the deep learning (DL) and transfer learning (TL) algorithms to find the COVID using X-rays, CT scans, and ultrasound images. The authors identified an appropriate CNN model which is utilized with VGG19 model to explore the scarcity and challenging issues in the COVID-19. Jiang et al. [12] presented a synthesis approach for CT images which is based on the generative adversarial network (GAN) which effectively creates the great quality and optimistic CT images for imaging tasks.

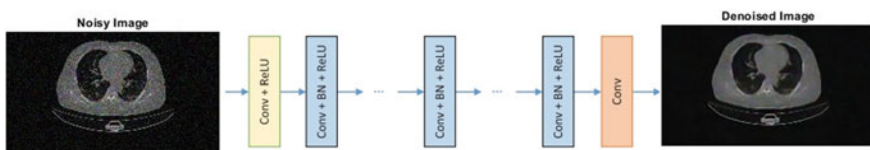
Liu et al. [13] introduced a matrix profile technique for detecting the anomaly of CT image in two levels. Here, the DenseNet was trained by using the anomaly weighted CT images to variate the COVID-19 and non-COVID-19 CT images. Yan [14] tried to explore the novel DCNN for segmenting the chest CT images to find out the COVID-19 infections. The COVID-19 infection had been segmented by feature variation block which adaptively adjusts the global characteristics of the features. The feature fusion was done by multiple scaling of the Progressive Atrous Spatial Pyramid Pooling which is useful for handling the convenient infected area with different shapes and appearances. Hu [15] presented the weakly supervised DL algorithm to identify and classify the COVID-19 cases from input CT image datasets. Rahaman et al. [16] have employed deep transfer learning algorithms in which the



authors have carefully examined the 15 pre-trained CNN models to come out with appropriate network to perform the classification task.

Tareh et al. [17] aimed to evaluate the effectiveness of the pre-trained CNN for automatic diagnosis of the COVID-19 from the chest X-ray image dataset. The method was explored by AI effectiveness and detailed discovery of COVID from input X-ray image dataset. Kusakunniran et al. [18] presented a solution to detect the COVID-19 automatically from the CX-ray images. The system adapted with ResNet-101 as a main architecture along with 44 million of parameter set. Sarkar et al. [19] have used the VisionPro Deep Learning™ and COGNEX's deep learning software to separate the Chest X-rays images from the COVIDx dataset. Ahmed et al. [20] presented an automated model for classifying the COVID-19 by utilizing available dataset of COVID and non-COVID X-ray images. The feature extraction process was performed by high-resolution network (HRNet). Narin et al. [21] have implemented variety of three binary classifications along with four classes criteria (bacterial pneumonia, viral pneumonia, COVID-19, and healthy) by using the five-fold cross-validation. ResNet101, Inception-ResNetV2, InceptionV3, ResNet50, and ResNet152 are pre-trained NN models to detect the coronavirus from the infected X-ray radiograph reports.

Swapnarekha et al. [22] presented VGG16 to make the use of RMSprop and ADAM optimizers for autonomous finding of the COVID-19. Zoabi et al. [23] established a trained machine learning algorithm to extract the eight sets of features such as known-infected contact, age, sex, and five initial symptoms (cough, shortness of breath, etc.). Pandit et al. [24] and made use of the DL model to automatically classify the COVID-19 using chest X-rays. The pre-trained deep learning VGG-16 was used to take the charge of classification task. The fine-tuned transfer learning (TL) was used to train the network effectively with less number of chest radiographs. Hence, deep learning and CNN approaches were used by the many of the authors [25–27] for detecting the COVID cases.



**Fig. 1** Denoising of CT images using DnCNN

### 3 Deep Learning Algorithm for COVID-19 Detection

#### 3.1 DnCNN Algorithm for Denoising of CT Images

The DnCNN architecture for denoising the input CT images is shown in Fig. 1 [28]. The noised image in the proposed DnCNN denoising algorithm is termed as  $y = x + v$ . The mapping function  $F(y)$  is learned by discriminative denoising models like cascade of shrinkage fields (CSF) and multilayer perceptron (MLP). Residual learning formulation is adapted to train the  $R(y) \approx v$  residual mapping, and it can be rewritten as  $R(y) = y - x$ . Initially, the average mean square error (MSE) between the expected residual image and noisy image can be considered as the loss function. The  $\theta$  trainable parameter is learnt with the help of this loss function in DnCNN. For  $N$  noisy trained images, the pair  $\{(y_i, x_i)\}_{i=1}^N = 1$  is assumed to be unity. The following steps are used for summarizing the process involved to denoising the input CT images.

**Step 1:** The noisy image feature 'Y' is analyzed by active convolution layer and covariate offset, and it can be selected by batch normalization.

**Step 2:** The selected features are extracted with the help of nonlinear mapping.

**Step 3:** The final residual image is calculated with batch normalization active convolution.

#### 3.2 Feature Extraction of COVID CT Images

Conventional ML algorithms require huge man power to detect and extract the features. Some of the manually extracted features are Haar cascades, scale-invariant feature transform, features from accelerated segment test (FAST), histogram of oriented gradient (HOG), speeded up robust feature, and min eigenfeatures.

In DL algorithms, the image features are automatically extracted from the input medical images. The network learns the importance of the extracted features on the output end by introducing the weights. The raw input CT images are first passed through the network layers, and it is identified the image patterns to make the feature sets. The neural network comprises of both feature extractor and classifier unit which are trained under end-to-end approach. This mannerism of the NN model is totally contrasted to the conventional ML model which holds the hand crafted features.

### 3.3 Classification of COVID-19 Images by MobileNet

The traditional ML algorithm needs the basic information about the source and target data. The feature set is extracted from the initial source data and used in the training phase to classify the target data. The previous version of MobileNet-V1 was intended through orientation to the conventional VGG structure that is used to construct the network model by mounding convolution layers (CL) in order to increase the accuracy rate.

Compared to the MobileNet-V1, the major development of MobileNet-V2 has two points, the initial point is the implementation of the linear bottleneck and the second point is complemented residual block. The MobileNet-V2 has main building block of depth separable convolution block and linear bottleneck with complemented residuals. Table 1 shows the characteristic implementation of the MobileNet-V2 in which the factors  $N$  to  $M$  represents the channels,  $E$  and  $S$  denote the expansion factor and stride, respectively. The linear restricted access sums the  $1 \times 1$  convolutional layer to the depth-wise convolutional layer, and also, it uses the activation in linear nature rather than the nonlinear. The point-based convolutional layer, performs the down-sampling to set the  $s$  parameter in the depth-wise convolutional layer. Table 2 shows the whole architecture of the MobileNet-V2, whereas the standard convolutional layer is denoted as conv-2d, the average pooling is termed as avg pool, number of output channels is indicated by term  $O_c$ , and it is recurrent in  $N_r$  times. There are

**Table 1** MobileNet-V2's bottleneck

Input	Layer-wise operator	Output
$X \times Y \times N$	$1 \times 1$ conv-2d, RLU6	$X \times Y \times tN$
$X \times Y \times Nt$	$3 \times 3$ dwise $s = s$ , RLU6	$X/s \times Y/s \times Nt$
$X/s \times Y/s \times Nt$	Linear $1 \times 1$ conv-2d	$X/s \times Y/s \times M$

**Table 2** Network configuration of MobileNet-V2 [29]

Shape of the input	Operation	$E$	$O_c$	$N_r$	$S$
$224 \times 224 \times 3$	Conv-2d	–	32	1	2
$112 \times 112 \times 32$	Bottleneck	1	16	1	1
$112 \times 112 \times 16$	Bottleneck	6	24	2	2
$56 \times 56 \times 24$	Bottleneck	6	32	3	2
$28 \times 28 \times 32$	Bottleneck	6	64	4	2
$14 \times 14 \times 64$	Bottleneck	6	96	3	1
$14 \times 14 \times 96$	Bottleneck	6	160	3	2
$7 \times 7 \times 160$	Bottleneck	6	320	1	1
$7 \times 7 \times 320$	conv2d $1 \times 1$	–	1280	1	1
$7 \times 7 \times 1280$	Avg pool	–	–	1	–
$1 \times 1 \times 1280$	conv2d	–	$k$	–	–

totally 19 layers associated with this network; the intermediate layers are mainly placed to excerpt the features. The final layer is mainly employed for classification purpose.

In transfer learning algorithms, the MobileNet-V2 is previously trained by prominent ImageNet which is considered to be a standard feature extractor. The training process is done in the two added layers and then fine-tuning is performed by locating the few of the layer instead of training of all layers.

## 4 Implementation Results

### 4.1 Experimental Dataset

The authorized COVID dataset is collected from [30, 31]. The dataset contains the different medical images such as severe pneumonia cases, lung segmented images, COVID-19 subjects, and non-COVID-19 images. This dataset simultaneously processed in DnCNN denoising algorithm and deep learning classifier models. This section deals the performance of each level of medical image processing algorithms. The sample images in the dataset are shown in Figs. 2, 3, and 4.

Figure 2a–h illustrates the disease-affected CT images, and healthy CT images are represented in Fig. 3a–h. The presented MobileNet-V2 not only focused on the CT images but also took the same concern on the processing of X-ray images for identifying the COVID-19. Figure 4a–h represents the diseased X-ray image sets.

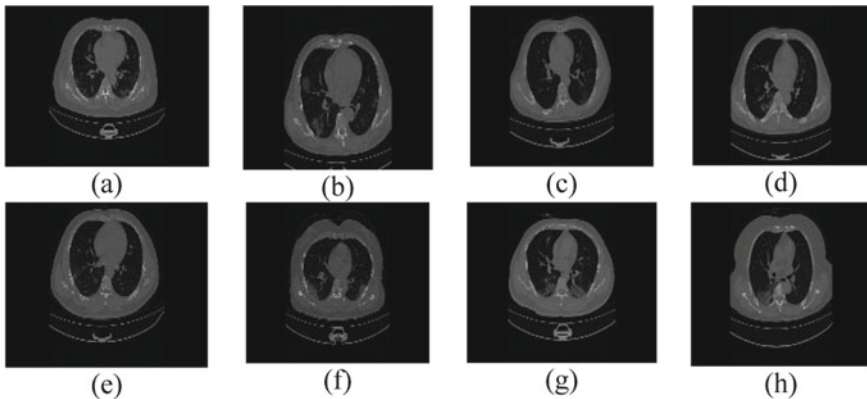


Fig. 2 a–h Diseased input CT images

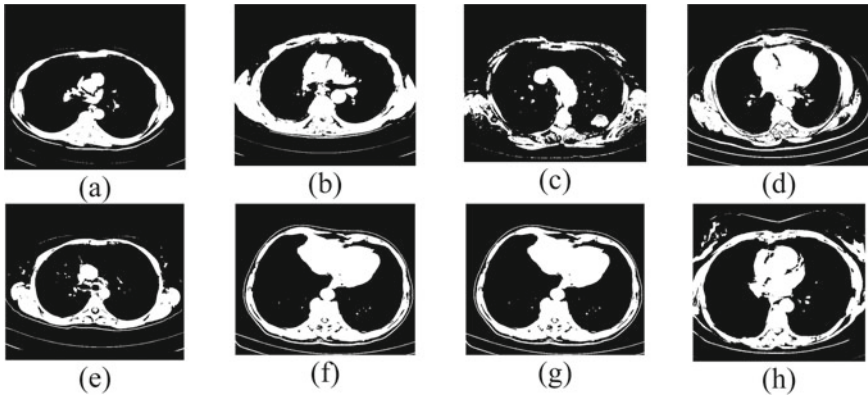


Fig. 3 a–h Healthy input CT images

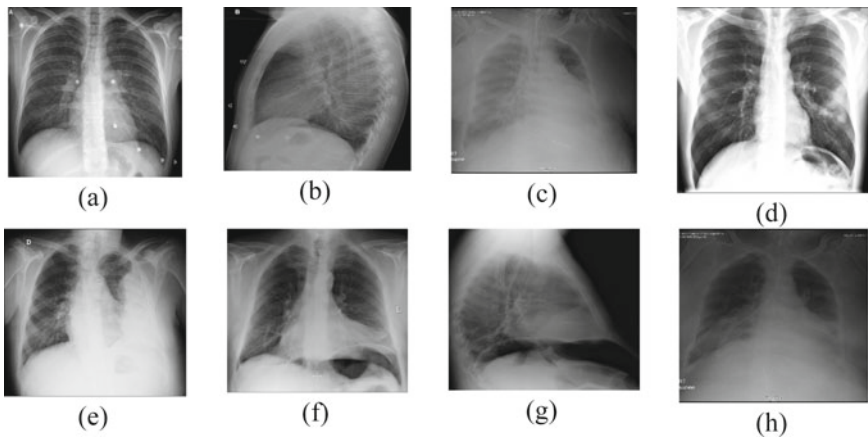


Fig. 4 a–h Diseased X-ray images

### 4.2 DnCNN-Based Denoising of Medical Dataset

The input CT/X-ray input dataset is initially passed through the DnCNN denoising algorithm in which the denoised image got from the residual images. The quality of the denoising algorithm is validated by measuring the PSNR and SSIM values. Table 3 shows the PSNR and SSIM values of the noised and denoised medical images while considering the Gaussian noise.

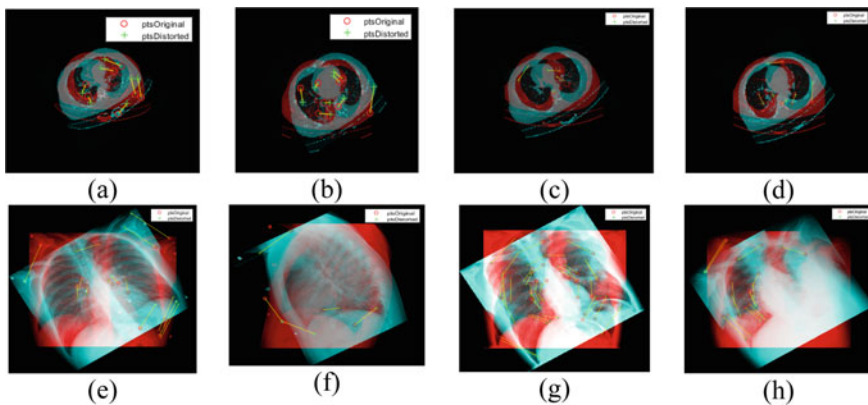
**Table 3** Denoising of CT/X-ray images using DnCNN algorithm

Images	Noised PSNR	Denoised PSNR	Noised SSIM	Denoised SSIM
CT-1	22.1011	27.0405	0.2276	0.3724
CT-2	21.8832	26.8667	0.2497	0.3983
CT-3	22.5037	28.1314	0.1035	0.2816
CT-4	22.0751	27.1368	0.2480	0.3945
CT-5	22.0510	27.0324	0.2623	0.4079
X-ray-1	20.2270	32.6682	0.1726	0.7938
X-ray-2	20.3437	32.4030	0.1389	0.7381
X-ray-3	20.0949	34.3288	0.1283	0.8353
X-ray-4	20.8911	29.9641	0.2650	0.8152
X-ray-5	20.4414	34.2872	0.1303	0.8317

### 4.3 Classification of COVID-19 Cases by Deep Learning Models

In conventional image processing algorithm, the feature extraction is performed after the denoising process. Here, some of the features are manually extracted with the help of speeded up robust features (SURF), features from accelerated segment test (FAST), and min eigenfeature extractors. Each feature extractor exhibits the matching points between the original and distorted medical images. Hence, Figs. 5a–d, 6a–d, and 7a–d illustrate the matching points of the SURF, FAST, and min eigenfeatures of the CT images. Similarly, Figs. 5e–h, 6e–h, and 7e–h illustrate the matching points of the SURF, FAST, and min eigenfeatures of the X-ray image sets.

The pre-trained MobileNet-V2 model is employed in the deep learning classifier. In classifier stage, the process is carried in the two main phases: the first training phase



**Fig. 5** a–h Matching points extraction from SURF feature extractor

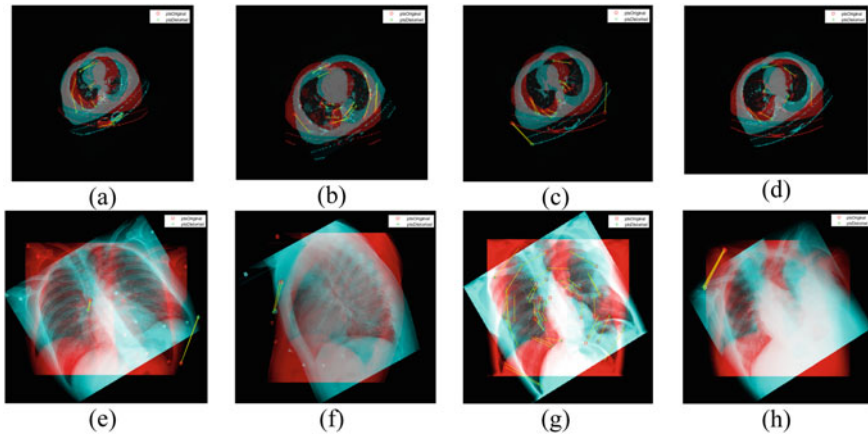


Fig. 6 a–h Matching points extraction from FAST feature extractor

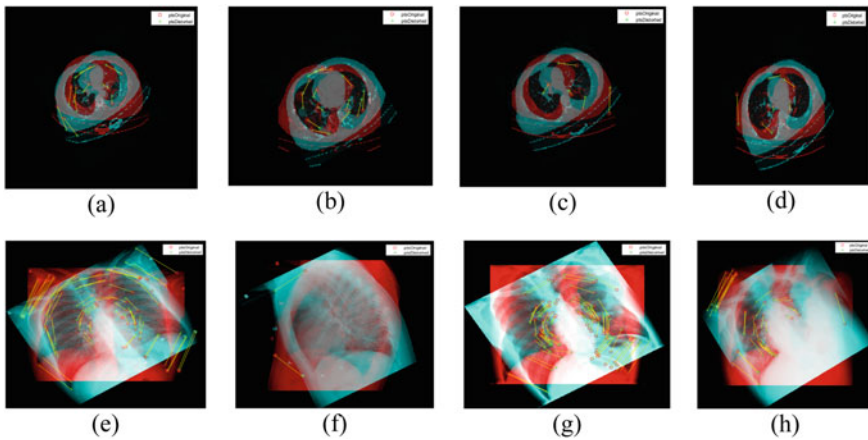


Fig. 7 a–h Matching points extraction from min eigenfeature extractor

trains the classifier with whole dataset and the second testing phase classifies the input image based on the training phase. Here, the MobileNet-V2 is fully pre-trained with entire training images. Table 4 represents the characteristics of the proposed classifier.

During the testing phase, each input medical image in the dataset is processed in variety of processing stages and finally classify the images. Table 5 illustrates the type of the image, entropy, and accuracy rate. Different machine learning and optimization techniques are used to improve the performance [32–34].

**Table 4** Characteristics of the proposed classifier

Specification	Used
Deep learning model	MobileNet-V2
Number of layers	155
Number of connections	164
Size of the input	$224 \times 224 \times 3$
Number of classes	4

**Table 5** Classification of COVID-19

Images	Type	Entropy	Accuracy (%)
I1	Pneumonia/COV-19	3.91492	85.4
I2	Pneumonia	7.70163	81.7
I3	Pneumonia	2.87851	79.4
I4	Pneumonia	4.03344	82.7
I5	Pneumonia/COV-19	4.1544	76.4
I6	Pneumonia/COV-19	4.24487	86.4
I7	Pneumonia/SARS	6.87146	87.9
I8	Non-COV-19	7.15035	83.7
I9	Pneumonia	7.16573	79.8
I10	Non-COV-19	7.61714	81.7

## 5 Conclusion

In this paper, a DL MobileNet-V2-based COVID-19 classification is done by processing the CT/CXR images. In the earlier stage, the medical images are preprocessed by DnCNN algorithm to remove the artifacts that are presented in the medical images. The preprocessing stage is validated by measuring the SNR and SSIM values of the noised and denoised medical images. Next to the preprocessing stage, feature extraction is used to detect and extract the important features in the denoised images. Finally, the training phase of the classifier is done to train the classifier with COVID-19 datasets. The testing phase is followed next to the training phase. In this testing phase, the presented classifier finds the non-COVID-19, pneumonia, pneumonia/SARS, COVID-19 and shows the 85% of accuracy rate. In the future, the presented approach will be further developed by taking the variety of COVID-19 types such as delta and delta+ variants.

## References

1. Mousavizadeh L, Ghasemib S (2021) Genotype and phenotype of COVID-19: their roles in pathogenesis. *J Microbiol Immunol Infect* 54:159–163



2. Sahin AR, Erdogan A, Agaoglu PM, Dineri Y, Cakirci AY, Senel ME (2019) Novel coronavirus (COVID-19) outbreak: a review of the current literature. *Eur J Med Oncol* 4:1–7
3. Zu ZY, Jiang MD, Xu PP, Chen W, Ni QQ, Lu GM (2020) Coronavirus disease 2019 (COVID-19): a perspective from China. *Radiology* 200490
4. Abbasi-Oshaghi E, Mirzaei F, Farahani F, Khodadadi I, Tayebinia H (2020) Diagnosis and treatment of coronavirus disease 2019 (COVID-19): laboratory, PCR, and chest CT imaging findings. *Int J Surg* 143–153
5. World Health Organization (2020) Clinical management of severe acute respiratory infection when novel coronavirus (2019-nCoV) infection is suspected: interim guidance, 28 Jan 2020
6. Ng MY, Lee EY, Yang J, Yang F, Li X, Wang H (2020) Imaging profile of the COVID-19 infection: radiologic findings and literature review. *Radiol Cardiothorac Imaging* 2
7. Ozsahin I, Sekeroglu B, Musa MS, Mustapha MT, Ozsahin DU (2020) Review on diagnosis of COVID-19 from chest CT images using artificial intelligence. *Comput Math Methods Med*
8. Tabik S (2020) COVIDGR dataset and COVID-SDNet methodology for predicting COVID-19 based on chest X-ray images. *IEEE J Biomed Health Inform* 24:3595–3605
9. Karadayi Y, Aydin MN, Öğrenci AS (2020) Unsupervised anomaly detection in multivariate spatio-temporal data using deep learning: early detection of COVID-19 outbreak in Italy. *IEEE Access* 8:164155–164177
10. Dong D (2021) The role of imaging in the detection and management of COVID-19: a review. *IEEE Rev Biomed Eng* 14:16–29
11. Horry MJ (2020) COVID-19 detection through transfer learning using multimodal imaging data. *IEEE Access* 8:149808–149824
12. Jiang Y, Chen H, Loew M, Ko H (2021) COVID-19 CT image synthesis with a conditional generative adversarial network. *IEEE J Biomed Health Inform* 25:441–452
13. Liu Q, Leung CK, Hu P (2020) A two-dimensional sparse matrix profile DenseNet for COVID-19 diagnosis using chest CT images. *IEEE Access* 8:213718–213728
14. Yan Q (2021) COVID-19 chest CT image segmentation network by multi-scale fusion and enhancement operations. *IEEE Trans Big Data* 7:13–24
15. Hu S (2020) Weakly supervised deep learning for COVID-19 infection detection and classification from CT images. *IEEE Access* 8:118869–118883
16. Rahaman MM, Li C, Yao Y, Kulwa F, Rahman MA, Wang Q, Qi S, Kong F, Zhu X, Zhao X (2020) Identification of COVID-19 samples from chest X-ray images using deep learning: a comparison of transfer learning approaches. *J X-Ray Sci Technol* 28:821–839
17. Tareh MM, Zhu N, Ali TAA, Hameed AS, Mutar ML (2021) Transfer learning to detect COVID-19 automatically from X-ray images using convolutional neural networks. *Int J Biomed Imaging* 1–9
18. Kusakunniran W, Karnjanapreechakorn S, Siriapisith T, Borwarnginn P, Sutassananon K, Tongdee T, Saiviroonporn P (2021) COVID-19 detection and heatmap generation in chest X-ray images. *J Med Imaging* 8
19. Sarkar A, Vandenhirtz J, Nagy J (2021) Identification of images of COVID-19 from chest X-rays using deep learning: comparing COGNEX VisionPro deep learning 1.0™ software with open source convolutional neural networks. *SN Comput Sci* 2:130
20. Ahmed S, Hossain T, Hoque OB, Sarker S, Rahman S, Shah FM (2021) Automated COVID-19 detection from chest X-ray images: a high-resolution network (HRNet) approach. *SN Comput Sci* 2
21. Narin A, Kaya C, Pamuk Z (2021) Automatic detection of coronavirus disease (COVID-19) using X-ray images and deep convolutional neural networks. *Pattern Anal Appl* 24:1207–1220
22. Swapnarekha H, Behera HS, Roy D, Das S, Nayak J (2021) Competitive deep learning methods for COVID-19 detection using X-ray images. *J Inst Eng (India) Ser B*
23. Zoabi Y, Deri-Rozov S, Shomron N (2021) Machine learning-based prediction of COVID-19 diagnosis based on symptoms. *Digit Med* 4
24. Pandit MK, Banday SA, Naaz R, Chishti MA (2021) Automatic detection of COVID-19 from chest radiographs using deep learning. *Radiography* 27:483–489

25. Ozturk T, Talo M, Yildirim EA, Baloglu UB, Yildirim O, Acharya UR (2021) Automated detection of COVID-19 cases using deep neural networks with X-ray images. *Comput Biol Med* 121:103792
26. Bhattacharya S, Maddikunta PKR, Pham Q-V, Gadekallu TR, Siva Rama Krishnan S, Chowdhary CL, Alazab M, Piran MJ (2021) Deep learning and medical image processing for coronavirus (COVID-19) pandemic: a survey. *Sustain Cities Soc* 102589
27. Liang S, Liu H, Gu Y, Guo X, Li H, Li L, Wu Z, Liu M, Tao L (2021) Fast automated detection of COVID-19 from medical images using convolutional neural networks. *Commun Biol* 4
28. Zhang K, Zuo W, Chen Y, Meng D, Zhang L (2017) Beyond a Gaussian denoiser: residual learning of deep CNN for image denoising. *IEEE Trans Image Process* 26:3142–3155
29. Sandler M, Howard A, Zhu M, Zhmoginov A, Chen LC (2018) MobileNetV2: inverted residuals and linear bottlenecks. In: 2018 IEEE/CVF conference on computer vision and pattern recognition, pp 4510–4520
30. Cohen JP, Morrison P, Dao L, Roth K, Duong TQ, Ghassemi M. COVID-19 image data collection: prospective predictions are the future. [arXiv:2006.11988](https://arxiv.org/abs/2006.11988)
31. <https://www.kaggle.com/andrewmvd/covid19-ct-scans>
32. Mahendran N, Shankar S, Mekala T (2019) EMA-PRBDS: efficient multi-attribute packet rank based data scheduling in wireless sensor networks for real time monitoring systems. *Int J Electron* 106. <https://doi.org/10.1080/00207217.2019.1692244>
33. Mahendran N, Shankar S, Mekala T (2020) LSAPSP: load distribution-based slot allocation and path establishment using optimized substance particle selection in sensor networks. *Int J Commun Syst*. <https://doi.org/10.1002/dac.4343>
34. Natarajan M, Subramanian S (2019) A cross-layer design: energy efficient multilevel dynamic feedback scheduling in wireless sensor networks using deadline aware active time quantum for environmental monitoring. *Int J Electron* 106(1):87–108. <https://doi.org/10.1080/00207217.2018.1501615>

# A Novel Method to Improve the Resolution of FLASH ADC for High-Resolution and High-Speed Applications



Asma Parveen I. Siddavatam, J. M. Nair, and P. P. Vaidya

## 1 Introduction

Analog to digital converters are the heart of many devices in electronic instruments such as oscilloscopes, physical parameter measuring devices, etc., and find use in many fields such as communication systems (optical fiber communication, radar communication, wireless networks, etc.), IoT, embedded systems, spectroscopy systems, instrumentation for data acquisition.

ADC plays a very important role in the processing of all real-world signals such as video, sound, images, sonar, radar, and physical signals like temperature, pressure, etc. A large variety of signals have led to the development of ADC for applications in such diverse fields. The accuracy of analog to digital conversion decides the accuracy of the signal measured.

ADCs are implemented using different types of architectures, but the performance parameters such as sampling rate, resolution, power consumption, etc., that make them suitable for use in various applications. There is also a trade-off between different parameters that depend on the type of architecture and the technology used in fabrication of ICs.

Most commonly sigma-delta, dual slope, and single slope type of ADCs are used in applications where high resolution and the low sampling rate are adequate. Pipeline and SAR type of ADCs are preferred in applications where low to medium resolution and medium to high conversion rate is required [1–4].

---

A. P. I. Siddavatam (✉)

Department of Information Technology, VESIT, University of Mumbai, Mumbai, India  
e-mail: [asma.parveen@ves.ac.in](mailto:asma.parveen@ves.ac.in)

J. M. Nair

VESIT, University of Mumbai, Mumbai, India

P. P. Vaidya

Research and Development, VESIT, University of Mumbai, Mumbai, India  
e-mail: [pp.vaidya@ves.ac.in](mailto:pp.vaidya@ves.ac.in)

Flash ADCs are the fastest and simplest converters and are used as building blocks in other ADC architectures and are preferred in high-speed applications [5].

## 2 Literature Survey

In the last couple of years, the need for digitization in various applications has led to extensive research in analog converters such as ADCs and DACs. The growth in wireless surveillance systems such as healthcare monitoring systems, agriculture, etc., that has led to research of low power, low-resolution flash ADC as battery life in such systems is very important. In surveillance applications, audio and video signals are processed that demand the need of 4–6 bits of ADC resolution [6].

In the past few decades, flash ADCs are commonly used in all three ultrawideband (UWB) frequency ranges. For low data rate applications such as UWB (lower band of UWB) receivers, wireless sensor networks, and RFID, low-resolution flash ADCs are used with a sampling rate of a few hundred MHz [7].

In wireless and mobile applications, variable resolution and variable power flash ADCs are preferred the most as the resolution of the ADC can be varied as per the need that controls power dissipation [8].

In optical, magnetic storage systems, and satellite receivers, low-resolution high-speed analog to digital conversion requires sampling frequency from few hundreds of MHz to a few GHz [12].

The development of new technologies such as driverless automobiles, augmented reality, the Internet of Things, and virtual reality has put a demand on high-speed data transmission and communication that works in the upper band of UWB (24–29 GHz) [7, 11]. The requirement of ADC with high bandwidth, high resolution, and high speed has increased drastically. The cost of special integrated ADCs used in IF/RF/5G mm wave applications is very high [13–21].

At sampling rates greater than approximately 2MS/s to a few GS/s, the resolution decreases by 1 bit every time sampling rate doubles [11]. The biggest challenge in the design of ADCs for such high-frequency applications greater than 500MHz is power consumption. To overcome the problem of power dissipation, various ADC architectures and IC technologies are proposed and implemented [6–14].

Many versions of flash ADC architectures such as combinations of flash-SAR, time-interleaved flash-SAR, time-interleaved flash ADCs, etc. ADCs and improvement in IC technology have resulted in fulfilling the need for UWB applications [3, 12–21].

Flash ADCs are very useful because of their high conversion rate. With increase in the resolution of flash ADC, the number of components gets doubled for improving the resolution by 1 bit that in turn increases the complexity of the circuit and power dissipation [15].

In flash ADC, analog components like comparators and resistors are the major components. The traditional versions of high-speed flash ADCs are half flash or semihalf flash, two-step, multistep flash converters, and pipeline ADCs. The principal

power dissipating elements are resistors, comparators, and the interstage amplifier [5, 11]. However, the reduction of power dissipation in conventional analog to digital converters is achieved by doing modifications to traditional implementation methods and technology.

### 3 Proposed Method to Improve Resolution of ADC

The increase in resolution of flash ADC from 8 bits to 12 bits requires a total of 4096 comparators, 4096 resistors, etc., which increases the power dissipation and is difficult to fabricate using IC technology.

Normally, flash ADCs up to 8–10 bits maximum are available commercially. To improve the resolution of flash ADC, a proposed method does not double the requirement of the number of comparators and other components for every increase in resolution by 1 bit as that of conventional flash ADC.

The new proposed method to improve the resolution of ADC is based on the estimation method described in the literature [22].

The new method requires a  $2^M$  number of comparators, an equal number of resistors, analog switches, and buffers for increasing the resolution by M-bits. Thus to improve the resolution of flash ADC from 8–12 bits, this method will require an additional 16 resistors, 16 comparators, 16 buffers as compared to that of conventional flash ADC that otherwise requires an extra 3840 number of comparators and an equal number of other components. Hence, new method does not add to the complexity of the circuit and power dissipation as that of traditional flash ADC. However, in a new method, the conversion time is slightly increased depending upon the response time of comparators, analog switches, and an amplifier.

A proposed block diagram of the scheme for improvement in resolution of conventional ADC using the proposed method is shown in Fig. 1. The structure of the proposed method is similar to that of flash ADC. For improving the resolution by M-bits, most significant bits (MSBs) of the ADC are decided externally by the proposed method, and the rest of the least significant bits (LSBs) of ADC can be generated by any commercially available flash ADC.

The proposed method consists of a resistor string, and the number of resistors in a resistor string decides the M number of MSB bits. The number of resistors in a resistor string is given by  $2^M$ .

One end of the resistor string is connected with the input voltage,  $V_a$ , and the other end of the resistor string is connected with the difference between input and the reference voltage,  $V_a - V_{ref}$  as shown in Fig. 1. In this technique, the input is tracked continuously because the  $2^M$  resistors of a string act functionally as if  $2^M$  DACs are working simultaneously.

The voltage level at successive resistor taps is equal to  $V_a$ -LSB,  $V_a$ -2LSB, and so on as given in Table 1. By connecting the two ends of the string with  $V_a$  and  $V_a - V_{ref}$ , the voltage at successive resistor tap is equal to input voltage subtracted

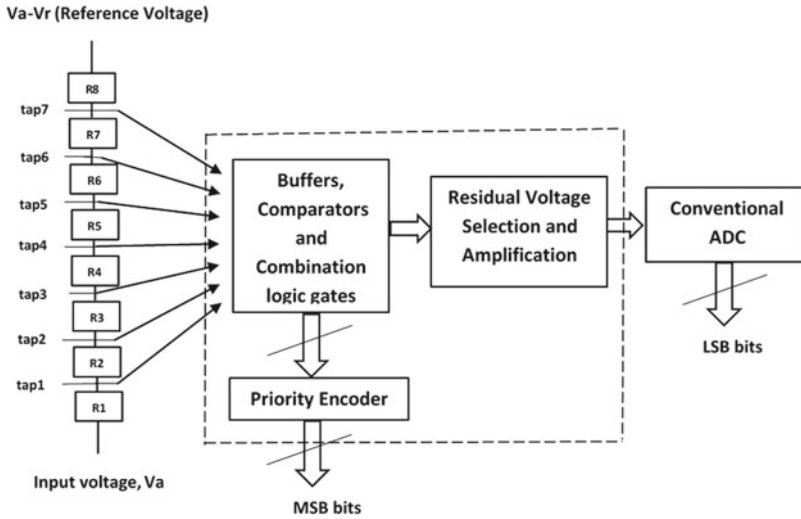


Fig. 1 Block diagram of proposed method to improve the resolution of flash ADC

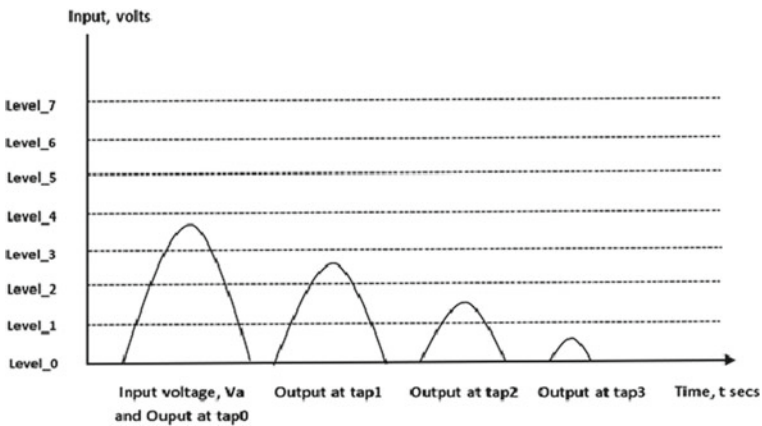


Fig. 2 Diagram to show voltage at resistor taps

by DC voltage equivalent to LSB, 2LSB, and so on, respectively, corresponding to the quantization levels as shown in Fig. 2.

The voltage at successive resistor tap with respect to the input voltage is shown in Fig. 2. The voltage obtained at successive resistor tap is applied to the associated buffer, comparator, and digital logic circuit for selection of voltage from proper resistor tap. The voltage obtained is always less than 1 LSB. It is selected by the corresponding comparator connected to the respective resistor tap. For this purpose, other inputs of the comparator which are not connected to resistor taps are grounded, because of which zero crossover comparator is used.

**Table 1** Voltage level at each resistor tap

Tap no.	Voltage at each resistor tap, when $(V_a - V_r)$ is grounded)	Voltage at each resistor tap, when $(V_a)$ is grounded)	Voltage at each resistor tap, with $(V_a - V_r)$ and $(V_a)$ connected
1	$V_a/8$	$7/8 * (V_a - V_r)$	$V_a - (7/8) * V_r$
2	$V_a/4$	$3/4 * (V_a - V_r)$	$V_a - (3/4) * V_r$
3	$3/8 * V_a$	$5/8 * (V_a - V_r)$	$V_a - (5/8) * V_r$
4	$V_a/2$	$(V_a - V_r)/2$	$(V_a) - (V_r)/2$
5	$5/8 * V_a$	$3/8 * (V_a - V_r)$	$V_a - (3/8) * V_r$
6	$3/4 * V_a$	$(V_a - V_r)/4$	$(V_a) - (V_r)/4$
7	$7/8 * V_a$	$(V_a - V_r)/8$	$(V_a) - (V_r)/8$

The selected voltage is amplified by an interstage amplifier so that it satisfies the full-scale range of the flash ADC. The amplified voltage is finally given to flash ADC for further digitization to get LSBs.

### 3.1 Design Considerations

Important components of the proposed system are resistors, comparators, analog switches, amplifiers, etc.

Selection of components is a very important step in circuit design cycle. The minimum, maximum, and absolute ratings of various parameters of the components selected have to be studied thoroughly.

Basic tolerance and matching of the resistors decide the error associated with A–D conversion and the temperature drift. Hence, low-temperature coefficient metal film resistors are selected.

To get a resolution of 11 bits, 3 bits are decided by the proposed method and rest of the 8 bits by conventional flash ADC. The reference voltage for both the resistor string and conventional ADC can be same or different, but correct relationship has to be maintained.

#### 3.1.1 Selection of Resistors

Since the equivalent resistance is different at each tap, voltage drop due to bias current is different at each tap that adds to static error. This error should be less than  $\pm 1/2$  LSB of the overall resolution. In most of the applications, analog to digital converters with DNL and INL errors of  $\pm 1/2$  LSB are adequate.

In the proposed method, resistors in a resistor string are selected with 0.1% tolerance. The value of least significant bit with 11-bit resolution is 4.88mV as given by Eq. 1.

$$\text{LSB} = \frac{V_{\text{ref}}}{2^M} \quad (1)$$

The maximum tolerable error is  $\pm 1/2 \text{ LSB}$ , and for 11-bit resolution, it is 4.88 mV. For a reference voltage  $V_{\text{ref}}$  of 10 V and  $N$  equal to 3 that is with eight resistors in a resistor string, the value of 3rd bit using Eq. 1 gives 1.25 V. 0.1% of 1/8th full-scale voltage is equal to 1.25 mV. Therefore, resistors in a resistor string with 0.1% tolerance will give a tolerable error of 1.25 mV, which is less than the maximum tolerable error of 4.88 mV that is less by a factor of 42.9%.

The selection of resistors in a resistor string depends on the factors like power dissipation, switching time, and bias current of an operational amplifier used as a buffer connected to each resistor tap. The value of the resistor, “R”, is selected in such a way that it should take care of all the three parameters mentioned. The minimum value of the resistor in a string resistor is dependent on the maximum source current available from the output of high-speed operational amplifier. The maximum source current for most of the high-speed operational amplifiers is in the range of 100mA.

The minimum value of the resistor,  $R_{\text{min}}$ , in a resistor string comprising of eight resistors with a reference voltage of 10 V is 12.5  $\Omega$ . The value of the resistor is maximum at the center tap of the resistor string. There is no upper limit on the value of the resistor with respect to the source current of an operational amplifier. The intrinsic capacitance associated with resistors decides the speed up to which the value of resistors can be used. This problem is addressed by selecting sufficiently low values of resistors which results in a smaller RC time constant, thus improving the speed of operation.

The value of equivalent resistance is maximum at the center tap of resistor string and which is equal to  $2R$ . The small amount of parasitic capacitance in picofarad and large value of the resistor in few  $k\Omega$  make switching time too large and low power dissipation, and on the other hand, small value of the resistor in few ohms results in large power dissipation and small switching time of few nanoseconds. For a value of resistor equal to 25  $\Omega$  and parasitic capacitance of 10 pF, switching time is 2.5 ns.

### 3.1.2 Selection of Comparators and Amplifier

Comparators with low response time and low bias current are used. The analog switches with low on-off time and amplifiers with low settling time is required.

The additional conversion time for the modified ADC is decided by the switching time of resistors, propagation time of comparators, combinational circuits, and inter-stage amplifier. Using presently available high-speed components, the total increase in the conversion time can be limited to 10 ns. The resolution of modified ADC increases by 3 bits at the cost of the increase in conversion time by few nanoseconds.



### 3.1.3 Gain of Interstage Amplifier

The gain requirement of an interstage amplifier depends on the reference voltage of the resistor string and conventional flash ADC.

The interstage gain is given by Eq. 2.

$$G = \frac{V_{r1}}{V_r} 2^M \tag{2}$$

where,  $M$  represents number of bits decided by resistors in a string.  $V_r$  be the reference voltage of resistor string.  $V_{r1}$  be the reference voltage of conventional flash ADC.

This method can be used with any ADC architecture to increase the overall resolution without much increase in conversion time.

## 4 Results of Simulation

The circuit is designed to achieve a resolution of 11 bits, 3 MSBs are given by external peripheral estimation circuit and remaining 8 LSBs by conventional flash ADC. As shown in Fig. 3, it is a resistor string applied with input voltage of 1 V and difference between input voltage and reference voltage of 5 V. To prove the concept, the circuit diagram is simulated using Multisim 14.0. Figures 4, 5, and 6 show buffer, zero crossover comparator, logic gates, and analog multiplexers used for selection of proper voltage from one of the resistor taps. Only one set of buffer, comparator, and logic gates is shown for simplicity. Figure 7 shows interstage amplifier for amplification of residual voltage selected from one of the resistor taps. The obtained amplified residual output after amplifying by a factor of 4 is given to 8-bit ADC for further digitization to get the desired resolution of 11 bits without much increase in the number of components. Figures 9 and 10 show 8-bit ADC and priority encoder for obtaining digital output (Fig. 8). Table 2 shows the results obtained for input voltage of 1 and 2 V peak to peak amplitude, but the circuit is designed only for positive voltages. The reference voltages for resistor string and ADC are 5 V and  $\pm 5$  V, respectively.

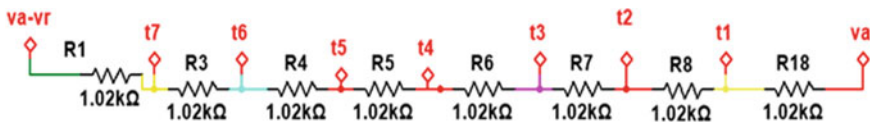


Fig. 3 Resistor string connected with voltage, Va and Va-Vr

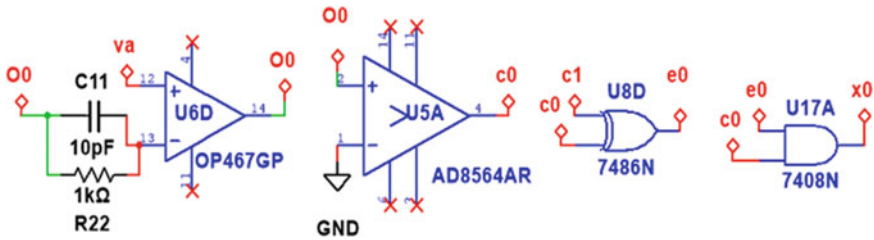


Fig. 4 Buffer, comparator, and logic gates

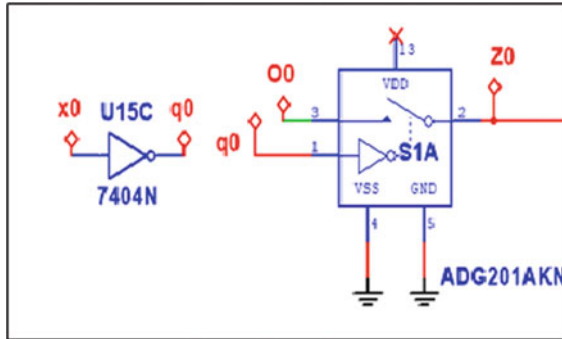


Fig. 5 Circuit for proper voltage selection

The increase in resolution from 8-bit to 11-bit of conventional flash ADC requires additional of 1792 comparators and other equal number of logic gates, whereas the new proposed method requires only 8 comparators and other equal number of logic gates. The requirement of number of comparators/other logic gates is reduced by a factor of 224. Hence, almost by the same factor, there is a reduction in power dissipation, chip area, cost, and complexity. The sub-ranging and pipeline type of ADC rely on ADC–DAC and subtractor for residual voltage generation adding to the conversion time, whereas the new proposed method can be used for residual voltage generation without much increase in the conversion time as it depends on the settling time of RC time constant of resistors in a resistor string, analog switches, priority encoder, etc. The new proposed method increases the conversion time by few nanoseconds by increasing the resolution by  $2^M$ .

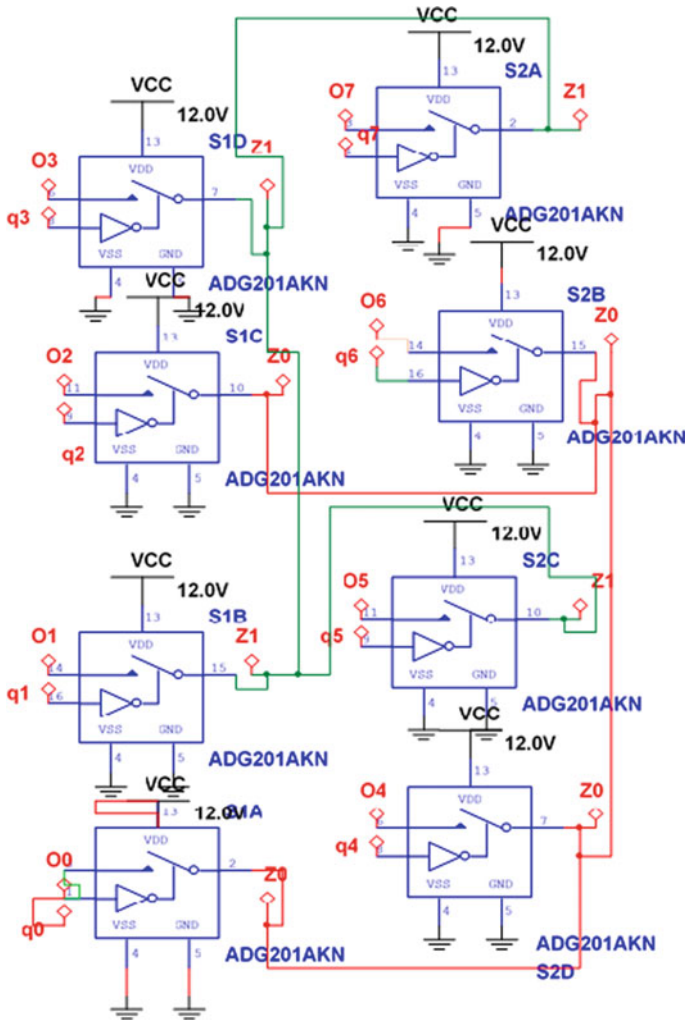


Fig. 6 Circuit for proper voltage selection

### 5 Conclusion and Future Scope

An improvement in resolution by M-bits does not increase exponentially the number of components, decreasing the complexity and power dissipation as compared to traditional flash ADC. There is a reduction in the number of comparators, and other components by a factor of 224 if 8 bit of flash ADC are modified to 11 bit using the proposed method. It eliminates the use of multiple DACs. It does not add up significantly to the conversion time as it is not dependent on DAC for residue generation; hence, it can be used in sub-ranging and pipeline type of ADC for residue generation.

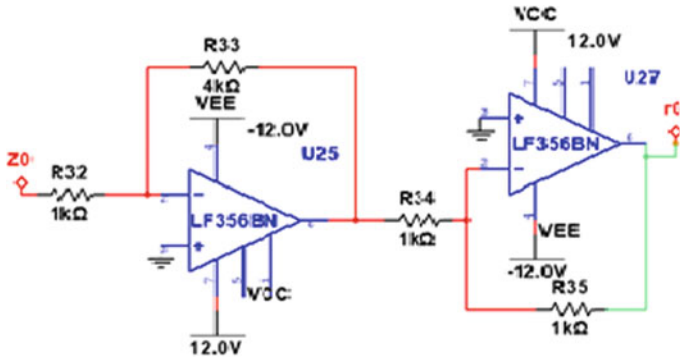


Fig. 7 Interstage amplifier

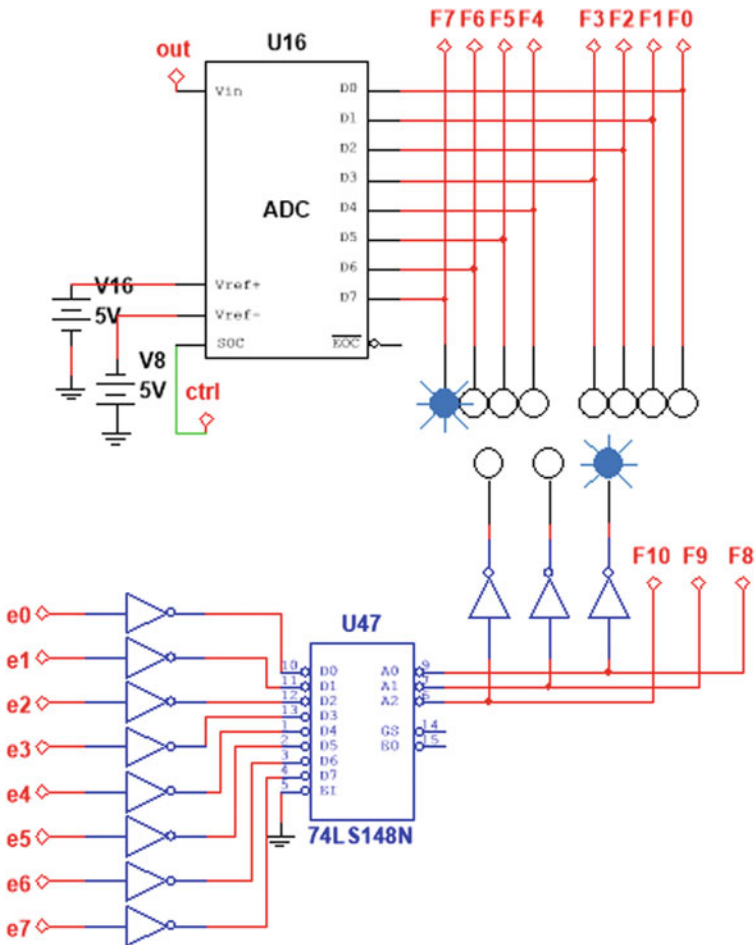


Fig. 8 8-bit ADC and priority encoder

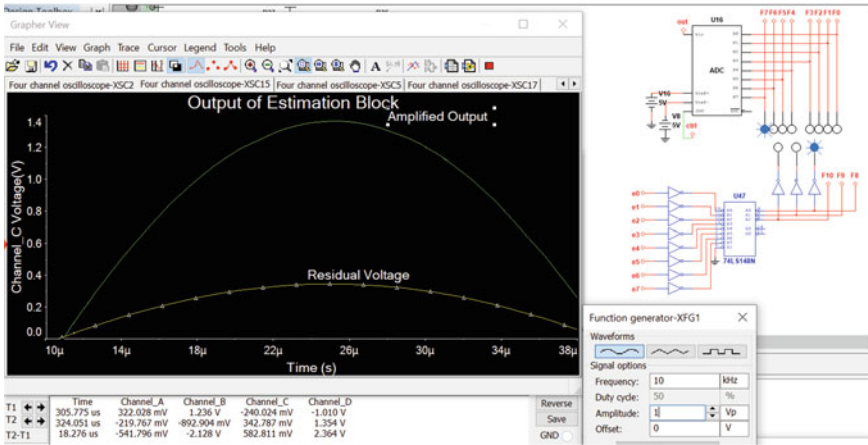


Fig. 9 Residual voltage, amplified voltage, and digital output for 1 Vp-p input

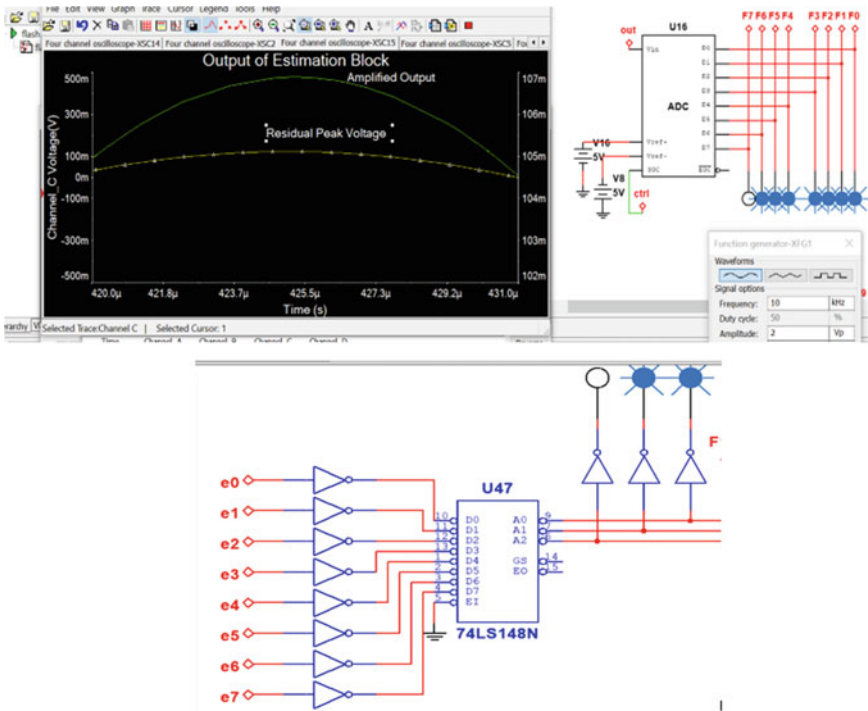


Fig. 10 Residual voltage, amplified voltage, and digital output for 2 Vp-p input

**Table 2** Summary of Figs. 9 and 10

S. no.	Input voltage (Vpp) in volts	Residual voltage, volts	Amplified residual voltage (input to flash ADC), volts	MSBs	LSBs
1	1 V	122.467 mV	476.404 mV	1 h	80 h
2	2 V	344.650 mV	1.364 V	3 h	7E h

This new method can be used as interleaved flash ADC and interleaved flash-SAR ADC. The proposed method along with time-interleaved flash ADCs helps to get high resolution and speed from medium to high frequency of UWB to be used in high speed and high-resolution applications. With the advancement in VLSI technology, it is possible to fabricate  $2^M$  comparators, resistors, and analog switches in a single chip for high values of M, and if the value of M is increased to eight, then the overall resolution of modified flash ADC will be 16 bit.

## References

1. Texas Instruments. <https://www.ti.com/lit/slyt423>. Last accessed 31 Aug 2021
2. MAXIM integrated. <https://www.maximintegrated.com/en/design/technical-documents/tutorials/8/810.html>
3. Zjajo A (2010) Design and debugging of multi-step analog to digital converters. Technische Universiteit Eindhoven. ISBN:978-90-386-2156-2. <https://doi.org/10.6100/IR657787>
4. Adsul J, Nair JM, Vaidya PP (2019) Design and simulation of a new reconfigurable analog to digital converter based on multisim: souvenir of IEEE technically co-sponsored 3rd Biennial international conference on Nascent Technologies in Engineering, ICNTE, January 4-2019
5. Razavi B (2017) The Flash ADC [A Circuit for All Seasons]. IEEE Solid-State Circuit Mag 9(3):9–13. <https://doi.org/10.1109/MSSC.2017.2712998>
6. Siddharth RK, Kumar YBN, Vasantha MH (2020) Power saving scheme for process corner calibrated standard cell based flash ADC in wireless surveillance applications. SN Comput Sci 1:310. <https://doi.org/10.1007/s42979-020-00328-3>
7. Patnaik SK, Banerjee S (2014) Design of a 500 MHz, 4-Bit low power ADC for UWB application. Int J Electron Signal Syst 4(2). ISSN: 2231-5969
8. Veeramachanen S, Mahesh Kumar A, Tummala V, Srinivas MB (2009) Design of a low power, variable-resolution flash ADC. In: 22nd international conference on VLSI design, 2009, pp 117–122. <https://doi.org/10.1109/VLSI.Design.2009.62>
9. <https://dewesoft.com/daq/types-of-adc-converters>. Types of A/D converters [Updated 2020] | Dewesoft
10. Tan PBY, Kordesch AV, Sidek O (2004) Simplified half-flash CMOS analog-to-digital converter. NSTI-Nanotech 2. [www.nsti.org](http://www.nsti.org), ISBN 0-9728422-8-4
11. Walden RH (1999) Analog-to-digital converter survey and analysis. IEEE J Select Areas Commun 17(4):539–550. <https://doi.org/10.1109/49.761034>
12. Sung BRS, Cho S, Lee C, Kim J, Ryu S (2009) A time-interleaved flash-SAR architecture for high speed A/D conversion. In: IEEE international symposium on circuits and systems, pp 984–987. <https://doi.org/10.1109/ISCAS.2009.5117923>

13. Talekar SG, Ramasamy S, Lakshminarayanan G, Venkataramani B (2009) A low power 700MSPS 4bit time interleaved SAR ADC in 0.18um CMOS: TENCON 2009–2009 IEEE Region 10 conference, 2009, pp 1–5. <https://doi.org/10.1109/TENCON.2009.5395819>
14. Jia H, Guo X, Zheng X, Xu X, Wu D, Zhou L, Wu J, Liu X (2020) A 4-bit 36 GS/s ADC with 18 GHz analog bandwidth in 40 nm CMOS process: electronics 2020, vol 9, no 10, p 1733. <https://doi.org/10.3390/electronics9101733> (Special Issue Analog/Digital Mixed Circuit and RF Transceiver Design)
15. Synopsys. <https://www.synopsys.com/designware-ip/technical-bulletin/5g-data-converters>
16. <https://www.everythingrf.com/News/details/8351-ADI-s-RF-Converters-for-4G-Base-Stations-Triple-Call-Capacity-Data-Throughput>
17. Analog. <https://www.analog.com/en/parametricsearch/11502>
18. <https://www.everythingrf.com/News/details/8277>
19. <https://www.analog.com/en/applications/technology/5g>
20. <https://www.design-reuse.com/news/47744/vervesemi-data-converters-5g-samsung-8nm>
21. <https://www.edn.com/the-abcs-of-interleaved-adcs/>
22. Siddavatam API, Vaidya PP, Nair JM (2019) A New Method for improving resolution of nuclear ADC for high resolution spectroscopy system. In: International conference on nascent technologies in engineering (ICNTE) pp 1–4. <https://doi.org/10.1109/ICNTE44896.2019.8945973>

# FPGA Implementation of Radix-2 Pipelined FFT Algorithm for High-throughput Applications



Rajasekhar Turaka, S. Ravi Chand, Tavanam Venkata Rao,  
and V. Kumara Swamy

## 1 Introduction

Discrete Fourier Transform (DFT) converts a finite length sequence of regularly spaced samples of a function into a same length sequence of equally-spaced samples of the discrete time Fourier transform, which is a complex valued function of frequency. Fourier transform is commonly used for filtering. The multiply operation can be performed by multiplying the  $x[n]$  and  $h[n]$  and take inverse Fourier transform of product which can also be used for correlation by reversing one of the sequences. We require  $N*N$  operations to calculate Fourier transform by direct evaluation of DFT formula. FFT algorithm has the advantage of reducing number of operations to  $N \log N$  due to its symmetry properties [1]. Discrete transform technique plays a vital role in digital signal processing (DSP) applications particularly in many advanced digital communication systems. The real-time computation of DFT is very important for applications like orthogonal frequency division multiplexing (OFDM), long-term evolution (LTE), digital video and audio broadcasting systems (DVB) and (DAB), digital subscriber line (DSL), asymmetric DSL, and very high-speed DSL [2]. There are two well-known architectures, namely memory-based and pipelined are used to design the FFT algorithm efficiently on hardware. Low area and low power are the two major advantages of memory-based architecture, but it suffers from the memory conflicts and long latency [3]. In order to overcome this long latency, a number of pipelined architectures were proposed, as pipelined architectures have a high-throughput rate [4]. Parallelism is a main thing that offers the

---

R. Turaka (✉) · S. Ravi Chand  
ECE, Nalla Narasimha Reddy Education Society's Group of Institutions, Hyderabad, Telangana,  
India  
e-mail: [valleyraj7311@gmail.com](mailto:valleyraj7311@gmail.com)

T. V. Rao · V. Kumara Swamy  
ECE, Sreenidhi Institute of Science and Technology, Hyderabad, Telangana, India



solution to the VLSI DSP architectures. It can be applicable in different stages of the implementation of DSP architectures [5].

In this paper, a high-throughput pipeline-based single delay feedback fast Fourier transform (SDF FFT) architecture is proposed. FPGA implementation of SDF is made very clear interpretation. The SDF FFT core not only reduces the critical path but also reduces the hardware usage when the core is mapped to two different Xilinx FPGA families [6]. This architecture consists of a sequence of blocks where each block process one sample per every clock cycle. Each block in the core has a butterfly unit and a rotator. The butterfly unit computes additions and multiplications whereas the rotator allocates the rotations in the plane by the use of twiddle factors [7]. The proposed design is very flexible and generic. The proposed design is implemented for vertex-4, Spartan 3A DSP, Artix-7 low voltage, and Zynq boards. For both area efficient and high throughput, the vertex-4 and Spartan 3A DSP devices can be used.

The organization of the paper is divided into following sections. Section 2 deals with brief description of the FFT algorithm. Basic structure of pipelined FFT architecture is discussed in Sect. 3. Section 4 deals with the proposed pipelined FFT architectures. Results and discussion are discussed in Sect. 5, and finally, Sect. 6 concludes the paper.

## 2 Overview of FFT Algorithm

The  $N$ -point discrete time signal  $x[n]$  can be represented in the frequency domain as

$$X(k) = \sum_{n=0}^{N-1} x(n)W_N^{nk}, 0 \leq k \leq N - 1$$

where  $W_N = e^{-j2\pi/N}$  is called the twiddle factor. Here,  $X(K)$  and  $x(n)$  are called the output data and input data, respectively.

It is known that by the use recursive factorization DFT, the radix-2 FFT can be derived. To compute the  $N$ -point DFT of a sequence, it requires  $O(N^2)$  number of complex multiplications, and for the same sequence, FFT algorithm requires  $O(N \log N)$  multiplications. The decimation-in-time (DIT) radix-2 FFT butterfly architecture is shown in Fig. 1.

To overcome the difficulty of calculation of the DFT, it is known to utilize the properties of the twiddle factor. By the usage of the FFT algorithm, we can subsequently lessen the computational complexity from order  $(N * N)$  to order  $(N \log N)$ . This regularity of the FFT algorithm is the main reason for the vast implementation in VLSI. It is also well known to be able to be used in the OFDM systems. The major difference between the DIT and decimation-in-frequency (DIF) would be the order of input and output data. In the case of DIT, the input data entered in bit reverse order, and the output data are in regular order whereas for the DIF works opposite to DIT.

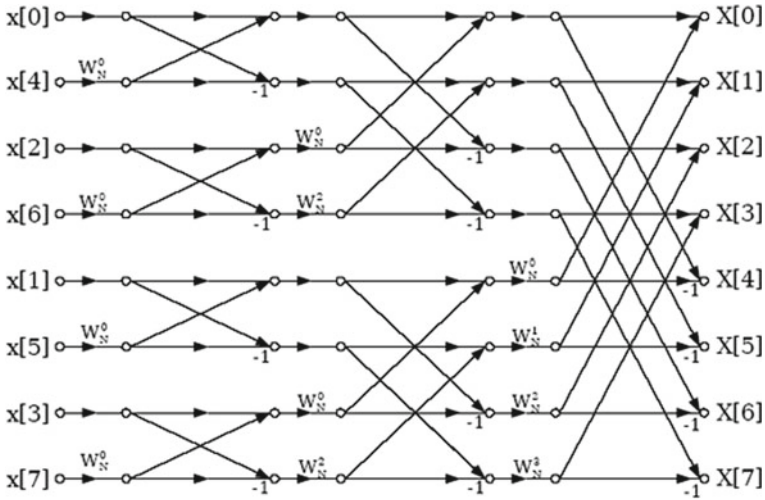


Fig. 1 DIT radix-2 FFT butterfly

### 2.1 High-radix FFT

The main reason for the existence of the high-radix FFT algorithm is to decrease the computational complexity of the algorithm. It will serve in the way of reducing the amount of data and multiply and add operations. It can be able to perform the reverse of the regular conventional pipelined implementation. In the figure, the input sequence is in bit reverse order. The radix-4 implementation of the DFT can greatly reduce the number of multiplications from  $N * N$  to  $N(\log N - 1)$  that is even lesser than that of the radix-2 FFT. But the main drawback that is within it is that the complexity of the architecture is greatly increased.

### 2.2 Split Radix FFT

The split radix has the main advantage in terms of the complex multiplications for  $N$ -point DFT. The split radix works in a way that it is divided into different radix that can be radix-4 or radix-8. So as a result of this their shape gets irregular and due to this main reason, it would be very hard to design the pipelined architecture. The split radix contains lesser number of multipliers and adders compared to that to the other radix. Spilt radix algorithm is good for low-power FFT processor applications [8]. A 16- point DIT FFT butterfly architecture is shown in Fig. 2.

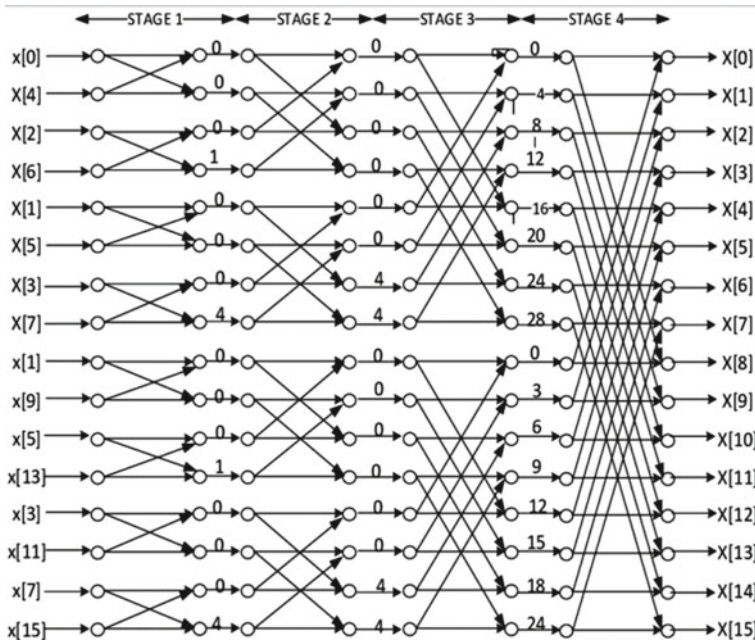


Fig. 2 16-point radix-2 DIT FFT butterfly architecture

### 3 Pipelined FFT Architectures

Due to the tremendous growth in the modern digital communications and television broadcasting systems, there is a need of high-performance dedicated FFT processor [9]. The pipeline FFT processor is an effective architecture that can satisfy the real-time spectral analysis requirements of many sonar and radar systems. The method of operation of the FFT can be interpreted using the basic mathematical operations [10]. They also have their own pros and cons, respectively.

To meet the various performance requirements, the processors are highly pipelined. In order to have the high throughput, pipelined architectures consumes more area. So that, the hardware cost is increased. The each stage of the architecture consists of butterfly element, delay element, and the complex multiplier as shown in Fig. 3.

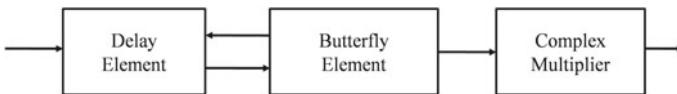
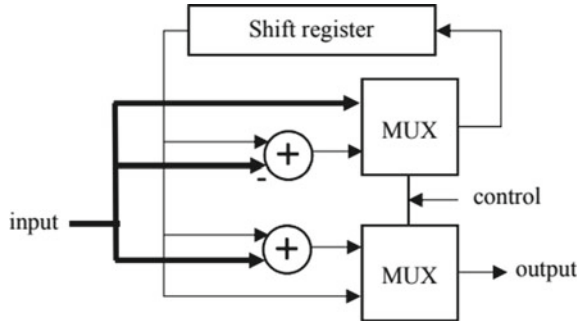


Fig. 3 Basic framework of pipelined FFT

**Fig. 4** Working modes of radix-2 PE



### 3.1 Processing Element

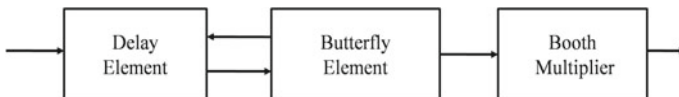
The processing element (PE) is the main part of the FFT architecture, and the main data processing is done in this stage. It is also known as the butterfly element. The computation of the PE is shown in Fig. 4. The PE works in two modes of operation, one is the operation mode, and the second is commutator mode. When the outputs of an adder are directly passed to multiplexers, then PE works in operation mode, and when the inputs are directly given to the multiplexers, then it operates in commutator mode, and the mode of operation is decided by the control signal.

In the operation mode, the real processing of data takes place that can be mainly the addition or subtraction. Whereas in the commutator mode, it only passes the data from inputs to the outputs.

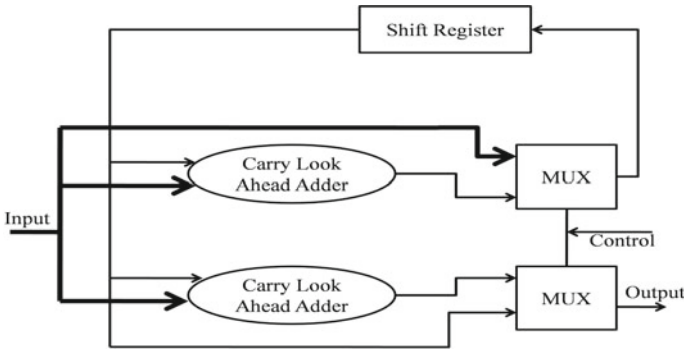
## 4 Proposed Pipelined FFT Architectures

In proposed FFT architecture, we have used booth multiplier for performing the multiplication operations instead of complex multiplier which is shown in Fig. 5.

The multiplication operations are done by the booth multiplier, whereas the addition operations are performed by the adders. Carry look-ahead adder is being used in the proposed method for the better results. The processing element is replaced by a carry look-ahead adder with MUX and shifting operations then forwarded to the complex multiplier for further operations.



**Fig. 5** Modified framework of pipelined FFT



**Fig. 6** Modified working modes for a radix-2 PE

Figure 6 shows the implementation of the carry look-ahead adder in radix-2 processing element. Thus, the change of the adder and multiplier adds an advantage to the processing element with high throughput.

### 5 Results and Discussion

In this, Table 1 makes a clear interpretation of the proposed architecture and its complete functionality when there are different types of devices used. The proposed model was modeled in the Verilog module and its primary functional verification, and synthesis was done by using the Xilinx software, i.e., Xilinx ISE 14.5 suite by considering different FPGA boards such as Virtex-4 XC4VFX12, Spartan-3A XC3SD1800A, Artix-7 low voltage XC7A100TL, and Zynq XC7Z010 devices. The simulation results of proposed architecture are shown in Fig. 7.

From Table 1, we can make note that the proposed model outperformed the existing model in terms of both the time delay and also the area of utilization of the device. In

**Table 1** Comparison of existed and proposed architectures

Device name	Architecture	Slice LUTs	Delay (ns)	Frequency (MHz)
Virtex-4	Existed	296 out of 5472	21.745	45.98
	Proposed	178 out of 5472	17.114	58.43
Spartan 3A DSP	Existed	300 out of 16,640	36.822	27.15
	Proposed	179 out of 16,640	33.691	29.68
Artix-7 low voltage	Existed	1331 out of 63,400	18.017	55.50
	Proposed	1715 out of 63,400	15.827	63.18
Zynq	Existed	1331 out of 17,600	12.067	82.87
	Proposed	1715 out of 17,600	10.552	94.76

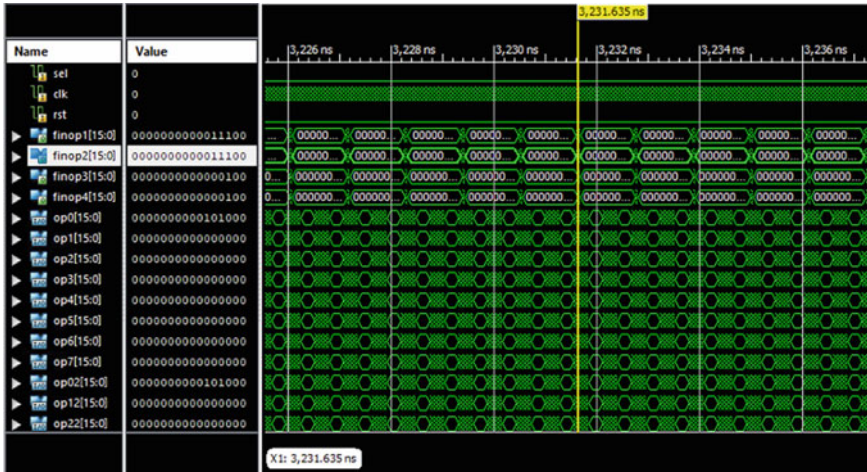


Fig. 7 Simulation results for proposed pipelined FFT architecture

the FPGA implementation, frequency improved of 27% for Virtex-4, 9% for Spartan-3A, 13% for Artix-7, and 14% for Zynq devices as compared with the existing architecture. To check the compatibility and applications, the proposed model is simulated in different Xilinx FPGA boards.

## 6 Conclusion

In this, new configurable radix-2 single path delay feedback pipelined architecture was proposed. The proposed architecture gives the best performance in terms of throughput. This is obtained by making the use of adders and multipliers that can have the lesser time delay than the existing conventional model. As such that it can be widely applicable in the high-throughput applications. But the main drawback observed is the higher consumption of the area in some cases.

## References

1. Cooley JW, Tukey JW (1965) An algorithm for the machine calculation of complex Fourier series. *Math Comput* 19(90):297–301
2. Turaka R, Satya Sai Ram M (2019) Low power VLSI implementation of real fast Fourier transform with DRAM-VM-CLA. *Microprocess Microsyst* 69:92–100
3. Wold EH, Despain AM (1984) Pipeline and parallel-pipeline FFT processors for VLSI implementation. *IEEE Trans Comput* C-33(5):414–426
4. Bergland GD (1969) Fast Fourier transform hardware implementations—an overview. *IEEE Trans Audio Electroacoust* AU-17:104–108

5. Qureshi F, Gustafsson O (2017) Efficient FPGA mapping of pipeline SDF FFT cores. *IEEE Trans Very Large Scale Integr (VLSI) Syst* 25(9):2486–2497
6. Garrido M, Andersson R, Qureshi F, Gustafsson O (2016) Multiplierless unity-gain SDF FFTs. *IEEE Trans Very Large Scale Integr (VLSI) Syst* 24(9):3003–3007
7. Duhamel P, Hollmann H (1984) Split-radix FFT algorithm. *Electron Lett* 20(1):14–16
8. Hung CP, Chen SG, Chen KL (2004) Design of an efficient variable-length FFT processor. In: *International symposium on circuits and systems IEEE (ISCAS)*, Vancouver, BC, Canada, vol 2, pp 833–836
9. Groginsky HL, Works GA (1970) A pipeline fast Fourier transform. *IEEE Trans Comput* C-19(11):1015–1019
10. Shokry B, Dessouky M (1892) A dynamically configurable-radix pipelined FFT algorithm for real time applications, 2017. *J. Clerk Maxwell, A treatise on electricity and magnetism*, 3rd edn., vol 2. Clarendon, Oxford, pp 68–73

# Design of High Speed Approximate Carry Select Adders Using RCPFA



T. Phaneendra, A. Anitha, and Rajasekhar Turaka

## 1 Introduction

Adders are the key blocks in any electronics devices and circuits. Adders are used for several applications like ALU, microprocessors, design of filters in DSP applications, etc. In case of precise results, cannot compromise on the design parameters like delay and power. Due to these difficulties in the delay and power requirements, sometimes have to compromise in the summation for unreliable results. It may lead to approximate results. This type of approximation is called approximate computation [1]. Approximate computing is also used in several applications like deep learning, image and DSP applications. This technique is used in adders to approximate either sum or carry to improve the performance of the adder according to the design parameters. Approximate adders were designed with mixing of several cells or gates for efficient outputs. By using approximate computing, reverse carry propagation full adder (RCPFA) is one method to improve the characteristics of the approximate adders. Based on improvements in different parameters and errors, these RCPFAs are used to design approximate carry select adders for high speed applications.

---

The original version of this chapter was revised: Author provided affiliation corrections have been incorporated. The correction to this chapter is available at [https://doi.org/10.1007/978-981-19-1111-8\\_87](https://doi.org/10.1007/978-981-19-1111-8_87)

T. Phaneendra · A. Anitha  
ECE, Velagapudi Ramakrishna Siddhartha Engineering College, Vijayawada, Andhra Pradesh, India  
e-mail: [phaneendra199705@gmail.com](mailto:phaneendra199705@gmail.com)

A. Anitha  
e-mail: [anithaa@vrsiddhartha.ac.in](mailto:anithaa@vrsiddhartha.ac.in)

R. Turaka (✉)  
ECE, Nalla Narasimha Reddy Education Society's Group of Institutions, Hyderabad, Telangana, India  
e-mail: [valleyraj7311@gmail.com](mailto:valleyraj7311@gmail.com)

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2022, 761 corrected publication 2022

T. Sengodan et al. (eds.), *Advances in Electrical and Computer Technologies*, Lecture Notes in Electrical Engineering 881, [https://doi.org/10.1007/978-981-19-1111-8\\_58](https://doi.org/10.1007/978-981-19-1111-8_58)



## 2 Prior Approximate Adders

In RCPFA, carry propagation is opposite direction to normal propagation and errors that occurred by approximation are reduced. Based on the approximation, RCPFA adders are three models (RCPFA-I, II and III). AOI and OAI gates are used to construct the RCPFA adders for reduced count of transistors. RCPFA adders have additional forecasting signal ( $F$ ) which is used to decide the output sum and carry in some states of the three RCPFA adders proposed. Truth tables, logic implementations of RCPFA adders and error comparison of error analysis, design parameters to prior approximate adders are shown in [2]. In this work, several existing adders are used as approximate computing to approximate either sum or carry to reduce the logical level or gate level implementation. In Error Tolerant Adder (ETA), carry was ignored in lower part of the summation [3]. In Lower part OR Adder (LOA) lower part of the adder blocks were replaced with OR gates to approximate the summation [4]. In Approximate Mirror Adder (AMA), approximation was achieved by removing the selective adder for reducing the gate count [5]. In Approximate XOR/XNOR-based Adder (AXA), XOR or XNOR logic was used to perform the summation operation [6]. In Transmission Gate Adder (TGA), transmission gates are used for logic implementation for approximate summation [7]. In Inexact Adder (InXA), approximation was done by either in carry or sum based on the summation [8]. In Speculative Adder, carry was predicted by prediction circuits to gain the carry propagation delay [9]. In Generic Accuracy Configurable Adder (GeAr), an error detection and correction units are placed with adder blocks for accurate results [10]. All these approximate adders are composite adders to upgrade results in terms of structural, error weight and probability. In that RCPFA is new logical implementation based on weightage of the inputs and outputs which results in reduced gate count. Different types of errors that occurred in approximate adders and comparison of those errors were discussed in [2–10]. Power, propagation delay, transistor count and power delay product are the key characteristics of the approximate adders to describe the efficiency and performance of adders.

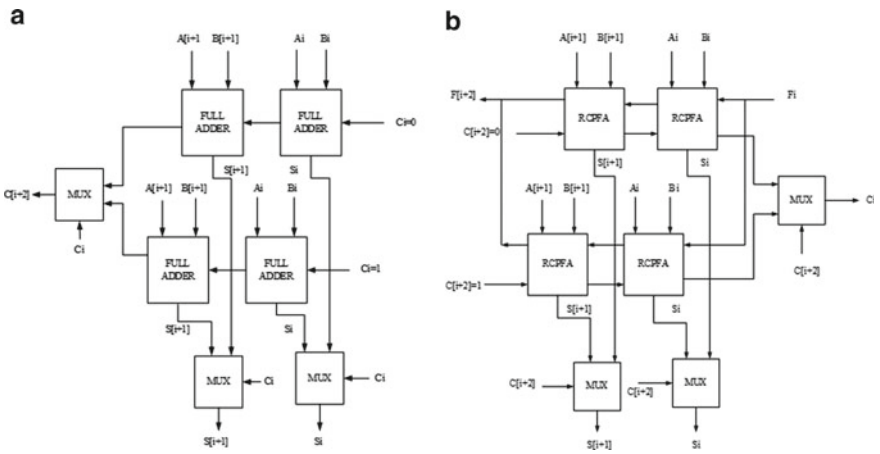
## 3 Conventional CSLA with 28Transistor Full Adder

Carry select adder (CSLA) is one of the high speed adders compared to other exact adders. In CSLA structure output carry and sums are calculated with two Ripple Carry Adder (RCA) blocks with carry input as 0 and carry input as 1. Multiplexer is used to select these results either 0 s result or 1 s result based on the selection line of the multiplexer. This selection line is the output carry of the previous block. Therefore output carry of the previous block will decide the outputs (Sums and carry) of the current block. Based on the 1-bit adder delay and multiplexer delay (From Table 1), number of selection blocks of the CSLA will be decided. For 8-bit CSLA, 4 blocks are required with 2-bit RCA at initial block. Each block performs 2-bit addition. The

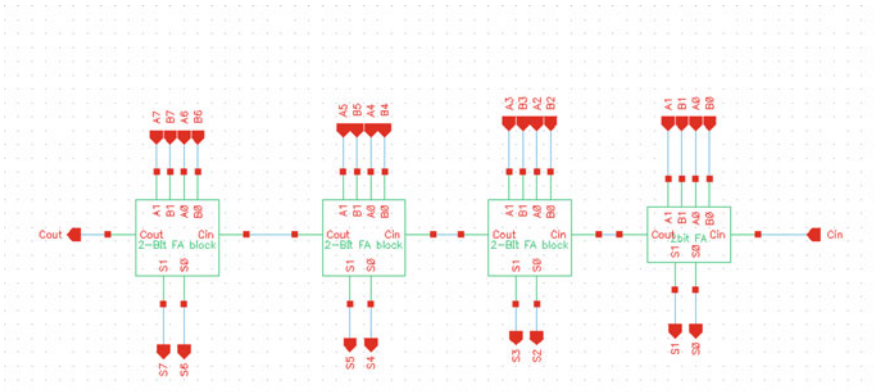
**Table 1** Delays of 1-bit adders and multiplexer

Cells	Delay (ps)
1-bit 28 T Full adder	45
1-bit RCPFA-I	76
1-bit RCPFA-II	33
1-bit RCPFA-III	50
Multiplexer	14.5

block diagram of 2-bit carry select block of CSLA is shown in Fig. 1. The 8-bit carry select adder is depicted in Fig. 2. Equation for finding the number of blocks in CSLA is given by



**Fig. 1** a Conventional and b proposed (with RCPFA) 2-bit carry select blocks of CSLA



**Fig. 2** 8-bit Carry select adder (CSLA)

$$B = \sqrt{\frac{nT_1}{T_2}} \tag{1}$$

where  $B$  is the number of selection blocks.  $n$  is the size of the adder.  $T_1$  is the delay of 1-bit adder block.  $T_2$  is the delay of multiplexer (Fig. 2).

### 4 Approximate Carry Select Adders

Reverse carry propagate full adder (RCPFA) is an approximate adder where the carry propagation happens in reverse direction. RCPFA introduces the forecasting signal  $F$ . RCPFA cell generates output forecasting signal ( $F_{i+1}$ ) based on the inputs ( $A$  and  $B$ ) which is input forecasting signal ( $F_i$ ) for next adder cell. The input  $F_i$  used to generate the output sum and carry of current block. Propagation of forecasting signal is normal (LSB to MSB) and output forecasting signal is taken from either of the two RCA blocks because both adders (RCPFA) have same inputs. Reverse carry propagation is used for reduction of the errors in approximate adders as well as reduction in some of the design parameters compared to the exact 1-bit adder. From Table 1, delays of the 1-bit adder cells and multiplexer are shown. The selection blocks of the conventional and proposed carry select adders are decided based on the delays of 1-bit adder cell and multiplexer. The proposed 2-bit selection blocks of CSLA with RCPFA. In that structure, two RCA with RCPFA blocks are used. One of the two RCA blocks operate with carry input as 0 and another RCA block operate with carry input as 1. Output carry of previous blocks ( $[i + 2]$  block) is given as selection line to the multiplexer.

#### 4.1 Carry Select Adder with RCPFA-I

Using RCPFA adder cell instead of exact 1-bit full adder cell, direction of carry propagation is reverse. Carry select adder with RCPFA-I which is shown in Fig. 3 need 6 selection blocks based on the delays from Table 1. Selection blocks are divided

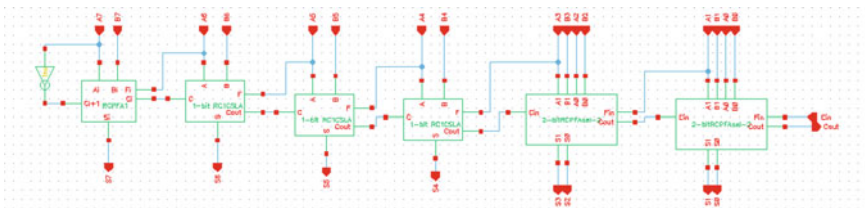


Fig. 3 8-bit CSLA with RCPFA-I

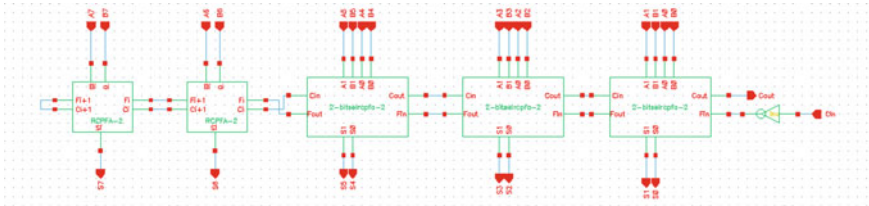


Fig. 4 8-bit CSLA with RCPFA-II

into three 1-bit selection blocks and two 2-bit selection blocks. At most significant part, 1-bit RCPFA-I placed. Because of the reverse carry propagation, input carry of most significant bit adder is taken as input of the  $A$  as per the operation of RCPFA-I adder. This is propagated from  $A_7$  to  $S_0$  (critical path) and  $C_{out}$  is ignored because there are no further blocks to perform addition.

### 4.2 Carry Select Adder with RCPFA-II

Carry select adder with RCPFA-II shown Fig. 4 needs 4 selection blocks based on the delays from Table 1. Selection blocks are divided as three 2-bit selection blocks and a 2-bit RCA block with RCPFA-I adder. Because of the reverse carry propagation, input carry of most significant bit adder is taken as output forecasting signal ( $F_8$ ).  $F_8$  is AND function of input  $A_7$  and  $B_7$  as per the operation of RCPFA-II adder. The output forecasting signal  $F_8$  is propagated to  $S_0$  (critical path).

### 4.3 Carry Select Adder with RCPFA-III

Carry select adder with RCPFA-III shown in Fig. 5, the selection blocks are 4 based on the delays from the Table 1. Selection blocks are divided as three 2-bit selection blocks and 2-bit RCA with RCPFA-II adder. Because of the reverse carry propagation,

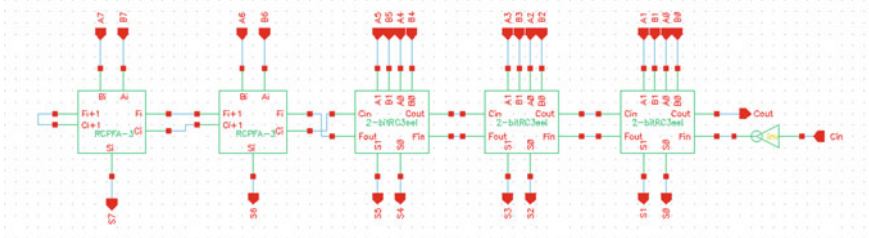


Fig. 5 8-bit CSLA with RCPFA-III

input carry of most significant bit adder is taken as output forecasting signal ( $F_8$ ).  $F_8$  is OR function of input  $A_7$  and  $B_7$  as per the operation of RCPFA-III adder. The output forecasting signal  $F_8$  is propagated to  $S_0$  (critical path) and  $C_{out}$  is ignored because there are no further blocks to perform addition.

## 5 Results and Discussion

Some of the parameters that define the performance of the adder are compared to 8-bit conventional and proposed CSLA adders. From Table 1 delays of 1-bit adder cells and multiplexer are discussed. The delay of RCPFA-II adder cell is reduced by 26% of delay compared to conventional FA and delay of the RCPFA-III is almost equal to conventional adder. But RCPFA-I has more delay compared to conventional one because of the critical path of the carry propagation. Whenever the delay of adder cell is closer to multiplexer delay then, more efficient carry select adder can be designed. RCPFA-II is twice the delay of the multiplexer delay where conventional is thrice the delay of the multiplexer delay. So, 8-bit CSLA adders are designed for high speed operations.

From Table 2, it can be observe that proposed 8-bit CSLA with RCPFA-I and RCPFA-II are improvement of 18 and 36% in delay compared to conventional 8-bit CSLA. Total power of proposed adder cells are high compared to 8-bit conventional CSLA because twice of the adder cells are required to design to improve delay. 8-bit CSLA with RCPFA-I has closer value of power compared to conventional. Power delay product specifies the efficiency of the adder and in 8-bit CSLA with RCPFA-I give power delay product with 12% improvement when compared to that of 8-bit

**Table 2** Font sizes of headings

	Average transient power (uW)	Average DC power (uW)	Total power (uW)	Delay	Power delay product	Average transient power (uW)
8-bit CSLA	0.325	0.88	1.2	211	2.5	446
8-bit CSLA with RCPFA-I	0.81	0.46	1.27	172	2.18	464
8-bit CSLA with RCPFA-II	135.1	0.35	135.45	135	182	336
8-bit CSLA with RCPFA-III	17.64	0.26	71.9	250	44	336

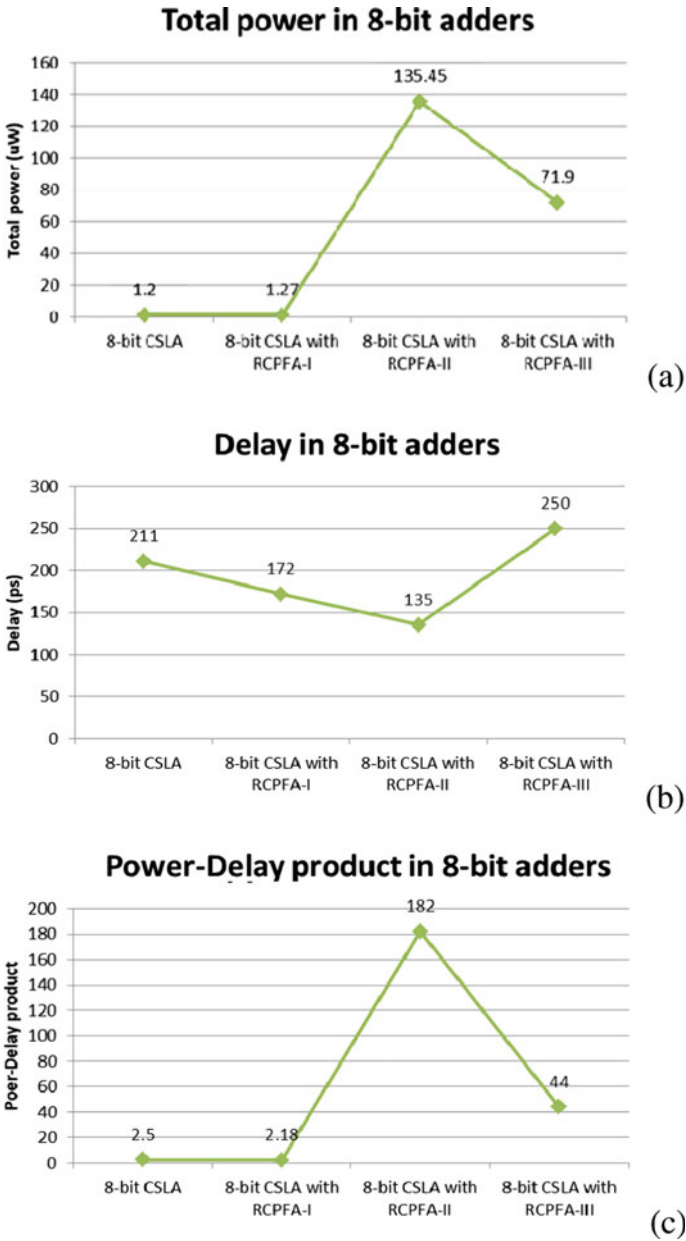
Table captions should always be positioned *above* the tables comparison table of 8-bit conventional and proposed CSLA

CSLA. Transistor count is reduced to 24% in 8-bit CSLA with RCPFA-II and III when compared to that of conventional 8-bit CSLA.

Comparison graphs of 8-bit proposed adders and conventional adder with the parameters mentioned in Table 2 are shown in Fig. 6b. Observation from Fig. 6 is 8-bit approximate reverse carry select adders had better delay performance compared to the conventional CSLA. From Fig. 6c, 8-bit CSLA with RCPFA-I has better efficiency in terms of power delay product when compared to that of conventional CSLA.

## 6 Conclusion

Performance of 8-bit CSLA with RCPFA approximate adders is tested and the simulation results were showed in this work. For high speed applications CSLA with RCPFA-I and CSLA with RCPFA-II can be used. For better energy CSLA with RCPFA-I is used. CSLA with RCPFA-II and CSLA with RCPFA-III adders can be used for small area applications because of less transistor count. These inexact adders can be used in several applications like deep neural networks, image processing and DSP applications.



**Fig. 6** Comparison graphs of parameters, **a** total power, **b** delay and **c** power delay product of 8-bit conventional and proposed CSLA

## References

1. Moreau T, Sampson A, Ceze L (2015) Approximate computing: Making mobile system more efficient. *IEEE Pervasive Comput* 14:1070–1081
2. Pashaeifar M, Kamal M, Kusha AA, Pedram M (2018) Approximate reverse carry propagate adder for energy-efficient DSP applications. *IEEE Trans Very Large Scale Integr (VLSI) Syst* 26:2530–2541
3. Zhu N, Goh WL, Zong W, Yeo KS, Kong ZH (2010) Design of low-power high-speed truncation-error-tolerant adder and its application in digital signal processing. *IEEE Trans Very Large Scale Integr (VLSI) Syst* 18:1225–1229
4. Mahdiani HR, Ahmadi A, Fakhraie SM, Lucas C (2010) Bio-inspired imprecise computational blocks for efficient VLSI implementation of soft-computing applications. *IEEE Trans Circuits Syst I Regul Pap* 57:850–862
5. Gupta V, Mahapatra D, Raghunathan A, Roy K (2013) Low power digital signal processing using approximate adders. *IEEE Trans Comput Aided Des Integr Circuits Syst* 32:124–137
6. Yang Z, Jain A, Liang J, Han J, Lombardi F (2013) Approximate XOR/XNOR-based adders for inexact computing. In: *Proceedings of 13th IEEE international conference on nanotechnology, China*, pp 690–693
7. Yang Z, Han J, Lombardi F (2015) Transmission gate-based approximate adders for inexact computing. In: *Proceedings of IEEE/ACM international symposium on nanoscale architectures, USA*, pp 145–150
8. Almurib HAF, Kumar TN, Lombardi F (2016) Inexact designs for approximate low power addition by cell replacement. In: *Proceedings of DATE conference on design, automation & test, Europe*, pp 660–665
9. Lin IC, Yang YM, Lin CC (2015) High-performance low-power carry speculative addition with variable latency. *IEEE Trans Very Large Scale Integr (VLSI) Syst* 23:1591–1603
10. Shafique M, Ahmad W, Hafiz R, Henkel J (2015) A low latency generic accuracy configurable adder. In: *Proceedings of 52nd ACM/EDAC/IEEE conference on design automation, USA*, pp 1–6



# Data Driven Approach to Achieve Coordinated Charging Among Electric Vehicles



Venkata Nikitha Machineni, Korada Sri Vardhana,  
and Vaithiyanathan Dhandapani 

## 1 Introduction

The global EV market is estimated to grow from approximately 80 lakhs in 2019 to 5 crores by 2025 and nearly 14 crore EVs by 2030, equivalent to an annual average growth rate close to 30%. Because of this continuous increase, it is predicted that EVs will account for as good as 7% of the global vehicle fleet by 2030. EVs will reach more or less 1.4 crore in 2025 and 2.5 crore vehicles in 2030, constituting 10% and 16% of all road vehicle sales [1]. Hence, it is really important to protect the power grid and its components by charging EVs coordinately. Currently in the practice of EVs uncoordinated charging, the user is given a choice of fast charging or regular charging. This process of allowing users to choose without appropriate check for potentially harmful consequences can lead to damage of power grid and transformer overloads. The presently available solutions for the coordinated EV charging are standard optimization methods, such as dynamic programming and quadratic programming. But the high computational resources required and the complex mathematical formulation necessary in these techniques possess many practical challenges.

Some of the solutions like wireless dynamic charging, turning street lamps into EVs charging stations are available but they have their own set of problems. Wireless dynamic charging is a popular solution due to its ability to charge the vehicles but the infrastructure required to set up the wireless charging is proven to be very costly [2]. Even in the case of the above said public lighting system usage, a costly infrastructure

---

V. N. Machineni · K. S. Vardhana · V. Dhandapani (✉)  
National Institute of Technology Delhi, New Delhi 110036, India  
e-mail: [dvaithiyanathan@nitdelhi.ac.in](mailto:dvaithiyanathan@nitdelhi.ac.in)

V. N. Machineni  
e-mail: [161220020@nitdelhi.ac.in](mailto:161220020@nitdelhi.ac.in)

K. S. Vardhana  
e-mail: [161220016@nitdelhi.ac.in](mailto:161220016@nitdelhi.ac.in)

modification, in the form of converting street lamps into electric vehicle charging stations is required [3]. Considering the above said problems, a data driven approach in the form of machine learning is proposed in our paper. The key advantage of this process, it does not require much computational power. Also, it has an added advantage of limited human intervention, thus allowing for a much more agile process in total. The random forest (RF) and K-nearest neighbors (KNN) are the machine learning techniques used in this paper to analyze the power usage in any given location and suggest the best mode of charging electric vehicles.

In the upcoming sections, Sect. 2 determines what data driven approach is, Sect. 3 describes the machine learning techniques used in this paper and their background. Section 4 presents the dataset used to construct the models in this paper. Section 5 discusses the results obtained for both models. Section 6 presents the conclusion of the paper.

## 2 Data Driven Approach

There were many studies conducted on the topic of coordinated charging of electric vehicles. Considering the Bi-Level theory of programming, Yao et al. [4] proposed a coordinated approach to charging and discharging procedures. In comparison, Jian et al. [5] used quadratic programming as the primary optimization approach to optimize and minimize the variation in load. The downside of this model is its assumptions taken for simplifying the optimization process.

There are three main viewpoints when charging EVs is under consideration. They are the customer's point of view, operator's point of view and their combinatorial viewpoint. Rawat and Niazi [6] compared these viewpoints. From the customer's point of view, the outcome depicted that the charging cost was reduced, and the load curve was improvised in the case of slow charging, whereas the undesirable effects were shown in the load curve in the case of fast charging. But from the operator's point of view, the outcome depicted a scenario that levels the load curve and circumvents power losses is preferred. The game theory for coordination of charging EVs is explored by Li et al. [7]. To establish the Nash equilibrium in the constructed game model, a Newton-type algorithm was presented in this study. On the other hand, Wang et al. [8], researched ways for maximizing the accommodation capacity of distributed energy. His model tries to determine the best place for charging electric vehicles considering them as mobile loads.

Almost all of these studies used quadratic programming or dynamic programming algorithms for achieving coordinated charging. These conventional optimization techniques require complex problem formulation and solving methodologies. The use of machine learning is less complicated since the only need is a proper dataset. The consideration of optimization constraints and problem formulation is not required in the case of machine learning.

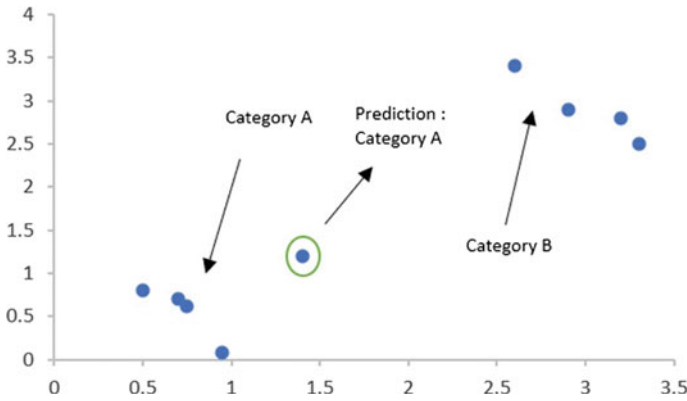
### 3 Background of Machine Learning Techniques Used

Machine learning provides the ability to computer systems to learn from past experiences without the explicit need of programming [9]. This provides the chance for the systems to be agile without much effort. The models are constructed and trained using historical data. If any of the parameters are changed then the model can be simply retrained with the new data. Machine learning algorithms are categorized into supervised and unsupervised learning models. Another category of categorization is dependent on the type of the variable to be predicted using the algorithms. This variable is termed as a response variable, if it is continuous, then it is a regression problem. Alternatively, if the response variable is categorical then it is a classification problem.

In supervised learning techniques, we use labeled data. This means the training set includes the inputs and correct output, which will help the model to learn and predict the correct result for the new data. These techniques recursively learn the mappings between input variables and the response variable. In the context of electric vehicle charging, we could train the model using the dataset containing the area name, density of electric vehicles in the area at a given time and the amount of power they used for charging over a period of time, the mode of charging they used. The mode of charging can be the response variable. Using the training data, the model will map the mode of charging with other input variables and test data can be used to test the accuracy. KNN, Random Forests, support vector machines (SVM), decision trees (DT), linear regression. KNN and random forest algorithms which are used in this paper are discussed below.

#### 3.1 KNN

K-nearest neighbor (KNN) is one of the supervised machine learning techniques that can be used for both regression and classification. The K-nearest neighbor algorithm compares the similitude among the test data or new data and training data and categorizes the test data into the category which is likely to the categories formed from training data. The KNN algorithm is mostly used in classification but is also capable of performing regression. Being a lazy learner algorithm, it only learns when the test cases are being executed against the model. Since it does not take any assumptions about the dataset, it can be categorized as a non-parametric algorithm. In the training phase it just stores the dataset and when it gets new data, then it classifies that data into a category that is much similar to the new data. The KNN [10] algorithm is easy to implement and is tolerant to noisy training data. It will be efficacious if the dataset is huge but the computation costs will be significantly higher (Fig. 1).



**Fig. 1** Illustration of KNN algorithm prediction

### 3.2 *Random Forest*

Random forest is a popular machine learning algorithm which is a supervised learning technique. It can be used for solving both classification and regression problems. The concept of ensemble learning is used here which means multiple classifiers are combined to solve a complex problem, thus improving the performance of the model. The dataset under consideration is broken into many subsets, and a decision tree is constructed on each subset. The decision from all the above said decision trees is averaged out to make sure the best possible accuracy is achieved. The averaging out is done through the voting procedure, where the outcome of all trees is divided into weighted categories. The category with the highest number of votes is decided to be the category of the entire random forest. The random forest's main advantage is that as the number of decision trees constructed increases, the accuracy goes higher and higher without the issue of overfitting. It is capable of handling large datasets with high dimensionality and enhances the accuracy of the model, preventing the overfitting issue. It first selects  $N$  random data points from the training set and builds the decision trees associated with the selected data points. It then chooses the number of decision trees to be built [11]. For each new data set, it runs the decision tree algorithm for each of the decision trees built and then assigns the category based on the above results (Fig. 2).

## 4 Model Used

The electric vehicle charging station's power rating (PR) is the dependent variable for the model proposed in this paper. It can be obtained from the power usage data of the locality under consideration. The model uses historical load data to predict the PR. The date and the time of day are the independent variables used. Table 1

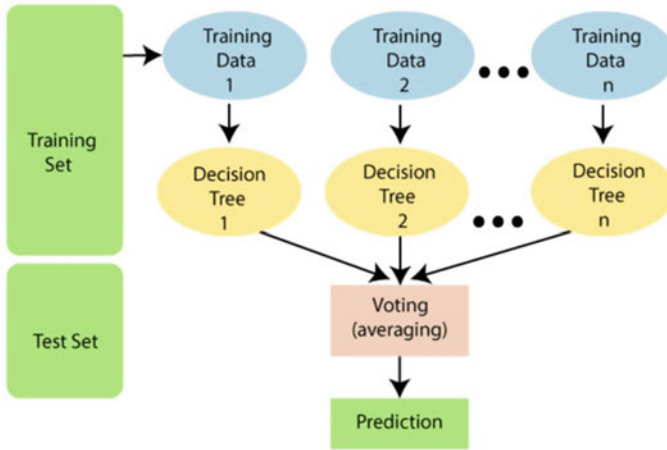


Fig. 2 Illustration of random tree algorithm working

Table 1 Dataset’s attributes and their ranges

S. No.	Attribute name	Data type	Data range
1	Day	Numeric	1–31
2	Month	Numeric	1–12
3	Hour	Numeric	0–23
4	Minute	Numeric	0–50 (multiples of 10)
5	Total power usage	Numeric	>0, dependent
6	Power rating (PR)	Categorical	High, normal, low

summarizes various parameters and their ranges used in the model. The dataset [12] used, is the collection of power usage from a local household society in the USA. It has 52,560 records with each record having 201 attributes. The power usage is collected for every 10 min. Due to the absence of historical EVs charging data, the model in this paper utilizes normal load data from that locality. From this dataset, a new dataset containing five attributes such as power, month, day, hour and minute is derived. Comma Separated Values (CSV) file is used to store the above derived data and is fed to the model. The model is constructed in python, with the help of scikit library. The matplotlib library is used to plot the accuracy graphs comparing actual results of test data and the predicted result. The model would be able to anticipate three classes such as high, normal and low of EV charging, based on the month, day, hour and minute. Hence, if the power rating is low, the slow charging method is used, and if the power rating is high, the fast charging method is used, else conventional method will be used.

## 5 Results

Metrics like accuracy, precession, recall, f1-score can be used to evaluate the results of test data fed to any machine learning model. Accuracy gives us the proportion between the number of data points that were labeled correctly and the total number of data points given to the model.

Accuracy can be determined by using the below formula

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}$$

Precision is the ratio of correctly predicted positive observations to the total predicted positive observations for each category. Recall is the ratio of correctly predicted positive observations to all observations in that category.

Precision and recall are determined using below formulae.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

F1-score is the weighted average of precision and recall. Hence, this score takes both false positives and false negatives into account which is a better metric than accuracy in this case because of uneven distribution. In all the above formulae, TP = True positive, TN = True negatives, FP = False positives and FN = False negatives.

The random forest model got an accuracy of 86.84%, with 5 decision trees while when using 10 decision trees we got an accuracy of 88.67%. The high category has a precision of 92%, low category has a precision of 82% and the normal category has a precision of 80% when 5 decision trees are used in the random forest model (Figs. 3 and 4).

	precision	recall	f1-score	support
HIGH	0.92	0.93	0.92	7433
LOW	0.82	0.80	0.81	1786
NORMAL	0.80	0.78	0.79	3921
accuracy			0.87	13140
macro avg	0.84	0.84	0.84	13140
weighted avg	0.87	0.87	0.87	13140

Fig. 3 Classification report of random forest model with 5 decision trees

	precision	recall	f1-score	support
HIGH	0.92	0.94	0.93	7433
LOW	0.85	0.85	0.85	1786
NORMAL	0.83	0.79	0.81	3921
accuracy			0.89	13140
macro avg	0.87	0.86	0.86	13140
weighted avg	0.89	0.89	0.89	13140

Fig. 4 Classification report of random forest model with 10 decision trees

The KNN model got an accuracy of 82.07%, with a K value of 5. The high category has a precision of 88%, low category has a precision of 75% and the normal category has a precision of 73% when K value is 5. With a K value of 10 we got an accuracy of 82.36%.

Figure 7 illustrates the accuracy variation as the no. of decision trees used in the random forest algorithm increases. It can be clearly seen that as the no. of decision trees increases, the accuracy increases. Figure 8 illustrates the accuracy variation versus the K value increases. The maximum accuracy is obtained when the K value is 13. As the K value increases, the boundary between the three categories becomes smoother but as evident from the following graph, error rate is minimum when K = 13 (Figs. 5 and 6).

	precision	recall	f1-score	support
HIGH	0.88	0.90	0.89	7433
LOW	0.75	0.72	0.74	1786
NORMAL	0.73	0.71	0.72	3921
accuracy			0.82	13140
macro avg	0.79	0.78	0.78	13140
weighted avg	0.82	0.82	0.82	13140

Fig. 5 Classification report of K-NN model with K = 5

	precision	recall	f1-score	support
HIGH	0.87	0.92	0.90	7433
LOW	0.73	0.74	0.73	1786
NORMAL	0.76	0.68	0.72	3921
accuracy			0.82	13140
macro avg	0.79	0.78	0.78	13140
weighted avg	0.82	0.82	0.82	13140

Fig. 6 Classification report of K-NN model with  $K = 10$

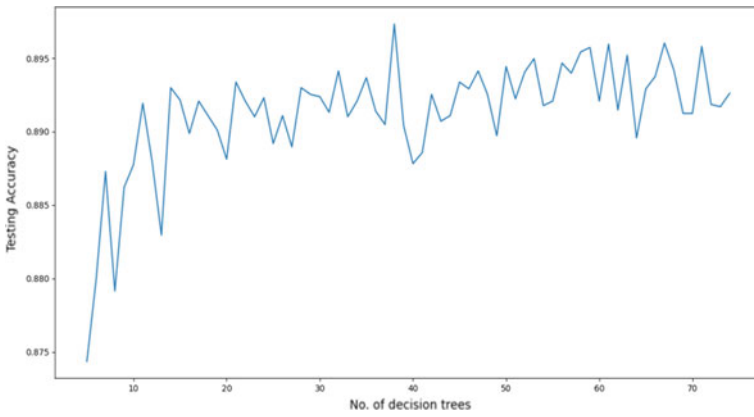


Fig. 7 Testing accuracy versus no. of decision trees used in the RF model

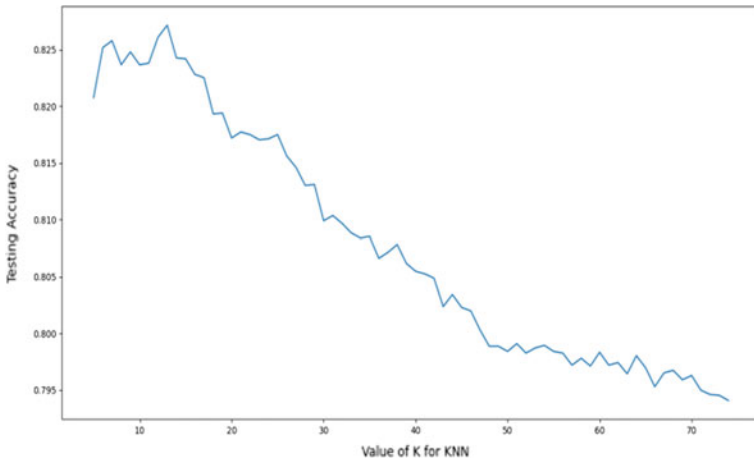


Fig. 8 Graph illustrating the testing accuracy versus the  $K$  value



## 6 Conclusion

This model provides a way for choosing the most power grid friendly method for charging Electric Vehicles. Hence, it provides solutions to many problems faced due to implementation of EVCS like grid overloading and power outages. The model thus helps in protecting the power grid from sudden power surges and consequently all the components of the power grid like transformers and extends their life period, reducing the need for frequent maintenance. From the results, we can conclude that the random forest model with an accuracy of 89.64% can be used for choosing the charging mode of EVs. Since this is not a computationally intensive task, we can deploy them with comparatively small costs.

## References

1. International Energy Agency's; Global EV Outlook 2020; available online at [https://iea.blob.core.windows.net/assets/af46e012-18c2-44d6-becd-bad21fa844fd/Global\\_EV\\_Outlook\\_2020.pdf](https://iea.blob.core.windows.net/assets/af46e012-18c2-44d6-becd-bad21fa844fd/Global_EV_Outlook_2020.pdf)
2. Maglaras LA, Jiang J, Maglaras A, Topalis FV, Moschogiannis S (2015) Dynamic wireless charging of electric vehicles on the move with mobile energy disseminators. *Int J Adv Comput Sci Appl (IJACSA)* 6:239–251
3. Mario AR, Fadi AA, Gagnaire M, Lascaux Y (2013) Telewatt: an innovative electric vehicle charging infrastructure over public lighting systems. In: Proceedings of the 2nd international conference on connected vehicles and expo (ICCVE), Las Vegas, NV, USA, 2–6 Dec 2013
4. Yao W, Zhao J, Wen F, Xue Y, Ledwich G (2013) A hierarchical decomposition approach for coordinated dispatch of plug-in electric vehicles. *IEEE Trans Power Syst* 28:2768–2778
5. Jian L, Xue H, Xu G, Zhu X, Zhao D, Shao Z (2013) Regulated charging of plug-in hybrid electric vehicles for minimizing load variance in household smart microgrid. *IEEE Trans Ind Electron* 60:3218–3226
6. Rawat T, Niazi K (2017) Comparison of EV smart charging strategies from multiple stakeholders' perceptions. *J Eng* 2017:1356–1361
7. Li J, Li C, Xu Y, Dong Z, Wong K, Huang T (2018) Noncooperative game-based distributed charging control for plug-in electric vehicles in distribution networks. *IEEE Trans Ind Inform* 14:301–310
8. Wang Z, Fan S, Liu B, Liu X, Wei Z (2017) Coordinated charging strategy of plug-in electric vehicles for maximising the distributed energy based on time and location. *J Eng* 13:1740–1744
9. Russel S, Norvig P (2018) Artificial intelligence: a modern approach. Pearson India Education Services Pvt. Ltd., Noida, India
10. Géron A (2017) Hands-on machine learning with scikit-learn and tensorflow. O'Reilly Media, California
11. Ho T (1998) The random subspace method for constructing decision forests. *IEEE Trans Pattern Anal Mach Intell* 20:832–844
12. Shibl M (2020) Bus Load Data 2020. Available online at: <https://www.kaggle.com/mostafashibl/electric-vehicles-charging-stations-power-rating>

# A 4-element Dual Band MIMO Antenna with L-Shaped Decoupling Stub for 5G Applications



R. Seetharaman , B. Varsha Poorani, Kavya Santha Kumar, M. Tharun, S. S. Sreeja Mole, and K. Anandan

## 1 Introduction

In Mobile phones which are mandatory part of well-being and it consists of an antenna which aid communication. Other devices in which antenna plays a significant role are wireless communication, Radars, GPS devices, WIMAX, MIMO, laptops, transceivers, space vehicle navigation, airplanes and many other devices. With the advent of 5 g, antenna development for such cases seeks extensive attention majorly MIMO antennas. It will be able to provide wireless connectivity for a wide range of applications like smart homes, traffic protection and control, critical infrastructure and industry applications and for very high-speed data delivery. It can be able to provide up to 10 Gbps data rate (10–100 times the speed improvement over 4G and 4.5G networks) and about 1-ms latency.

Also, Miniaturization and multi-cell array antennas offer the possibility of high-speed data transmission. Also, here the use of dual band is considered to be more advantageous over the single band, it is unique in its ability to provide a strong, stable wireless connection in often difficult to reach locations. A dual band antenna can use both frequencies at once or switch between the two frequencies depending on which option provides a stronger connection in the given area. The dual band also provides us an alternative to avoid signal interference among different devices operating in

---

R. Seetharaman (✉) · B. Varsha Poorani · M. Tharun  
Department of ECE, CEG Campus, Anna University, Chennai 600025, India  
e-mail: [rseetharamanece@annauniv.edu](mailto:rseetharamanece@annauniv.edu)

K. S. Kumar  
Department of Computer Engineering, Texas A&M University, College Station, TX 77801, USA

S. S. Sreeja Mole  
Department of ECE, Christu Jyothi Institute of Technology & Science, Jangaon 506167, India

K. Anandan  
PS Division of E & E, TUV SUD South Asia Pvt. Ltd., Chennai 600041, India

the similar frequency range and therefore dual band antennas are a stable, easy way to connect between our day-to-day things.

An “8-element Dual band MIMO Antenna with Decoupling Stub for 5G Smartphone Applications” operating over two desired range of frequencies (3.3–4.2 GHz) and 5 GHz band (4.8–5.0 GHz) is proposed with much lesser ECC value [1]. Next a “Multi-Band 10-Antenna Array for Sub-6 GHz MIMO Applications in 5G Smartphones” works at the sub-6 GHz spectrum is proposed with desired ECC and Isolation value and study of User Hand Effects in case of Multiple antennas [2]. “A 4-Port 2-Element MIMO Antenna for 5G Portable Applications” for small portable devices, IOT and cellular applications as a MIMO antenna is proposed [3].

The metal frame is the mandatory component for a smartphone also it would leave negative impact on the antenna radiation performance. For addressing this problem several techniques like etching open slot on the metal frame is proposed [4, 5]. In addition, for the purpose of reducing coupling various techniques are employed like introducing parasitic elements, slots etc., but an evident and promising technique which is the using of frequency selective surface technique is proposed for the purpose of reducing coupling between antennas in mm-wave MIMO [6, 7]. Envelope Correlation Coefficient serves as the major parameter for evaluating Multiple Input and Multiple Output (MIMO) systems.

## 2 Materials and Methods

The advantages of the MIMO systems are the higher capacity, increased data rates, reduced multi-path fading effects, more link reliability, and wider coverage area. Use of reactive loading is a very popular technique for obtaining a dual band behavior is to introduce a reactive loading to a single patch. The easiest method is connecting a stub to one radiating edge, so as to introduce an additional resonant length that is facilitate the second operating frequency. As per the above method, the height of the mobile phone is only 5 mm. The antenna elements have the same structure and dimensions. The side frames are orthogonal to the system ground plane, and the width and length of each antenna on the side frames is 3.9 mm and 17 mm, respectively, length and width of substrate is selected to have a size of 130 mm and 74 mm, respectively. Both the side-edge frame and the system circuit board are fabricated using 0.8 mm thick and FR4 substrate of relative permittivity 4.4 and loss tangent 0.02 [8]. The antenna is composed of two structures the monopole and L-shaped stub. Monopole which is bent normally for miniaturization is employed along with the L-shaped short circuit stub for the purpose of achieving dual band operation. After designing of single antennas, it is just replicated and placed along the two longer edges of the side substrates.

## 2.1 MIMO Considerations

Good diversity performance between two antennas depends on several factors:

1. The antennas which are placed in larger number should have independent directivity that must be captured.
2. Minimum coupling is one of the criteria.
3. Multiple antennas should more or less match with the gain of individual antenna.
4. Miniaturization reduce interaction between antennas considerably.
5. The multiple antennas should possess well matched terminations.

Coupling is the major issue in MIMO. Antenna coupling occurs when two or more antennas are placed in such close physical proximity to one another so there occurs a necessity to take care about spacing between antennas also which is generally greater than one half wavelength. So, in order to avoid coupling, spacing is the criteria that takes the lead. Also, there are lots of techniques used for achieving dual frequency bands like use of orthogonal mode, multiple patches and reactive loading. But here reactive loading concept is employed. The design parameters are given as follows,

1. The single antenna designed can be operated at the bands of; 3.3–3.6 GHz and 4.8–5.0 GHz.
2. The system circuit board is selected to have an length and width of 130 mm and 74 mm, respectively.
3. The substrate used here is FR-4 lossy with a thickness of only 0.8 mm (loss tangent 0.02).
4. The ground plane and the radiating structure (MIMO antenna and L-shaped short circuit stub) is copper annealed.
5. The height of the edge frame is kept 5 mm.
6. The area of the side frames is 3.9 mm × 17 mm.

The equivalent flowchart of the design is shown in Fig 1. The initial step is creating a ground plane and then substrate with the suitable substrate and ground material with optimisable dimensions. Construct side substrates which depict the one similar to the edge of a mobile phone. Along the longer sides, place four antennas in such a way that each longer side occupies a pair of antenna with sufficient spacing. After the design is over, assign discrete port to each individual antenna and verify design model, parameters then followed by simulating it using time domain solvers and obtain results and plot it.

## 2.2 Antenna Setup

The single antenna has a pair of radiating elements monopole and stub. The single antenna is designed and can be operated over the frequency range of 3.3–3.6 GHz and 4.8–5.0 GHz, the four antennas are constructed on the side substrates created [9].

**Fig. 1** Flow chart of the design

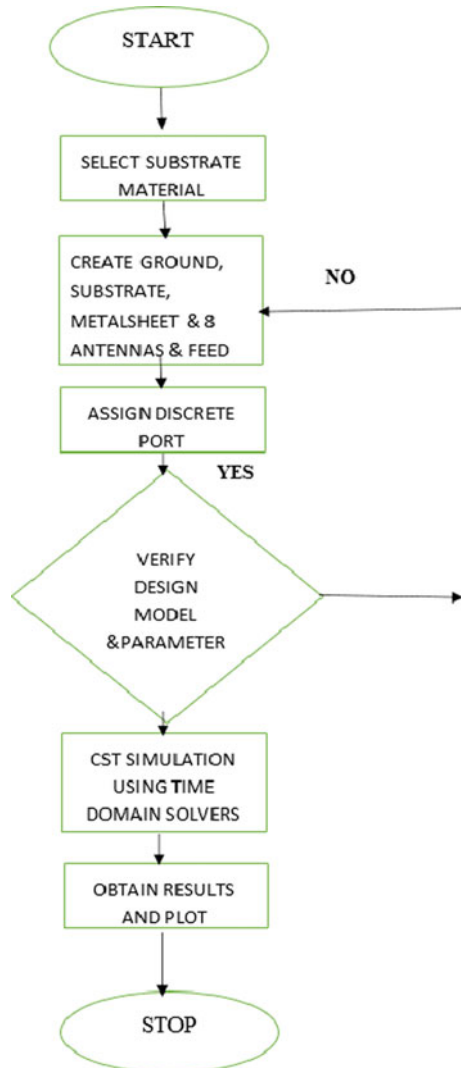


Figure 2 is the antenna setup which is to be placed on four extreme corners with appropriate spacing. Figure 3 is the 3-D view of antenna. Figure 4 is the side view of antenna. The monopole is the long bent structure, and the coupling capacitance generated by the L-shaped branch which acts as a stub cancelling the reactance part by making the antenna to operate in the desired frequency by ensuring of matching the impedance of the low frequency band and four antennas are constructed along the two longer edge substrates and once the antenna design is over, coaxial feed is given and discrete port is assigned for the four antennas and finally the design is simulated for the evaluation of the parameters as proposed.

Fig. 2 Antenna setup

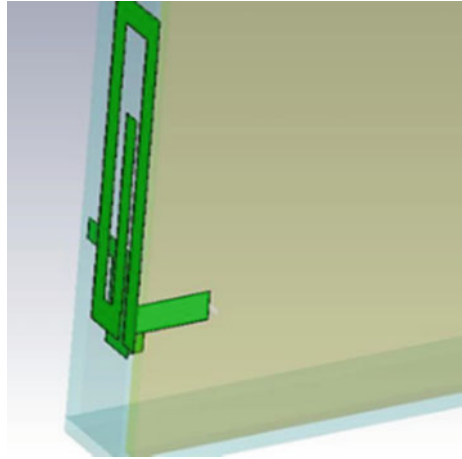
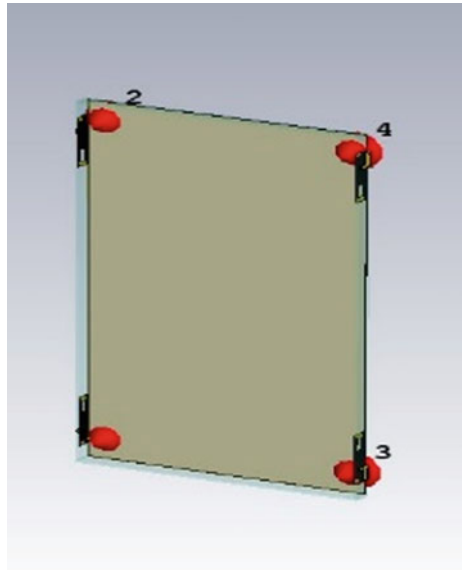


Fig. 3 3-D view



### 3 Results and Discussion

#### 3.1 Reflection Coefficient

The reflection coefficients ( $S_{11}$ ,  $S_{22}$ ,  $S_{33}$ ,  $S_{44}$ ) of four antennas are less than  $-6$  dB in the desired frequency range of 3.3–3.6 GHz and 4.8–5.0 GHz inferred from Figs. 5 and 6, respectively, indicating that suitable impedance matching for our application is obtained [10].

Fig. 4 Side view

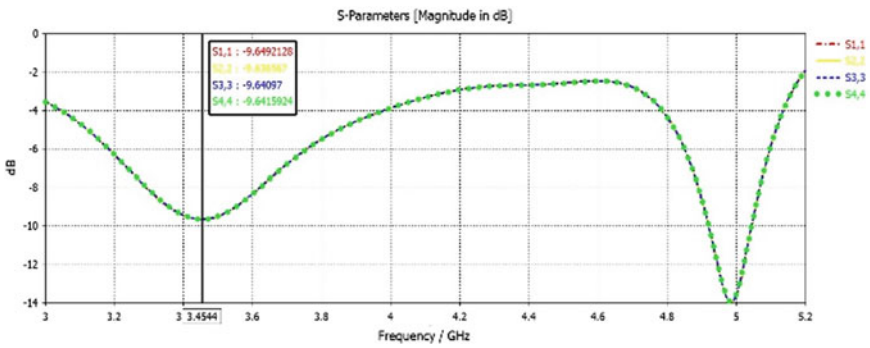
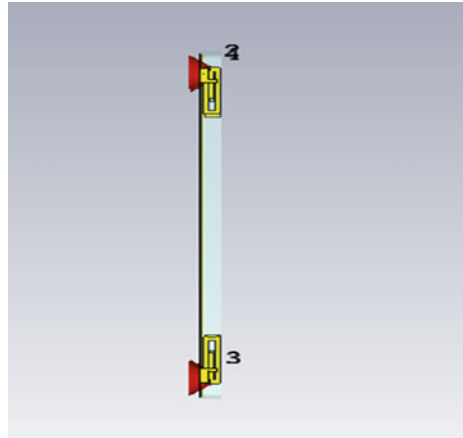


Fig. 5 Reflection coefficient in 3.3–3.6 GHz

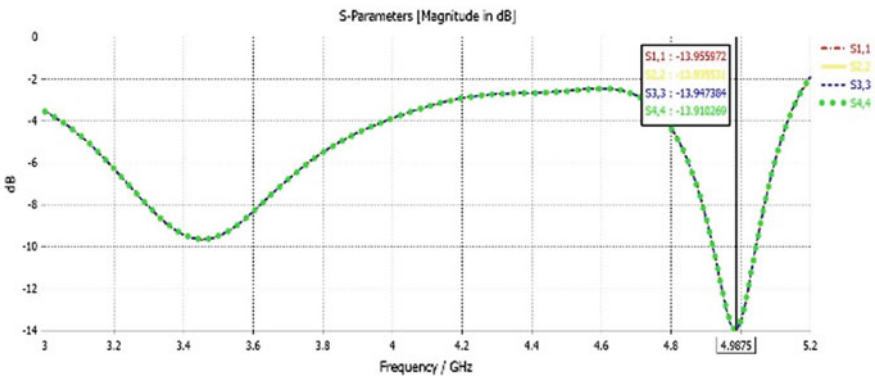


Fig. 6 Reflection coefficient in 4.8–5.0 GHz band

### 3.2 Transmission Coefficient

In this work, the dip was achieved at 3.45 GHz for (3.3–3.6 GHz band) and 4.99 GHz for (4.8–5 GHz) below  $-6$  dB. The transmission coefficients for all the four MIMO antennas are achieved below  $-15$  dB inferred from Figs. 7 and 8, respectively. Also, from the transmission coefficients, we can arrive at the isolation between the antennas which is possibly measured using Vector Network Analyzer (VNA). From the figures, it is evident that all the transmission coefficients are below  $-15$  dB. So, the isolation is also less than  $-15$  dB. For MIMO applications there should be the sufficient

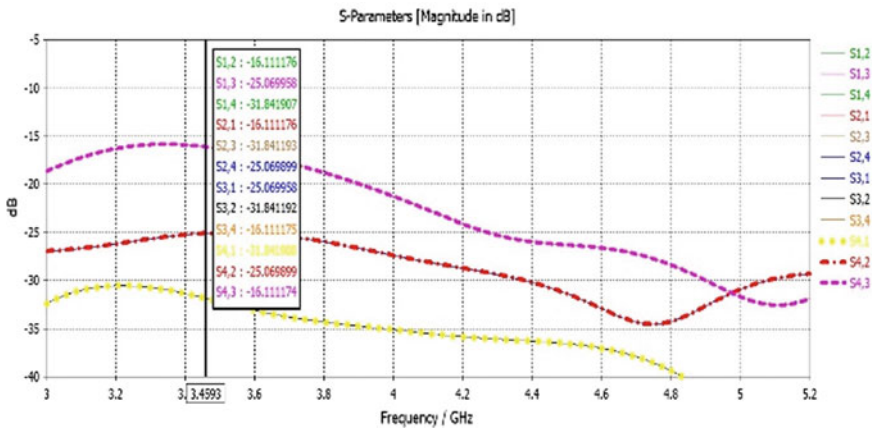


Fig. 7 Transmission coefficient in 3.3–3.6 GHz band

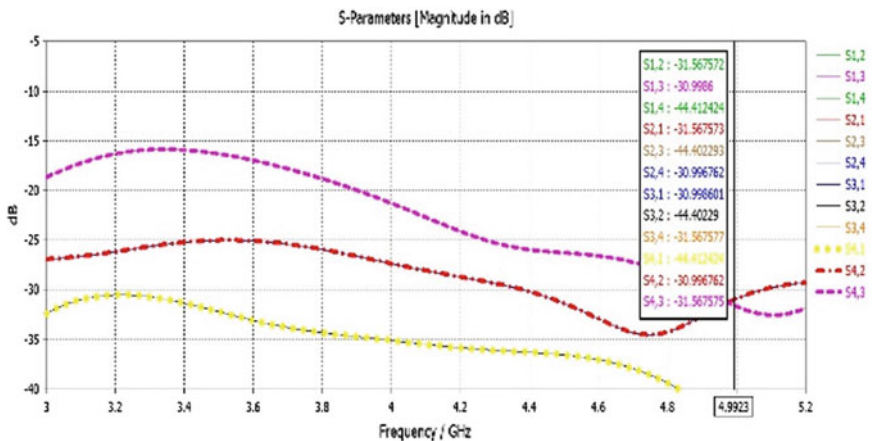


Fig. 8 Transmission coefficient in 4.8–5.0 GHz band



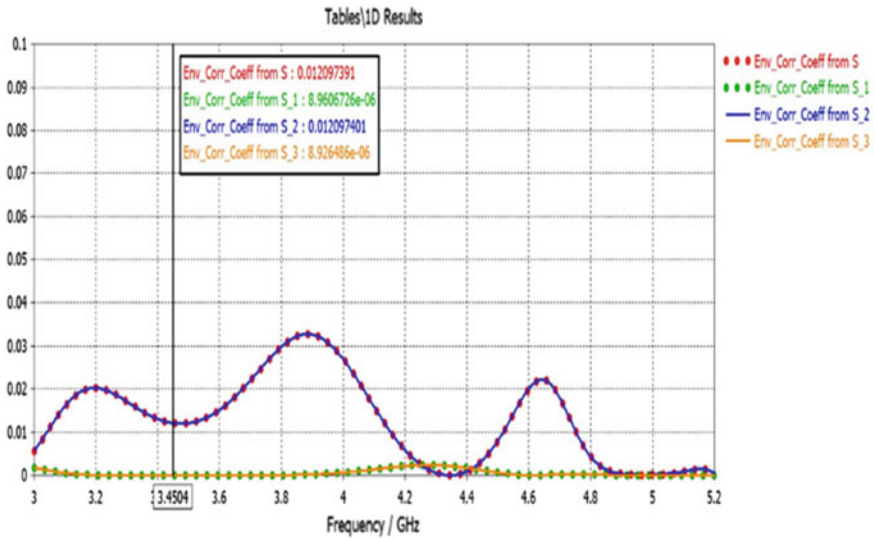


Fig. 9 ECC in 3.3–3.6 GHz band

isolation between the antennas in order to avoid mutual coupling. Normally, the mutual coupling of MIMO antenna should be less than 20 dB.

### 3.3 Envelope Correlation Coefficient

From the above observed results, it is inferred that the ECC for the band 3.3–3.6 GHz in Fig. 9 is found to be less than 0.01, when ECC is calculated for MIMO antennas by checking the effect of port 1–2, 2–3, 3–4 and 1–4. From Fig. 10, it is inferred that the ECC value for the band 4.8–5 GHz is also less than 0.01 which satisfies the condition for MIMO operation as mentioned in paper which is to be less than 0.1. In general, ECC is calculated from far field components. Here, it is calculated from S parameters.

### 3.4 Efficiency

Efficiency is measured in the two frequency bands (3.3–3.6 GHz) and (4.8–5 GHz) at each center frequencies 3.45 GHz and 4.99 GHz (Fig. 11), the efficiency was found to be 81.33% and 31.06%, respectively. It is the most efficient for Lower Bands, but the Higher Band efficiency can be improved by simple design changes following the 6 × 6 antenna design.

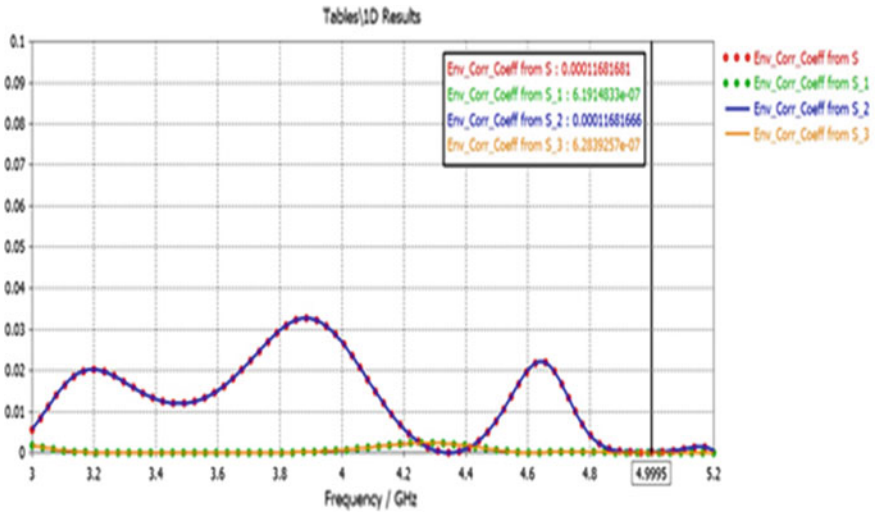


Fig. 10 ECC in 4.8–5 GHz band

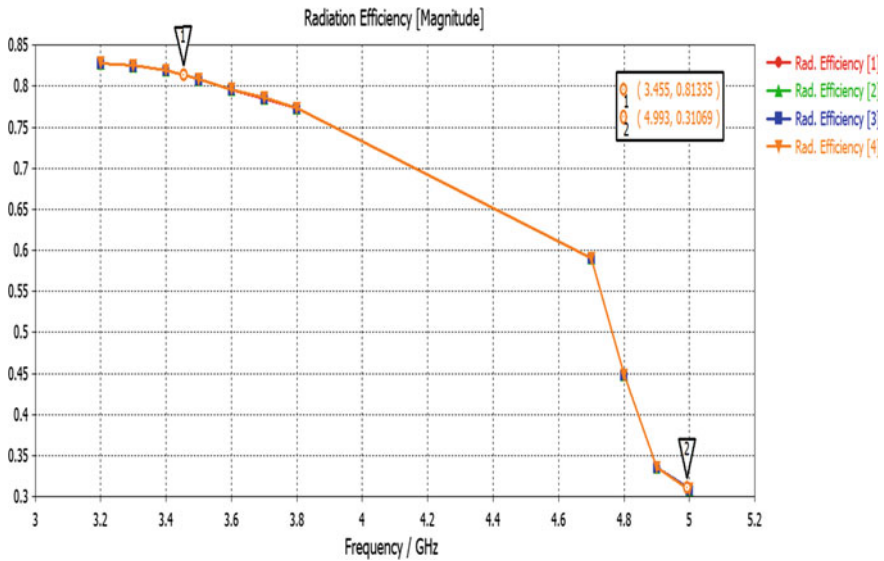


Fig. 11 Efficiency

### 4 Conclusion

The antenna is located on the side frame, to create a full screen smartphone antenna design. The work outcomes are ECC less than 0.1, isolation less than 12 dB, the

reflection coefficients less than  $-6$  dB and antenna efficiency of 81.33% in lower band (3300–3600 MHz) and 31.06% for higher band (4800–5000 MHz) are obtained. Its ECC offers the best results. The outcomes are ECC less than 0.1, isolation less than 12 dB, the reflection coefficients less than  $-6$  dB and antenna efficiency of 81.33% in lower band (3300–3600 MHz) and 31.06% for higher band (4800–5000 MHz) are obtained. Its ECC offers the best results.

## 5 Future Work

The antenna design is the most efficient for Lower Bands, but the Higher Band efficiency can be improved by simple design changes following the  $6 \times 6$  antenna design. And also further more higher designs like  $8 * 8$ ,  $10 * 10$ , etc., But the thing is that we should compensate for User hand effects while considering opting for higher array designs. To achieve a relatively high isolation, the antenna size is relatively small, novel for today's ultra-thin smartphone applications. Future work relies on antenna efficiency which in high band is low so our moto is to increase it in higher band. By working on tri band, quad band and also focus on increasing the antenna number and decreasing size more it is possible to improve efficiency of antenna further in higher bands.

## References

1. Cui L, Guo J, Liu Y, Sim CYD (2019) An 8-element dual-band MIMO antenna with decoupling stub for 5G smartphone applications. *IEEE Antennas Wirel Propag Lett* 18(10):2095–2099
2. Ren Z, Zhao A, Wu S (2019) MIMO antenna with compact decoupled antenna pairs for 5G mobile terminals. In: *IEEE Antennas Wireless Propag Lett*
3. Chen S-C, Chou L-C, Hsu C-IG, Li S-M (2020) Compact sub 6 GHz Four element MIMO Slot antenna system for 5G tablet devices. *IEEE Access* 8:154652–154662
4. Li Y, Sim CYD, Luo Y, Yang G (2019) High-isolation 3.5 GHz eight-antenna MIMO array using balanced open-slot antenna element for 5G smartphones. *IEEE Trans Antennas Propag* 67(6):3820–3830
5. Jiang W, Cui Y, Liu B, Hu W, Xi Y (2019) A dual-band MIMO antenna with enhanced isolation for 5G smartphone applications. *IEEE Access* 7:112554–112563
6. Li Y, Sim CYD, Luo Y, Yang G (2018) Multiband 10-antenna array for Sub-6 GHz MIMO applications in 5-G smartphones. *IEEE Access* 6:28041–28253
7. Chattha HT (2019) 4-Port 2-Element MIMO Antenna for 5G Portable Applications. *IEEE Access* 7:96516–96520
8. Hussain R, Alreshaid AT, Podilchak SK, Sharawi MS (2017) Compact 4G MIMO antenna integrated with a 5G array for current and future mobile handsets. *IET Microwaves Antennas Propag* 11(2):271–279
9. Qin Z, Geyi W, Zhang M, Wang J (2016) Printed eight-element MIMO system for compact and thin 5G mobile handset. *Electron Lett* 52(6):416–418
10. Chen S, Hsu M (2019) LTE MIMO closed slot antenna system for laptops with a metal cover. *IEEE Access* 7:28973–28981

# An Efficient Hybrid Genetic Algorithm and Whale Optimization Algorithm for Benefit Maximization of Optimal Sitting and Sizing Problem of DG in Distribution Systems



K. Banumalar and B. V. Manikandan

## 1 Introduction

The worldwide development of various methodologies in the integration of distributed generation (DG) has several environmental and economic benefits, and they enforce some operational problems. These may comprise, not limited to, problems of voltage drop and power losses. Most of research works carried out on optimal placement of DG-based optimization method and objective functions enhance voltage profile, reduce loss, and short circuit [1].

Later, numerous optimization algorithms like PSO [2], CSA [3], and LOA [4] had been proposed to solve optimal placement problem in distribution networks. Combined GA-fuzzy method [5] and PBIL-PSO algorithm [6] had been proposed for solving optimal placement problem.

To this direction, this paper presents a hybrid genetic algorithm and whale optimization algorithm for the optimum size and placement of DG based on minimum system losses, voltage magnitude, stability and costs of operation. The performance of proposed hybrid genetic algorithm and whale optimization algorithm is compared with GA and WOA based on fitness value and convergence rate. This work is organized as follows: Sect. 2 discusses the problem formulation. Section 3 presents algorithm implementation. Sections 4 and 5 elaborate simulations results and conclusions.

---

K. Banumalar (✉) · B. V. Manikandan  
Mepco Schlenk Engineering College, Sivakasi, Tamil Nadu, India  
e-mail: [kbanumalar@mepcoeng.ac.in](mailto:kbanumalar@mepcoeng.ac.in)

## 2 Problem Formulation

### 2.1 Objective Function

The aim is to real power loss minimization, and it is given in Eq. (1).

$$\text{Minimization of } P_{\text{Loss}} \quad (1)$$

$$P_{\text{Loss}} = \sum_{i=1}^n I_i^2 R_i \quad (2)$$

where  $n$  is the number of branch,  $I_i$  and  $R_i$  are current and resistance of  $i$ th branch. Subjected to the following constraints.

With DG subject to real power constraints given by:

$$\sum_{k=1}^N P_{DG} = P_D + P_{\text{Loss}} \quad (3)$$

The inequality constraints on  $V_i$  of DG are given by:

$$V_{i,\min} \leq V_i \leq V_{i,\max} \quad (4)$$

The inequality constraints on  $P_{DG}$  of DG are given by:

$$P_{DG,\min} \leq P_{DG} \leq P_{DG,\max} \quad (5)$$

### 2.2 Optimal Location of DG

The proper setting of DG units diminishes the losses and increases voltage profile.

The steps for optimum location of DG are:

1. Run system line flow.
2. Estimate the index vector (IV) [7] using Eq. (6).

$$\text{Index}[n] = \frac{1}{V_n^2} + \frac{I_q[k]}{I_p[k]} + \frac{Q_{\text{eff}}[n]}{\text{total } Q} \quad (6)$$

where  $V[n]$  and  $Q_{\text{eff}}[n]$  are voltage and effective load at  $n$ th bus.  $I_p[k]$  and  $I_q[k]$  are real and imaginary  $k$ th branch current.

3. Arrange index vector in descending order.

4. Voltage normalization by  $V(i)/0.95$ .
5. Bus normalized value below 1.01 appropriate for sizing DG.
6. Now place the DG and perform load flow. The DG with minimum loss selected as optimal size.

### 2.3 Economic Analysis

The yearly energy loss cost [8] calculated by Eqs. (7 and 8)

$$C_L = (T_{RL}) * (k_d + k_e + L_s * 8760) \tag{7}$$

$$L_s = k * L_f + (1 - k) * L_f^2 \tag{8}$$

where  $L_s$  and  $T_{RL}$  are loss factor and real power losses;  $k_d$  and  $k_e$  are yearly demand power and energy loss cost.

Real and reactive power of DG is calculated by Eqs. (9–11).

$$C(P_{DG}) = aP_{DG}^2 + bP_{DG} + c \$/MWh \tag{9}$$

$$C(Q_{DG}) = \left[ \text{Cost}(Sg \max) - \text{Cost}\left(\sqrt{Sg \max^2 - Qg^2}\right) \right] * k \tag{10}$$

where

$$Sg \max = \frac{Pg \max}{\cos \Phi} \tag{11}$$

## 3 Optimal Location and Sizing of Distribution Generation

Here obtained hybrid GA/WOA for optimal location problem result is compared separately with WOA and GA.

### 3.1 Genetic Algorithm

Figure 1 shows the flowchart for implementation of proposed problem using GA.

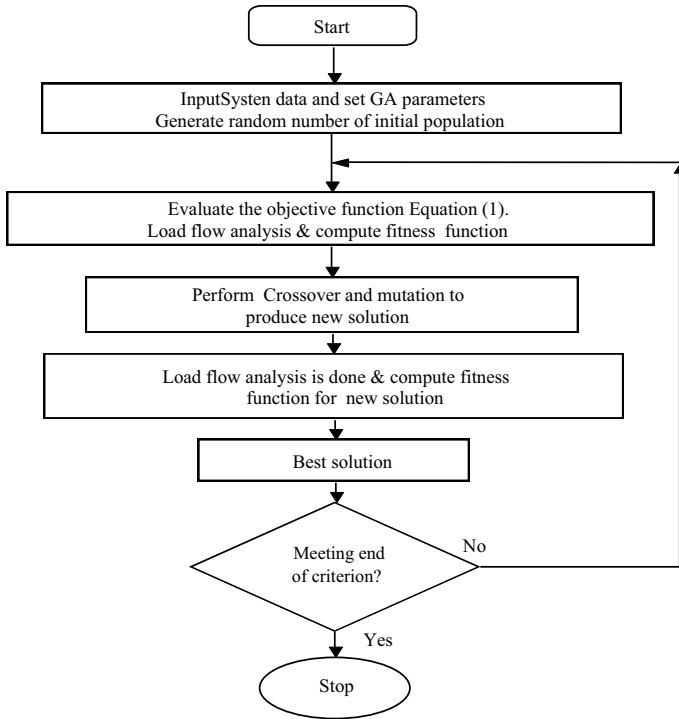


Fig. 1 Proposed problem using GA

### 3.2 Whale Optimization Algorithm

Whale optimization algorithm (WOA) [9] has been proposed by Mirjalili in 2016. Mathematical model of WOA is as follows.

- **Encircling prey**

Prey position modeled by the following equations:

$$\vec{D} = \left| \vec{C} \cdot \vec{X}^*(t) - \vec{X}(t) \right| \tag{12}$$

$$\vec{X}(t + 1) = \vec{X}^* - \vec{A} \cdot \vec{D} \tag{13}$$

$$\vec{A} = 2\vec{a} \cdot \vec{r} - \vec{a} \tag{14}$$

$$\vec{C} = 2 \cdot \vec{r} \tag{15}$$

where  $X^*$  and  $X$  are optimal solution and vector position,  $t$  is status of run,  $A$  and  $C$  are adjustment factor,  $a$  [2-0] and  $r$  [0, 1].

- **Bubble net hunting method**

Spiral actions of whales are as follows:

$$\vec{X}(t+1) = \vec{D}' \cdot e^{bl} \cdot \cos(2\pi l) + \vec{X}^*(t) \quad (16)$$

$$\vec{X}(t+1) = \begin{cases} \vec{X}^*(t) - \vec{A} \cdot \vec{D} & \text{if } p < 0.5 \\ \vec{D} \cdot e^{bl} \cdot \cos(2\pi l) + \vec{X}^* & \text{if } p \geq 0.5 \end{cases} \quad (17)$$

$$D' = \vec{X}^* - \vec{X}(t) \quad (18)$$

where  $D'$  is the whale and prey distance,  $b$   $[-1, 1]$ ,  $p$   $[0, 1]$ , and  $l$   $[-1, 1]$  are used.

- **Search for prey**

Mathematical model is to obtain the global solution as follows.

$$\vec{D} = \left| \vec{C} \cdot \vec{X}_{\text{rand}} - \vec{X} \right| \quad (19)$$

$$\vec{X}(t+1) = \vec{X}_{\text{rand}} - \vec{A} \cdot \vec{D} \quad (20)$$

Figure 2 shows the flowchart of optimization of DG using WOA.

### 3.3 Proposed Methodology

The proposed problem resolved by using novel hybrid GA/WOA. This problem is divided into two works. Optimal location of DG result is decimal decides the suitable bus for DG installation and solved using GA. The result was attained through GA then performed by WOA for DG sizing. Figure 3 shows the flowchart of proposed hybrid GA/WOA.

## 4 Simulations and Results

To represent the efficiency of proposed hybrid GA/WOA algorithm, it is tested on IEEE 15, 33, 69, and 85 bus systems [10] using MATLAB software and determine the optimal location of DG unit. Table 1 shows the optimum DG position, sizes, energy loss cost,  $P_{\text{DG}}$ , and  $Q_{\text{DG}}$  cost. From Table 1,  $V_{\text{min}}$  is more and losses is low in case of type III DG at 0.9 pf as compared to DG operating at unity power factor in all the test system, and Figs. 4, 5, 6, and 7 represent voltage profile of all the test systems.



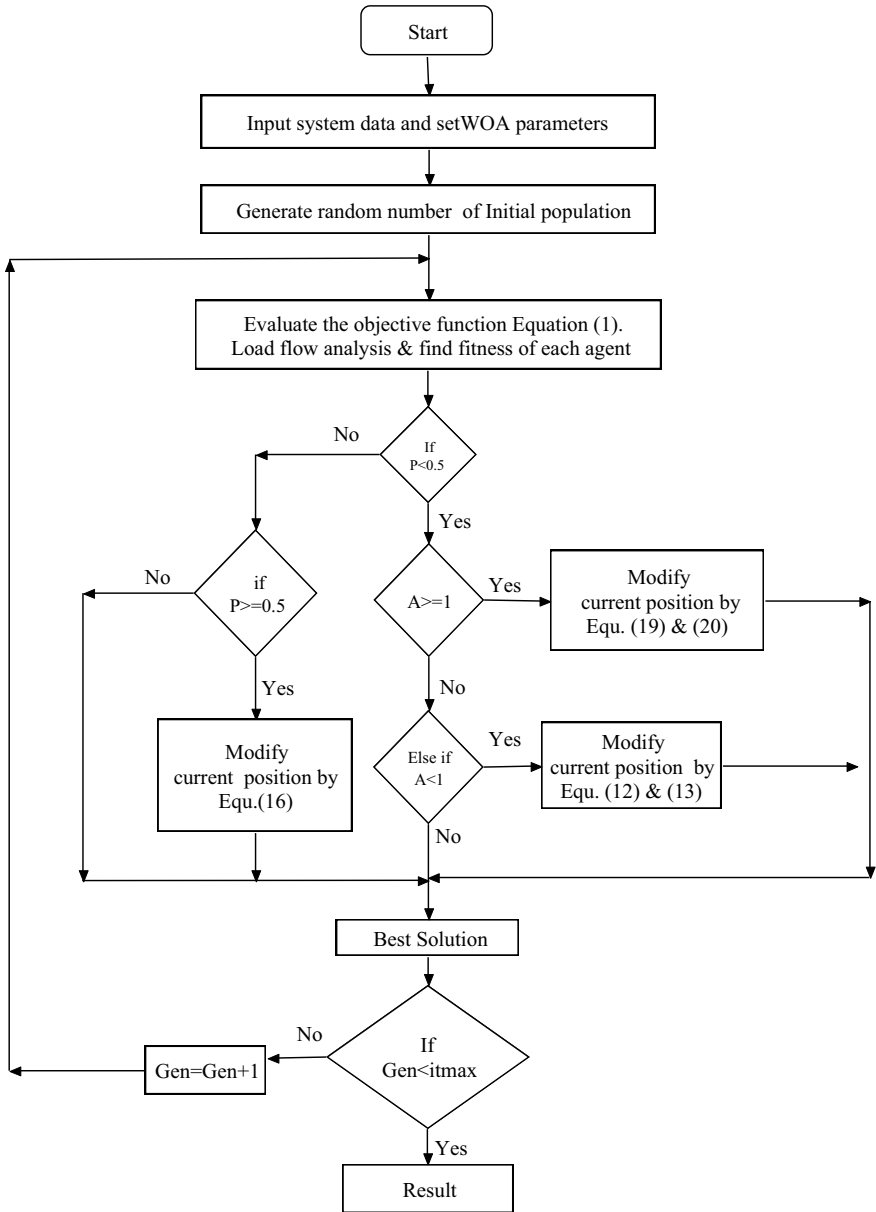


Fig. 2 Proposed problem using WOA

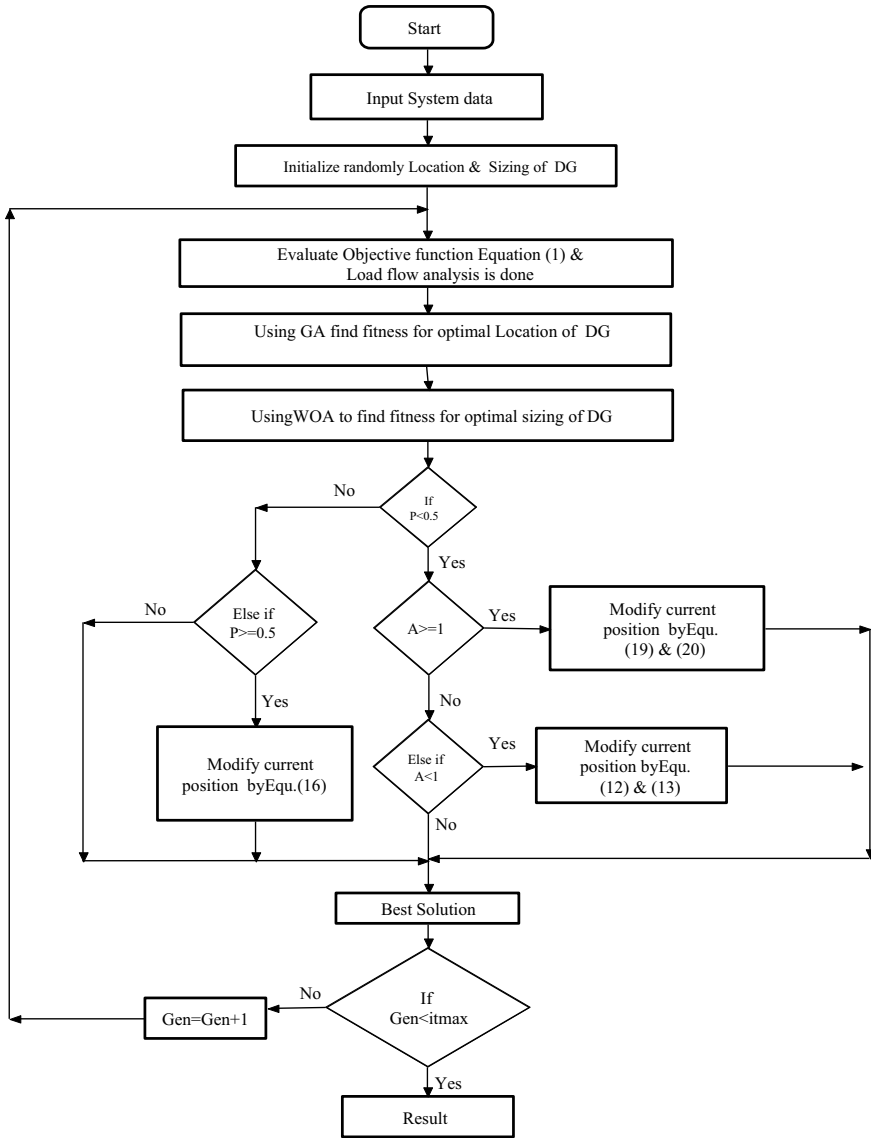


Fig. 3 Proposed problem using hybrid GA/WOA

**Table 1** DG placement by GA, WOA, and GA/WOA

IEEE bus system	Techniques	DG location	DG size (kVA)	TLP (kW)	TLR (kVAR)	Vmin (p.u.)	Energy losses cost (\$)	P <sub>DG</sub> cost (\$/MWh)	Q <sub>DG</sub> cost (\$/MVARh)	
15 bus system	Without DG	-	-	61.7933	57.2969	0.9445	4970.3	-	-	
	With DG at 0.9 pf									
	GA	6	906.991	34.973	29.99	0.966	2701.29	16.599	1.8612	
	WOA	6	907.785	33.385	29.89	0.959	2685.31	16.540	1.8656	
	GA/WOA	15	910.513	27.99	22.89	0.97	2255.99	16.599	1.871	
	With DG at UPF									
	GA	6	672.344	45.3228	41.26	0.956	3687.71	13.91	-	
	WOA	6	675.248	45.8035	41.88	0.9527	3684.18	13.754	-	
	GA/WOA	15	681.101	41.56	40.03	0.96	3396.91	13.855	-	
	Without DG	-	-	210.997	143.032	0.9038	16,982.57	-	-	
33 bus system	With DG at 0.9 pf									
	GA	15	1258	112.7	77.449	0.937	5700.01	55.5	6.2	
	WOA	15	1255.89	108.406	74.7726	0.939	5615.251	54.9	5.92	
	GA/WOA	30	1939.9	78.4	58.97	0.938	5491.93	52.41	5.41	
	With DG at UPF									
	GA	6	1053.7	141.37	93.82	0.935	8730.65	47.05	-	
	WOA	15	1061	133.503	90.74	0.933	8704.6	51.4	-	
	GA/WOA	30	1542.66	125.16	89.28	0.927	8650.16	31.32	-	
	Without DG	-	-	225.023	102.176	0.9092	18,101.76	-	-	
	With DG at 0.9 pf									
GA	61	2216.3	27.9636	16.49	0.9728	2249.2	40.091	4.482		

(continued)

**Table 1** (continued)

IEEE bus system	Techniques	DG location	DG size (kVA)	TLP (kW)	TLR (kVAR)	Vmin (p.u.)	Energy losses cost (\$)	P <sub>DG</sub> cost (\$/MWh)	Q <sub>DG</sub> cost (\$/MVARh)	
	WOA	61	2217.39	27.9649	16.4606	0.9724	2249.2	40.1	4.48	
	GA/WOA	61	2217	27.95	16.41	0.97	2249.4	40.09	4.4824	
	With DG at UPF									
	GA	61	1872.7	83.22	40.57	0.9685	6694	37.7	-	
	WOA	61	1872.82	83.2279	40.538	0.9683	6694	37.7	-	
	GA/WOA	61	1873.6	83.2210	40.54	0.97	6693.8	37.73	-	
85 bus system	Without DG	-	-	315.7	198.356	0.8714	20,520.71	-	-	
	With DG at 0.9 pf									
	GA	55	1283.67	157.43	90.98	0.9257	2541.6	41.23	4.178	
	WOA	55	1289	157.485	90.9812	0.9255	2541.2	41.1	4.18	
	GA/WOA	55	1289	157.401	90.98	0.926	2540.91	40.99	4.18	
	With DG at UPF									
GA	55	945.16	224.7	137.14	0.91	6791.6	38.81	-		
WOA	55	946.347	224.049	136.299	0.9109	6791	38.7	-		
GA/WOA	55	946.35	224.013	136.29	0.91	6791	38.7	-		

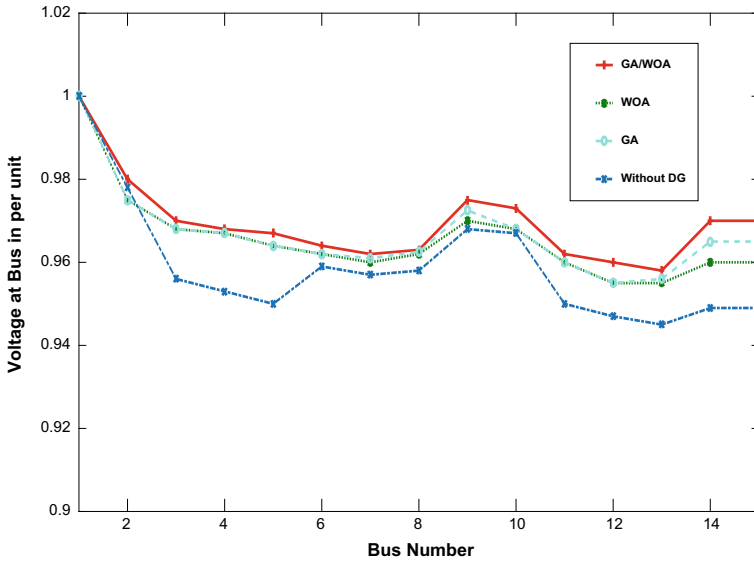


Fig. 4 Voltage magnitudes-15 bus

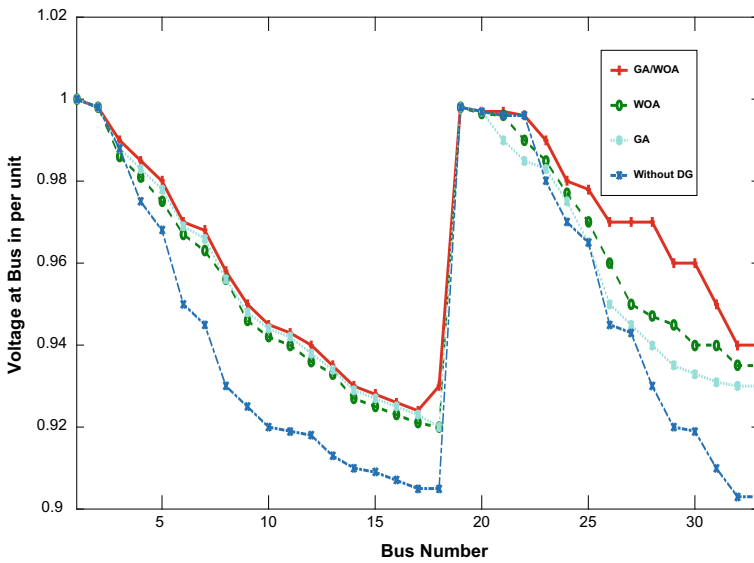


Fig. 5 Voltage magnitudes-33 bus

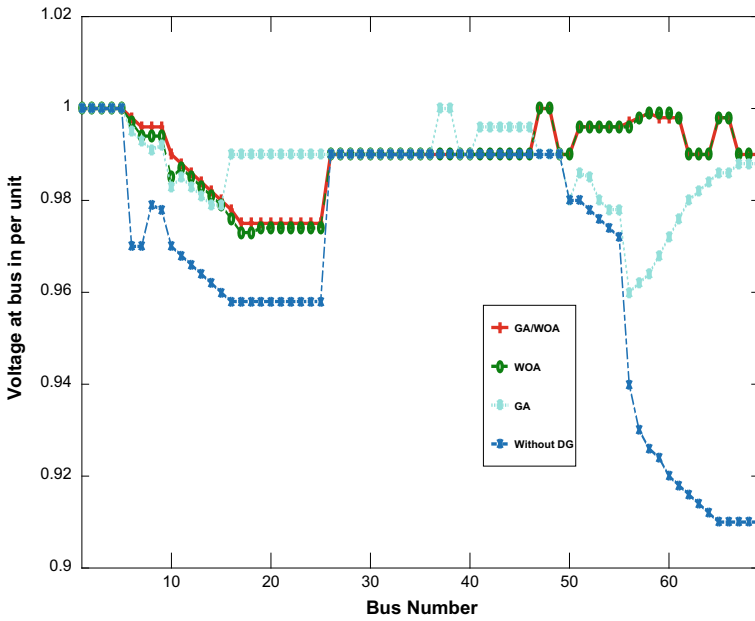


Fig. 6 Voltage magnitudes-69 bus

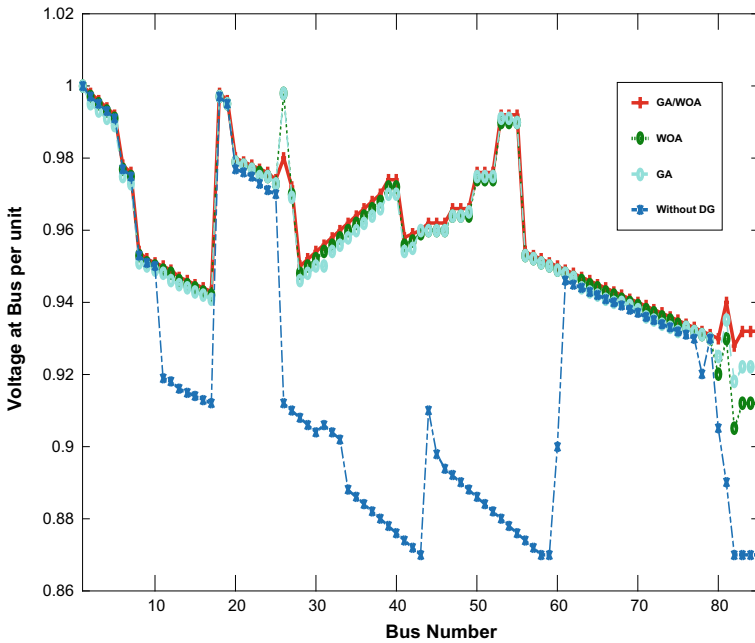


Fig. 7 Voltage magnitudes-85 bus

### 4.1 Performance Analysis

The performance analysis is done based on the solution quality and computational efficiency for all the test system. At the same time, initial random population for all the three (GA, WOA, and hybrid GA/WOA) techniques is kept the same during comparison analysis, and its performances are given in Figs. 8, 9, 10, and 11.

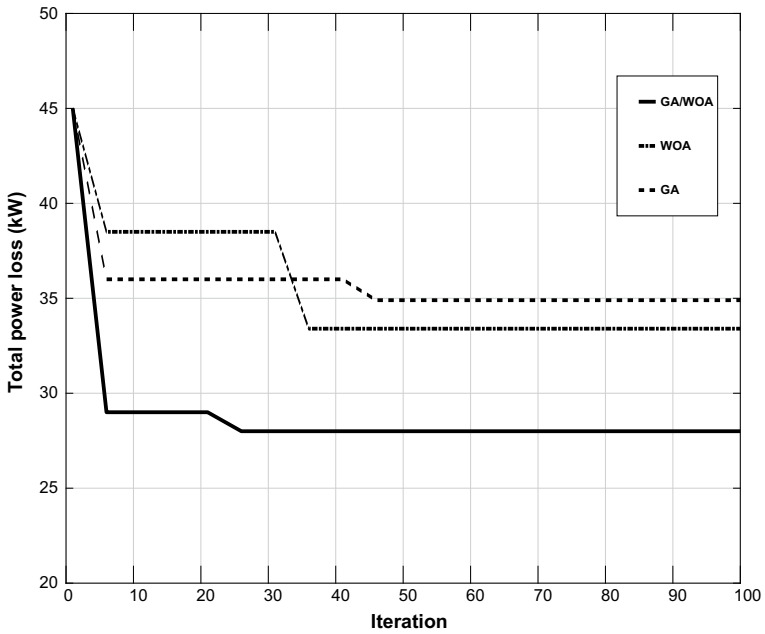


Fig. 8 Computational graph-15 bus

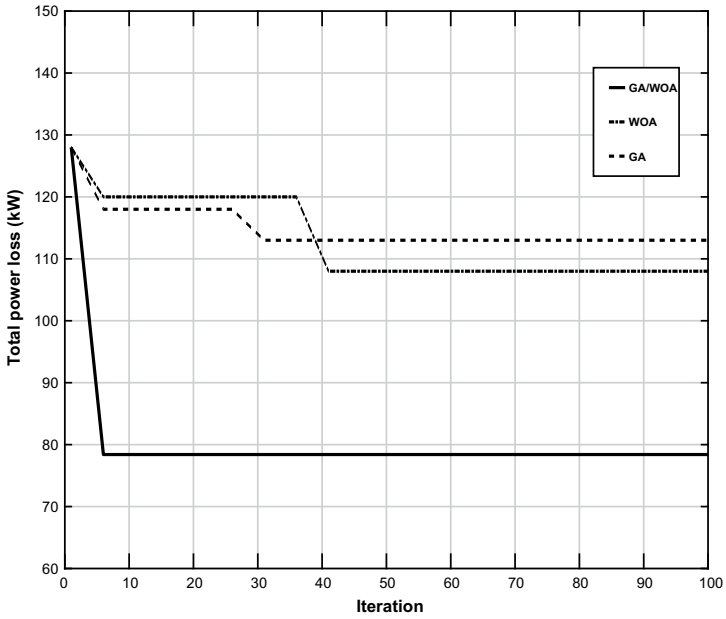


Fig. 9 Computational graph-33bus

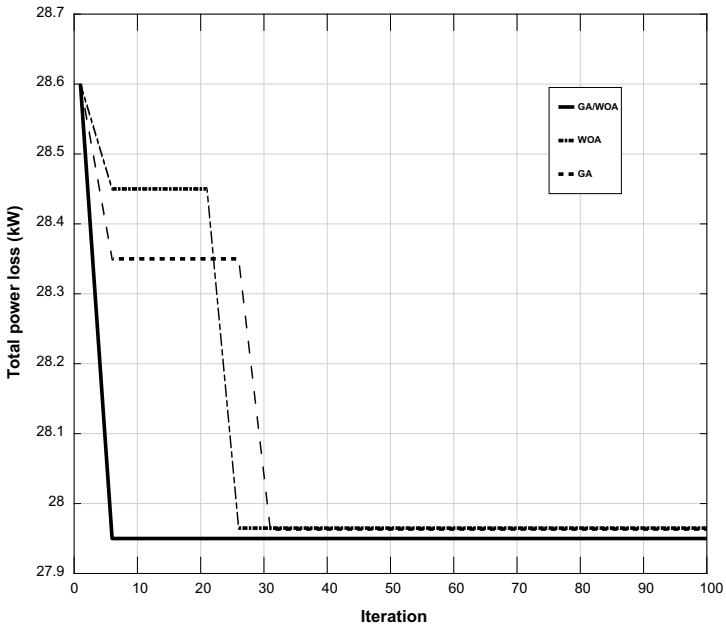


Fig. 10 Computational graph-69 bus



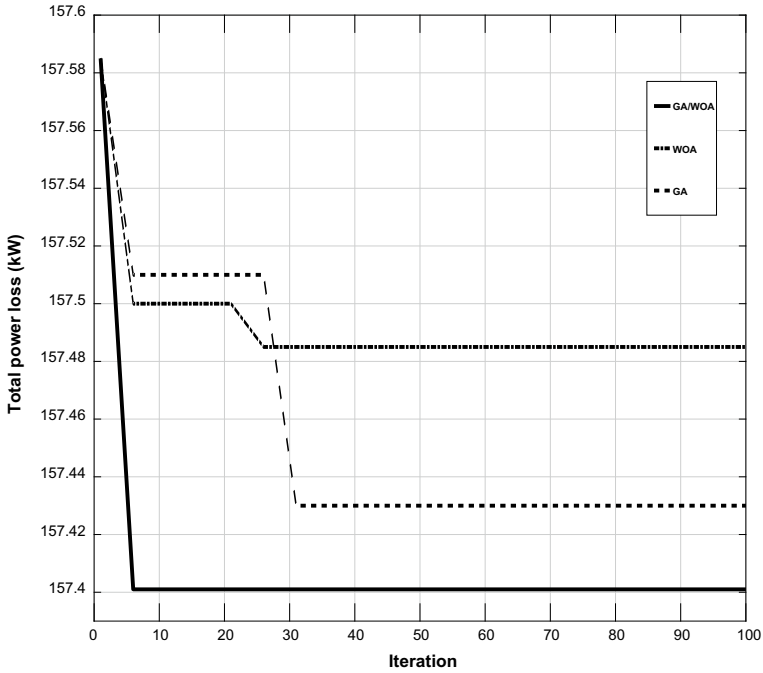


Fig. 11 Computational graph 85 bus

### 5 Conclusion

In this paper, a hybrid genetic algorithm and whale optimization algorithm are proposed to determine the optimal location and size of DGs and implemented on IEEE 15, 33, 69, and 85 bus systems. From the simulation results and considering objective function, active power losses, and reactive power, it can be concluded that the proposed hybrid algorithm has more probability in finding optimum solutions.

### References

1. Colmenar-Santos A, Reino-Rio C, Borge-Diez D, Collado-Fernández E (2016) Distributed generation: a review of factors that can contribute most to achieve a scenario of DG units embedded in the new distribution network. *Renew Sustain Energy Rev* 59:1130–1148
2. El-Zonkoly AM (2011) Optimal placement of multi-distributed generation units including different load models using particle swarm optimization. *Swarm Evol Comput* 1:50–59
3. Nguyenab TT, Truonga AV (2015) Distribution network reconfiguration for power loss minimization and voltage profile improvement using cuckoo search algorithm. *Int J Electr Power Energy Syst* 68:233–242

4. Dinkara Prasad Reddy P, Veera Reddy VC, Gowri Manohar T (2018) Ant lion optimization algorithm for optimal sizing of renewable energy resources for loss reduction in distribution systems. *J Electr Syst Inf Technol* 5:663–680
5. Javidtash N, Jabbari M, Niknam T, Nafar M (2017) A novel mixture of non-dominated sorting genetic algorithm and fuzzy method to multi-objective placement of distributed generations in Microgrids. *J Intell Fuzzy Syst* 33:1–8
6. Grisales-Noreña LF, Montoya DG, Ramos-Paja CA (2018) Optimal sizing and location of distributed generators based on PBIL and PSO techniques. *Energies* 11:1–27
7. Murthy VVSN, Kumar A (2013) Comparison of optimal DG allocation methods in radial distribution systems based on sensitivity approaches. *Int J Electr Power Energy Syst* 53:450–467
8. Suresh MCV, Belwin EJ (2018) Optimal DG placement for benefit maximization in distribution networks by using dragonfly algorithm. *Renew: Wind Water Solar* 5:1–8
9. Mirjalili S (2016) The whale optimization algorithm. *Adv Eng Softw* 95:51–67
10. Baran ME, Wu FF (1989) Optimal sizing of capacitors placed on a radial distribution system. *IEEE Trans Power Delivery* 4:735–743

# The Tensor of Conductivity with the External Fields in Rectangular Quantum Wire for the Case of Electrons—Acoustic Phonon Scattering



Hoang Van Ngoc

## 1 Introduction

Low-dimensional materials have been interested in researchers for a long time, but research on low-dimensional materials is still a research direction that brings many applications in modern science and technology today [1–3]. Application in the manufacture of electronic components brings great achievements. Application in science and technology creates products with scientific value and reduces the cost of the product, making the product smaller, lighter, and more efficient [4–6].

Low-dimensional semiconductor limits the motion of the carriers in the directions. Quantum wires are one-way semiconductors in which electrons are confined in two directions, thin films, superlattices, quantum wells in which electrons are confined in one direction and quantum dots. This work studies rectangular quantum wire, a kind of wire that is quite popular with many scientists interested. In a rectangular quantum wire, the electron moves freely in the  $Oz$  direction and is limited to moving in the  $xOy$  plane. The appearance of the external field causes the electromagnetic properties of the material to change and new effects appear. The photons of fields interact with the atom in the crystal lattice, and these interactions produce different results with the different wavelength of the photons. The shorter the radiation wavelength, the deeper the photon penetrates the atom, creating different effects.

Conductivity tensor is a much-talked-about concept in materials that characterize the electrical properties of materials, so this is an important concept. Conductivity tensor calculation makes more sense when the field is present. In this work, we are interested in electrical conductivity in rectangular wire and consider with optical phonons.

---

H. Van Ngoc (✉)

Institute of Applied Technology, Thu Dau Mot University, Thu Dau Mot, Binh Duong Province, Vietnam

e-mail: [ngochv@tdmu.edu.vn](mailto:ngochv@tdmu.edu.vn)

## 2 Calculating of Conductivity

The expression for the energy and the wave function has the form:

$$\varepsilon_{n,l,p_z} = \frac{p_z^2}{2m} + \frac{\pi^2}{2m} \left( \frac{n^2}{L_x^2} + \frac{l^2}{L_y^2} \right) \tag{1}$$

$$V = 0 \Leftrightarrow \begin{cases} 0 \leq y \leq L_y; \\ 0 \leq x \leq L_x \end{cases}; \quad \psi_{n,l,\bar{p}}(x, y, z) = \sqrt{\frac{1}{L_z}} e^{i\bar{p}z} \sqrt{\frac{1}{L_x}} \sin\left(\frac{n\pi x}{L_x}\right) \sqrt{\frac{2}{L_y}} \sin\left(\frac{l\pi y}{L_y}\right) \tag{2}$$

$$V = \infty \Leftrightarrow \begin{cases} x < 0; x \geq L_y; \\ y < 0; y \geq L_x \end{cases}; \quad \psi_{n,l,\bar{p}}(x, y, z) = 0 \tag{3}$$

Here  $L_x, L_y$  are the size of the wire in the two dimensions of the confined;

$$I_{n,l,n',l'}(\vec{q}) = \frac{32\pi^4 (q_x L_x n n')^2 \left(1 - (-1)^{n+n'} \cos(q_x L_x)\right)}{\left[(q_x L_x)^4 - 2\pi^2 (q_x L_x)^2 (n^2 + n'^2) + \pi^4 (n^2 - n'^2)^2\right]^2} \times \frac{32\pi^4 (q_y L_y l l')^2 \left(1 - (-1)^{l+l'} \cos(q_y L_y)\right)}{\left[(q_y L_y)^4 - 2\pi^2 (q_y L_y)^2 (l^2 + l'^2) + \pi^4 (l^2 - l'^2)^2\right]^2}$$

is form factor.

The density of current:

$$\begin{aligned} \vec{j}(t = 0) &= \int (\vec{R}(\varepsilon) + \vec{R}^*(\varepsilon)) d\varepsilon = \vec{j}_1 + \vec{j}_1^* \\ &= \frac{4e^2 n_0}{m} \frac{\tau(\varepsilon_F)}{1 + \omega^2 \tau^2(\varepsilon_F)} \left\{ \left[ \varepsilon_F - \frac{\pi^2}{2m} \left( \frac{n^2}{L_x^2} + \frac{l^2}{L_y^2} \right) \right] I \right. \\ &\quad \left. + \lambda \frac{\tau(\Omega) [1 - \omega^2 \tau(\Omega) \tau(\varepsilon_F)]}{1 + \omega^2 \tau^2(\Omega)} \right. \\ &\quad \left. - A \frac{\tau(\varepsilon_F) [1 - \omega^2 \tau^2(\varepsilon_F)]}{1 + \omega^2 \tau^2(\varepsilon_F)} \right\} \vec{E} \end{aligned} \tag{4}$$

here

$$A = \frac{e^2 F^2}{2m\Omega^3} \frac{\xi^2 k_B T}{2\rho v_s V} I_{n,l,n',l'}^2 \sqrt{2m \left( \varepsilon_F - \frac{\pi^2}{2m} \left( \frac{n^2}{L_x^2} + \frac{l^2}{L_y^2} \right) \right)} \quad (5)$$

$$\lambda = \frac{e^2 F^2}{2m\Omega^3} \frac{\xi^2 k_B T}{2\rho v_s V} I_{n,l,n',l'}^2 \left\{ \sqrt{2m \left( \varepsilon_F - \frac{\pi^2}{2m} \left( \frac{n^2}{L_x^2} + \frac{l^2}{L_y^2} \right) \right)} \right. \\ \left. \left[ \sqrt{2m \left( \Omega - \frac{\pi^2}{2m} \left( \frac{n^2}{L_x^2} + \frac{l^2}{L_y^2} \right) \right)} - 1 \right] \right\} \quad (6)$$

$\vec{R}(\varepsilon) = -\frac{e}{m} \sum_{n,\vec{p}_z} \vec{p}_z f_1(\vec{p}_z) \delta(\varepsilon - \varepsilon_{n,\vec{p}_z})$  and  $\vec{R}^*(\varepsilon) = -\frac{e}{m} \sum_{n,\vec{p}_z} \vec{p}_z f_1^*(\vec{p}_z) \delta(\varepsilon - \varepsilon_{n,\vec{p}_z})$  are the partial current densities;  $f_1(\vec{p}_z) = -\vec{p}_z \vec{\chi} f'_0(\varepsilon_{n,\vec{p}_z})$ ;  $f'_0(\varepsilon_{n,\vec{p}_\perp}) = \frac{\partial f_0(\varepsilon_{n,\vec{p}_\perp})}{\partial \varepsilon_{n,\vec{p}_\perp}}$ ;  $f_0(\varepsilon_{n,\vec{p}_z}) = \theta(\varepsilon_F - \varepsilon_{n,\vec{p}_z})$ ;  $\vec{\chi} = \frac{e}{m} \vec{E} \frac{\tau(\varepsilon)}{1 - i\omega\tau(\varepsilon)}$ .

We have:

$$\vec{j}(t=0) = \sigma \vec{E}(t=0) = \sigma 2\vec{E} \quad (7)$$

so:

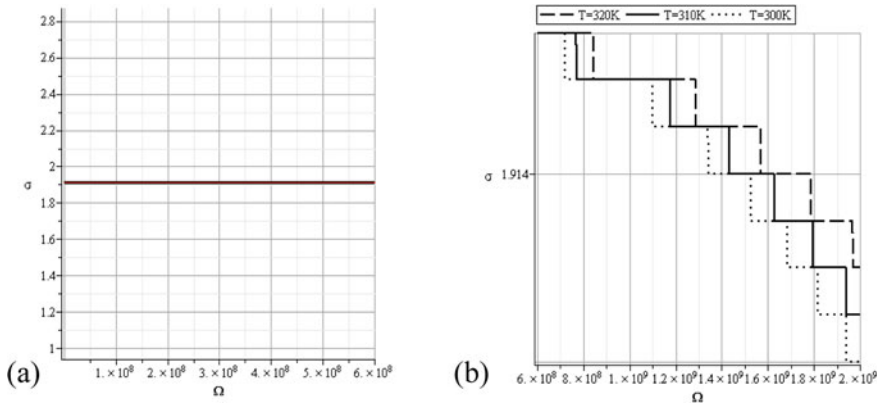
$$\sigma_{ik} = \frac{2e^2 n_0}{m} \frac{\tau(\varepsilon_F)}{1 + \omega^2 \tau^2(\varepsilon_F)} \left\{ \left[ \varepsilon_F - \frac{\pi^2}{2m} \left( \frac{n^2}{L_x^2} + \frac{l^2}{L_y^2} \right) \right] \delta_{ij} \right. \\ \left. + \frac{\tau(\Omega) [1 - \omega^2 \tau(\Omega) \tau(\varepsilon_F)]}{1 + \omega^2 \tau^2(\Omega)} \lambda + -\frac{\tau(\varepsilon_F) [1 - \omega^2 \tau^2(\varepsilon_F)]}{1 + \omega^2 \tau^2(\varepsilon_F)} A \right\} \quad (8)$$

Equation (8) is the required conductivity expression of the system.

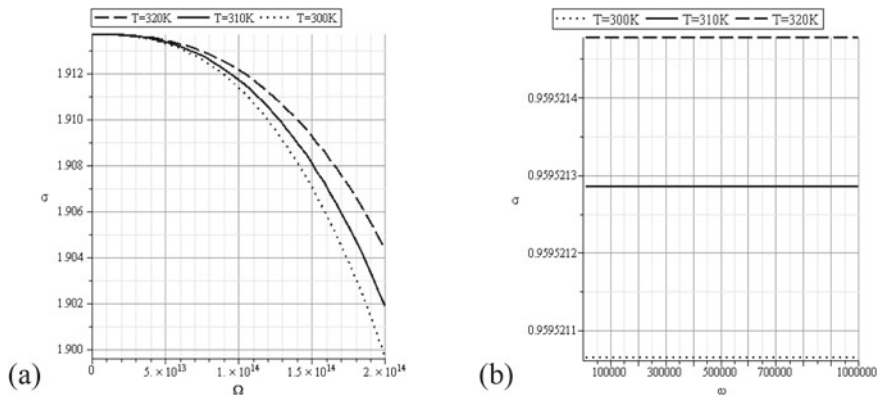
### 3 Discussion

Figure 1a shows the variation of conductivity with laser frequency when it is less than  $6 * 10^8$ , the graph shows that the conductivity is independent of frequency when the frequency is small. When the frequency changes from  $6 * 10^8$  to  $2 * 10^9$ , the conductivity decreases with step function (Fig. 1b). For three different temperatures, we have three separate plots, and for the same frequency, we see that the higher the temperature, the greater the value of the conductivity. This is perfectly suited because the higher the temperature, the greater the kinetic energy of the charged particle and the increase in electrical conductivity. In laser field, frequencies are greater than  $10^{13}$ , conductivity decreases as faster as frequency increases (Fig. 2a).

The variation of conductivity with plane-polarized wave frequency is shown in Fig. 2b. At low frequencies (less than  $10^5$ ), we see that the electromagnetic wave

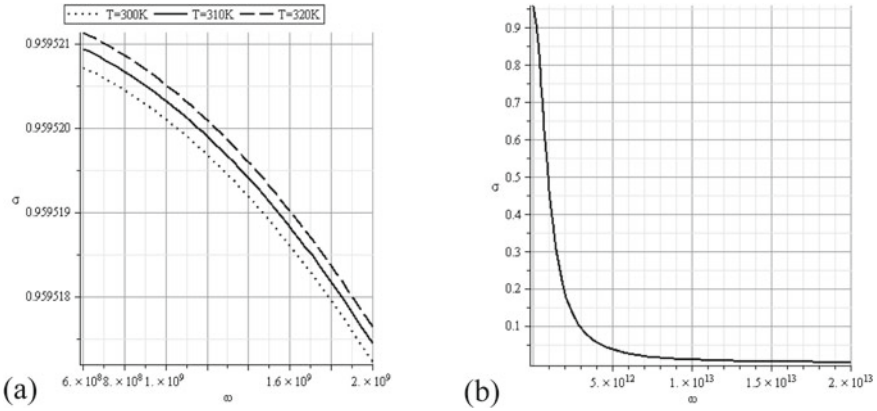


**Fig. 1** Variation of  $\sigma$  on  $\Omega$  when  $\Omega < 6 * 10^8$  (a) and  $6 * 10^8 < \Omega < 2 * 10^9$  (b)

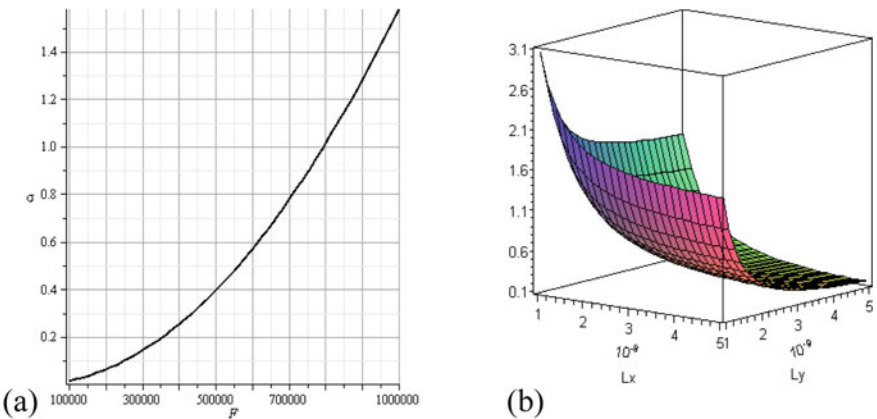


**Fig. 2** Variation of  $\sigma$  on  $\Omega$  when  $\Omega > 10^{13}$  (a) and the dependence of  $\sigma$  on  $\omega$  when  $\omega < 10^5$  (b)

energy is small, so there is no effect on the conductivity. At this frequency, the interaction between electromagnetic waves and matter is negligible. At frequencies from  $6 * 10^8$  to  $2 * 10^9$ , this dependence is shown in Fig. 3a, small decrease in conductivity. When the frequency is greater than  $10^{12}$  tensor conductivity starts to decrease rapidly (Fig. 3b). The conductivity continues to decrease until it reaches an almost constant value when frequency is about  $1.5 * 10^{13}$ . Looking at Fig. 4a, we see that the variation of conductivity with amplitude is very clear. The conductivity increases rapidly as the amplitude increases. Figure 4b shows the variation of conductivity on size, and the conductivity tensor is inversely proportional to the size of the wire.



**Fig. 3** d variation of  $\sigma$  on  $\omega$  when  $6 * 10^8 < \omega < 2 * 10^9$  (a) and  $\omega > 10^{12}$  (b)



**Fig. 4** Variation of  $\sigma$  on  $F$  (a) and the dependence of  $\sigma$  on  $L_x, L_y$  (b)

### 4 Conclusions

The work calculated the conductivity in a rectangular wire and considered the problem applied to acoustic phonons. The work has shown the variation of electrical conductivity according to the parameters of the external field. The graphs are also studied and surveyed. The main conclusion here is that the conductivity varies rapidly with the amplitude of the fields, and the diameter of the wire we are studying.

## References

1. Trien HD, Nhan NV (2011) The nonlinear absorption of a strong electromagnetic waves caused by confined electrons in cylindrical quantum wire. *Prog Electromagn Res Lett* 20:87–96
2. Bau NQ, Trien HD (2009) The nonlinear absorption coefficient of strong electromagnetic waves caused by electrons confined in quantum wires *J Korean Phys Soc* 56(1):120–127
3. Van Ngoc H, Nhan NV, Vương DQ (2017) The photon—drag effect in rectangular quantum wire with an infinite potential. *VNU J Sci Math-Phys* 33(1):53–59
4. Van Ngoc H, Nhan NV, Vuong DQ (2017) The photon—drag effect in cylindrical quantum wire with an infinite potential for the case of electrons—acoustic phonon scattering. *VNU J Sci Math Phys* 33(4):80–86
5. Matula RA (1979) Electrical resistivity of copper, gold, palladium, and silver. *J Phys Chem Ref Data* 8:1147
6. Pawar SD, Murugavel P, Lal DM (2009) Effect of relative humidity and sea level pressure on electrical conductivity of air over Indian Ocean. *J Geophys Res* 114:D02205



# An Implementation of Hybrid CNN-LSTM Model for Human Activity Recognition



Kolla Sai Krishna and Surekha Paneerselvam

## 1 Introduction

Human activity recognition (HAR) aims to recognize, understand, and predict the actions of human(s) from a set of information collected from sensors. Human activity recognition is required in several areas namely, sports, game controls, health care, detection and identification. Monitoring and classification of human action has been a popular area of research in the recent years because of the increasing demands in the field of healthcare specifically in monitoring and caring for elderly people. Historically, collection of data to predict the actions of humans was challenging and expensive until the evolution of smart phones. The use of smart phones and personal tracking devices has become an inalienable part of everyday life. Recent survey shows that 91% of the people around the world own a mobile device, smart in nature. In addition to the basic features, smart phones are capable of computing, sensing and multitasking due to the availability of various sensors embedded in them. In the previous work, about two decades ago, camera was used as the primary source to collect information for human activity recognition. This approach was suitable for applications operating in controlled area and requires infrastructure support for the installation of video cameras in areas under surveillance [1, 2].

During the recent years, the availability of smart phones, wearable devices, personal tracking devices, etc., enables collection of a large amount of data due to the embedded inertial sensors in these devices. Body-worn sensors also known as wearable sensors are suitable for uncontrolled indoor and open-air conditions.

---

K. S. Krishna (✉) · S. Paneerselvam

Department of Electrical and Electronics Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Bengaluru, India

e-mail: [bl.en.p2ebs19005@bl.students.amrita.edu](mailto:bl.en.p2ebs19005@bl.students.amrita.edu)

S. Paneerselvam

e-mail: [p\\_surekha@blr.amrita.edu](mailto:p_surekha@blr.amrita.edu)

The data is typically gathered from a group of body-worn sensors, which are set at different body parts, for example, wrist, chest, lower leg, etc. The signals from the accelerometer and gyroscope can be collected when these devices are in contact with human body. Using the collected data, analysis can be performed to predict and recognize the activities performed by a human (e.g., sit, stand, walk, etc.,) which further helps in decision making.

Deep learning has developed and grown so rapidly in several research areas, particularly in image processing, time-series classification, and natural language processing [3]. Wang et al. [4] used a composed network consists of CNN and LSTM to predict human activities from the data obtained through tri-axial accelerometer and gyroscope. Over the last decade deep learning has grown extraordinarily in the areas of visual feature classification. Several researchers attempted to analyze deep learning models to recognize features based on visual information. In the previous work of action recognition, Fergus et al. [5] utilized CNN-based Boltzmann machine to distinguish visual information and to predict different features. Ji et al. [6] used a CNN layer in 3-D to obtain more activity data through time sequence. Liu et al. [7] recommended that CNN along with conditional random fields can be consolidated for activity separation. CNN is capable of naturally performing space-time attributes, whereas a conditional random field is capable of collecting the relation among activities. Existing deep learning strategies that are generally used are, RNN [8] and LSTM network.

Zeng et al. [9] presented a model considering the single-axis sensor information as 1-D images and transported them to CNN for recognition. Jiang et al. [10] consolidated the sign arrangements of accelerometer and gyroscope into a visual model, empowering deep convolution neural network (DCNN) to consequently gain proficiency with the minimal features from the dynamic image. Chen et al. [11] modified convolution kernel of CNN to change according to the behavior of 3-axial sensor information. Ronao et al. [12] used a convNet, which showcased efficient information for versatile behavior recognition using robust and advanced smart phone sensors.

Murad et al. [13] and Zhou et al. [14] presented three DCNN structures dependent on LSTM to set up recognition models to obtain the shape and time relation on input signal that could achieve more exact recognition. Guan [15] and Qi and Zhang [16] modified LSTM and used a combination model, coordinating various LSTM learners to a single model. The performance analysis detailed that the combined model with LSTM is better than a single LSTM network. Aloysius et al. [17] and Jayadeep et al. [18] presented a CNN-based work for the recognition of finger spelling and detecting Indian sign language, respectively. Aparna et al. [19] presented a CNN-LSTM-based research for detecting Indian sign language.

In this paper, a combined deep learning model CNN-LSTM is proposed to recognize the human activities. The model learns the behavior of the activities based on the sensor data. Upon tuning the parameters, the model provides accurate solution for classification of human activities. The paper is divided into three sections. Section 2 describes the proposed methodology which includes data preprocessing, and feature extraction using CNN-LSTM. Section 3 discusses the hyperparameters of the model

and their influence on the performance of the model for HAR. Section 4 concludes the work.

## 2 Methodology

The description of the dataset used in this work and the preprocessing strategies applied on the data set are discussed in this section. The working of convolutional neural networks (CNN) and the methodology involved in constructing a hybrid end-to-end CNN-LSTM model to perform the HAR classification is also presented.

### 2.1 Dataset

The major aim of human activity recognition (HAR) is to build an optimal machine learning model to classify the human activities accurately. Features and targets from the dataset are fed to the deep learning algorithms not only for predicting activities performed by users but also for training and evaluating the algorithms to produce the final output.

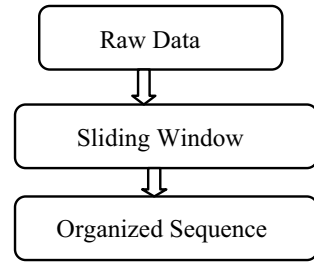
In this work, Human Activities and Postural Transitions [20] HAPT dataset and Human Activity Recognition [21] HAR dataset from UCI machine learning repository are used for training and testing of the model. The HAPT dataset has a total of twelve activities. These twelve activities are further divided into three static, three dynamic and six transition activities. The static activities are namely, sitting, standing and lying, while dynamic activities are walking, walking upstairs and walking downstairs. The transition activities include the changeover between any two static activities [20].

The data is collected from 30 volunteers in the age range of 19–48, the volunteers wore a sensible phone on their body. The data is obtained from the accelerometer and gyroscope in smart phones [20, 21]. The transition activities are of very short range and the basic activities are of high frequency and considerable duration, hence there is a difference in the amount of data between transition activities and other activities. The total amount of data for the six transition activities accounts only to 8% of the total amount of data [4].

### 2.2 Input Stage “Data Preprocessing”

Since the amount of data from the sensors is large, it becomes difficult to enter all the data into a deep learning model at once to perform human activity recognition. The data processing model is shown in Fig. 1. The first stage is the raw data processing stage, which consists of data obtained from accelerometer and gyroscope.

**Fig. 1** Data processing



**Table 1** Data sequence to the corresponding activity

Exp	Label	Start	End
1	5	250	1228
2	7	1229	1764
3	10	1765	2311
4	4	2312	2765

To process this data without any information loss, the sliding window technique is used. This technique splits the data in a manner such that the second split data does not lose the data connectivity from the first split and the third split data does not lose the connectivity from the second split and so on. This continuity pattern is illustrated in Table 1 in start and end columns. Finally the resulting organized sequence of data is used as the input to the deep learning model.

### 2.3 Convolutional Neural Networks

Convolutional neural networks (CNN’s) are deep learning models that take in an input vector, assign importance (learnable weights and biases) to the features in the data such that one feature can be differentiated from the other. Unlike the primitive feature processing methods, the filters and characteristics are not hand-engineered, instead the model is trained to learn and adapt these features to perform the desired functionality. The architecture portrayed by CNN is analogous to that of the connectivity design of neurons in the human brain. CNNs are designed to capture the spatial and temporal feature dependencies in the data through the application of relevant filters, these models are better trained to understand the complexity and sophistication involved in feature set of data.

#### 1. Feature Learning for CNN

A CNN can extract features from the input signal through a convolution operation performed on the signal and a kernel. Generally, this can be achieved by activation function. The activation function is able to detect features captured by the kernels. ReLU activation function is able to neutralize the negative values to zero. The pooling



**Fig. 2** Model for convolutional neural network

**Table 2** The convolution and pooling layer selection

Layers	C1	C2	C3
Channel size	8	18	36
Stride	$1 \times 1 \times 1$	$1 \times 1 \times 1$	$1 \times 1 \times 1$
Channel number	18	36	72
Type of layer	P1	P2	P3
Channel size	18	36	72
Channel number	18	36	72

layer receives local characteristics from the map to perform the convolution with a local sampling operation to reduce the size of the network. The pooling layer and convolution layer form a deep structure that can automatically retrieve activity data from the original activity information. The CNN model for the activity recognition is shown in Fig. 2, where C represents convolution layer and P represents the pooling layer.

The CNN network comprises of three convolution layers and three pooling layers and outputs a number of activity features with feature map. Table 2 show the parameter selection for CNN layer.

### 2.4 Long Short-Term Memory Layer

LSTM layer is a type of recurrent neural network (RNN) has many applications in deep learning. It has feedback connections in the form of cells consisting of a memory element to remember time intervals of the sequence. A common LSTM unit consists of an input gate, an output gate and a forget gate [4]. These three gates regulate the information processing in and out of cell. In this work, LSTM layer is used in combination with CNN to obtain improved recognition rate of the transition activities. The LSTM layer presented in this paper is shown in Fig. 3.

The output of the LSTM layer is an array consisting of time and activity related data that is fed to fully connected layer to obtain global activity characteristics. The



**Fig. 3** LSTM architecture for feature extraction

learning process of the deep learning model is complicated because the information associated at the input of each layer changes with the parameters of the previous layer. In order to prevent the information from changing rapidly, there is a need to decrease the training speed, which can be achieved by using learning rate. Then further, batch normalization (BN) is applied to normalize the values of each layer in LSTM (to improve the activity recognition). Further, these normalized features are presented to Softmax activation function to obtain the activity features for activity classification.

### 2.5 Implementation of the Model

The overall architecture of the proposed hybrid model based on CNN and LSTM is shown in Fig. 4. To improve recognition rate, small batch sizes of data is used during training and testing of the model. A dropout layer is also introduced with the input layer, which sets the improperly selected data to zero during the initial training of the model and access neurons to build a stronger model by not completely depending on specific neurons.

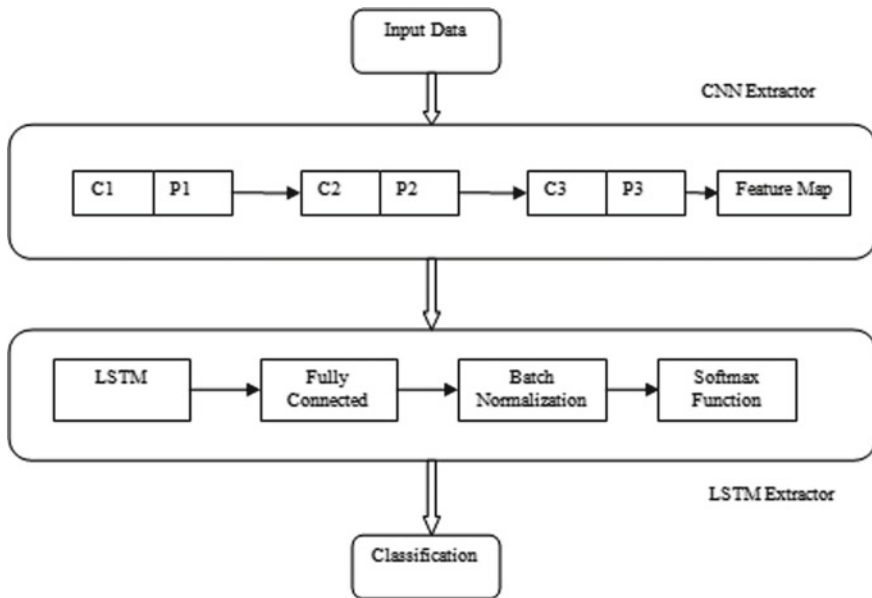


Fig. 4 Overall architecture of the proposed model

**Table 3** Effect of number of neurons on recognition rate

Number of neurons	Recognition rate (HAPT)	Recognition rate (HAR)	Recognition time (min)
8	89.56	90.06	11
16	92.01	91.34	14
32	91.92	91.42	19
64	94.21	92.60	20

### 3 Analysis and Results

The proposed model is implemented in Tensor flow [4] using Python libraries. The training of the model was performed on a conventional computer with a 2.4 GHz central processor unit and 8 GB memory. The deep learning model is evaluated based on variables called as hyperparameters. Choosing these hyperparameters play an important role in producing a better recognition rate and less recognition time. An analysis was performed on the hybrid network by tuning the hyperparameters such as neurons, batch normalization, learning rate, batch size and dropout. The performance of these parameters on the recognition rate and recognition time are monitored and discussed in the following sections.

#### 3.1 Neurons in LSTM

The performance of the number of neurons in LSTM layer determines complexity of the model. As the neurons are less, the network does not have the necessary training and information processing capabilities, and the recognition rate is low. For increased number of neurons, the recognition rate increases and training ability of the network slows down. In the proposed model, the neurons are varied from 8 to 64 and the recognition rate for both HAPT and HAR dataset is evaluated along with the recognition time. The results are presented in Table 3. It is observed that the recognition rate is less as the neurons is 8 and the recognition rate reaches 94.21% when the number of neurons is 64 for HAPT dataset. The recognition time also increases with increase in the number of neurons. Compromising the recognition time, and giving more emphasis on the recognition rate, the number of neurons in the model is selected as 64 for further analysis.

#### 3.2 Learning Rate

The learning rate parameter controls the step size of the model with respect to the estimated error of the model. Choosing a specific learning rate is a challenge, as too low a value can result in a long training time, while too high a value can lead

**Table 4** Performance of learning rate on recognition rate

Learning rate	Recognition rate (HAPT)	Recognition rate (HAR)
0.001	91.10	90.19
0.0015	91.14	92.60
0.002	92.41	90.49
0.0025	94.21	91.81
0.003	91.07	90.27
0.0035	91.31	90.13
0.004	91.12	91.02
0.0045	91.87	90.84

**Table 5** Effect of BN on recognition rate

S. No.	Recognition rate (HAPT)	Recognition rate (HAR)
Without BN	91.18	90.21
With BN	94.21	92.60

to learning an unstable training process. Experiments are carried out with a range of learning rates from 0.001 to 0.0045. As demonstrated in Table 4, it is observed that with the increase in learning rate, the recognition rate of the model increases. However the model reaches a maximum recognition rate of 94.21% at a learning rate of 0.0025 for the HAPT dataset. For the HAR dataset, a maximum recognition rate of 92.60% at a learning rate of 0.0015.

### 3.3 Batch Normalization (BN)

BN is used for training a neural network to standardize the input information presented to a layer as mini-batches. It stabilizes the learning process and reduces the number of training sessions required to train deep neural networks. Table 5 shows the influence of recognition rate with and without batch normalization. An improved recognition rate is observed while adding the BN, since the learning process is stabilized while learning in batches.

### 3.4 Batch Size

The batch size parameter refers to the number of training samples used in iteration. To feed the entire data in one epoch is computationally expensive and hence, the dataset is divided into several parts which are called batches. In this work, the batch



**Table 6** Performance of batch size on recognition rate

Batch size	Recognition rate (HAPT)	Recognition rate (HAR)
25	90.21	90.01
50	90.83	90.21
75	91.20	91.15
100	90.89	90.89
150	94.21	92.60
200	92.60	92.56
250	92.54	92.54
300	91.20	91.20
350	92.19	92.12
400	91.93	91.73

**Table 7** Effect of dropout on recognition rate

Dropout	Recognition rate (HAPT)	Recognition rate (HAR)
0.20	91.89	90.49
0.25	92.25	91.89
0.50	94.21	92.60
1.00	93.12	91.53

size is varied from 25 to 400. It is demonstrated from Table 6, for a batch size of 150 the recognition rate of the model is 94.21%. Choosing a batch size generally depends on the size of dataset and with minimum chance of over fitting.

### 3.5 Dropout

Dropout layer helps to reduce over fitting in the training by randomly selecting the neurons. Dropout layer can be used with input layers and not with output layers. In this work dropout is varied between 0.20 and 1.00 and the observations are recorded as shown in Table 7. It is observed that the recognition rate is 94.21% when the dropout is 0.50.

### 3.6 Hyperparameters

Based on the experiments carried out by varying the hyperparameters, the optimal choice of these parameters for the CNN-LSTM for both HAPT and HAR dataset are listed in Table 8.

**Table 8** Best parameters of the model

Parameters	HAPT	HAR
Input array size	150	150
Input Channels	8	8
Dropout	0.5	0.5
Activation function	ReLU, Softmax	ReLU, Softmax
Conv Kernel size	3	3
Pooling size	2	2
LSTM layer	1	1
Neuron number	64	64
Learning rate	0.0025	0.0015
Batch size	150	150
Epoch	400	400

### 3.7 Classification Report

The measure of correct predictions for a classification algorithm is given by the classification report. Metrics such as accuracy, precision, recall and F1\_score are evaluated using true positives (TP), true negatives (TN), false positives (FP) and false negatives (FN).

#### 1. Accuracy ( $A$ )

It is characterized as the quantity of right expectations over the complete number of forecasts made by the model. This is a decent pointer, particularly if the objective variable classes in the information are coordinated.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (1)$$

#### 2. Precision

The total true positive results to the total number of predicted results calculated by the classifier.

$$\text{Precision } (P) = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (2)$$

#### 3. F1\_Score

The F1 score is considered to be good if the overall number of false positives and false negatives is low on test set. It is defined as the harmonic mean of precision and recall ( $R$ ).

$$\text{F1\_Score} = \frac{2 * P * R}{P + R} \quad (3)$$

**Table 9** Classification report of activities

Activity	Precision	Recall	F1_Score
Walking	0.96	0.95	0.95
Walking downstairs	0.91	0.88	0.90
Walking upstairs	0.95	0.96	0.95
Standing	0.90	0.88	0.89
Sitting	0.94	0.93	0.94
Lying	0.94	0.90	0.92
Stand to sit	0.89	0.86	0.88
Sit to stand	0.90	0.88	0.89
Stand to lie	0.88	0.86	0.87
Sit to lie	0.81	0.79	0.80
Lie to sit	0.83	0.84	0.84
Lie to stand	0.77	0.71	0.74

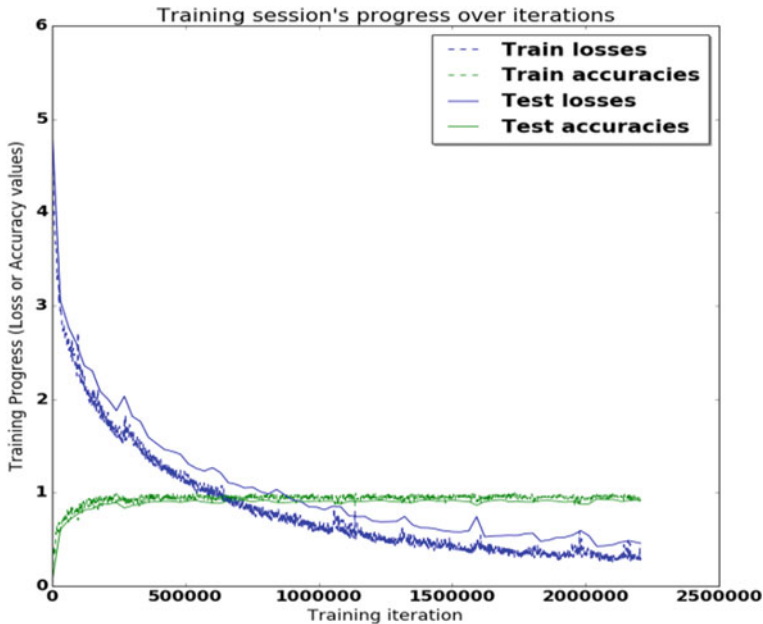
The classification report is presented as shown in Table 9 which consists of metrics such as precision, F1\_score and recall computed using Eqs. (1)–(3).

### 3.8 Train and Test Verification

The train and test verification of the model explains about the loss and accuracy of the activities for improper training or missed data points of the activity. Figure 5, shows the train and test loss for HAPT dataset. In the beginning stage, as the number of iterations is less, train and test losses of activities are higher, which is approximately equal to 2.5 times the error for a single predicted activity. Gradually, as the number of iterations is increased train and test losses are minimized and only few transition activities are not correctly predicted. The reason for this incorrect prediction is due to the less amount of data related to transition activity in the HAPT dataset.

## 4 Conclusion

An analysis of deep learning model for human activity recognition has been experimented on HAPT and HAR dataset. In the CNN-LSTM model, CNN obtains the features of the data and LSTM model obtains the relationship between global features and local features. From the results it can be concluded that the model is trained and learned to meet the objective of human activity recognition. It is observed that maximum recognition rate of 94.21% is obtained for a batch size of 150. The other hyperparameters of the model are almost constant for both the datasets except for the learning rate; 0.0025 for HAPT dataset, and 0.0015 for HAR dataset. The evaluation



**Fig. 5** Loss of train and test accuracies

metrics are also computed for the static, dynamic and individual transition activities. In future, more complex activities such as driving, cooking and swimming can be added to the dataset and the performance of the model can be evaluated. Also recent research shows that the integration of the deep learning model and shallow learning model can perform better classification of activities. The deep learning model can decrease the training time and the shallow learning model can achieve improved accuracy. The integration between the two models could be explored in future for more accurate and reliable recognition.

## References

1. Golestani N, Moghaddam M (2020) Human activity recognition using magnetic induction-based motion signals and deep recurrent neural networks. *Nat Commun* 11(1):1–11
2. Jain A, Kanhangad V (2017) Human activity classification in smartphones using accelerometer and gyroscope sensors. *IEEE Sens J* 18(3):1169–1177
3. LeCun Y, Bengio Y, Hinton G (2015) Deep learning. *Nature* 521(7553):436. <https://doi.org/10.1038/nature14539>
4. Wang H, Zhao J, Li J, Tian L, Tu P, Cao T, An Y, Wang K, Li S (2020) Wearable sensor-based human activity recognition using hybrid deep learning techniques. *Secur Commun Netw*
5. Taylor GW, Fergus R, LeCun Y, Bregler C (2010) Convolutional learning of spatio-temporal features. In: *European conference on computer vision*. Springer, Berlin, pp 140–153

6. Ji S, Xu W, Yang M, Yu K (2012) 3D convolutional neural networks for human action recognition. *IEEE Trans Pattern Anal Mach Intell* 35(1):221–231
7. Liu C, Liu J, He Z, Zhai Y, Hu Q, Huang Y (2016) Convolutional neural random fields for action recognition. *Pattern Recogn* 59:213–224
8. Cho K, Van Merriënboer B, Gulcehre C, Bahdanau D, Bougares F, Schwenk H, Bengio Y (2014) Learning phrase representations using RNN encoder-decoder for statistical machine translation. *arXiv preprint arXiv:1406.1078*
9. Zeng M, Le TN, Bo Y et al (2015) Convolutional neural networks for human activity recognition using mobile sensors. In: Proceedings of the 2014 6th international conference on mobile computing, applications and services, Austin, TX, USA, pp 197–205
10. Jiang W, Yin Z (2015) Human activity recognition using wearable sensors by deep convolutional neural networks. In: Proceedings of the 23rd ACM international conference on Multimedia, pp 1307–1310
11. Chen Y, Xue Y (2015) A deep learning approach to human activity recognition based on single accelerometer. In: 2015 IEEE international conference on systems, man, and cybernetics. IEEE, pp 1488–1492
12. Ronao CA, Cho SB (2016) Human activity recognition with smartphone sensors using deep learning neural networks. *Expert Syst Appl* 59:235–244
13. Murad A, Pyun JY (2017) Deep recurrent neural networks for human activity recognition. *Sensors* 17(11):2556
14. Zhou J, Sun J, Cong P et al (2019) Security-critical energy-aware task scheduling for heterogeneous real-time MPSoCs in IoT. *IEEE Trans Serv Comput (TSC)* 12:99
15. Guan Y, Plötz T (2017) Ensembles of deep lstm learners for activity recognition using wearables. *Proc ACM Interact Mob Wearable Ubiquitous Technol* 1(2):1–28
16. Qi L, Zhang X, Dou W, Hu C, Yang C, Chen J (2018) A two-stage locality-sensitive hashing based approach for privacy-preserving mobile service recommendation in cross-platform edge environment. *Futur Gener Comput Syst* 88:636–643
17. Aloysius N, Geetha M (2020) A scale space model of weighted average CNN ensemble for ASL finger spelling recognition. *Int J Comput Sci Eng* 22(1):154–161
18. Jayadeep G, Vishnupriya NV, Venugopal V, Vishnu S, Geetha M (2020) Mudra: convolutional neural network based Indian sign language translator for banks. In: 2020 4th International conference on intelligent computing and control systems (ICICCS). IEEE, pp 1228–1232
19. Aparna C, Geetha M (2019) CNN and stacked LSTM model for Indian sign language recognition. In: Symposium on machine learning and metaheuristics algorithms, and applications. Springer, Singapore, pp 126–134
20. Weiss GM, Lockhart JW, Pulickal TT, McHugh PT, Ronan IH, Timko JL (2016) Actitracker: a smartphone-based activity recognition system for improving health and well-being. In: 2016 IEEE international conference on data science and advanced analytics (DSAA). IEEE, pp 682–688
21. Anguita D, Ghio A, Oneto L, Parra X, Reyes-Ortiz JL (2013) A public domain dataset for human activity recognition using smartphones. In: Esann, vol 3, p 3

# Sliding Mode Control-Based Standalone Wind Energy System



R. M. Meenakshi and K. Selvi

## 1 Introduction

The conflict in growing need of electricity and age old way of long distribution lines paves way for an economic alternative, namely the microgrids. A microgrid is a low-voltage network with small-scale power generation units, energy storage devices and loads [1]. The present trend in microgrid is the power generation from renewable energy sources. Microgrid with renewable resources can help to eradicate the environmental concerns due to fossil fuels. As the nature of renewable sources and the modern day load are predominantly DC, the need for a common DC interface arises. The integration of present AC utility grid and the DC sub-grids evolves a complex network of hybrid AC/DC microgrids [2]. The power conversion between the DC and AC interface is performed with the interlinking converters. These power converters act as a brain of the microgrids by controlling the voltage and current levels. The efficiency of power transfer in microgrid relies on the proper control strategy to control the operation of switching converters. Since the last decade, many researchers have suggested different control strategies for stable operation of microgrid [3, 4]. The conventional cascaded linear control and the droop control techniques are widely discussed for various configuration and operating scenarios of microgrid [5–8]. These techniques suffer from slow response, complex tuning of control parameters and inefficiency in accommodating multiple constraints. One of the main issues in employing renewable energy sources for generation is the intermittent nature of power output.

---

R. M. Meenakshi (✉) · K. Selvi  
Department of Electrical and Electronics Engineering, Thiagarajar College of Engineering,  
Madurai, TN 625015, India  
e-mail: [meenakshirm1@student.tce.edu](mailto:meenakshirm1@student.tce.edu)

K. Selvi  
e-mail: [kseee@tce.edu](mailto:kseee@tce.edu)

The linear controllers fail to dynamically adapt for the varying operating condition and the environment. Following the linear controllers, many works report the performance of the fuzzy [9] and neural network [10] controllers. These works majorly represents microgrid with single sources. Also, they emphasize on providing a clear knowledge about different operating scenarios to the controller. This makes the controller inefficient during uncertainties. As the renewable resources do not provide continuous power output, it is necessary to take care of reliability in power supply. The energy storage systems and other reliable energy sources can ensure the continuity of power supply in microgrid. The controller must promptly respond under all operating conditions, handling multiple sources and loads. In view of these necessities, the research studies have shifted the focus on nonlinear controllers. Certain nonlinear controllers, namely backstepping controller [11], sliding mode controllers, etc., have been widely discussed for both grid interfaced and standalone microgrids.

Among these controllers, the sliding mode control (SMC) strategy shows a fast dynamic response and ensures system stability. The SMC is featured by its robustness against any disturbances. In a previous research work [12], this controller is employed for power extraction in DC sources by tracking the parameters at device level of renewables. The control of power flow between the system and the battery is less reported. Though the SMC proves better when two wind energy conversion systems interact with the grid [13], both the machines will fail to supply the loads under less windy conditions. This can increase the burden on the grid. Also these works partly still rely on PI controllers for error tracking of parameters. This can deteriorate the performance of SMC. This controller is further reported for solar photovoltaic systems, and the performance is analyzed for standalone mode [14]. The power balance between the DG source and energy storage system with the help of SMC can prove to be a better solution in handling the uncertainties of the environment and other operating disturbances.

This paper aims to analyze the performance of SMC in a Wind PMSG-battery based standalone microgrid. The aim of our work is to employ the sliding mode control approach to obtain: (1) optimal power output from wind energy according to wind speed variations, (2) bidirectional power flow control between the battery storage system and the microgrid and (3) ensure stable operation of the microgrid under all operating conditions. Further sections present the description of components in microgrid, the modeling of each component with the corresponding power electronic converters, the application of sliding mode strategy and the discussion on results under a few operating scenarios.

## 2 Description of Proposed Wind-Battery Microgrid

In this paper, a standalone microgrid based on PMSG-battery system is presented, as shown in Fig. 1. The proposed system consists of PMSG-driven wind energy system, a switched mode rectifier, battery storage system with a bidirectional DC-DC converter, DC link capacitor, voltage source inverter, LC filters and AC-DC loads.

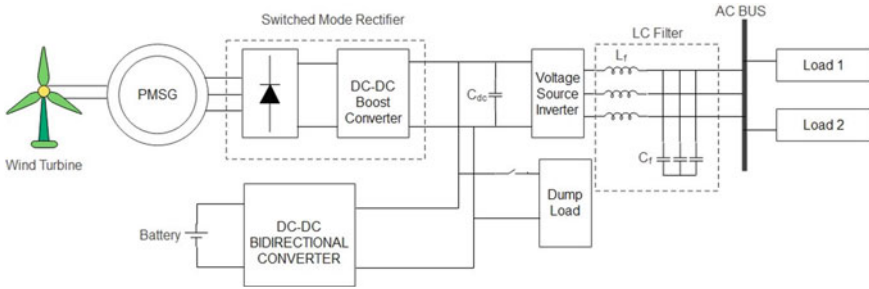


Fig. 1 Proposed standalone microgrid

A SMC-based controller is proposed to control the power electronics converters in the proposed system. Appendix I gives the detailed specifications of each component.

### 3 Modeling of Wind Energy Conversion Systems

A wind energy system as in Fig. 2 is modeled with a wind turbine driving the PMSG machine directly without any gear box. The output of the PMSG machine is controlled by employing a switched mode rectifier built with diode rectifier and DC-DC boost converter. A wind turbine system converts kinetic energy into mechanical energy to drive the electrical machine to generate power. The mechanical power extracted by the wind turbine is given by,

$$P_m = 0.5 * \rho * C_p(\lambda, \beta) * A * v_w^3 \tag{1}$$

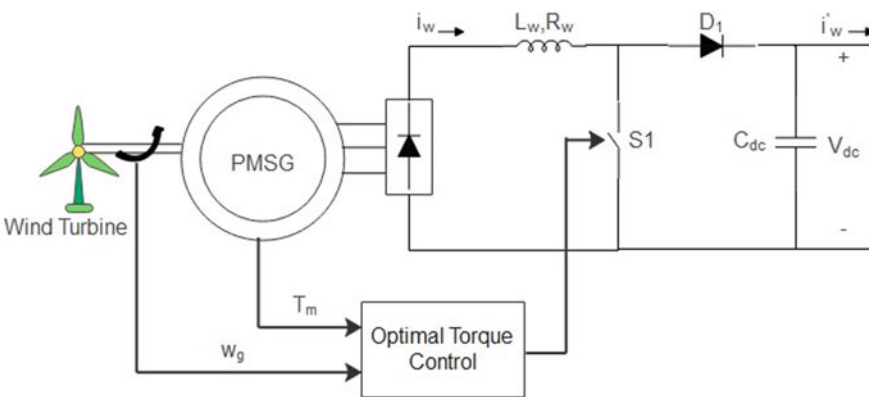


Fig. 2 Wind energy system



where  $C_p$  is the coefficient of performance,  $\lambda$  is the tip speed ratio (TSR),  $\beta$  is the pitch angle,  $A$  is the area of turbine blade ( $m^2$ ),  $\rho$  is the air density ( $kg/m^3$ ), and  $v_w$  is the wind velocity ( $m/s^2$ ).

To extract maximum power from the wind turbine, the optimum tip speed ratio required is,

$$\lambda_{opt} = \frac{\omega_w * r}{v_w} \tag{2}$$

where  $\omega_w$  is the angular speed of the rotor (rad/s) and  $r$  is the turbine radius (m).

By substituting for wind speed  $v_w$  in (1), the optimum mechanical power from the turbine is obtained as,

$$P_m = 0.5\rho\pi r^5 \frac{\omega_m^3}{\lambda_{opt}^3} C_{pmax} \tag{3}$$

where  $C_{pmax}$  is the maximum power coefficient to achieve MPPT. Equation (3) can be written as,

$$P_{m-op} = K_{p-op}\omega_m^3 \tag{4}$$

where

$$K_{p-op} = 0.5\rho\pi r^5 \frac{C_{pmax}}{\lambda_{op}^3}$$

As per the torque power relationship,  $P_{m-op} = \omega_m T_{m-op}$ , and from Eq. (4), the expression for optimum torque can be expressed as

$$T_{m-op} = K_{p-op}\omega_m^2 \tag{5}$$

To extract the maximum power, by using optimum torque control (OTC) technique, a reference torque related to the reference current  $I_{wref}$  is determined by the following equation:

$$I_{wref} = \frac{T_{m-op} * \omega_m}{V_w} \tag{6}$$

where  $V_w$  is the output DC voltage measured from the DC side of diode rectifier. The voltage and current equations for the DC-DC converter can be derived as follows:

$$\begin{bmatrix} \frac{dI_w}{dt} \\ \frac{dV_{dc}}{dt} \end{bmatrix} = \begin{bmatrix} -\frac{R_w}{L_w} & -(1-d_1)\frac{1}{L_w} \\ (1-d_1)\frac{1}{C_{dc}} & 0 \end{bmatrix} \begin{bmatrix} I_w \\ V_{dc} \end{bmatrix} + \begin{bmatrix} \frac{1}{L_w} \\ 0 \end{bmatrix} V_w + \begin{bmatrix} 0 \\ -\frac{1}{C_{dc}} \end{bmatrix} I'_w \tag{7}$$

where  $I_w$ ,  $V_{dc}$  and  $d_1$  are the wind current, DC link voltage and the pulse signal for boost converter switch  $S_1$ , respectively.

## 4 Modeling of Battery Storage System

A lead acid battery is designed to balance the power in the microgrid system with proper discharging and charging using a bidirectional buck boost converter as shown in Fig. 3. The battery is coupled to the PMSG system at the DC link through the converter. The voltage and current equation of the converter can be written as:

$$\begin{bmatrix} \frac{dI_{\text{bat}}}{dt} \\ \frac{dV_{\text{dc}}}{dt} \end{bmatrix} = \begin{bmatrix} -\frac{R_{\text{bat}}}{L_{\text{bat}}} & -d_{23} \frac{1}{L_{\text{bat}}} \\ d_{23} \frac{1}{C_{\text{dc}}} & 0 \end{bmatrix} \begin{bmatrix} I_{\text{bat}} \\ V_{\text{dc}} \end{bmatrix} + \begin{bmatrix} \frac{1}{L_{\text{bat}}} \\ 0 \end{bmatrix} V_{\text{bat}} + \begin{bmatrix} 0 \\ -\frac{1}{C_{\text{dc}}} \end{bmatrix} I'_{\text{bat}} \quad (8)$$

where  $I_{\text{bat}}$ ,  $V_{\text{dc}}$  and  $d_{23}$  are the battery current, DC link voltage and the control signal for the switches  $S_2$  and  $S_3$  of the bidirectional converter.

## 5 Sliding Mode Control of Wind and Battery Converters

Figure 4 shows the schematic implementation of sliding mode control. The main purpose of the controller discussed in this paper is:

1. DC bus voltage regulation at point of common coupling of wind and energy storage system.
2. Reliable supply to the AC loads in standalone mode.

For proper control operation, the error tracking trajectories are given as:

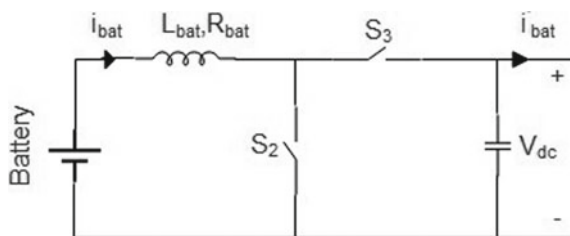
$$\begin{Bmatrix} e_1 \\ e_2 \end{Bmatrix} = \begin{Bmatrix} x_1 - x_1^* \\ x_2 - x_2^* \end{Bmatrix} \quad (9)$$

where  $x_1^*$  and  $x_2^*$  are the wind and battery reference currents. The basic necessity of a sliding mode controller is to define a sliding surface for each control parameter. The sliding surfaces are given by the error surfaces as:

$$S_1 = e_1 \quad (10)$$

$$S_2 = e_2 \quad (11)$$

**Fig. 3** Bidirectional battery converter



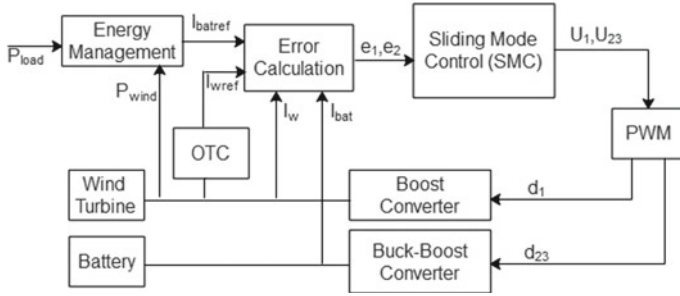


Fig. 4 Proposed controller structure

Next, the derivative of the sliding surface is given as

$$\dot{S}_1 = \dot{e}_1 \tag{12}$$

$$\dot{S}_2 = \dot{e}_2 \tag{13}$$

Substituting for  $\dot{e}_1$  in (12) gives:

$$\dot{S}_1 = \frac{V_w}{L_w} - \frac{R_w}{L_w}x_1 - (1 - d_1)\frac{x_3}{L_w} - x_1^* \tag{14}$$

To achieve stable operation, the derivative of surface must be selected as:

$$\dot{S}_k = -\rho_k \text{sgn}(S_k) \tag{15}$$

where  $\rho_k > 0$  and  $k = 1, 2$ . The  $\text{sgn}$  represents the signum function and can be written as:

$$\text{sgn}(x) = \begin{cases} \frac{x}{|x|} & x \neq 0 \\ 0 & x=0 \end{cases} \tag{16}$$

By applying (15) in (14), the control law  $d_1$  can be obtained as

$$d_1 = \frac{L_w}{x_3}[-\rho_1 \text{sgn}(S_1) - \frac{V_w}{L_w} + \frac{R_w}{L_w}x_1] + \frac{L_w}{x_3}[x_1^*] + 1 \tag{17}$$

The  $d_1$  is the duty cycle signal applied to the boost converter of wind energy system connected at the common DC bus. It is bounded by  $0 < d_1 < 1$ . For a battery storage system, the derivative of sliding surface and the control law can be obtained as:

$$\dot{S}_2 = \frac{V_{bat}}{L_{bat}} - \frac{R_{bat}}{L_{bat}}x_2 - x_3^* \tag{18}$$

$$d_{23} = \frac{L_{bat}}{x_3} [\rho_2 \text{sgn}(S_2) + \frac{V_{bat}}{L_{bat}} - \frac{R_{bat}}{L_{bat}} x_2] + \frac{L_{bat}}{x_3} - x_2^* \tag{19}$$

where  $d_{23}$  is the duty cycle signal for the bidirectional converter between the battery and DC bus system. It is bounded between 0 and 1.

The control signals for the voltage source inverter are provided by the conventional DC control according to the implementation in [15].

## 6 Results and Discussion

The proposed control strategy was implemented in the MATLAB/SIMULINK platform, with detailed design of each system components in standalone microgrid. The wind speed and load disturbances are applied similar to the practical scenarios, in order to analyze the robustness of the sliding mode controller. Such variations are employed to show how effectively the proposed controller regulates the DC bus voltage. The controller performance is assessed under two conditions, namely (a) variable load and (b) fluctuating wind conditions. Combining these two scenarios, the stability of DC link, power balance by the energy storage system and optimal power extracted from wind were investigated. Figure 5 shows the variations in wind profile, and step changes in load are presented. The AC linear loads of 3, 2 and 2 kW are each stepped in at  $t = 0, 7$  and 14s. The wind speed varies continuously withing the range of 11–13m/s.

Initially, the battery SOC is 65%. The constant load prevailing in the microgrid is 3 kW. The wind power generation is around 6 kW as per the PMSG machine rating. The excess power is employed to charge the battery. Figures 6a–b show the wind power generation and increase in battery SOC. At  $t = 7$  s, a load disturbance of 2kW is introduced. The battery still charges at a slower rate, while wind energy supplied to the change in load. Followed by this, another load of 2 kW is again introduced at  $t = 14$  s, where the load demand becomes higher than the generation. At this point, the battery starts to discharge and supports to maintain the power balance in the microgrid.

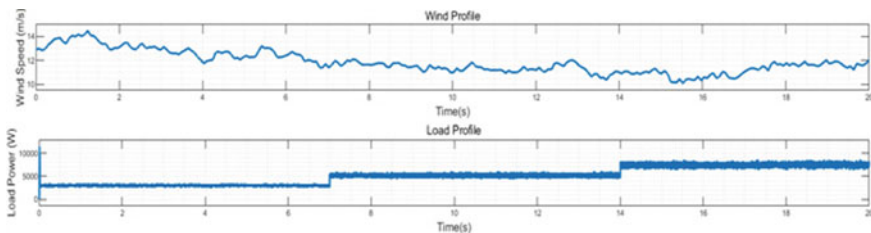
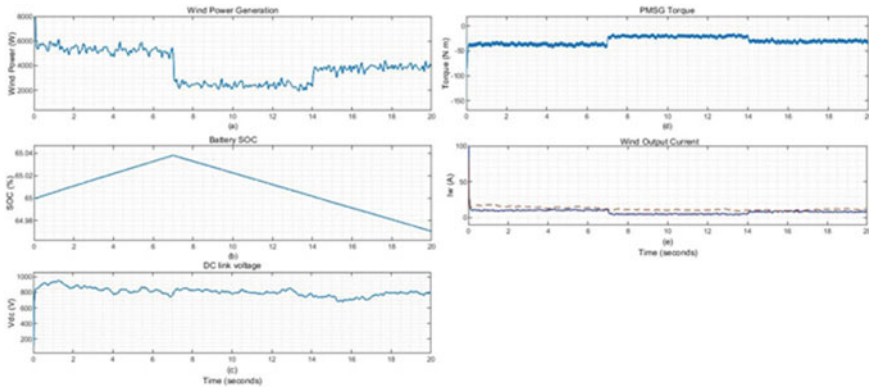


Fig. 5 Wind and load profile



**Fig. 6** a Wind power. b Battery SOC. c DC link voltage. d PMSG torque. e Wind output current

**Table 1** PI and SMC controller specifications

Specifications	SMC controller	PI controller
Overshoot	18.0457%	24.1356%
Settling time (s)	19.9391	19.9360
Peak time (s)	1.2739	1.2732

During such continuous wind variations and step changes in load, the DC link voltage as shown in Fig. 6c is maintained constant. The stabilization of DC link voltage achieved by the sliding mode control strategy is compared with the conventional PI controller. From this comparison, it is seen that the overshoot due to proposed SMC controller is lower compared to the conventional controller. The table shows the time-domain specification comparison of the two controllers (Table 1).

Figure 6d confirms the optimal torque control showing constant torque for the PMSG machine, which is achieved by tracking the diode rectifier output current (Fig. 6e) as per Eq. 6.

Under similar conditions of wind speed and load changes, the case of battery SOC of 95% is examined. In this scenario, during excess power generation, a dump load is introduced to balance the DC link voltage. Figure 7a shows that battery only discharges when there is a demand in power at  $t = 14$  s with a 2 kW load disturbance. Until then, the battery remains ideal, diverting the power to the dump load. Figure 7b shows that the DC link voltage remains constant as balanced by the dump load. Figure 7c shows the dump load power support during excess generation duration of 0–14s. Figure 8 shows that the AC side load voltage is maintained with the load current variations according to the load profile.

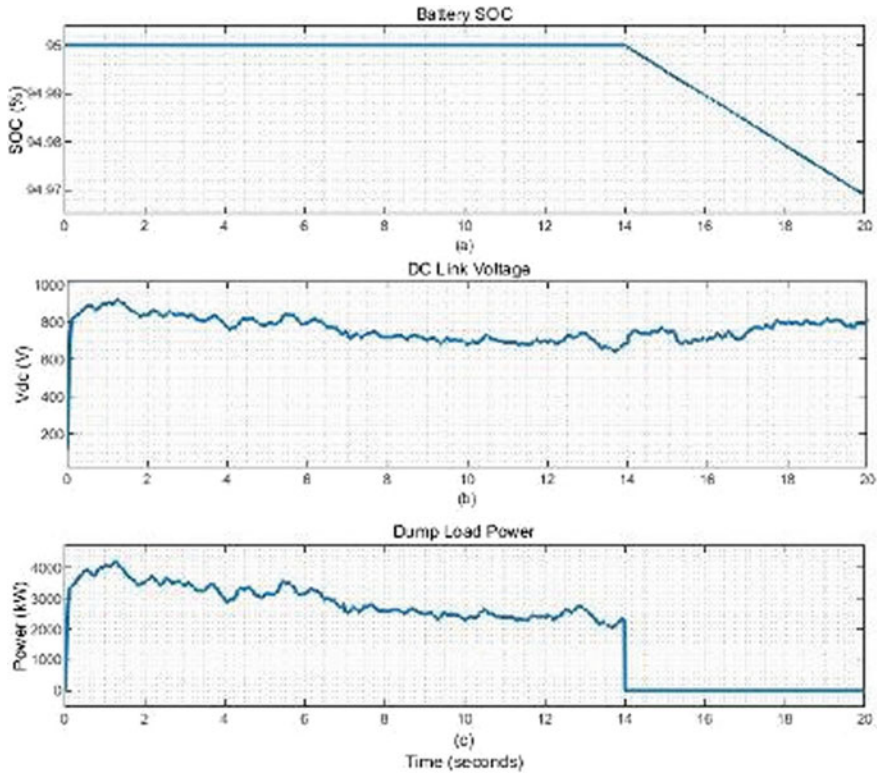


Fig. 7 a Battery SOC. b DC link voltage. c Dump load power

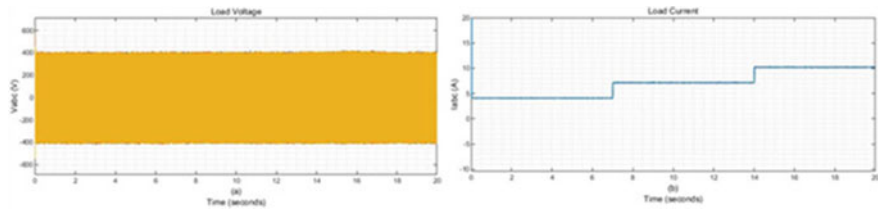


Fig. 8 a Load voltage and b Load current

## 7 Conclusion

This paper presents the design and analysis of sliding mode controller for maintaining DC link voltage and power balance in the standalone hybrid AC/DC microgrid. The performance of the SMC is evaluated under certain scenarios in the SIMULINK environment. It is seen that the dynamics in wind profile and load changes are better addressed by the sliding mode controller in comparison to the PI controller. Also,

the battery storage system effectively participates under excess and deficit power demands. The parameters of SMC can be tuned easier in comparison to the PI controllers. It is seen that though the SMC provides a better adaptation to disturbances, still further improvements are much necessary to tightly control the DC link voltage. To overcome such limitations, one or more reliable generation source could be further introduced for proper power balance. Also, the controller could be better designed to handle the higher order dynamics. Further, the control approach could be extended to the interaction of microgrid with utility grid and transition between islanded and grid-connected modes.

**Acknowledgements** This research work was supported by the Research Excellence strand of Savitha Project-Thiagarajar Research Fellowship under Ref: TCE/Dean-RD/TRF (2020–2023).

## Appendix

### *Design Specifications*

#### Wind turbine model

Air density	1.225 kg/m <sup>3</sup>
Rated wind speed	12 m/s
Rotor diameter	1.16 m
$\lambda_{opt}$	7.5
$C_p$	0.43
<i>PMSG</i>	
Stator phase resistance	0.425 $\Omega$
Armature inductance	8.5 mH
Flux linkage	0.433 Wb
Pole pairs	5
Rated speed	153 rad/s
<i>Lead acid battery</i>	
Voltage	300 V
Rated capacity	35 AH
Maximum charge current	7.2 A
Maximum discharge current	10 A
<i>AC/DC microgrid parameters</i>	
DC grid voltage	800 V
AC grid voltage	400 V
Frequency	50 Hz
DC bus capacitor	1500 $\mu$ F
AC bus LC filter	$L_f = 1$ mH, $C_f = 100$ $\mu$ F
AC loads	Load 1 = 3 KW , Load 2= 2kW, 2kW

## References

1. Mahmoud MS, Azher Hussain S, Abido MA (2014) Modeling and control of microgrid: an overview. *J Franklin Inst* 351(5):2822–2859. <https://doi.org/10.1016/j.jfranklin.2014.01.016>
2. Gupta A, Doolla S, Chatterjee K (2018) Hybrid AC-DC microgrid: systematic evaluation of control strategies. *IEEE Trans Smart Grid* 9(4):3830–3843. <https://doi.org/10.1109/TSG.2017.2727344>
3. Nikam V, Kalkhambkar V (2021) A review on control strategies for microgrids with distributed energy resources, energy storage systems, and electric vehicles. *Int Trans Electr Energy Syst* 31(1):1–26. <https://doi.org/10.1002/2050-7038.12607>
4. Olivares DE, Mehrizi Sani A, Etemadi AH, Cañizares CA, Iravani R, Kazerani M, Hajimiragha AH, Gomis Bellmunt O, Saeedifard M, Palma-Behnke R, Jiménez-Estévez GA, Hatziargyriou ND (2014) Trends in microgrid control. *IEEE Trans Smart Grid* 5(4):1905–1919. <https://doi.org/10.1109/TSG.2013.2295514>
5. Fazeli M, Holland P (2017) Universal and seamless control of distributed resources–energy storage for all operational scenarios of microgrids. *IEEE Trans Energy Convers* 32(3):963–973. <https://doi.org/10.1109/TEC.2017.2689505>
6. Eghtedarpour N, Farjah E (2014) Power control and management in a Hybrid AC/DC microgrid. *IEEE Trans Smart Grid* 5(3):1494–1505. <https://doi.org/10.1109/TSG.2013.2294275>
7. De Matos JG, Silva E, Ribeiro LAS (2015) Power control in AC isolated microgrids with renewable energy sources and energy storage systems. *IEEE Trans Indus Electron* 62(6):3490–3498. <https://doi.org/10.1109/TIE.2014.2367463>
8. Tayab UB, Bin Roslan MA, Hwai LJ, Kashif M (2017) A review of droop control techniques for microgrid. *Renew Sustain Energy Rev* 76:717–727. <https://doi.org/10.1016/j.rser.2017.03.028>
9. Muyeen SM, Al-Durra A (2013) Modeling and control strategies of fuzzy logic controlled inverter system for grid interconnected variable speed wind generator. *IEEE Syst J* 7(4):817–824. <https://doi.org/10.1109/JSYST.2013.2239893>
10. Singh P, Lather JS (2020) Artificial neural network-based dynamic power management of a DC microgrid: a hardware-in-loop real-time verification. *Int J Ambient Energy* 1–19. <https://doi.org/10.1080/01430750.2020.1720811>
11. Boujmil MH, Badis A, Nejib Mansouri M (2018) Nonlinear robust backstepping control for three-phase grid-connected PV systems. *Math Probl Eng*. <https://doi.org/10.1155/2018/3824628>
12. Baghaee HR, Mirsalim M, Gharehpetian GB, Talebi HA (2017) A decentralized power management and sliding mode control strategy for hybrid AC/DC microgrids including renewable energy resources. *IEEE Trans Indus Inform* 3203(c):1–1. <https://doi.org/10.1109/tii.2017.2677943>
13. Errami Y, Ouassaid M, Maaroufi M (2015) A performance comparison of a nonlinear and a linear control for grid connected PMSG wind energy conversion system. *Int J Electr Power Energy Syst* 68:180–194. <https://doi.org/10.1016/j.ijepes.2014.12.027>
14. Rezkallah M, Hamadi A, Chandra A, Singh B (2015) Real-time HIL implementation of sliding mode control for standalone system based on PV array without using dumpload. *IEEE Trans Sustain Energy* 6(4):1389–1398. <https://doi.org/10.1109/TSST.2015.2436333>
15. Bhende CN, Mishra S, Malla SG (2011) Permanent magnet synchronous generator-based standalone wind energy supply. *System* 2(4):361–373



# Instantaneous Power Estimation Algorithm in Hybrid Power Filter for Power Quality Improvement



S. Sindhu and M. R. Sindhu

## 1 Introduction

Increased use of power electronics devices and nonlinear loads has resulted in more power quality challenges, such as current harmonics and low power factor [1]. Since then, a great deal of study has been done into providing clean power to essential devices connected to the point of common coupling (PCC), as well as solving the various power quality issues caused by the widespread use of loads that demand nonlinear current from the supply. Power quality compensators ensure that power quality at the PCC remains within IEEE standards [2]. Tuned passive filters were employed in the beginning to eliminate power quality issues. It consists of tuned LC filters for filtering out harmonic components produced by the nonlinear loads. The merit of passive filters is that the implementation of these filters was not expensive. The installation of passive filters can lead to resonance between the system impedance and passive filter impedance which may result in amplifying the source current harmonics. With APF, the drawbacks of passive filters such as fixed harmonic compensation, resonance, and huge size could be overcome. An APF injects equal and opposite harmonics as well as reactive components into the system. This guarantees that the power drawn from the supply has a power factor of unity. The rating, the size, and therefore the cost of the APF vary depending on the nature, and the amount of compensation this filter has to provide [3].

---

S. Sindhu (✉)

Department of Electronics and Communication Engineering, Jyothi Engineering College, Thrissur, Kerala, India

e-mail: [sindhus2478@gmail.com](mailto:sindhus2478@gmail.com)

M. R. Sindhu

Department of Electrical and Electronics Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Coimbatore, India

e-mail: [mr\\_sindhu@cb.amrita.edu](mailto:mr_sindhu@cb.amrita.edu)

When low-cost passive filters are used in conjunction with APFs, rating of the APF used is reduced, and hence, it is possible to reduce the size of the APF. As a result, hybrid filter topologies [4] have been developed. A parallel combination of shunt passive filter and shunt APF has been devised in this study of a shunt hybrid power filter. The hybrid power filter performance is determined by the compensating signals generated by the APF operation. There are numerous control methods for APFs which are published in the literature [1]. Frequency-domain or time-domain compensation techniques can be used in control algorithms. The APF in this work is controlled using time-domain-based control algorithms because frequency-domain approaches need tedious computation and so have a delayed response [5].

The control algorithms for APF produce voltage or current compensation signals in real time. Instantaneous reactive power theory (IRPT) [6], synchronous detection (SD), Icos $\phi$  algorithm [7], exponential composition algorithm (ECA) [8, 9], and instantaneous power estimation algorithm (IPEA) [10] are some of the time-domain compensation strategies available. The IRPT, SD algorithms do not perform well with unbalanced source voltages, Icos $\phi$  algorithm uses sample/hold circuit and zero-crossing detector leading to a delay in response, and ECA involves complex calculations [10]. This paper uses IPEA as the control for the APF as it involves simple computation steps and fast response. The topology of hybrid filter is briefly explained in Sect. 2. Section 3 discusses the effectiveness of the hybrid filter with simulation results. The experimental setup and the results are analyzed in Sect. 4, and Sect. 5 summarizes the conclusion.

## 2 Shunt Hybrid Filter Topology

To reduce current harmonics and reactive power in the topology considered in this study, hybrid filters consist of shunt APFs and shunt passive filters that are connected in parallel with the nonlinear load as in Fig. 1. The control design for the shunt APF and the passive filter is detailed in the sub-sections that follow.

## 3 Control Strategy for Shunt APF

The IPEA control approach is used in this research to generate compensation signals to compensate for higher-order harmonics in the APF component of the hybrid power filter. This control algorithm for shunt APF is based on simple mathematical computations. The load currents and the balance source voltage are sensed. The in-phase and quadrature parts of the fundamental load currents are extracted using low pass and band pass part of the biquad filter. A unit template for each phase is created from the balanced supply voltage. The quadrature component and the in-phase components of each phase are also retrieved from the unit template with the help of biquad filters. Figure 2 depicts an IPEA diagram for phase a. The output is obtained by multiplying

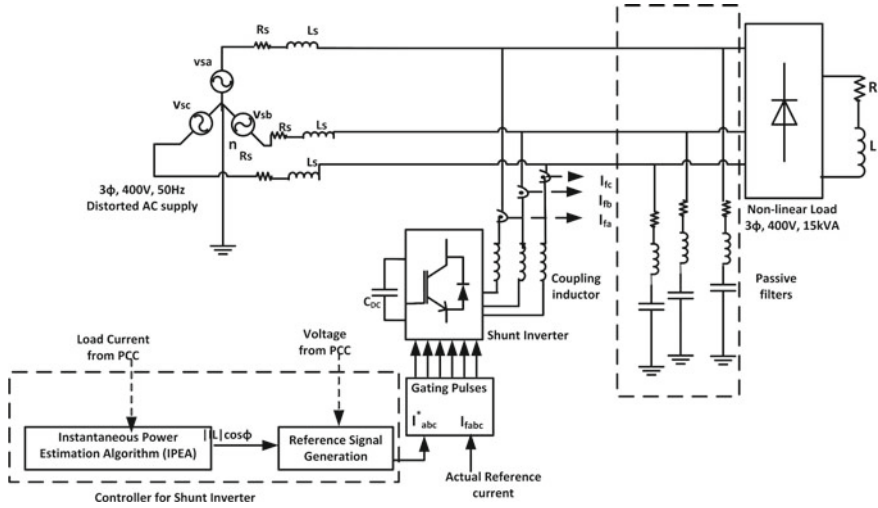


Fig. 1 Schematic of shunt hybrid power filter

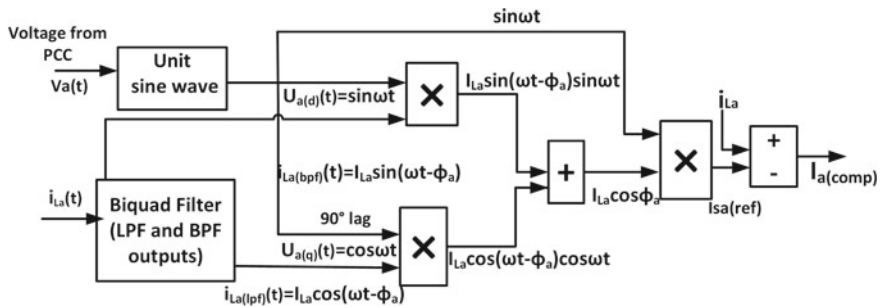


Fig. 2 IPEA for shunt APF for phase a

the quadrature components of load current with that of the unit template.

$$\begin{aligned}
 i_{La(pf)}(t)U_{a(q)}(t) &= I_{La} \cos(\omega t - \phi_a) \cos \omega t \\
 i_{Lb(pf)}(t)U_{b(q)}(t) &= I_{Lb} \cos\left(\omega t - \frac{2\pi}{3} - \phi_b\right) \cos\left(\omega t - \frac{2\pi}{3}\right) \\
 i_{Lc(pf)}(t)U_{c(q)}(t) &= I_{Lc} \cos\left(\omega t + \frac{2\pi}{3} - \phi_c\right) \cos\left(\omega t + \frac{2\pi}{3}\right) \quad (1)
 \end{aligned}$$

The result of product of the in-phase components obtained by passing the load current through a biquad filter and the respective phase of unit template is as shown in (2)

$$\begin{aligned}
i_{La(bpf)}(t)U_{a(d)}(t) &= I_{La} \sin(\omega t - \phi_a) \sin \omega t \\
i_{Lb(bpf)}(t)U_{b(d)}(t) &= I_{Lb} \sin\left(\omega t - \frac{2\pi}{3} - \phi_b\right) \sin\left(\omega t - \frac{2\pi}{3}\right) \\
i_{Lc(bpf)}(t)U_{c(d)}(t) &= I_{Lc} \sin\left(\omega t + \frac{2\pi}{3} - \phi_c\right) \sin\left(\omega t + \frac{2\pi}{3}\right)
\end{aligned} \quad (2)$$

The magnitude of the fundamental part of the load current in each phase is calculated by adding the results obtained in (1) and (2) and, then, applying the trigonometric identity.

$$\begin{aligned}
I_{La} \cos \phi_a &= i_{La(lpf)}(t)U_{a(q)}(t) + i_{La(bpf)}(t)U_{a(d)}(t) \\
I_{Lb} \cos \phi_b &= i_{Lb(lpf)}(t)U_{b(q)}(t) + i_{Lb(bpf)}(t)U_{b(d)}(t) \\
I_{Lc} \cos \phi_c &= i_{Lc(lpf)}(t)U_{c(q)}(t) + i_{Lc(bpf)}(t)U_{c(d)}(t)
\end{aligned} \quad (3)$$

The reference source current in each phase obtained by multiplying the value obtained in (3) and the unit amplitude wave is as in (4)

$$\begin{aligned}
i_{sa(\text{ref})} &= I_{La} \cos \phi_a U_a \\
i_{sb(\text{ref})} &= I_{Lb} \cos \phi_b U_b \\
i_{sc(\text{ref})} &= I_{Lc} \cos \phi_c U_c
\end{aligned} \quad (4)$$

The compensating currents are calculated by subtracting the source reference current from the load current.

$$\begin{aligned}
i_{sa(\text{comp})} &= i_{La} - i_{sa(\text{ref})} \\
i_{sb(\text{comp})} &= i_{Lb} - i_{sb(\text{ref})} \\
i_{sc(\text{comp})} &= i_{Lc} - i_{sc(\text{ref})}
\end{aligned} \quad (5)$$

In the case of unbalanced load conditions, the average of the values of (5) is calculated, before calculating the reference source current.

### 3.1 Design of Shunt Passive Filter

Fixed passive filters tuned to the fifth and seventh harmonic frequencies are connected in shunt at the load side to minimize the rating of the shunt APF [11]. The passive filter is designed based on the % VAR to be met by the fixed shunt passive filters. The maximum three-phase reactive power load demand is calculated. Also, the fifth- and seventh-order passive filters not only provide the designed reactive power but also sink the fifth- and the seventh-order harmonics. VAR to be supplied by the passive filter per phase is calculated as

$$Q_{\text{per phase}} = \frac{\text{VAR}}{3} \quad (6)$$

Capacitive reactance ( $X_c$ ) and capacitance ( $C$ ) per phase are calculated as

$$X_c = \frac{V^2}{Q_{\text{per phase}}} \text{ and } C = \frac{Q}{\omega \cdot V^2} \quad (7)$$

Inductance ( $L$ ) per phase is calculated as

$$L_5 = \frac{1}{\omega_5^2 C} \text{ and } L_7 = \frac{1}{\omega_7^2 C} \quad (8)$$

Resistance ( $R$ ) per phase for a given quality factor (QF) is calculated as

$$R_5 = \frac{\omega L_5}{\text{QF}} \text{ and } R_7 = \frac{\omega L_7}{\text{QF}} \quad (9)$$

The  $R$ ,  $L$ , and  $C$  values are mutually dependent on one another as well as on the reactive power demand to be supplied and the quality factor.

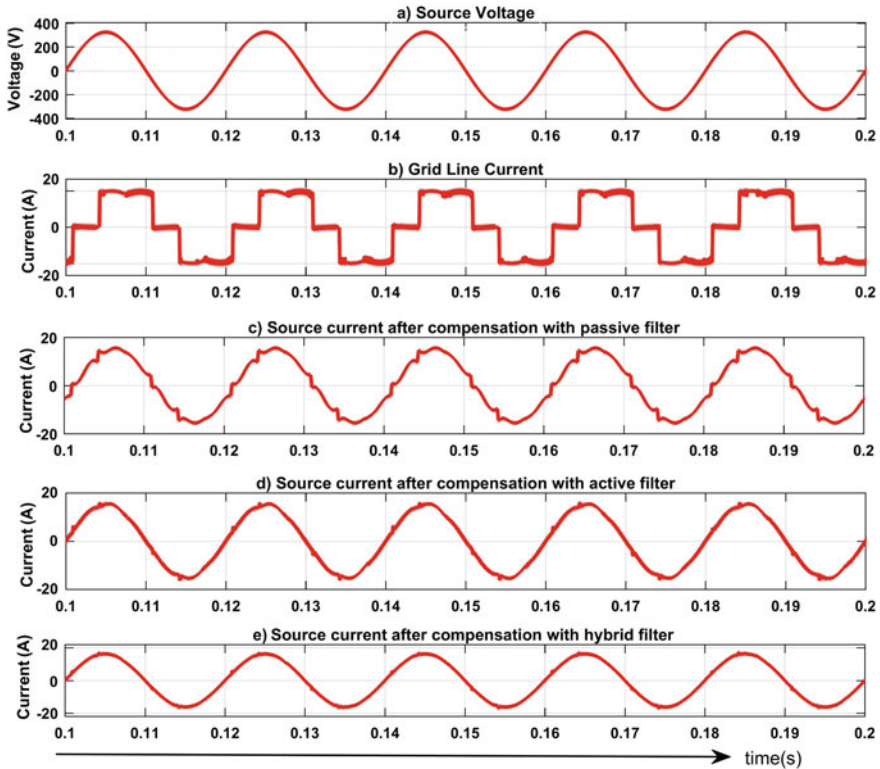
## 4 Simulation Results for Hybrid Filter

A test system for simulation study is modeled with a 400 V, 50 Hz three-phase balanced AC voltage source with line impedance  $R_s = 0.1 \Omega$ ,  $L_s = 0.2 \text{ mH}$  feeding a 7 kVA nonlinear load with a power factor of 0.6. Voltage sensors are used to sense the phase voltage, while the current sensors sense the load and source currents. A three-phase voltage source inverter (VSI) of suitable rating with DC-link capacitance of 45  $\mu\text{F}$ , 700 V acts as the shunt APF. The switching frequency of the VSI is set to 20 kHz.

The shunt passive filters sink the fifth and the seventh harmonics and are designed as discussed in Sect. 2. The load reactive power requirement is 6kVAR. The capacitance value for the fifth and seventh harmonics is calculated as 20  $\mu\text{F}$ . The inductance is calculated as 20 and 10 mH for providing minimum impedance for fifth and seventh harmonics. For a quality factor 40, the resistance value for the fifth and seventh harmonics is approximately 0.8  $\Omega$  and 0.5  $\Omega$ , respectively.

Thus, the passive filters are designed to meet specified reactive power demand of the load. The shunt APF with the IPEA control algorithm generates the compensation currents to compensate for the higher-order harmonics and reactive power. Simulink in MATLAB is used to conduct the system simulation analysis.

Figure 3 depicts the simulation results, which show that the hybrid filter compensates for current harmonics and power factor better than the passive or APF alone. The grid voltage and the grid line current are as in Figs. 3a and b. The compensation of



**Fig. 3** Simulation results of phase a **a** Grid/source voltage **b** Grid line current **c** Current after compensation with passive filter **d** With APF **e** With hybrid filter

source current with passive filter is shown in Fig. 3c. The passive filter compensates for the designed reactive power and filters out the fifth- and seventh-order harmonics. The compensation with the APF is shown in Fig. 3d. The shunt APF compensates for the reactive power as well as the distortions in the current waveform. The compensation with hybrid filter shown in Fig. 3e also compensates for reactive power and the harmonics. The passive filter of the hybrid filter provides a low impedance path for the fifth and seventh harmonics, and the shunt APF provides compensation for higher-order harmonics. This will help in reducing the burden on the APF and thus reduce the APF rating and size. This topology of this hybrid filter is effective in reducing current distortion and reactive power.

From Table 1, it can be observed that passive filter can filter out only the tuned frequencies and provides partial reactive power compensation. The shunt APF mitigates most of the harmonics and meets the reactive power demand of the load. But, the rating requirement of the APF is very high. So, the hybrid filter with reduced APF rating gives a better performance with reduced current harmonics which is within the IEEE standards. The total harmonic distortion (THD%) without filter is 28.96%, and

**Table 1** Comparison of passive filter, shunt APF, and hybrid filter

Parameters	Without filter	With passive filter	With shunt APF	With hybrid power filter
Fundamental current (rms in A)	15	10.82	10.80	11.54
THD%	28.96	5.11	1.97	0.70
Reactive power (kVAR)	3.491	1.608	0.058	0.003

with passive filter, the fifth- and the seventh-order frequencies reduced from 21.80% and 12.87% to 0.58% and 0.43%, respectively. Hence, the overall THD% is 5.11% which is beyond the IEEE limits. With APF alone, the THD% is reduced to 1.97%, and with hybrid power filter, the THD% is further reduced to 0.70%.

## 5 Experimental Verification

For the laboratory prototype, a three-phase supply and the nonlinear load –3kVA thyristor converter feeding a resistive load are connected through dSPACE DS1104 controller with a sampling period set to 120  $\mu$ s. The IPEA algorithm is implemented using DS1104. To avoid resonance and filter out the fifth- and seventh-order harmonics, the passive filters are designed to sixth-order harmonics. As an optimum value, the passive filter is designed with an inductor value 5 mH and capacitor value 80  $\mu$ F. A three-phase SEMIKRON IGBT inverter is operated as a shunt inverter and connected through a coupling inductor of 5mH. FLUKE power quality analyzer is used to analyze the experimental results. The experimental results in Fig. 4 show that the hybrid power filter compensates for current distortions and makes the power factor close to unity.

## 6 Conclusion

The topology discussed in this paper can be implemented in applications with existing passive filters. Tuned passive filters were designed and simulated in MATLAB/Simulink along with a shunt APF. The passive filter was designed to filter out the lower-order harmonics, mainly the fifth and the seventh harmonics thereby reducing the rating of the shunt APF. The higher-order harmonics and the reactive power were compensated by the IPEA controlled shunt-connected APF. The implementation results show that the hybrid power filter effectively supplies the reactive power demand of the load and, also, reduces the current harmonics.

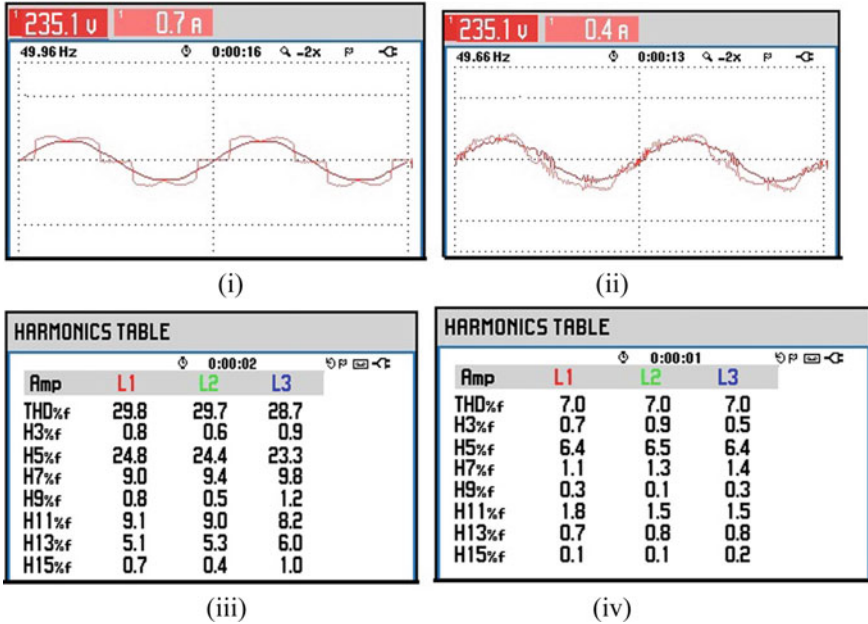


Fig. 4 a Source/grid voltage and current waveforms without filter b Voltage and current waveforms with hybrid filter c THD% without filter d THD% with filter

### References

1. Singh B, Chandra A, Haddad KA (2015) Power quality: Problems and mitigation techniques. Wiley, London
2. IEEE recommended practice and requirements for harmonic control in electric power systems (2014). IEEE Standards 519–2014:1–29
3. De Olivindo JAS, Machado IR (2017) Shunt active power filters for energy quality improvement in distributed generation system. In: 26th International Symposium on industrial electronics (ISIE). Edinburgh, pp 146–15
4. Sindhu MR, Manjula GN, Sindhu S (2016) Photovoltaic based adaptive shunt hybrid filter for power quality enhancement. In: IEEE international conference on power electronics, drives and energy systems. Trivandrum, pp 1–6
5. Jalil M, Amiri A (2020) An effective structure of three-phase parallel hybrid active power filter to accurate harmonic elimination. In: 15th International conference on protection and automation of power systems (IPAPS). Iran, pp 123–129
6. Harirchi F, Simoes MG (2018) Enhanced instantaneous power theory decomposition for power quality smart converter applications. IEEE Trans Power Electron 33(11):9344–9359
7. Raveendran V, Nair MG (2014) Smartpark as shunt active filter using modified Icos $\phi$  controller. In: Power and energy systems conference towards sustainable energy. PESTSE, Bangalore
8. Sindhu S, Sindhu MR, Nambiar TNP (2015) An exponential composition algorithm based UPQC for power quality enhancement. Procedia Technol 21:415–422
9. S Sindhu, Sindhu MR, Nambiar TNP (2017) Comparative study of exponential composition algorithm under dynamic conditions. In: International conference on technological advancements in power and energy. Kollam, pp 1–6



10. Sindhu S, Sindhu MR, Nambiar TNP (2019) Design and implementation of instantaneous power estimation algorithm for unified power quality conditioner. *J Power Electron* 19(3):815–826
11. Beres NR, Xiongfei W, Liserre M, Blabjerg F, Bak CL (2015) A review of passive power filter for three phase grid-connected voltage source converter. *IEEE J Emerging Sel Top Power Electron* 4(1):54–69

# Multi-objective Optimal Power Flow Using Whale Optimization Algorithm Consists of Static VAR Compensator



G. Karthik Varma and Bathina Venkateswara Rao

## 1 Introduction

The daily requirement for electric power is increasing day by day. This wants to generate more power in order to meet the increased demand and all the generation should be connected properly. On a group of interconnected systems, losses take place due to overloading and contingency conditions. Overloading can be minimized by creating new lines [1]. The major problems that occur due to upgrading new transmission lines are additionally expensive, and it takes extra period such as years. In order to lessen losses and increase growth power passage through line enhances the voltage profile in the power system via placing shunt-connected FACTS device called SVC [2]. One of the significant concerns of power system planning and operations is optimal power flow (OPF), which can be controlled by optimization problems with an objective function and a number of equality & inequality constraints [3]. In this OPF problem, the variable control indicates generators' real power, generator bus voltage, and reactive power generation of VAR sources [4].

Authors in [5] apply differential evolution (DE) on the problem of OPF. The methods are tested on standard IEEE 118, IEEE 57 and IEEE 30-bus systems for several OPF. In [6], authors use the gravitational search algorithm (GSA) to solve the optimal solution for the problem of OPF in a power system. In [7], cuckoo search algorithm (CSA) applied on IEEE 57 bus system and the obtained results are compared between the CSA and FPA, in that CSA is found to be better. In [8], authors present the backtracking optimization algorithm (BSA), to solve the OPF problem. In this paper, there are different objective functions used such as fuel cost, valve-point effect, and voltage profile enhancement.

---

G. K. Varma · B. V. Rao (✉)

Department of EEE, VR Siddhartha Engineering College, Vijayawada 520007, India

e-mail: [bvrao.eee@gmail.com](mailto:bvrao.eee@gmail.com)

The above-told methods have lack in efficiency when result the solution in the global search space. The above-said authors did not incorporate the SVC in the system. In this paper, the authors propose the whale optimization algorithm not only for solving the OPF problem but also for identifying the proper size for SVC to enhance the system performance. WOA enhances the performance by providing optimal results and is passed out in MATLAB background for IEEE 57 bus system.

## 2 Mathematical Formulation of Multi-Objective Optimal Power Flow (MOOPF) Problems

The MOOPF main aim is to optimize the control variables by considering two or more objective [4].

Objective 1: Minimization of cost.

The total fuel cost function for the number of generating units is denoted by below equation.

$$F_1 = \left( \sum_{i=1}^{NTG} \alpha_i + \beta_i P_{TGi} + \gamma_i P_{TGi}^2 \right) \text{ \$/Hr} \quad (1)$$

Objective 2: Minimization of emission.

An emanation of SO<sub>x</sub>, NO<sub>x</sub> gases [9] is considered in ton/Hr utilizing beneath equation,

$$F_2 = \sum_{i=1}^{NTG} 10^{-2} (a_i + b_i P_{TGi} + c_i P_{TGi}^2) + d_i \exp(e_i P_{TGi}) \quad (2)$$

Objective 3: True power losses minimization.

By using below equation power loss is calculated,

$$F_3 = \sum_{k=1}^{NT} G_{k(i,j)} [V_i^2 + V_j^2 - 2V_i V_j \cos(\delta_{ij})] \quad (3)$$

Objective 4: Enhancement of voltage stability of the network.

The L-record is determined for all heap transports and the greatest worth out of those goes about as the worldwide pointer for the framework steadiness [10]. In this manner, the target of framework stability is given by,

$$F_4 = \left| 1 - \sum_{i=1}^{NG} F_{ji} \frac{V_i}{V_j} \right| \rightarrow \text{where } \rightarrow j = 1, 2, \dots, NL$$

$$\text{and } \rightarrow F_{ji} = -\text{inv}[Y_{LL}][Y_{LG}] \quad (4)$$

By considering the above single objectives, multi-objectives have been formulated as various case studies.

Case 1: fuel cost and emission minimization.

It consists of fuel cost, and emission with selected weight factor is 100.

$$\text{Obj}_1(x, u) = F_1 + W_1 * F_2 \quad (5)$$

Case 2: Minimization of fuel cost and losses.

It consisting of both fuel cost and losses with selected weight factor is 100.

$$\text{Obj}_2(x, u) = F_1 + W_2 * F_3 \quad (6)$$

Case 3: Fuel cost and voltage stability minimization.

It consists of fuel cost along with voltage stability.

$$\text{Obj}_3(x, u) = F_1 + W_3 * F_4 \quad (7)$$

weight factor is 100 taken from [11].

## 2.1 Equality Constraints

$$P_{Gi} - P_{Di} - V_i \sum_{j=1}^{N_b} V_j \begin{pmatrix} G_{ij} \cos \theta_{ij} \\ + B_{ij} \sin \theta_{ij} \end{pmatrix} = 0 \quad (8)$$

$$Q_{Gi} - Q_{Di} - V_i \sum_{j=1}^{N_b} V_j \begin{pmatrix} G_{ij} \sin \theta_{ij} \\ B_{ij} \cos \theta_{ij} \end{pmatrix} = 0 \quad (9)$$

### 2.2 Inequality Constraints

$$V_{Gi}^{\min} \leq V_{Gi} \leq V_{Gi}^{\max}, i \in N_g \tag{10}$$

$$V_{Li}^{\min} \leq V_{Li} \leq V_{Li}^{\max}, i \in N_l \tag{11}$$

$$Q_{Gi}^{\min} \leq Q_{Gi} \leq Q_{Gi}^{\max}, i \in N_t \tag{12}$$

$$T_i^{\min} \leq T_m \leq T_i^{\max}, i \in N_c \tag{13}$$

$$Q_{ci}^{\min} \leq Q_{ci} \leq Q_{ci}^{\max}, i \in N_c \tag{14}$$

### 2.3 Mathematical Model of SVC

SVC is a shunt-connected device and provides or absorber reactive power to adjust capacitive or inductive flow to keep up with or control explicit boundaries of the electrical force framework [12]. It is demonstrated as a flawless reactive power booster at the load ends.

$$I_{SVC} = j B_{SVC} V_k \tag{15}$$

$$Q_{SVC} = Q_k = -V_k^2 B_{SVC} \tag{16}$$

## 3 Whale Optimization Algorithm (WOA)

Whales are extravagant animals. The most fascinating thing about the humpback whales is their exceptional chasing strategy. This rummaging conduct is called the bubble-net taking care of technique [13]. They discovered two moves related to bubbles and named them ‘up twisting’ and ‘double circles’. In the previous move, humpback whales jump around 12 m down and afterward begin to make bubbles in a twisting shape around the prey and swim up toward the surface [14]. The later move incorporates three unique stages: coral circle, lob tail, and catch circle. Itemized data about this conduct can be found in. It is worth focusing on here that bubble-net taking care of is one-of-a-kind conduct that must be seen in humpback whales [15].

$$\vec{D} = |\vec{C} \cdot \vec{X} * (t) - \vec{X}(t)| \tag{17}$$

$$\vec{X}(t) = \vec{X} * (t) - \vec{A} \cdot \vec{D} \tag{18}$$

$$\vec{A} = 2\vec{a} \cdot \vec{r} - \vec{a} \tag{19}$$

$$\vec{c} = 2\vec{r} \tag{20}$$

Every whale  $r$  can be represented as a vector with real values and is shown as

$$Y_i = (Y_{i,1}, Y_{i,2}, \dots, x_{i,k})^T \text{ subjected to } 0 < x_{i,1} \dots < x_{i,k} < L \tag{21}$$

The location of each whale is identified randomly by the following equations.

$$Y_{i,j} = g_{\min} + \text{rand}(0, 1) * (g_{\max} - g_{\min}) \tag{22}$$

### 4 Simulation Results

This paper presents the optimization of minimization of real power generation cost along with losses, emission, and voltage stability. In this voltage at the generator buses, the rating of reactive power compensation devices and transformer tap settings are controlled. The results are given below; Table 1 provides the details of IEEE57 bus system.

Table 2 presents the results of all the multi-objective functions considered in three different cases without and with SVC. It has been observed that by placing

**Table 1** Main characteristics of the studied system

Characteristics	IEEE57	
	Value	Details
Buses	57	–
Branches	80	–
Generators	7	Buses: 1, 2, 3, 6, 8, 9, 12
Load voltage limits	50	[0.94–1.06]
Shunt VAR compensation	3	Buses: 18, 25 and 53
Transformers with off-nominal tap ratio	17	Branches: 19, 20, 31, 35, 36, 37, 41, 46, 54, 58, 59, 65, 71, 73, 76 and 80
Control variable	33	–

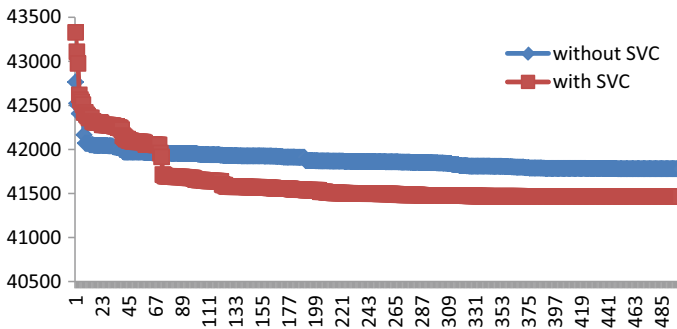
**Table 2** Optimal solutions obtained by WAO for IEEE57 bus system without and with SVC at bus32

S. No	Variables	Case 1		Case 2		Case 3	
		Without SVC	With SVC	Without SVC	With SVC	Without SVC	With SVC
1	PG2	91.0297	69.0126	93.3127	94.1402	94.1570	83.0637
2	PG3	46.2823	51.4824	48.1760	53.1693	47.0469	47.0980
3	PG6	91.6180	91.6829	93.5377	94.3636	42.7357	67.9470
4	PG8	453.9019	476.5641	421.641	434.952	469.052	470.0030
5	PG9	82.2757	51.7606	72.7472	43.7779	100.000	74.4714
6	PG12	356.7256	369.8139	390.242	387.573	366.118	368.1963
7	V1 (PU)	1.1000	1.1000	1.1000	1.1000	1.1000	1.1000
8	V2 (PU)	1.1000	1.1000	1.1000	1.1000	1.1000	1.1000
9	V3 (PU)	1.1000	1.1000	1.1000	1.1000	1.1000	1.1000
10	V6 (PU)	1.1000	1.1000	1.1000	1.1000	1.1000	1.1000
11	V8 (PU)	1.1000	1.1000	1.1000	1.1000	1.1000	1.1000
12	V9 (PU)	1.1000	1.1000	1.1000	1.1000	1.1000	1.1000
13	V12	1.1000	1.1000	1.1000	1.1000	1.1000	1.1000
14	Qc18	18.0225	18.3635	17.1849	19.0519	6.9345	5.0284
15	Qc25	13.6720	17.4910	10.7271	19.0621	0.6030	3.6655
16	Qc53	17.6337	17.7955	17.5084	19.1041	2.0376	18.2359
17	T19 (PU)	0.9868	1.0338	0.9141	1.0193	0.9805	1.0508
18	T20 (PU)	0.9868	1.0216	0.9177	1.0187	0.9808	1.0245
19	T31 (PU)	0.9163	1.0338	0.9502	1.0191	0.9775	1.0480
20	T35 (PU)	0.9868	1.0146	0.9139	1.0182	0.9808	1.1000
21	T36 (PU)	0.9285	1.0088	0.9521	1.0190	0.9806	1.1000
22	T37 (PU)	0.9739	1.0338	0.9464	1.0171	0.9759	1.0210
23	T41 (PU)	0.9868	1.0338	0.9274	1.0187	0.9806	1.0199
24	T46 (PU)	0.9549	1.0288	0.9163	1.0192	0.9806	1.0407
25	T54 (PU)	0.9331	1.0338	0.9205	1.0174	0.9807	1.0818
26	T58 (PU)	0.9868	1.0338	0.9332	1.0184	0.9576	1.0353
27	T59 (PU)	0.9868	1.0292	0.9185	1.0256	0.9806	1.0179
28	T65 (PU)	0.9687	1.0129	0.9328	1.0188	0.9649	1.0365
29	T66 (PU)	0.9738	0.9939	0.9202	1.0266	0.9806	1.1000
30	T71 (PU)	0.9257	1.0338	0.9272	1.0193	0.9806	1.0186
31	T73 (PU)	0.9246	1.0338	0.9321	1.0145	0.9806	1.0193
32	T76 (PU)	0.9868	0.9992	0.9349	1.0254	0.9806	1.1000
33	T80 (PU)	0.9868	1.0338	0.9195	1.0188	0.9808	1.0185
34	SVC size	–	8.4432	–	10.0000	–	10.0000

(continued)

**Table 2** (continued)

S. No	Variables	Case 1		Case 2		Case 3	
		Without SVC	With SVC	Without SVC	With SVC	Without SVC	With SVC
35	Fuel cost (\$/h)	41,650	41,324.89	41,714	41,313.8	41,681	41,241.96
36	Emission (ton/h)	1.3336	1.431	1.2873	1.33	1.4061	1.3984
39	PLoss (MW)	14.4148	14.27	12.1251	12.75	14.6563	14.4714
40	L-index	0.2524	0.1608	0.2172	0.1592	0.2756	0.1701
41	Fitness Value	41,780	41,464.79	42,925	42,589.2	41,707	41,258.96



**Fig. 1** Convergence curve for Case 1 by considering without and with SVC

SVC at bus number 32; in all the cases, fitness function value has been reduced. It is also observed that fitness value was low with the minimization of fuel cost along with voltage stability. Convergence curves for all three cases without and with SVC are shown in Figs. 1, 2, and 3. From figures, it has been observed that with SVC, optimal solution is achieved with less number of iterations. Obtained results with WOA compared with other methods available in literature and present in Table 3.

## 5 Conclusion

OPF with WOA outperforms the various cases considering multi-objectives of fuel cost along with emission, fuel cost with losses and fuel cost with voltage stability. It is evident that from the obtained results regulated variables with WOA yield better output compared to MPSO, MFO, MDE, and FPA. This method is tested on IEEE57 bus system. For the purpose of getting better results, SVC has been incorporated into



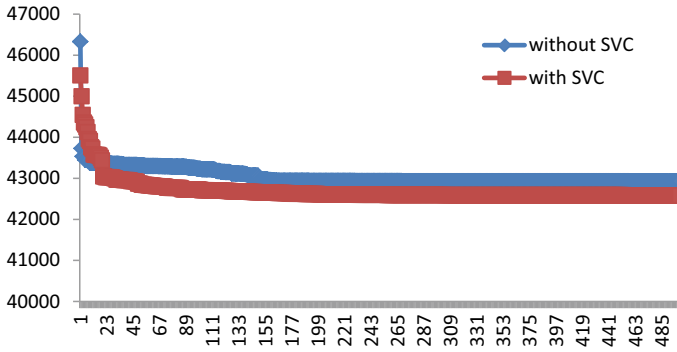


Fig. 2 Convergence curve for Case 2 by considering without and with SVC

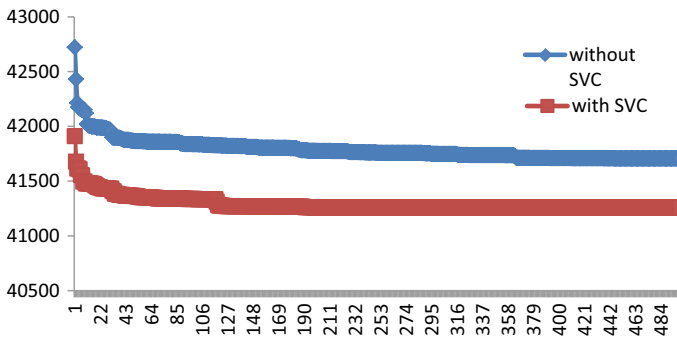


Fig. 3 Convergence curve for Case 3 by considering without and with SVC

the IEEE57 bus system at bus number 32. It is also observed that with SVC better results achieved. In the future, solar and wind generation can also be incorporated into system to obtain the best results.

**Table 3** Comparison of the WOA with MPSO, MDE, MFO and FPA for IEEE 57 bus system

Objective function	Objective	WAO	MPSO [11]	MDE [11]	MFO [11]	FPA [11]
Fuel Cost	Fuel cost (\$/h)	<b>41,666</b>	41,678.6	41,695.8	41,686.41	41,701.95
	Emission (ton/h)	1,4067	1,9443	2,0291	2,0036	2,0688
	P <sub>Loss</sub> (MW)	14,115	15,1271	15,9052	15,611	16,1543
	L-index	0,2278	0,28872	0,29232	0,29017	0,29183
Fuel Cost with L-index (Case 3)	Fuel cost (\$/h)	<b>41,681</b>	41,694,14	41,689,5	41,680,19	41,684,18
	Emission (ton/h)	1,4061	1,977	2,0	1,9192	1,9223
	P <sub>Loss</sub> (MW)	14,656	15,4554	15,7092	15,1026	15,2193
	L-index	<b>0,2756</b>	0,27918	0,27677	0,27467	0,27429
	Fitness value	<b>41,707</b>	41,722,05	41,717,2	41,711,61	41,707,66

## References

1. Pravallika DL, Rao BV (2016) Flower pollination algorithm based optimal setting of TCSC to minimize the transmission line losses in the power system. *Procedia Comput Sci* 92:30–35
2. Lakshmi P, Rao BV, Devarapalli R, Rai P (2020) Optimal power flow with bat algorithm for a power system to reduce transmission line losses using SVC. In: 2020 ICEFEET, pp 1–5. <https://doi.org/10.1109/ICEFEET49149.2020.9186964>
3. Bouktir LS (2004) GA for solving the optimal power flow problem 16
4. Kahourzade S, Mahmoudi A, Mokhlis HB (2015) A comparative study of multi-objective optimal power flow based on particle swarm, EP, and GA. *Electr Eng* 97:1–12
5. Biswas PP, Suganthan PN, Mallipeddi R, Amaratunga GA (2018) OPF solutions using differential evolution algorithm integrated with effective constraint handling techniques. *Eng Appl Artif Intel* 68:81–100
6. Duman S, Güvenç U, Sönmez Y (2012) Optimal power flow using GSA. *Energy Convers Manage* 59:86–95
7. Rao BV (2017) Real power generation optimization and optimal setting of UPFC for power loss minimization using cuckoo search algorithm. *J Electr Eng* 17(2):289–294
8. Chaib AE, Mehasni R, Abido MA (2016) OPF with emission and non-smooth cost functions using backtracking search optimization algorithm. *Int J Electr Power Energy Syst* 81:64–77
9. Ravi NC, Asir Rajan C (2013) Emission constraint OPF using differential evolution. *IJCA* 61:12–15
10. Oukennou A, Sandali A (2019) Novel voltage stability index for electric power system monitoring. *Int J Electr Comput Eng Syst (Online)* 10:1–9
11. Mohamed AAA, Mohamed YS, Hemeida AM (2017) OPF using moth swarm algorithm. *Electr Power Syst Res* 142:190–206
12. Rao BV, Kumar GVN (2014) Voltage collapse proximity indicator based placement and sizing of static VAR compensator using bat algorithm to improve power system performance. *BIJPSIC* 4:31–38
13. Mirjalili S, Lewis A (2016) The whale optimization algorithm. *Adv Eng Softw* 95:51–67
14. Tu J et al. (2021) Evolutionary biogeography-based WOA methods with communication structure: Towards measuring the balance. *Knowl Based Syst* 212:106642
15. Medani K, Sayah S, Bekrar A (2018) WOA based optimal reactive power dispatch: A case study of the Algerian power system. *Electr Power Syst Res* 163:696–705

# Standalone PMSG-Based Wind Energy Conversion System Under Unbalanced Load Conditions



P. N. Koteswara Rao, R. Mahalakshmi , and K. Sudarsana Reddy 

## 1 Introduction

In the recent days, renewable energy technology has become the trending and prominent topic due to increase in the concentration of atmospheric carbon dioxide. Despite COVID-19 prompted the economic losses, 260 GW capacity of renewables has been added globally in 2020 [1]. A report by International Renewable Energy Agency (IRENA) states that 127 GW of wind energy and 111 GW of solar energy have been added in particular. It infers that wind energy is generously generated [2]. Research in WECS has been popular since decades. The WECS can be categorically divided into four [3]. The type-3 and type-4 are being popular and mostly used generation technology. The type-3 WECS utilizes the doubly fed induction generator (DFIG) technology, and type-4 uses either a PMSG or squirrel cage induction generator (SCIG) as its generators. For standalone application, type-3 becomes complicated as it requires much complex controllers for its power electronic interfaces. Whereas for a low power application type-4 WECS suits best as it requires simple controllers along with low-rated power electronic devices. Moreover, PMSG [4] has its own points of benefits when compared to SCIG. Replacement of conventional DC excitation in synchronous generators (SG) with solar excitation or permanent magnets increases efficiency and promotes the greener technology [5].

The battery augmented to renewable systems is becoming popular and reliable because of its continuous power flow. Renewables being intermittent, a continuous power flow for a standalone system cannot be realized. A battery attached to the same system can overcome the issue [6]. The increase in the technology of battery

---

P. N. K. Rao · R. Mahalakshmi (✉) · K. S. Reddy  
Department of Electrical and Electronics Engineering, Amrita School of Engineering, Amrita  
Vishwa Vidyapeetham, Bengaluru, India  
e-mail: [d\\_mahalakshmi@blr.amrita.edu](mailto:d_mahalakshmi@blr.amrita.edu)

systems, applications pertaining to micro-grids [7, 8], smart grids are becoming the research topic in the literature.

The type-4 WEC generation technology requires power electronic interfaces like uncontrolled rectifiers, three-phase inverters and DC-DC converters like boost converter, SEPIC converter [9]. Boost converter is chosen in most applications because of its simple realization and less requirement of passive elements. A controller is required for each controlled power electronic switch utilized for appropriate operation in the system. Controllers range from simple PI controllers to intelligent controllers like fuzzy logic controllers (FLC) [10], artificial neural networks (ANN) [11] which utilizes embedded controller as its core to dump the logic. Maximum power from the renewable systems like WECS, solar can be achieved by adjusting or controlling the switch in the boost converter. Many literature proposes various maximum power point tracking (MPPT) techniques for achieving better efficiency like perturb and observe algorithm (P&O), incremental conductance (INC).

The primary objective of the paper is to continuously power three-phase unbalanced system using a standalone PMSG based WECS. The secondary objective is to control the switches in power electronic interfaces and logic for augmented battery using a single multi-input multi-output (MIMO)-based embedded controller. All simulations and analysis have been conducted using MATLAB/Simulink software. The paper has been organized into five subsections. Section 2 deals with analysis of the system which details the overview of system and designing parameters of the system. In Sect. 3, controller logic for maximum power extraction using boost converter, control pulses for PWM-based inverter, and switching logic for battery system has been discussed. Section 4 discusses the simulation model and its respective results. Through Sect. 5 work gets concluded.

## 2 System Description

### 2.1 Block Diagram

The block diagram for the proposed system has been depicted in Fig. 1. It can be inferred that the system contains PMSG-based WEC, diode rectifier, boost converter, IGBT-based PWM inverter, battery system attached to PWM switch, three-phase unbalanced load and embedded controller. The maximum power from PMSG-based WEC is extracted using the boost converter control logic. Maximum DC power from PMSG-based WEC connected diode rectifier is transferred along with the battery power when required to a three-phase PWM inverter. The battery is switched to the system when DC link power decreases under provided threshold. The IGBT-based PWM three-phase inverter is controlled using sinusoidal pulse width modulation technique so as to achieve less harmonic sine wave at the load terminals. The control logics mentioned above are provided by embedded controller. As a single controller is

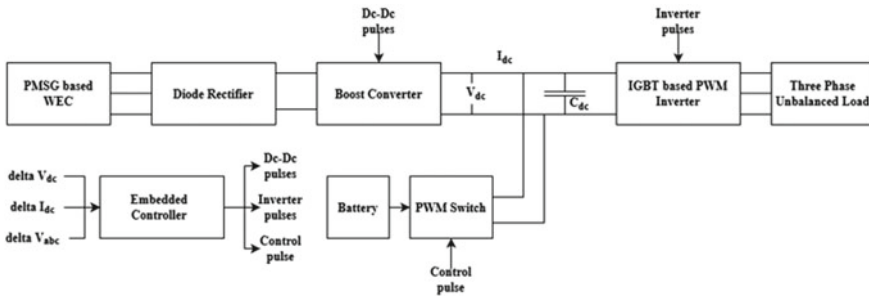


Fig. 1 Block diagram representation of proposed system

utilized in the system for three control techniques, the provided embedded controller functions as MIMO system.

### 2.2 Design

A 6 kW wind turbine has been considered for this work. The wind turbine characteristics (mechanical output power, generated torque from wind turbine) for various wind speed have been tabulated, and its respective plot has been displayed. Figure 2 shows the plot of  $P_{out}$  and  $T_{net}$  from wind turbine versus wind speed in per unit. It can be inferred that cut-in power is generated at speed of 0.2 pu, respectively. The governing equation for mechanical output power from wind turbine has been mentioned in Eq. (1). The optimum and maximum power can be achieved from wind turbine by accordingly choosing the constant ( $K_{opt} = \frac{1}{2} * \rho * \pi * r^2 * C_p$ ), where  $\rho$  is air density, r is the rotor blade radius, and  $C_p$  is coefficient of power. Thereby from Eq. (1), the equation for target optimum torque can be calculated and its respective equation is mentioned in (2).

$$P_{out} = \frac{1}{2} * \rho * \pi * r^2 * C_p * W_m^3 \tag{1}$$

$$T_{ref} = \frac{1}{2} * \rho * \pi * r^2 * C_p * W_m^2 \tag{2}$$

The wind turbine is attached to a 6 kW, 415 V PMSG. The three-phase output from PMSG is given to a diode rectifier to produce a constant DC power with terminal voltage of 200 V. To boost the voltage to desired 500 V and to retain a maximum power for varying wind speed, a boost converter has been designed. The battery of 6 kWh has been chosen for backup operation during low wind speeds. The design parameters of wind turbine and boost converter along with battery specifications have been mentioned in Table 1, respectively.

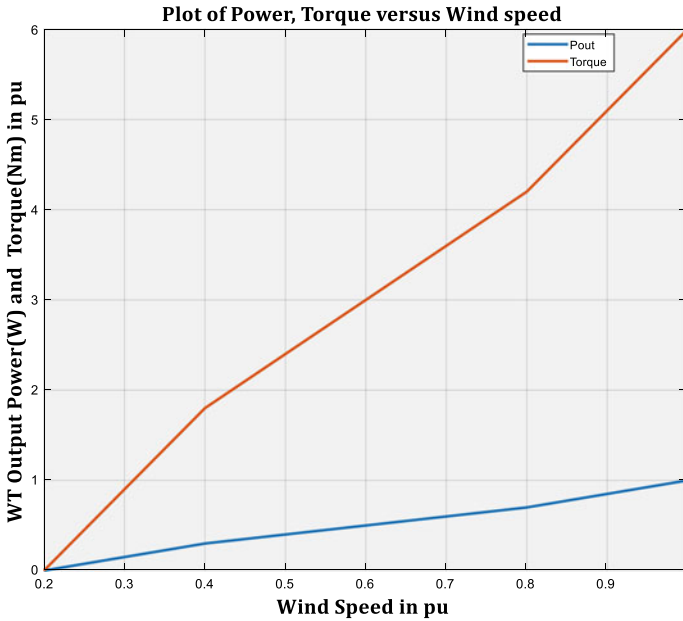


Fig. 2 Plot of wind turbine characteristics

Table 1 Specifications and parameters of wind turbine, battery along with boost converter

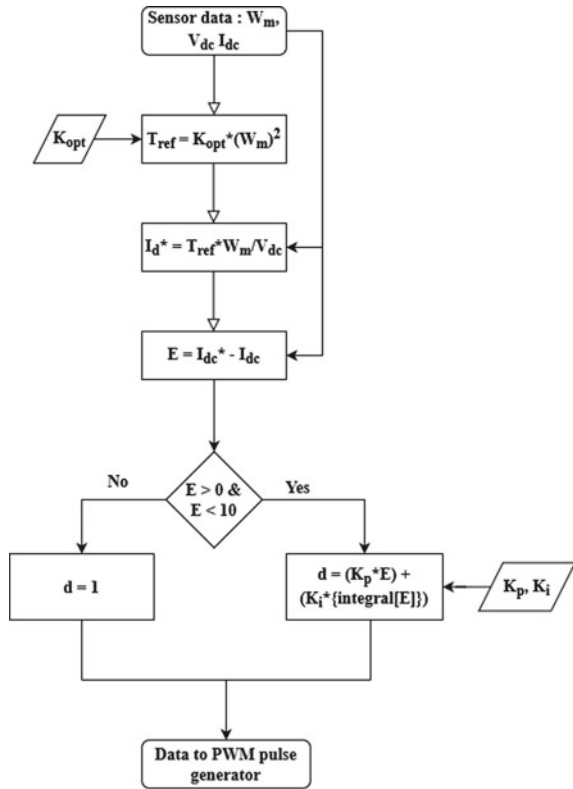
Parameters of wind turbine and backup battery used	Values	Parameters of boost converter	Boost converter values
WT rated power (kW)	6	Input voltage (V)	200
Rotor speed (rpm)	280	Output voltage (V)	500
Rotor blade radius (m)	1.73	Capacitor ( $\mu$ F)	900
Power coefficient ( $C_p$ )	0.3	Inductor (mH)	31.09
Battery ampere rating (Ah)	20	Duty ratio (%)	60
Initial SOC (%)	40	Switching frequency (kHz)	20

### 3 Control Logic

#### 3.1 Control Methodology for Boost Converter

The control logic for generating pulses to the switch in boost converter has been depicted in Fig. 3. The function of this controller is to extract maximum power. From the flowchart, it can be inferred that the required sensor data wind speed, DC link voltage and current would be sensed and would be provided to the embedded control. An optimum value of constant ( $K_{opt} = 0.11842$ ) has been calculated by using

**Fig. 3** Flowchart for controlling boost converter switches



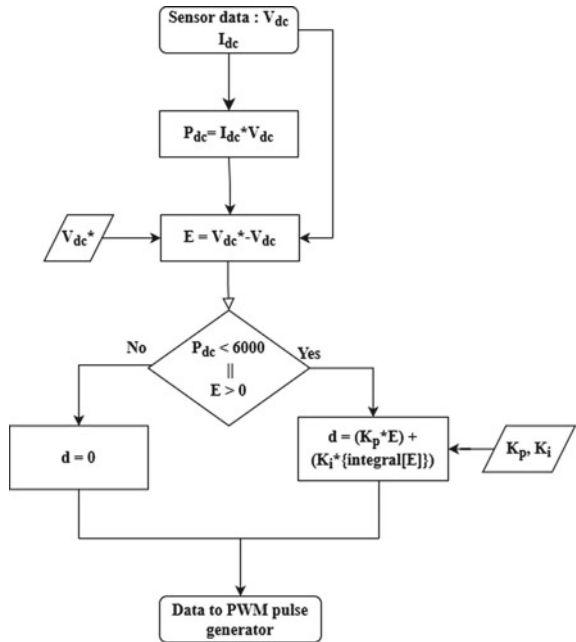
the specifications mentioned in Table 1, respectively. This value is fed as a constant to the controller. A reference torque and DC link current would be calculated in every iteration. The error in DC link current is checked, and accordingly, a duty ratio is produced. This duty ratio is utilized to generate PWM pulses for the converter in order to extract maximum power. The tuned  $K_p$  and  $K_i$  are fed as constant to the controller.

### 3.2 Control Methodology for Inverter

The control logic for generating pulses to the switch connected to backup has been depicted in Fig. 4. The function of this controller is to maintain constant DC link voltage and power (acting as backup) of the load during low wind speed (output mechanical wind turbine power). From the flowchart, it can be inferred that the required sensor data DC link voltage and current would be sensed and would be provided to the embedded control. A DC link power along with error (DC link voltage) would be calculated in every iteration. This error and power are compared



**Fig. 4** Flowchart for controlling a battery connected switch

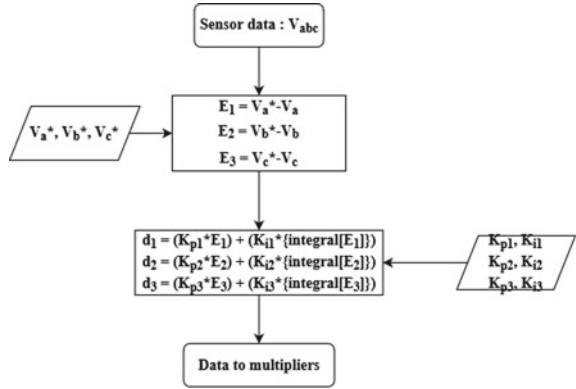


with the respective preset values, and accordingly, a duty ratio is produced. This duty ratio is utilized to generate PWM pulses by comparing it with sawtooth waveform. The tuned  $K_p$  and  $K_i$  are fed as constant to the controller.

### 3.3 Control Methodology for Battery Connected Switch

The control logic for generating pulses to the inverter switches has been depicted in Fig. 5. During unbalanced load condition, the line currents are not equal in each phase, and this will cause unequal voltage at the AC bus. This voltage unbalance will be leading to the line voltages at AC bus to become unbalanced. So, it is important to compensate the unbalance voltage at AC bus. The function of this controller is to compensate unbalanced voltage at AC bus. From the flowchart, it can be inferred that three-phase line voltage at AC bus would be sensed and would be provided to the embedded control. A voltage error at each phase (RMS value) is calculated in each iteration. A duty ratio is produced according to the error. The tuned  $K_p$  and  $K_i$  are fed as constant to the controller. This duty ratio or otherwise amplitude is multiplied by the sinusoidal waveforms, and then they are compared with respective sawtooth waveforms to generate PWM pulses. These three PWM pulses are again passed through a not gate to produce complement pulses. These six pulses would be provided to the three-phase inverter to balance the voltage at AC bus during unbalance loading.

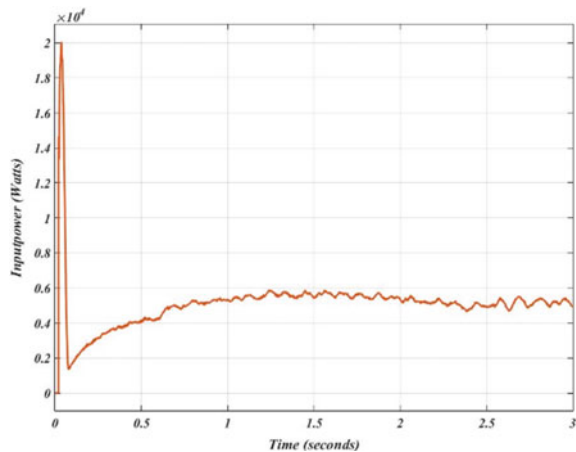
**Fig. 5** Flowchart for controlling inverter switches



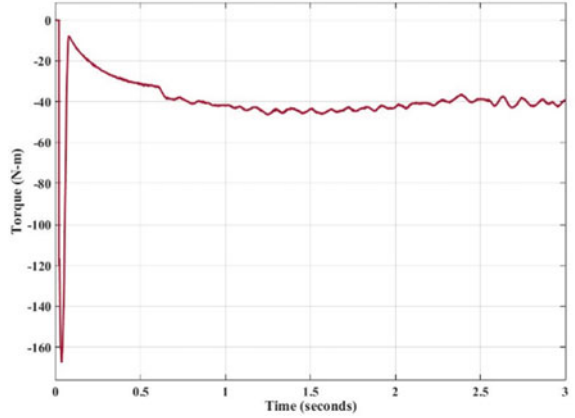
### 4 Simulation Results

The proposed block diagram shown in Fig. 1 has been simulated using MATLAB/Simulink software, and its respective results have been discussed in this section. The model has been tested for two case scenarios. Case 1 is the reduction of wind speed from 12 m/s to 10 m/s. Case 2 is the introduction of unbalanced load into the system. Figure 6 shows the output power from wind turbine at rated wind speed 12 m/s. It can be seen that 6 kW of mechanical power is getting injected into PMSG. Figure 7 shows the rated torque at rated speed. A constant -40 Nm torque is produced from wind turbine. Figures 8 and 9 show the single-phase RMS voltage of PMSG and DC link voltage variations during change of wind speed from 12 m/s to 10 m/s at  $t = 3$  s. From Fig. 8, it can be inferred that voltage gets dropped to 225 V from 260 V, respectively. Similarly, from Fig. 9 it can be depicted that DC link voltage is dropped from 460 to 400 V, respectively.

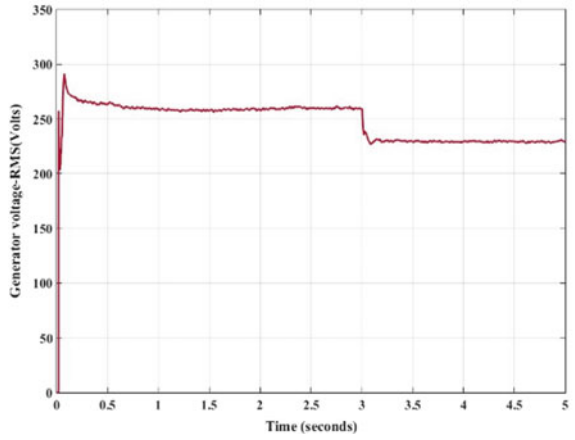
**Fig. 6** Input mechanical power to PMSG in watts



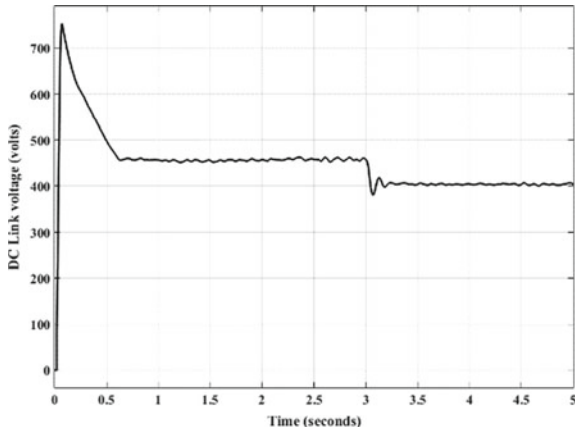
**Fig. 7** Input mechanical torque to PMSG in Nm



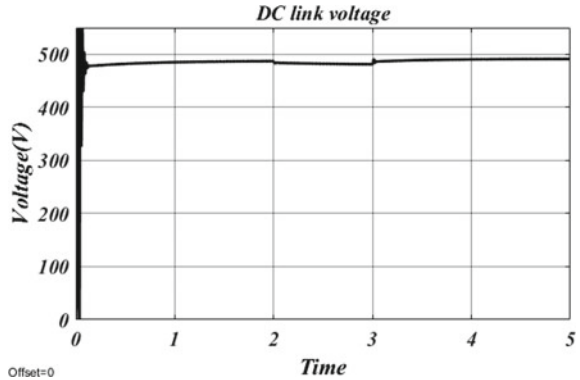
**Fig. 8** PMSG single-phase RMS voltage (volts) profile during case-1



**Fig. 9** DC link voltage (volts) profile during case-1 (without controller)



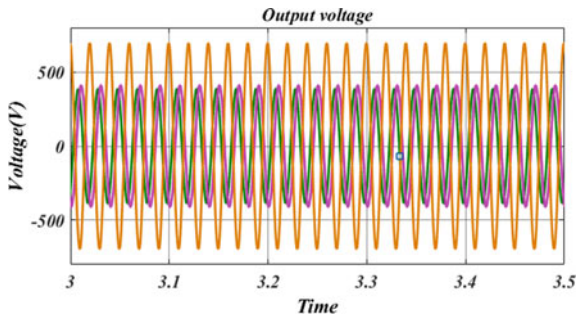
**Fig. 10** DC link voltage (volts) profile during case-1 (with controller)



After introduction of embedded controller, the backup battery compensates the remaining power during the transition of wind speed as overall DC link power is less than 6 kW and due to DC link voltage error. Figure 10 shows the DC link voltage profile after introduction of controller. It can be seen that a constant 500 V has been maintained at DC link even when wind speed reduces from rated speed.

Figures 11, 12 and 13 are the simulation results of the model during case-2. Figure 11 depicts the voltage profile at AC bus when inverter is feeding an unbalanced load. It can be seen that the unbalanced profile is maintained at AC bus. The embedded controller is even used to compensate the unbalanced profile by altering the PWM pulses of three-phase inverter [12]. The pulses for inverter after introduction of unbalanced load are shown in Fig. 12. After compensation, the AC bus voltage profile can be seen in Fig. 13, respectively. It can be inferred that a balanced voltage profile is maintained with the controller.

**Fig. 11** Inverter output voltage profile without controller



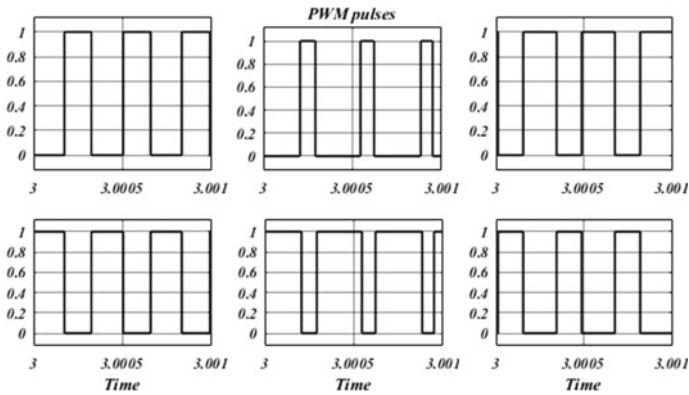
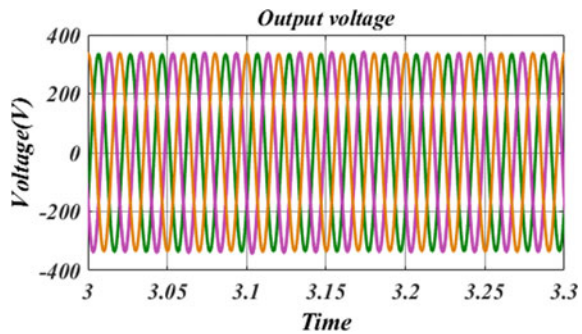


Fig. 12 PWM pulses for the inverter

Fig. 13 Inverter output voltage profile after introduction of controller



### 5 Conclusion

A reliable wind energy conversion system has been developed by connecting a backup battery deployed with a control logic. A 6 kW wind turbine has been connected to a 415 V, 6 kW PMSG for generating three-phase power. This power has been rectified, and the respective DC voltage has been boosted using designed boost converter. A maximum power from wind turbine has been achieved by boost converter. The logic for generating pulses to converter for achieving maximum power is by calculating optimum constant value ( $K_{opt}$ ) which is 0.11842. During low wind speeds, the backup battery (6 kW, 20 Ah) maintains DC link voltage and powers the load accordingly by the control logic deployed in switch connected to battery. This DC power is fed to three-phase AC loads (bus) using three-phase inverter. A converter logic is deployed in inverter switches for compensating the unbalanced voltage profile. Three different control logics have been programmed in single embedded controller so as to achieve multi-input and multi-output configuration. The simulated model has proven that the developed model is able to compensate unbalanced profile at AC bus and is able to power loads during low wind speeds.

## References

1. Renewable Capacity Statistics (2021) International Renewable Energy Agency, pp 1–64
2. Mahalakshmi R, Thampatty SKC (2020) Detailed modelling and development of a laboratory prototype for the analysis of subsynchronous resonance in DFIG-based wind Farm. In: International transactions on electrical energy systems, vol. 30(3), pp 1–22
3. Singh M, Muljadi E, Jonkman J, Gevorgian V (2014) A simulation for wind turbine generators with FAST and MATLAB-simulink modules. Natl Renew Energy Lab (NREL), pp 1–125
4. Mahalakshmi R, Reddy KS, Gautam M (2020) Control of standalone DFIG based wind turbine generator using machine learning algorithm. In: 2020 4th international conference on electronics, communication and aerospace technology (ICECA). pp 204–209
5. Gupta RA, Singh B, Jain BB (2015) Wind energy conversion system using PMSG. In: 2015 international conference on recent developments in control, automation and power engineering (RDCAPE), pp 199–203
6. Koteswa Rao PN, Mahalakshmi R (2020) Effective control co-ordination of standalone PMSG based wind energy conversion system with battery storage under different wind speed and load conditions. In: 2020 international conference on recent trends on electronics, information, communication and technology (RTEICT). Bangalore, India, pp 104–109
7. Dongmei Z, Nan Z, Yanhua L (2012) Micro-grid connected/islanding operation based on wind and PV hybrid power system. IEEE PES Innovative Smart Grid Technol, pp 1–6
8. Gayathri MR, Ramprabhakar J, Rajendran A (2019) Decentralized droop control in DC micro-grid. In: 2019 international conference on smart systems and inventive technology (ICSSIT). IEEE, Tirunelveli, India
9. Sudarsana Reddy K, Sai Teja Reddy B, Deepa K, Sireesha K (2021) An optimized controller for zeta converter-based solar hydraulic pump. Lect Notes Electr Eng 778:401–418
10. Mahalakshmi R, Ashwin Kumar A, Kumar A (2014) Design of fuzzy logic based maximum power point tracking controller for solar array for cloudy weather conditions. In: 2014 power and energy systems: Towards sustainable energy. Bangalore, pp 1–4
11. Bhatt GA, Gandhi PR (2019) Statistical and ANN based prediction of wind power with uncertainty. In: 2019 3rd international conference on trends in electronics and informatics (ICOEI), pp 622–627
12. Kumar AN, Joji S, Tangirala A, Sujith M (2018) A 31-level MLI topology with various level-shift PWM techniques and its comparative analysis. In: 2018 3rd international conference on communication and electronics systems (ICCES). IEEE, Coimbatore, India

# Application of Pythagorean Fuzzy Rough Distance Method in MCDM Problem for Buying Induction Motor



Revathi and Radhamani

## 1 Introduction

Fuzzy set was introduced by Zadeh. These concepts play important role in major fields. It is majorly used to handle the uncertainty problem with membership grade. Later, yager [1] introduced the concepts of Pythagorean fuzzy set. It also deals about the both membership and non-membership grade with the condition  $0 \leq (\mu_{PF})^2 + (\lambda_{PF})^2 \leq 1$ . IF topological spaces were introduced by Coker. IF rough sets and their topological structures were introduced by Lei Zhou et al. [2]. Distance and similarity measures for Pythagorean fuzzy sets were given by Ejegwa [3] that deals about the distance and similarities given for the Pythagorean fuzzy set and applied in the MCDM problems, respectively. This paper deals with Hamming distance, normalized Hamming distance, Euclidean distance, normalized Euclidean distance and similarity measure on Pythagorean fuzzy rough set. We proposed application in MCDM problems for buying induction motors based on the criteria.

## 2 Preliminaries

**Definition 2.1** Let universal set be  $Z$ .  $G$  be the Pythagorean fuzzy subset of  $Z$ , then membership function is defined as  $\mu_G : Z \rightarrow [0; 1]$  and non-membership function is defined as  $\lambda_G : Z \rightarrow [0; 1]$ .

**Definition 2.2** Let  $Z$  is a universal set. An PF set  $G$  in  $Z$  of the form  $G = \{ \langle m; \mu_P(m); \lambda_P(m) \rangle / m \in Z \}$  where the membership function is defined as  $\mu_{PF} :$

---

Revathi · Radhamani (✉)

Department of Mathematics, Annamalai University, Annamalai Nagar, Chidambaram, Tamil Nadu 608002, India

e-mail: [radhamani0195@gmail.com](mailto:radhamani0195@gmail.com)

$Z \rightarrow [0; 1]$  and the degree of non-membership function is defined as  $\lambda_{PF} : Z \rightarrow [0; 1]$  of an object  $z \in Z$  with the condition.

$$0 \leq (\mu_{PF})^2 + (\lambda_{PF})^2 \leq 1$$

and its denoted by  $PF(Z)$  as  $= \langle \mu_P, \lambda_P \rangle$

**Definition 2.3** Let  $Z$  be non-empty finite universe. PF equivalence relation is defined on  $Z \times Z$ : The pair  $(Z; PFR)$  is defined to be PF approximation space. For any  $G \in F(Z)$ , the lower  $PFR(G)$  and upper approximation  $PFR(G)$  respect to  $(Z; PFR)$  are defined as:

$$PF \underline{R}(G) = \{v; \mu_{PF \underline{R}(G)}(v); \lambda_{PF \underline{R}(G)}(v)/v \in Z\};$$

$$PF \overline{R}(G) = \{v; \mu_{PF \overline{R}(G)}(v); \lambda_{PF \overline{R}(G)}(v)/v \in Z\};$$

where membership function is defined by

$$\mu_{PF \underline{R}(G)}(v) = \bigwedge_{z \in Z} [(1 - PFR(v, z) \vee \mu_G(z)],$$

$$\mu_{PF \overline{R}(G)}(v) = \bigvee_{z \in Z} [(PFR(v, z) \wedge \mu_G(z)],$$

where non-membership function is defined by

$$\lambda_{PF \overline{R}(G)}(v) = \bigwedge_{z \in Z} [(1 - PFR(v, z) \vee \mu_G(z)],$$

$$\lambda_{PF \underline{R}(G)}(v) = \bigvee_{z \in Z} [(PFR(v, z) \wedge \lambda_G(z)],$$

The pair  $PFR(G) = [PF \underline{R}(G), PF \overline{R}(G)]$  is said to be Pythagorean fuzzy rough set of  $G$  with respect to  $(Z, PFR)$ .

**Definition 2.4** Let  $Z$  be non-empty set and  $A, B, C \in PFS(Z)$ . Distance measure ‘ $d$ ’ between  $A$  and  $B$  is a function  $d: PFS \times PFR \rightarrow [0, 1]$  satisfies.

- i.  $0 \leq d(X, Y) \leq 1$
- ii.  $d(X, Y) = 0$  iff  $X = Y$
- iii.  $d(X, Y) = d(Y, X)$
- iv.  $d(X, Z) + d(Y, Z) \geq d(X, Y)$

**Definition 2.5** Let  $Z = \{z_1, z_2, \dots, z_n\}$  be an universal set, and let criteria set be  $CT = \{ct_1, ct_2, \dots, ct_n\}$  of PFR sets on  $Z$ . Then Hamming distance between  $PFR(X)$  and  $PFR(Y)$  is defined as



$$d_{\text{PFRS}}(X, Y)_{\text{hd}} = \frac{1}{2} \sum_{k=1}^n \left\{ \begin{array}{l} \left| \mu_{\text{PF}\underline{R}(X)}(x_k) - \mu_{\text{PF}\underline{R}(Y)}(x_k) \right| \\ + \left| \lambda_{\text{PF}\underline{R}(X)}(x_k) - \lambda_{\text{PF}\underline{R}(Y)}(x_k) \right| \\ + \left| \mu_{\text{PF}\overline{R}(X)}(x_k) - \mu_{\text{PF}\overline{R}(Y)}(x_k) \right| \\ + \left| \lambda_{\text{PF}\overline{R}(X)}(x_k) - \lambda_{\text{PF}\overline{R}(Y)}(x_k) \right| \end{array} \right\}$$

**Definition 2.6** Let  $Z = \{z_1, z_2, \dots, z_n\}$  be an universal set, and let criteria set be  $\text{CT} = \{ct_1, ct_2, \dots, ct_n\}$  of PFR sets on  $Z$ . Then normalized Hamming distance between  $\text{PFR}(X)$  and  $\text{PFR}(Y)$  is defined as

$$d_{\text{PFRS}}(X, Y)_{\text{nhd}} = \frac{1}{2n} \sum_{k=1}^n \left\{ \begin{array}{l} \left| \mu_{\text{PF}\underline{R}(X)}(x_k) - \mu_{\text{PF}\underline{R}(Y)}(x_k) \right| \\ + \left| \lambda_{\text{PF}\underline{R}(X)}(x_k) - \lambda_{\text{PF}\underline{R}(Y)}(x_k) \right| \\ + \left| \mu_{\text{PF}\overline{R}(X)}(x_k) - \mu_{\text{PF}\overline{R}(Y)}(x_k) \right| \\ + \left| \lambda_{\text{PF}\overline{R}(X)}(x_k) - \lambda_{\text{PF}\overline{R}(Y)}(x_k) \right| \end{array} \right\}$$

**Definition 2.7** Let  $Z = \{z_1, z_2, \dots, z_n\}$  be an universal set, and let criteria set be  $\text{CT} = \{ct_1, ct_2, \dots, ct_n\}$  of PFR sets on  $Z$ . Then Euclidean distance between  $\text{PFR}(X)$  and  $\text{PFR}(Y)$  is defined as

$$d_{\text{PFRS}}(X, Y)_{\text{ed}} = \frac{1}{2} \sum_{k=1}^n \left\{ \begin{array}{l} \left| \mu_{\text{PF}\underline{R}(X)}(x_k) - \mu_{\text{PF}\underline{R}(Y)}(x_k) \right| \\ + \left| \lambda_{\text{PF}\underline{R}(X)}(x_k) - \lambda_{\text{PF}\underline{R}(Y)}(x_k) \right| \\ + \left| \mu_{\text{PF}\overline{R}(X)}(x_k) - \mu_{\text{PF}\overline{R}(Y)}(x_k) \right| \\ + \left| \lambda_{\text{PF}\overline{R}(X)}(x_k) - \lambda_{\text{PF}\overline{R}(Y)}(x_k) \right| \end{array} \right\}$$

**Definition 2.8** Let  $Z = \{z_1, z_2, \dots, z_n\}$  be an universal set, and let criteria set be  $\text{CT} = \{ct_1, ct_2, \dots, ct_n\}$  of PFR sets on  $Z$ . Then normalized Euclidean distance between  $\text{PFR}(X)$  and  $\text{PFR}(Y)$  is defined as

$$d_{\text{PFRS}}(X, Y)_{\text{ned}} = \frac{1}{2} \sum_{k=1}^n \left\{ \begin{array}{l} \left| \mu_{\text{PF}\underline{R}(X)}(x_k) - \mu_{\text{PF}\underline{R}(Y)}(x_k) \right| \\ + \left| \lambda_{\text{PF}\underline{R}(X)}(x_k) - \lambda_{\text{PF}\underline{R}(Y)}(x_k) \right| \\ + \left| \mu_{\text{PF}\overline{R}(X)}(x_k) - \mu_{\text{PF}\overline{R}(Y)}(x_k) \right| \\ + \left| \lambda_{\text{PF}\overline{R}(X)}(x_k) - \lambda_{\text{PF}\overline{R}(Y)}(x_k) \right| \end{array} \right\}$$

### 3 Application in MCDM Problem for Buying 3-Phase Induction Motor

**Algorithm:**

- (1) Construct the criteria set  $\text{CT} = \{\text{CT}_1, \text{CT}_2, \text{CT}_3\}$  for MCDMP
- (2) Case (1): Compute PF set  $\text{PF}(X)$  over  $U$  into PFRS using Definition (2.3) and find  $\mu_{\text{PF}\underline{R}(X)}(x)$ ,  $\mu_{\text{PF}\overline{R}(X)}(x)$  values and  $\lambda_{\text{PF}\overline{R}(X)}(x)$ ,  $\lambda_{\text{PF}\underline{R}(X)}(x)$  values based

on previous record data to the specific problem. Similarly, compute PFRS for  $Y$  and  $Z$  company.

- (3) Compute the Hamming distance for  $PFR(X)$  and  $PFR(Y)$

$$d_{PFRS}(X, Y)_{hd} = \frac{1}{2} \sum_{k=1}^n \left\{ \begin{array}{l} \left| \mu_{PF \underline{R}(X)}(x_k) - \mu_{PF \underline{R}(Y)}(x_k) \right| \\ + \left| \lambda_{PF \underline{R}(X)}(x_k) - \lambda_{PF \underline{R}(Y)}(x_k) \right| \\ + \left| \mu_{PF \overline{R}(X)}(x_k) - \mu_{PF \overline{R}(Y)}(x_k) \right| \\ + \left| \lambda_{PF \overline{R}(X)}(x_k) - \lambda_{PF \overline{R}(Y)}(x_k) \right| \end{array} \right\}$$

- (4) Compute the normalized Hamming distance for  $PFR(X)$  and  $PFR(Y)$

$$d_{PFRS}(X, Y)_{nhd} = \frac{1}{2n} \sum_{k=1}^n \left\{ \begin{array}{l} \left| \mu_{PF \underline{R}(X)}(x_k) - \mu_{PF \underline{R}(Y)}(x_k) \right| \\ + \left| \lambda_{PF \underline{R}(X)}(x_k) - \lambda_{PF \underline{R}(Y)}(x_k) \right| \\ + \left| \mu_{PF \overline{R}(X)}(x_k) - \mu_{PF \overline{R}(Y)}(x_k) \right| \\ + \left| \lambda_{PF \overline{R}(X)}(x_k) - \lambda_{PF \overline{R}(Y)}(x_k) \right| \end{array} \right\}$$

- (5) Compute the Euclidean distance for  $PFR(X)$  and  $PFR(Y)$

$$d_{PFRS}(X, Y)_{ed} = \sqrt{\frac{1}{2} \sum_{k=1}^n \left\{ \begin{array}{l} \left| \mu_{PF \underline{R}(X)}(x_k) - \mu_{PF \underline{R}(Y)}(x_k) \right|^2 \\ + \left| \lambda_{PF \underline{R}(X)}(x_k) - \lambda_{PF \underline{R}(Y)}(x_k) \right|^2 \\ + \left| \mu_{PF \overline{R}(X)}(x_k) - \mu_{PF \overline{R}(Y)}(x_k) \right|^2 \\ + \left| \lambda_{PF \overline{R}(X)}(x_k) - \lambda_{PF \overline{R}(Y)}(x_k) \right|^2 \end{array} \right\}}$$

- (6) Compute the normalized Euclidean distance for  $PFR(X)$  and  $PFR(Y)$

$$d_{PFRS}(X, Y)_{ned} = \sqrt{\frac{1}{2n} \sum_{k=1}^n \left\{ \begin{array}{l} \left| \mu_{PF \underline{R}(X)}(x_k) - \mu_{PF \underline{R}(Y)}(x_k) \right|^2 \\ + \left| \lambda_{PF \underline{R}(X)}(x_k) - \lambda_{PF \underline{R}(Y)}(x_k) \right|^2 \\ + \left| \mu_{PF \overline{R}(X)}(x_k) - \mu_{PF \overline{R}(Y)}(x_k) \right|^2 \\ + \left| \lambda_{PF \overline{R}(X)}(x_k) - \lambda_{PF \overline{R}(Y)}(x_k) \right|^2 \end{array} \right\}}$$

- (7) Case (2) Repeat steps 3, 4, 5, 6 and 7 for ‘ $X$  and  $Z$ ’ company  $PFR(X)$  and  $PFR(Z)$
- (8) Case (3) Repeat steps 3, 4, 5, 6 and 7 for ‘ $Y$  and  $Z$ ’ company  $PFR(Y)$  and  $PFR(Z)$
- (9) Checking the distance measure (definition 2.4)
- (10) Compare the values of  $X, Y, Z$  and ranking alternatives based on distance measure and conclude.

Suppose BR Textile Industry decides to buy a 3-phase induction motor among the three companies  $X, Y, Z$  (alternatives) for their huge production. They wish to invest in good company that the product may be judged with the attributes like overload capacity must be high, self-starting efficiency high, smooth speed control,

maintenance cost less, starting torque is high, etc. In order to take right decision about the company, BR production experts' opinion is a need for the criteria.

**Step1:**

Let the criteria set be

$$CT = \{CT_1, CT_2, CT_3\}$$

where

CT<sub>1</sub> Self-starting and smooth speed control,

CT<sub>2</sub> Maintenance cost less,

CT<sub>3</sub> Starting torque high, respectively.

**Step 2:**

Consider,  $U = \{X_1; X_2; X_3\}$ ;

	X company	Y company	Z company
X <sub>1</sub>	(0.4,0.6) (0.7,0.4)	(0.5,0.6) (0.5,0.6)	(0.4,0.6) (0.5,0.6)
X <sub>2</sub>	(0.4,0.7) (0.4,0.4)	(0.4,0.6) (0.5,0.6)	(0.6,0.4) (0.5,0.4)
X <sub>3</sub>	(0.6,0.6) (0.6,0.4)	(0.4,0) (0.4,0.6)	(0.4,0.4) (0.4,0.6)

**Step 3:**

Case (1)

The Hamming distance for PFR(X) and PFR(Y)

$$\begin{aligned}
 &= \frac{1}{2} \sum_{k=1}^n \left\{ \begin{array}{l} \left| \mu_{PF \underline{R}(X)}(x_k) - \mu_{PF \underline{R}(Y)}(x_k) \right| \\ + \left| \lambda_{PF \underline{R}(X)}(x_k) - \lambda_{PF \underline{R}(Y)}(x_k) \right| \\ \left| \mu_{PF \overline{R}(X)}(x_k) - \mu_{PF \overline{R}(Y)}(x_k) \right| \\ + \left| \lambda_{PF \overline{R}(X)}(x_k) - \lambda_{PF \overline{R}(Y)}(x_k) \right| \end{array} \right\} \\
 &= \frac{1.9}{2} \\
 d_{PFRS}(X, Y)_{hd} &= 0.95.
 \end{aligned}$$

**Step 4:**

The normalized Hamming distance for PFR(X) and PFR(Y)

$$\begin{aligned}
 &= \frac{1}{2n} \sum_{k=1}^n \left\{ \begin{array}{l} \left| \mu_{PF \underline{R}(X)}(x_k) - \mu_{PF \underline{R}(Y)}(x_k) \right| \\ + \left| \lambda_{PF \underline{R}(X)}(x_k) - \lambda_{PF \underline{R}(Y)}(x_k) \right| \\ \left| \mu_{PF \overline{R}(X)}(x_k) - \mu_{PF \overline{R}(Y)}(x_k) \right| \\ + \left| \lambda_{PF \overline{R}(X)}(x_k) - \lambda_{PF \overline{R}(Y)}(x_k) \right| \end{array} \right\} \\
 &= \frac{1}{6} \sum_{k=1}^3 \left\{ \begin{array}{l} 0.1 + 0 + 0.2 + 0.2 + 0 + 0.1 \\ + 0.1 + 0 + 0.2 + 0.6 + 0.2 + 0.2 \end{array} \right\} \\
 &= \frac{1.9}{6} \\
 d_{PFRS}(X, Y)_{nhd} &= 0.3166
 \end{aligned}$$

**Step 5:**

The Euclidean distance for PFR(X) and PFR(Y)

$$\begin{aligned}
 d_{PFRS}(X, Y)_{ed} &= \sqrt{\frac{1}{2n} \sum_{k=1}^n \left\{ \begin{array}{l} \left| \mu_{PF \underline{R}(X)}(x_k) - \mu_{PF \underline{R}(Y)}(x_k) \right|^2 + \\ \left| \lambda_{PF \underline{R}(X)}(x_k) - \lambda_{PF \underline{R}(Y)}(x_k) \right|^2 + \\ \left| \mu_{PF \overline{R}(X)}(x_k) - \mu_{PF \overline{R}(Y)}(x_k) \right|^2 + \\ \left| \lambda_{PF \overline{R}(X)}(x_k) - \lambda_{PF \overline{R}(Y)}(x_k) \right|^2 \end{array} \right\}} \\
 &= \frac{1}{6} \sum_{k=1}^3 \left\{ \begin{array}{l} 0.1 + 0 + 0.2 + 0.2 + 0 + 0.1 \\ + 0.1 + 0 + 0.2 + 0.6 + 0.2 + 0.2 \end{array} \right\} \\
 &= \sqrt{\frac{0.59}{2}} \\
 d_{PFRS}(X, Y)_{ed} &= 0.543
 \end{aligned}$$

**Step 6:**

The normalized Euclidean distance for PFR(X) and PFR(Y)

$$\begin{aligned}
 d_{PFRS}(X, Y)_{ned} &= \sqrt{\frac{1}{2n} \sum_{k=1}^n \left\{ \begin{array}{l} \left| \mu_{PF \underline{R}(X)}(x_k) - \mu_{PF \underline{R}(Y)}(x_k) \right|^2 \\ + \left| \lambda_{PF \underline{R}(X)}(x_k) - \lambda_{PF \underline{R}(Y)}(x_k) \right|^2 \\ \left| \mu_{PF \overline{R}(X)}(x_k) - \mu_{PF \overline{R}(Y)}(x_k) \right|^2 \\ + \left| \lambda_{PF \overline{R}(X)}(x_k) - \lambda_{PF \overline{R}(Y)}(x_k) \right|^2 \end{array} \right\}} \\
 &= \sqrt{\frac{1}{6} \sum_{k=1}^3 \left\{ \begin{array}{l} 0.01 + 0.04 + 0.04 + 0.01 + \\ 0.01 + 0.04 + 0.36 \\ + 0.04 + 0.04 \end{array} \right\}} \\
 &= \sqrt{\frac{0.59}{6}} \\
 d_{PFRS}(X, Y)_{ned} &= 0.31358
 \end{aligned}$$

**Step 7:**

Case(2)

The Hamming distance for PFR(X) and PFR(Z)

$$d_{\text{PFRS}}(X, Z)_{\text{hd}} = \frac{1.8}{2}$$

$$d_{\text{PFRS}}(X, Z)_{\text{hd}} = 0.9.$$

The normalized Hamming distance for PFR(X) and PFR(Z)

$$d_{\text{PFRS}}(X, Z)_{\text{nhd}} = \frac{1.8}{6}$$

$$d_{\text{PFRS}}(X, Z)_{\text{nhd}} = 0.3$$

The Euclidean distance for PFR(X) and PFR (Z)

$$d_{\text{PFRS}}(X, Z)_{\text{ed}} = \sqrt{\frac{0.38}{2}}$$

$$d_{\text{PFRS}}(X, Z)_{\text{ed}} = 0.435$$

The normalized Euclidean distance for PFR(X) and PFR (Z)

$$d_{\text{PFRS}}(X, Z)_{\text{ned}} = \sqrt{\frac{0.38}{6}}$$

$$d_{\text{PFRS}}(X, Z)_{\text{ned}} = 0.251$$

**Step 8:**

Case (3)

The Hamming distance for PFR(Y) and PFR(Z)

$$d_{\text{PFRS}}(Y, Z)_{\text{hd}} = \frac{1.3}{2}$$

$$d_{\text{PFRS}}(Y, Z)_{\text{hd}} = 0.65.$$

The normalized Hamming distance for PFR(Y) and PFR(Z)

$$d_{\text{PFRS}}(Y, Z)_{\text{nhd}} = \frac{1.3}{6}$$

$$d_{\text{PFRS}}(Y, Z)_{\text{nhd}} = 0.2166$$

The Euclidean distance for PFR(Y) and PFR(Z)

$$d_{\text{PFRS}}(Y, Z)_{\text{ed}} = \sqrt{\frac{1.3}{6}}$$

$$d_{\text{PFRS}}(Y, Z)_{\text{ed}} = 0.406$$

The normalized Euclidean distance for PFR( $Y$ ) and PFR( $Z$ )

$$d_{\text{PFRS}}(Y, Z)_{\text{ned}} = \sqrt{\frac{0.33}{6}}$$

$$d_{\text{PFRS}}(Y, Z)_{\text{ned}} = 0.2345$$

Distance measure	XY company	XZ company	YZ company
Hamming distance	0.95	0.9	0.65
Normalized hamming distance	0.3166	0.3	0.2166
Euclidean distance	0.543	0.435	0.406
Normalized Euclidean	0.31358	0.251	0.2345

**Step 9:**

It follows from Definition 2.3 that it satisfies all the below conditions.

- (1)  $0 \leq d(X, Y) \leq 1$  hold,

Since  $d(X, Y), d(X, Z), d(Y, Z) \in [0, 1]$ .

- (2)  $d(X, Y) = 0$  iff  $X = Y$  holds, its straight forward
- (3)  $d(X, Y) = d(Y, X)$  holds, use of absolute and square values.
- (4)  $d(X, Z) + d(Y, Z) \geq d(X, Y)$  holds for all four proposed distances.

**Step 10:**

Distance measure and alternatives of ranking.

Distance measure	XY company	XZ company	YZ company
Hamming distance	0.95	0.9	0.65
Normalized Hamming distance	0.3166	0.3	0.2166
Euclidean distance	0.543	0.435	0.406
Normalized Euclidean	0.31358	0.251	0.2345
Rank	1	2	3

## 4 Conclusion

In this paper, we introduce four types of distance measures on Pythagorean fuzzy rough sets. After that we apply the concept in MCDM problem to buy a 3-phase induction motor, we conclude that the *XY* company is the best company compared to other *X Z* company, *YZ* company. From alternatives of ranking, team expert decided to buy 3-phase induction motor from *XY* company. In further study, we will analyse the new distance measure on PFRS.

## References

1. Yager RR (2013) Pythagorean fuzzy subsets. In: IFSA world congress and NAFIPS annualmeeting. Edmontom, Canada, pp 57–61
2. Zhou L, Wei-Zhi W, Zhang W-X (2009) On intuitionistic fuzzy rough sets and their topological structures. *Int J Gen Syst* 38:589–616
3. Ejegwa PA (2018) Distance and similarity measures for pythagorean fuzzy sets. Published by Granular computing

# Real-Time Monitoring of Buses Utilizing Phasor Measurement Units for a 24-Bus System



G. Babu Naik

## 1 Introduction

In the year of 1970's has newly PMU's system create the SCDR in power system lines [1]. In the past years the microcomputers were incapacity of handling the power systems algorithm for distance relay. The symmetrical component distance relay (SCDR) was invented and give the linear components of voltage and current, which converts the three-phase system of transmission line into the single equation of six faults equations utilizing the symmetrical parameters. Over a time period, it turns into evident that microcomputers systems had advanced to the point where line relaying no longer required this innovation. The SCDR's use of very efficient approaches for measuring voltages and currents of the symmetrical components, on the other hand, it has been proved that it is very useful for other use of applications. Impact, the most support of the power system analysis is programmers typed such as to measure the voltage and currents positive sequence of the network used by methods are load flow analysis, stability of system, open and short network, the power flow techniques and other network systems. The synchronization process of real-time data of phasor angle and the phasor measurements techniques involved for particular systems distributed power network offer a great occasion to improve power system topology control and protection systems. The synchronized power network port, phasor measurements techniques is only recently developed a viable option, thanks to the introduction of fast microcontrollers and software. The GPS satellites deliver extremely accurate synchronization. The major development of GPS-based Phasor Measurement Units (PMU) represents the possible monitoring and directly observe system functioning status. [2] As final result, their system is under full possible for usage in power systems

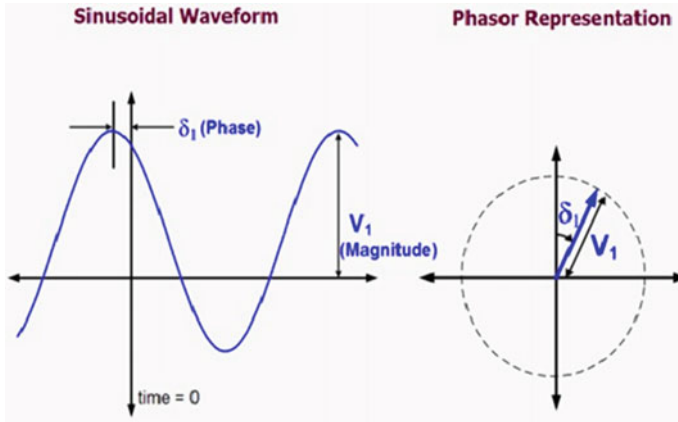
---

G. B. Naik (✉)

Department of Electrical and Electronics Engineering, BMS Institute of Technology and Management, Bengaluru, Karnataka, India

e-mail: [babunaikg@bmsit.in](mailto:babunaikg@bmsit.in)





**Fig. 1** The sinusoidal signal of Phasor diagram

network. Phasor Measurement Units (PMU) synchronize sample clocks with GPS transmissions calculated phasors are the single references. A sinusoidal signal is represented by a phasor, which is a quantity containing magnitude and phase [3–4]. The phase is representing the phase angle has the distance between a sinusoidal peak signal and a well-defined reference, and it is measured in angles units. The location in this case is a fixed point with respect to the time (Fig. 1).

## 2 Objective of This Paper

This paper describes a strategy for placing of PMU's that ensures that the monitor the vulnerable buses constructed on the overall system's transient and voltage stability analyses. If PMUs location are positioned at the buses, where transmission lines are more critical and it should monitor the in real-time situations for phase angles around the different nodes position of transmission lines, which also indicates the nearness to the instability that function will be properly define the problems. When transient and voltage stability predictions are considered, at that time it is very difficult to identify the buses where PMUs should be positioned due to the various disturbances occur in transmission lie. The ideally placing of such important buses identified through a transient and voltage stability research is bee determined by using of ILP technique with 'equality' and 'inequality' constraints.

### 3 Voltage Stability Analysis Approaches to Find Identification of Critical Buses

The main purpose of this work is to identify vulnerable buses using voltage stability analysis [5]. Monitoring voltage phasors at each node is critical for power system analysis. Both magnitude and phase are represented by phasor quantities. Unfortunately, only the magnitudes are measured in the current ‘Supervisory Control and Data Acquisition/Energy Management System’ (SCADA/EMS), and phases are inferred using state estimation algorithm.

#### 3.1 Voltage Stability-Index (L)

For power system network indicates as follows.

$n$ —represents the number of buses,

$g$ —represents the generator buses, and we can write equation for a certain system.

$$\begin{bmatrix} I_G \\ I_L \end{bmatrix} = \begin{bmatrix} Y_{GG} & Y_{GL} \\ Y_{LG} & Y_{LL} \end{bmatrix} \times \begin{bmatrix} V_G \\ V_L \end{bmatrix}$$

where  $I_G, I_L$  are the currents and  $V_G, V_L$  are voltages. Using above equations after simplifying we get in Matrix form

$$\begin{bmatrix} V_L \\ I_G \end{bmatrix} = \begin{bmatrix} Z_{LL} & F_{LG} \\ K_{GL} & Y_{GG} \end{bmatrix} \times \begin{bmatrix} I_L \\ V_G \end{bmatrix}$$

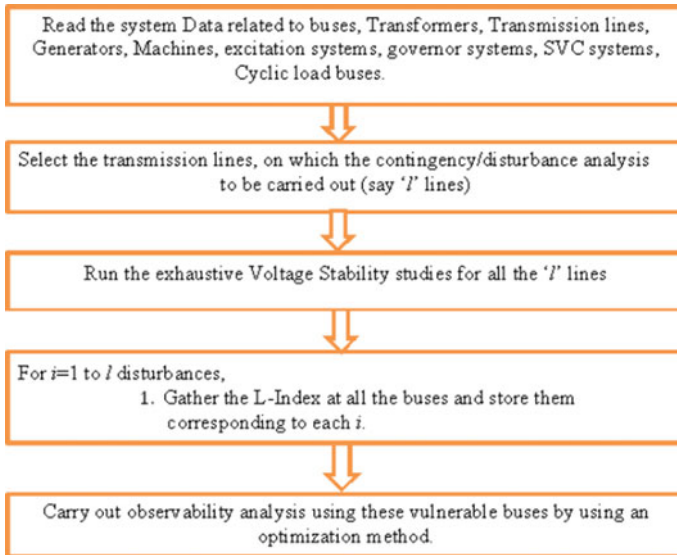
where  $[F_{LG}] = -\{[Y_{LL}]^{-1}[Y_{LG}]\}$

$[F_{LG}]$  are the generator buses referred to the columns and load buses are referred to the rows in matrix.

This matrix relates the load bus and source bus voltages. This matrix also provides the information for electrical distance of load and generator nodes, which is also can be defined in terms of loads and generators systems. The function of VSI(L) is calculated for a system operational situation using the state estimate or load flow findings as follows:

$$L_j = \left| 1 - \sum_{i=1}^{i=g} F_{ji} \frac{V_i}{V_j} \right|$$

The ‘L-index’ value can be obtained the ranges of ‘0’ and ‘1’. If the ‘L-index’ value is unity and zero indicates the system of voltage stability is improved.



**Fig. 2** Algorithm for the proposed approach

### 3.2 Proposed Approach

The suggested method for locating vulnerable buses entails reading the data of buses [6–8], power system transmission lines, and other sources, selecting the function of a transmission line for contingency analysis, performing a thorough a voltage stability study for the all lines of power system network, and determining the value of greatest ‘L-index’. To locate the critical buses, merit lists them in descending order and categorizes them into groups shown in Fig. 2.

### 3.3 24-Bus Equivalent Indian Southern Grid

The 24-bus system has generator buses are four and load buses are eight make the fleet. Buses 1, 2, 3, and 4 are equipped with generators, whereas buses 5, 6, 7, 8, 9, 10, 13, and 15 are equipped with load buses (Fig. 3).

The L-index range of the 24-bus southern grid scenario are shown in Fig. 4. Is the load of 100% at all the buses and the Fig. 5. Represents the 110% at bus 5 (Fig. 6).

By increasing the 10% load for L-index values at buses 5, 6, 7, 8, 9, 10, 13, 15 publicized in below figures accordingly (Figs. 7, 8 and 9).

For 24-bus system increasing the load by 10% at load buses and the depending upon the value of L-index placed by decreasing order [9–11]. The mostly critical buses are repeated appears in buses as well as appears in time shown in the group.

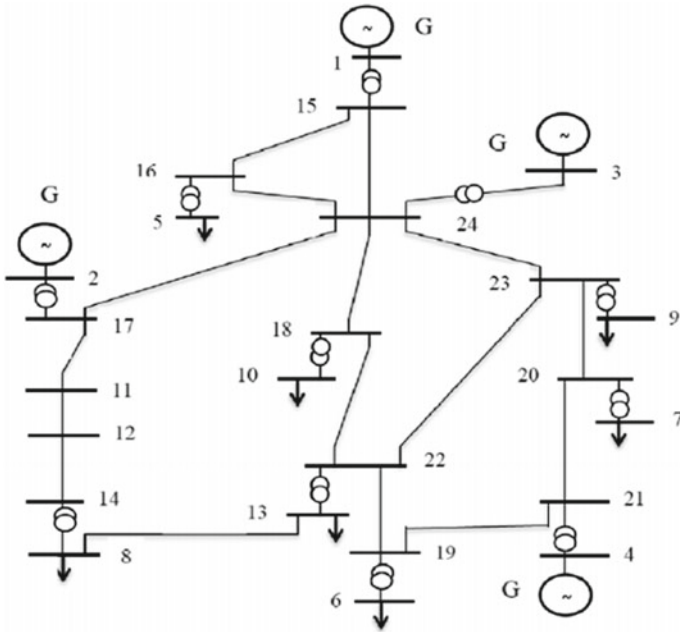


Fig. 3 Represents the 24-bus system of diagram single-line for EHV system

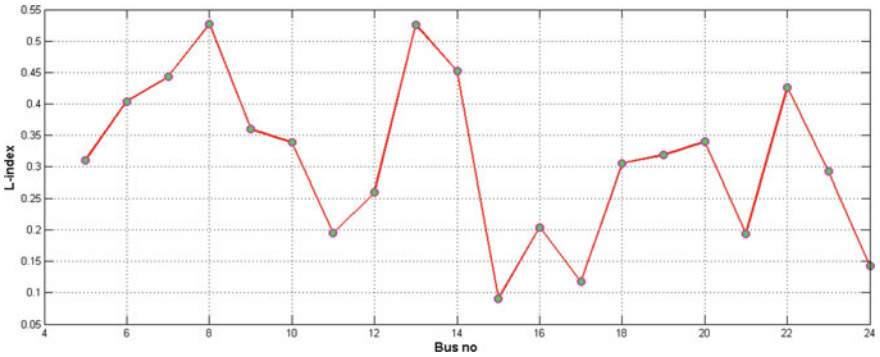


Fig. 4 24-bus system of L-index fully loaded.

The most critical buses, which appears as the heist time will be considered as the critical for faulty conditions in transmission lines (Table 1).

As a result, the most susceptible buses on the Indian southern grid’s 24-bus equivalent are 14, 22, 20, 19, 18, 23, 16, 11, 21, 24, 17, and 15, in decreasing order of risk.

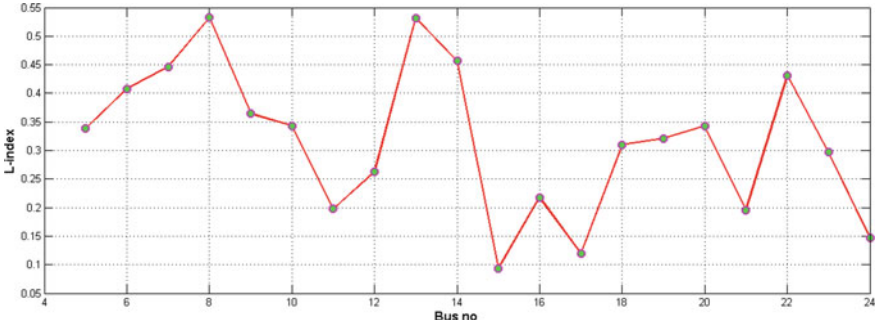


Fig. 5 24-bus system of L-index with a load factor of 110% at bus 5

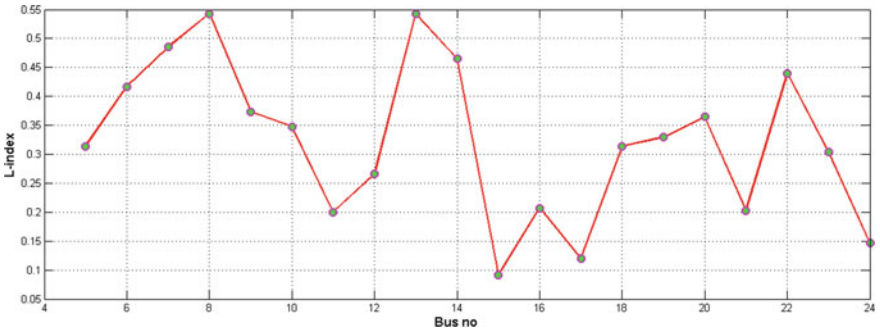


Fig. 6 L-index for a 24-bus system with a load factor of 110% at bus 7

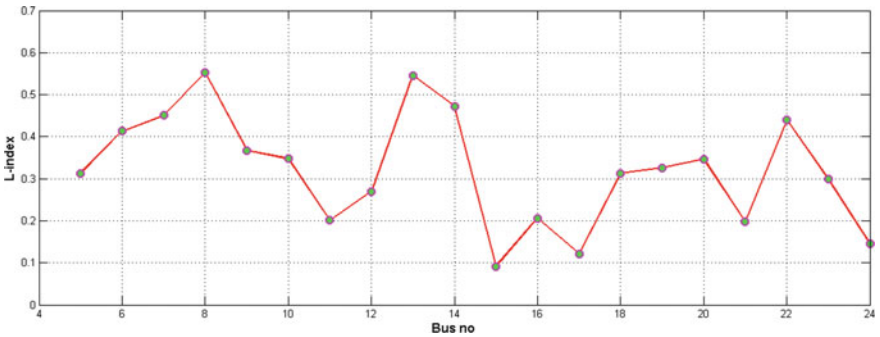


Fig. 7 L-index for a 24-bus system with a load factor of 110% at bus 8

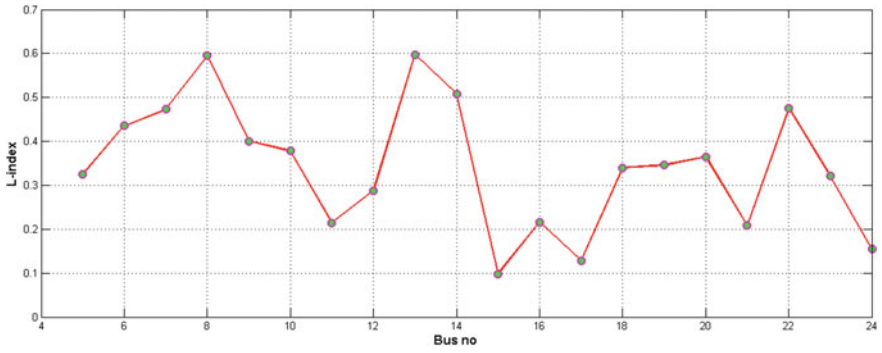


Fig. 8 L-index for a 24-bus system with 110% occupancy on buses 5, 6, 7, and 8

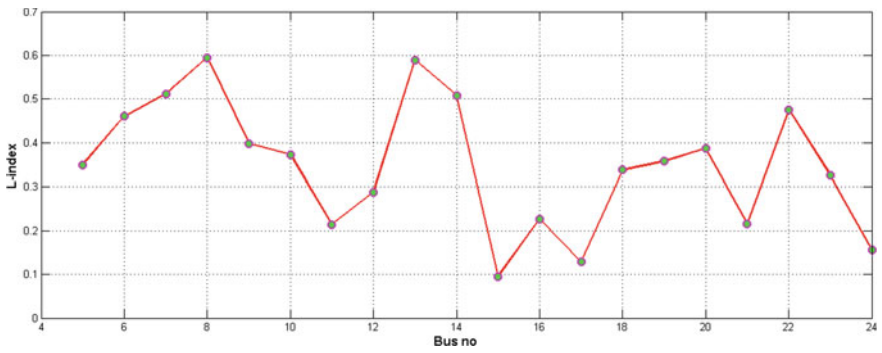


Fig. 9 L-index for a 24-bus system with 110% occupancy on bus 9, 10, 13, and 15

Table 1 24 bus system of L-index

Buses	Place 1	Place 2	Place 3	Place 4
14	10/10			
22		10/10		
20			10/10	
19				10/10

## 4 Observability Analysis

The observability analysis and phasing technique are discussed in this study [12–17]. The ability to observe the power system network is necessary for state estimation. All categories of electrical quantities measured in the power system are estimated using state estimation. The results can be applied for operation control of power system online and management system.

- (i) **Optimal PMU Placement**  
The use of synchronized phasor measurements raises the bar for power system monitoring, control, and protection. PMUs have already been implemented in a number of utilities throughout the world for a variety of applications.
- (ii) **24-Bus Equivalent Indian Southern Grid**  
To observe complete power system network for 24-bus system the PMUs required and must be minimized by considering most critical buses is the main objective. By determining the PMUs are required complete to be observability for power system network. Table 2: shows the minimal number of PMUs is required for 24 bus system of southern grid observability of the system.
- (iii) **Case studies using modified ILP technique**  
On a modified Indian southern grid 24 bus system, the modified ILP approach is used. The approach is put to the test in all four scenarios, and the PMU locations are determined by prioritizing the vulnerable buses. In this case, the all PMUs rates are almost is the equal for installation to monitor the power system network. For this scenario, a 24-bus comparable for Indian southern grid is required, and with the changed formulation in the optimization problem, a total of network 11 PMUs is required for all situations (Table 3).

**Table 2** Optimal PMU placement for 24 bus system

Minimum no PMUs	Optimal PMU locations
11	14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24

**Table 3** Optimal PMU placement for 24 bus system

Sl.No	Bus category	Number of PMUs required	Location at the buses
1	14	11	Buses: 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24
2	14, 22	11	Buses: 14, 15,16, 17, 18, 19, 20, 21, 22, 23, 24
3	14, 22, 22	11	Buses: 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24
4	14, 22, 20,19	11	Buses: 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24

## 5 Conclusions

A method based on voltage stability analysis has been proposed to identify vulnerable buses, and this category is applied to the Indian Southern Grid-24-bus system. To determine faults for power system transmission line at least required PMUs for total observability of the system was developed by the ILP technique. This been implemented for all the tested systems network and also considered for the power system protection. A newly method of optimally placing of PMUs has been devised, ensuring the essential nodes of a system being monitored. The framed optimization problem is solved using integer linear programming.

## References

1. Behera S, Dev Choudhury NB, Tripathy D (2021) Brief study on applications of phasor measurement units in smartgrid technology. In: 2021 1st international conference on power electronics and energy (ICPEE), 2021, pp 1–5. <https://doi.org/10.1109/ICPEE50452.2021.9358725>
2. Ahmed MM, Imran K (2019) An optimal PMU Placement Against N-1 contingency of PMU using integer linear programming approach. *Int Conf Appl Eng Math (ICAEM)* 2019:127–132. <https://doi.org/10.1109/ICAEM.2019.8853656>
3. Matavalam ARR, Singhal A, Ajarapu V (2020) Monitoring long term voltage instability due to distribution and transmission interaction using unbalanced  $\mu$ PMU & PMU measurements. *IEEE Power Energy Soc General Meet (PESGM)* 2020:1–1. <https://doi.org/10.1109/PESGM41954.2020.9282086>
4. Reis AW, Guarda FGK, Gastaldini CC (2019) Simulation of a phasor measurement unit in real time using typhoon virtual HIL. In: 2019 IEEE PES innovative smart grid technologies conference—Latin America (ISGT Latin America), pp 1–5. <https://doi.org/10.1109/ISGT-LA.2019.8895261>
5. Alzaareer K, Saad M (2018) Real-time voltage stability monitoring in smart distribution grids. *Int Conf Renew Energy Power Eng (REPE)* 2018:13–17. <https://doi.org/10.1109/REPE.2018.8657671>
6. Alzahawi T (2018) Modified application of phasor measurement unit in voltage monitoring of power system. *Third Sci Conf Electr Eng (SCEE)* 2018:230–234. <https://doi.org/10.1109/SCEE.2018.8682240>
7. Abdelkader SM, Hatata AY, Abdelbaky MM (2017) Dynamic stability analysis of synchronous generator using phasor measurement units. *Nineteenth Int Middle East Power Syst Conf (MEPCON)* 2017:1416–1423. <https://doi.org/10.1109/MEPCON.2017.8301369>
8. Gupta AK, Verma K, Niazi KR (2017) Contingency constrained optimal placement of PMUs for wide area low frequency oscillation monitoring. In: 2017 7th international conference on power systems (ICPS), pp 579–584. <https://doi.org/10.1109/ICPES.2017.8387360>
9. Wu Y, Musavi M, Lerley P (2016) Synchrophasor-based monitoring of critical generator buses for transient stability. *IEEE Trans Power Syst* 31(1):287–295. <https://doi.org/10.1109/TPWRS.2015.2395955>
10. Phadke G, Ibrahim M, Hlibka T (1997) Fundamental basis for distance relaying with symmetrical components. *IEEE Trans PAS, PAS-96(2):635–646, Mar/Apr 1977*
11. Phadke AG (1993) Synchronised phasor measurements in power systems. *IEEE Comput Appl Power* 6(2):1015
12. Baldwin TL, Mili L, Boisen MB Jr, Adapa R (1993) Power system observability with minimal Phasor measurement placement. *IEEE Trans Power Syst* 8(2):707–715



13. Xu B, Abur A (2004) Observability analysis and measurement placement for systems with PMUs. In: Proceeding of the 2004 IEEE PES conference and exposition, vol. 2, pp. 943–946, 10–13 Oct 2004
14. Dua D, Dambhare S, Gajbhiye RK, Soman SA (2008) Optimal multistage scheduling of PMU placement: An ILP approach. *IEEE Trans Power Delivery* 23(4):1812–1820
15. Chakrabarti S, Kyriakides E, Eliades DG (2009) Placement of synchronized measurements for power system observability. *IEEE Trans Power Deliv* 24(1):12–19
16. Sodhi R, Srivastava SC (2008) Optimal PMU placement to ensure observability of power system. In: Fifteenth national power systems conference (NPSC), Dec 2008
17. Mahipal B, Naik GB, Chander SP, Ravi G (2014) Optimal siting and sizing of DG and capacitor in distribution network using weight-improved particle swarm optimization algorithm (WIPSO). In: *IOSR journal of engineering (IOSRJEN)* [www.iosrjen.org](http://www.iosrjen.org), vol. X(X) (Month. 2014), II-VXII, PP 00–00, ISSN (e): 2250–3021, ISSN (p): 2278–8719

# State-Space Modelling for Two-Area Multiple-Source AGC Using PI, PID, PI-PD Controllers



P. Vidya Sagar and M. S. Krishna Rayalu

## 1 Introduction

Every power system's ultimate goal is to provide consumers with reliable power. However, accomplishing the set goal is challenging due to a mismatch between increasing load demand and scheduled generation. If frequency along with tie line power are appropriately managed, the imbalance between load demand and generation may be mitigated, allowing the intended goal to be achieved. In today's power systems, it is impossible to manually control these parameters. Modern control technologies, such as AGC, are employed to automatically reduce frequency along with tie line power anomalies for this purpose. AGC employs a wide range of advanced control techniques to keep frequency anomalies as well as tie line power variances in tolerable ranges.

Numerous papers have been published in the field of AGC, ranging from earlier design methodologies to more recent methodologies like fuzzy logic [1], PSO [2–4], genetic algorithms, bat algorithm and firefly algorithm etc. All of these methods are used to tune AGC gain parameters while simulating in MATLAB. AGC regulates the gain values of the controllers through a feedback mechanism. The gain levels can be manually changed using the trial and error approach, or they can be automatically adjusted using various types of current artificial methodologies.

The traditional technique called integral time absolute error (ITAE) [5] incorporated to enhance a standard PID controller. For adjusting the gain values of PID controller, an efficient artificial intelligence-based PSO technique [2–4] is proposed and implemented in two areas of multiple-source inter-linked power networks. The efficacy and robustness of the adopted approach were assessed using an injection of 1% step load disturbance in area 1.

---

P. Vidya Sagar · M. S. Krishna Rayalu (✉)  
Department of E.E.E, V.R. Siddartha Engineering College, Vijayawada, AP 520007, India  
e-mail: [mskr@vrsiddartha.in](mailto:mskr@vrsiddartha.in)

### 1.1 AGC Approach for Multiple Areas Integrated with Multiple-Source Power Systems

The multi-area with multiple sources in each area is depicted in Fig. 1. The generation in this case is made up of classical hydro, gas, and thermal sources. In recent years, the need for sustaining energy has increased, and in comparison, to existing energy sources, sustainable energy sources such as solar and wind have been incorporated. AGC is the means of controlling generator output power between various control zones in response to frequency aberrations and tie line power abnormalities. Hence, the AGC’s objective is to keep the area control error (ACE) [5] within tolerable limits. The frequency should be monitored within the prescribed limits in order to maintain load demand and allotted generation stability.

$$ACE_i = \Delta P_{tie,i} + B_i \Delta F_i \tag{1}$$

Each area has many power generations sources in this case. Each area employs conventional energy-producing sources such as gas, hydro, and thermal as well as incorporating renewable green power sources like solar and wind plants. The *i*th control area is linked to the *j*th control area by an AC tie line in this system.  $P_{ri}$  and  $P_{rj}$  MW are the power ratings of the *i*th and *j*th control areas, respectively.

Conventional plants such as gas, hydro, thermal, and sustainable energy-producing sources like wind and solar are used to generate load-sharing factors  $K_{ti}$ ,  $K_{hi}$ ,  $K_{gi}$ ,  $K_{si}$ , and  $K_{wi}$  [6, 7]. Load demand and scheduled economic load dispatch determine

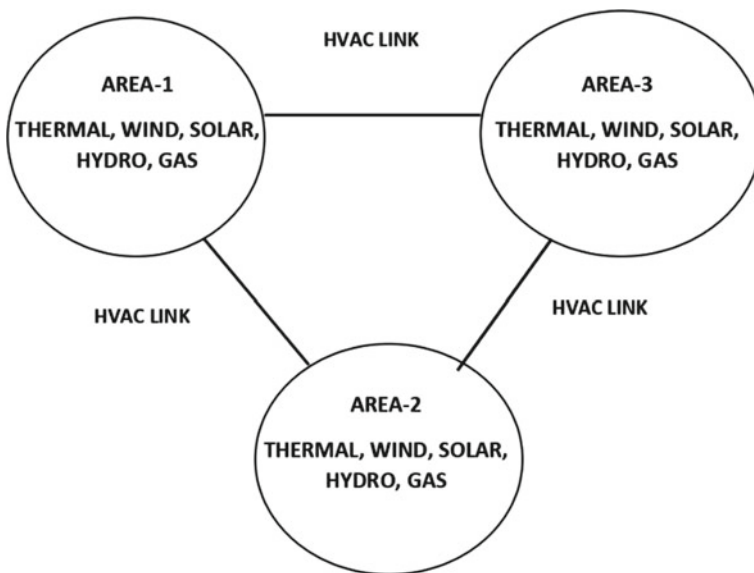


Fig. 1 Multi-area diverse source power system

the value of the sharing factor. For the  $i$ th area, the  $P_{Ggi}$ ,  $P_{Gti}$ ,  $P_{Ghi}$ ,  $P_{Gsi}$ , and  $P_{Gwi}$  are the powers generated by the gas, thermal, hydro, solar, and wind plants. The total power generation  $P_{Gi}$  may be expressed as

$$P_{Gi} = P_{Gti} + P_{Ghi} + P_{Gwi} + P_{Ggi} + P_{Gsi} \tag{2}$$

where  $P_{Gti} = K_{ti} P_{Gi}$ ,  $P_{Ghi} = K_{hi} P_{Gi}$ ,  $P_{Gwi} = K_{wi} P_{Gi}$ ,  $P_{Gsi} = K_{si} P_{Gi}$ ,  $P_{Ggi} = K_{gi} P_{Gi}$   $K_{ti} + K_{hi} + K_{wi} + K_{si} + K_{gi} = 1$ .

For small deviation in load change in the  $i$ th area, the Eq. (2) can be written as

$$\Delta P_{Gi} = \Delta P_{Gti} + \Delta P_{Ghi} + \Delta P_{Gwi} + \Delta P_{Ggi} + \Delta P_{Gsi} \tag{3}$$

## 2 PI-PD Approach for Two-Area AGC

The cascaded PI-PD controller is a combination of two processes, such as the inner and outer processes. In this scenario, the output of the inner process is transmitted to the inner process. This controller is primarily concerned with the minimization of outer process instabilities. System’s performance has increased by utilizing this controller. The PI-PD [8] controller’s mathematical representation is (Fig. 2),

$$G_{PI-PD}(s) = \left( K_1 + \frac{K_2}{s} \right) (K_3 + K_4 s) \tag{4}$$

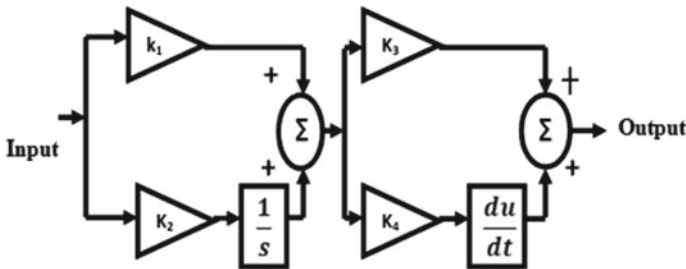


Fig. 2 Representation of the PI-PD controller block diagram

### 3 Two-Area Diverse Source AGC Power System Employed with PID and PI-PD Control

The Fig. 3 illustrates the simulation model of the two-area diversified source AGC integrated with PID and also simulated using a PI-PD controller [8]. The data for the above model is given in the appendix. The tie line as well as frequency anomalies must be minimized to zero to balance generation and load demand. These controllers are usually used as a backup. Any of today’s artificial intelligence tools may be used to automatically optimize a controller. The particle swarm optimization (PSO) strategy is employed to automatically update controller the gain values [2–4]. The controllers enable system control without requiring comprehensive knowledge of system properties or transfer functions, and they are simple to implement and execute.

#### 3.1 Representation of State-Space Model of Two-Area Integrated Multiple-Source Power System

**State variables for the Two-Area power system:** Fig. 4 shows the state variables developed for the system shown in Fig. 3. All of these state variable techniques are also covered in the literature of [1, 6, 9–12]

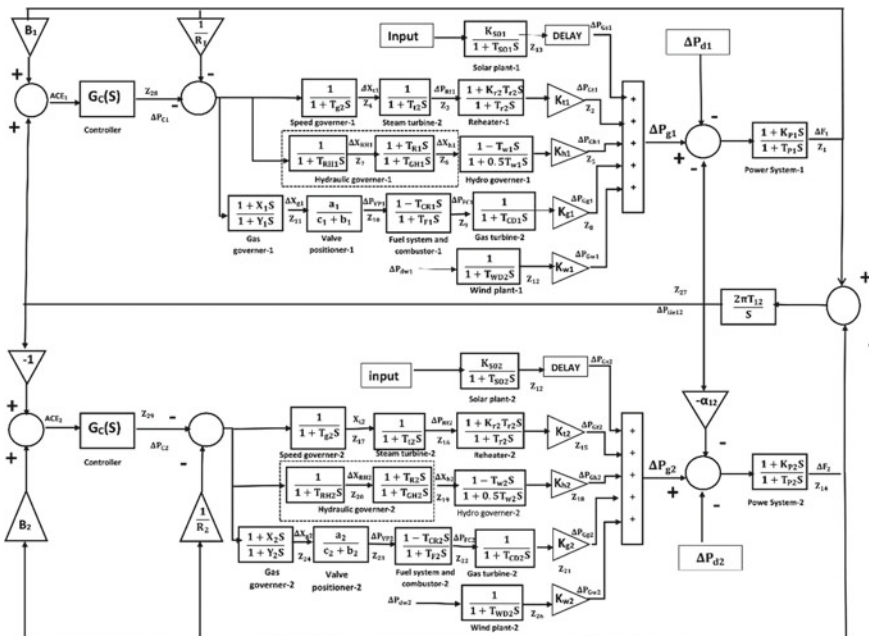


Fig. 3 Two-area AGC diverse source power system simulation diagram

Area-1	$Z_1=\Delta F_1, Z_2=\Delta P_{Gt1}, Z_3=\Delta P_{Rt1}, Z_4=\Delta X_{t1}, Z_5=\Delta P_{Gh1}, Z_6=\Delta X_{h1},$ $Z_7=\Delta X_{RH1}, Z_8=\Delta P_{Gg1}, Z_9=\Delta P_{FC1}, Z_{10}=\Delta P_{VP1}, Z_{11}=\Delta P_{GW1},$ $Z_{13}=\Delta P_{GS1}, Z_{28}=\int ACE_1 dt$
Area-2	$Z_{14}=\Delta F_2, Z_{15}=\Delta P_{Gt2}, Z_{16}=\Delta P_{Rt2}, Z_{17}=\Delta X_{t2}, Z_{18}=\Delta P_{Gh2}, Z_{19}=\Delta X_{h2},$ $Z_{20}=\Delta X_{RH2}, Z_{21}=\Delta P_{Gg2}, Z_{22}=\Delta P_{FC2}, Z_{23}=\Delta P_{VP2}, Z_{24}=\Delta X_{g2},$ $Z_{25}=\Delta P_{GW2}, Z_{26}=\Delta P_{GS2}, Z_{29}=\int ACE_1 dt, Z_{27}=\Delta P_{tie i,j}$

Fig. 4 The two-area system’s state variables

**Inputs of Control**  $[u_1 \ u_2] = [P_{c1}, P_{c2}]$  for area 1 and area 2.

**Inputs of disturbance**  $[d_1 \ d_2] = [\Delta P_{d1} \ \Delta P_{dw1} \ \Delta P_{ds1}; \ \Delta P_{d2}, \ \Delta P_{dw2} \ \Delta P_{ds2}]$  for 2 areas.

The below is a state equation [6], the matrix can be represented in the form,

$$\dot{Z} = Gz + Hu$$

$$y = Cz + Du$$

The dimensions of matrix G = (29 × 29)

The dimensions of matrix H = (29 × 2)

**Non-zero elements of Matrix G: order of (29 × 29)**

$g_{1,1} = \frac{-1}{T_{P1}}$	$g_{1,2} = \frac{K_{P1}}{T_{P1}}$	$g_{1,5} = \frac{K_{P1}}{T_{P1}}$
$g_{1,8} = \frac{K_{P1}}{T_{P1}}$	$g_{1,12} = \frac{K_{P1}}{T_{P1}}$	$g_{1,13} = \frac{K_{P1}}{T_{P1}}$
$g_{1,27} = -\frac{K_{P1}}{T_{P1}}$	$g_{2,2} = -\frac{1}{T_{r1}}$	$g_{2,3} = \left[ \frac{K_{r1}}{T_{r1}} - \frac{K_{r1}K_{t1}}{T_{t1}} \right]$
$g_{2,4} = \frac{K_{r1}K_{t1}}{T_{t1}}$	$g_{3,3} = -\frac{1}{T_{i1}}$	$g_{3,4} = \frac{1}{T_{i1}}$
$g_{4,1} = \frac{1}{T_{g1}R_1}$	$g_{4,4} = -\frac{1}{T_{g1}}$	$g_{5,1} = \frac{2K_{h1}T_{R1}}{K_{GH1}T_{RH1}R_1}$
$g_{5,5} = -\frac{2}{T_{W1}}$	$g_{5,6} = \left[ \frac{2K_{h1}}{T_{W1}} + \frac{2K_{h1}}{T_{GH1}} \right]$	$g_{5,7} = \left[ \frac{2K_{h1}T_{R1}}{T_{GH1}T_{RH1}} - \frac{2K_{h1}}{T_{GH1}} \right]$
$g_{6,1} = -\frac{T_{R1}}{T_{GH1}T_{RH1}R_1}$	$g_{6,6} = -\frac{1}{T_{GH1}}$	$g_{6,7} = \left[ \frac{1}{T_{GH1}} - \frac{T_{R1}}{T_{GH1}T_{RH1}} \right]$
$g_{7,1} = -\frac{1}{T_{RH1}R_1}$	$g_{7,7} = -\frac{1}{T_{RH1}}$	$g_{8,8} = -\frac{1}{T_{CD1}}$
$g_{8,9} = \frac{K_{g1}}{T_{CD1}}$	$g_{8,10} = -\frac{K_{g1}T_{CR1}}{T_{CD1}T_{F1}}$	$g_{9,9} = -\frac{1}{T_{F1}}$
$g_{9,10} = \left[ \frac{1}{T_{F1}} + \frac{T_{CR1}}{(T_{F1})^2} \right]$	$g_{10,1} = -\frac{X_1}{b_1R_1Y_1}$	$g_{10,10} = -\frac{1}{b_1}$
$g_{10,11} = -\frac{1}{b_1}$	$g_{11,1} = \left[ \frac{X_1}{R_1Y_1^2} - \frac{1}{R_1Y_1} \right]$	$g_{11,11} = -\frac{1}{Y_1}$
$g_{12,12} = -\frac{1}{T_{WD1}}$	$g_{13,13} = -\frac{1}{T_{st1}}$	$g_{14,14} = -\frac{1}{T_{P2}}$
$g_{14,15} = \frac{K_{P2}}{T_{P2}}$	$g_{14,18} = \frac{K_{P2}}{T_{P2}}$	$g_{14,21} = \frac{K_{P2}}{T_{P2}}$

(continued)



## 4 Optimal Quadratic Controller Design

The performance index ( $J$ ) [13] to be provided will be as follows:

$$J = \frac{1}{2} \int_0^{\infty} (z^T Q z + u^T R u) dt. \quad (5)$$

where

$Q$  is a matrix with a real symmetric positive semi-definite value.

$R$  is a matrix with a symmetric positive definite [6, 13].

And  $R$  and  $Q$  provide the control on state error and the quantity of energy consumed. According to the optimum control principle [13], the following is the optimum control for full state feedback:

$$U = -Kz.$$

By solving Ricatti equation,  $P$  can be obtained. Gain matrix  $K$  [13] would be obtained.

$$G'P + PG - PHR^{-1}H'P + Q \quad (6)$$

$$K = \text{Optimal gain matrix} = R^{-1}H'P \quad (7)$$

## 5 Results

The appendix includes data on the different parameters of the multiple-area multiple-source power system, and the simulation for the controllers of PI, PID as well as PI-PD is shown in Fig. 3. The PSO strategy, an AI tool, was used to automatically update the controller gain values. Figures 5, 6 and 7 show comparative frequency as well as tie line power deviance waveforms employing PI, PID, PI-PD, and LQR. Matrices  $Q$  and  $R$  have been created using the LQR technique via trial and error. Except for diagonal entries, the matrix  $Q$ 's elements are all zeros.

The gain values of various controllers are listed below in Table 1.

The non-zero matrix  $Q$  coefficients,

$$\begin{aligned} Q_{1,1} &= 0.1, Q_{2,2} = 0.1, Q_{3,3} = 0.1, Q_{4,4} = 0.1, Q_{5,5} = 0.1, \\ Q_{6,6} &= 0.1, Q_{7,7} = 0.1, Q_{8,8} = 0.1, Q_{9,9} = 0.1, Q_{10,10} = 0.1, \\ Q_{11,11} &= 0.1, Q_{12,12} = 10, Q_{13,13} = 1, Q_{14,14} = 1, Q_{15,15} = 1, \end{aligned}$$



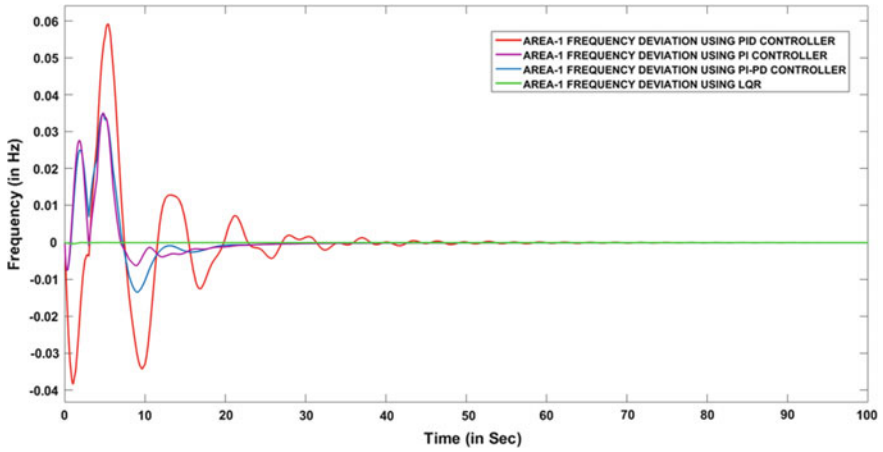


Fig. 5 Area-1 frequency deviations employing PI, PID, PI-PD, and LQR

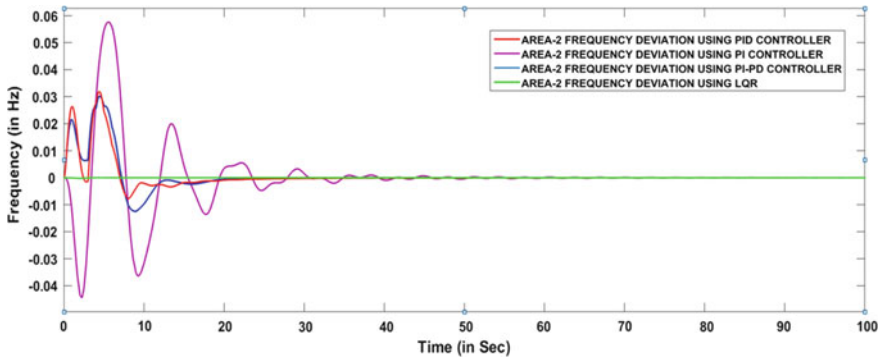


Fig. 6 Area-2 frequency deviations employing PI, PID, PI-PD, and LQR

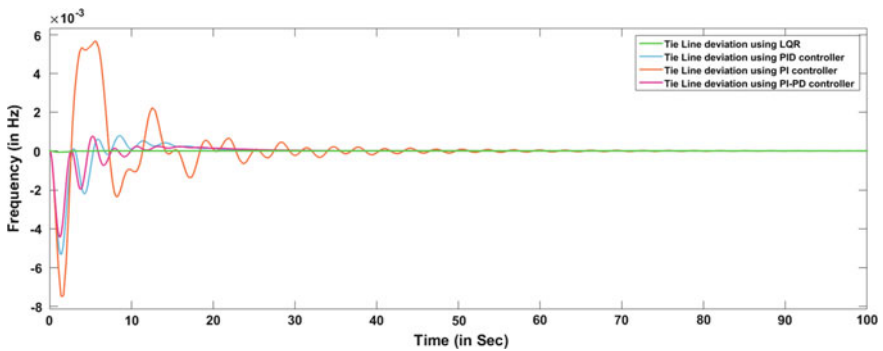


Fig. 7 Tie line power deviations employing PI, PID, PI-PD, and LQR

**Table 1** Two-area system's state variables

Controller	Gain values
PI	$K_{p1} = 4.443, K_{i1} = 3.8774, K_{p2} = 7.6786, K_{i2} = 7.9725$
PID	$K_{p1} = 1.0, K_{i1} = 1.00, K_{d1} = 0.6856, K_{p2} = 1.00, K_{i2} = 1.00, K_{d2} = 0.443$
PI-PD	$K_{p1} = 0.9336, K_{i1} = 0.9028, K_{p2} = 1.00, K_{d1} = 0.9315, K_{p3} = 0.9428, K_{i2} = 1.00, K_{p4} = 1.00, K_{d2} = 0.7065$

**Table 2** The comparative results of various controllers

Controller	Over shoot	Under shoot	Settling time (s)
PI	0.059	0.038	85
PID	0.0362	0.0132	30.8
PI-PD	0.0327	0.0064	20
LQR	$3.88e-5$	$4.54e-5$	10

$$\begin{aligned}
 Q_{16,16} &= 1, Q_{17,17} = 1, Q_{18,18} = 1, Q_{19,19} = 1, Q_{20,20} = 1, \\
 Q_{21,21} &= 1, Q_{22,22} = 1, Q_{23,23} = 1, Q_{24,24} = 1, Q_{25,25} = 1, \\
 Q_{26,26} &= 1, Q_{27,27} = 10, Q_{28,28} = 10, Q_{29,29} = 1
 \end{aligned}$$

$$\begin{aligned}
 R &= [0.00010 \\
 &\quad 0\ 0.0001]
 \end{aligned}$$

The waveforms in the following figures show how various types of controllers can be used to improve frequency as well as tie line anomalies (Table 2).

## 6 Conclusion

This study evaluates a two-area power system with multiple generating sources employed in every control area. For generation, each area has traditional hydro, thermal, and gas sources, as well as renewable sources like wind and solar plants. Through the integration of different controllers like PI, PID as well as PI-PD, the tie line as well as frequency changes for the load injection in area-1 were observed. To enhance the performance of state-space modelling, an optimal regulator like LQR was designed. And the results of all the controllers were compared, The PID controller showed better settling time than the PI controller, and PI-PD showed better results than PID, and LQR showed optimal results than the PI-PD controller. To get better results, different algorithms can be used.

## Appendix

Data for the simulation diagram showed in Fig. 3:

$P_{rj} = 2000$  MW,  $H_i = 5.0$  MW/MVA,  $F_r = 50$  Hz,  $B_j = 0.425$  p.u MW/Hz,  $T_{12} = 0.0433$ ,  $D_j = 8.33 \times 10^{-3}$  p.u MW/Hz,  $\alpha_{12} = -1$ ,  $KP_j = 120$ ,  $TP_j = 20$ ,  $K_{tj} = 0.60$ ,  $K_{hj} = 0.20$ ,  $K_{gj} = 0.10$ ,  $K_{wj} = 0.10$ ,  $K_{soj} = 1.8$  s,  $T_{gj} = 0.08$  s,  $T_{tj} = 0.30$  s,  $T_{rj} = 10.0$  s,  $K_{rj} = 0.3$  s,  $T_{RHj} = 41.6$  s,  $T_{Rj} = 5$  s,  $T_{GHj} = 0.513$  s,  $T_{Wj} = 1.0$  s,  $T_{WDj} = 5.0$  s,  $T_{soj} = 1.8$  s,  $X_j = 0.60$  s,  $Y_j = 1.0$  s,  $b_j = 0.05$  s,  $c_j = 1.0$  s,  $T_{Fj} = 0.23$  s,  $T_{CRj} = 0.01$  s,  $T_{CDj} = 0.20$  s.

## References

1. Kothari P, Nagrath IJ (1980) Modern power system analysis. TMH (1980)
2. Sharma D et al (2015) PSO optimized PID controller for load frequency control. Eur J Adv Eng Technol 2(11):43–48
3. Gaurav D et al (2018) PSO based automatic generation control of multi-area power systems with renewable energy sources. Int J Electron Electr Comput Syst 7(3), ISSN 2348-117
4. Gaing Z-L (2004) Particle swarm optimization approach for optimum design of PID controller in AVR system. IEEE Trans Energy Convers 19(2) (2004)
5. Charitha Reddy YVL, Krishnarayalu MS, AGC of multi source multi area deregulated power system using PID controller. J Electr Eng (JEE). Article 18.1.40, [www.jee.ro](http://www.jee.ro)
6. Hakimuddin N, Nasiruddin I, Bhatti TS, Arya Y (2020) Optimal automatic generation control with hydro, thermal, gas, and wind power plants in 2-area interconnected power system. Electr Power Compon Syst
7. Ibraheem et al, AGC of two area interconnected power system with diverse sources in each area. J Electr Eng (JEE)
8. Bhongade S et al, Automatic generation control scheme including photo-voltaic generating system for interconnected power system. [www.iitk.ac.in](http://www.iitk.ac.in)
9. Ogata K (2010) Modern control engineering 5/e. PHI
10. Ibraheem N, Bhatti TS (2014) AGC of two area power system interconnected by AC/DC links with diverse sources in each area. Int J Electr Power Energy Syst 55:297–304
11. Arya Y, Kumar N (2017) Optimal control strategy-based AGC of electrical power systems. A comparative performance analysis. Optimal Control Appl Methods 38(6):982–992
12. Talukdar BK et al (2018) Design of state space model and optimal controller for automatic generation control of two area thermal power system. Int J Adv Res Innovative Ideas Educ (IJARIIIE) 4(3)
13. Babu Nayak DHCP, Krishna Rayalu MS (2020) Automatic generation control of two area thermal power system with PI controllers using state space approach. SSRG Int J Electr Electron Eng 7(6):17–21
14. Arya Y (2017) AGC performance enrichment of multi-source hydrothermal gas power systems using new optimized FOPID controller and redox flow batteries. Energy 127:704–715
15. Kothari ML et al (1998) Application of optimal control strategy to automatic generation control of a hydrothermal system. IEE Proc 135(4)

# Optimization of Power Generation Costs Through Soft Computing Techniques



M. V. Suganyadevi  and A. R. Danila Shirly 

## 1 Introduction

Within the grid functioning, efficient load dispatch is a critical daily optimization effort. Economic load dispatch in electricity production has the primary goal of scheduling committed power system outputs to encounter load requirements at the lowest feasible operating cost while maintaining system equality and different limitations. The quantity of available power-consuming products profile varies from bus to bus in an uncontrolled way; also generating plan has to be able to handle it directly without affecting the cost of operation significantly [1]. The classical seek strategies present to carry out best generation schedules with non-convex fuel price capabilities contemplating the load sample versions that cannot attain worldwide most effective solution. Different solution methods were evolved to resolve ELD problems.

To overcome the ED problem, evolutionary solutions have recently been proposed. A genetic algorithm's (GA) [2] early convergence eliminates the overall quality and limits the seeking ability, giving it a higher chance of finding a local optimal value. PSO [3] is capable of producing excessive acceptable results in a brief computing time it also features further overall probing capacity on starting execution and then confined seek close to stopping execution. HPSO [4] provides two things to PSO: breeding and a GA subpopulation approach.

Using the breeding strategy and subpopulation strategy, it is possible of immediately leaping from the current seeking factor into the powerful area. As a result, HPSO is capable of achieving a more efficient most efficient response than the traditional PSO. The application of GA, EP, PSO, HPSO, MAPSO, PSA and HMAPSO

---

M. V. Suganyadevi (✉)

Department of EEE, Saranathan College of Engineering, Trichy, India

e-mail: [mvsuganyadevi@gmail.com](mailto:mvsuganyadevi@gmail.com)

A. R. Danila Shirly

Department of EEE, Loyola-ICAM College of Engineering & Technology, Chennai, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2022

901

T. Sengodan et al. (eds.), *Advances in Electrical and Computer Technologies*,

Lecture Notes in Electrical Engineering 881,

[https://doi.org/10.1007/978-981-19-1111-8\\_71](https://doi.org/10.1007/978-981-19-1111-8_71)

algorithms to solve the ED problem of three test systems with non-convex fuel cost functions is the paper's main contribution. Comparison of the GA, EP, PSO, HPSO, MAPSO, PSA and BF algorithms for generator cost minimization with the new HMAPSO approach. The presented technique is used to estimate the most suitable cost for producing power [5] for the IEEE bus topologies like 14 bus, New England 39 and 118.

## 2 Mathematical Representation

The main objective is to minimize the total fuel cost of generation which is given below

$$\text{Minimize } F = \sum_{j=1}^n F_{c_j}(P_j) \quad (1)$$

where  $F_{c_j}(P_j)$  denotes the fuel cost function of unit  $j$  and  $P_j$  denotes the real power generated by unit  $j$ , subject to power balance restrictions

$$\sum_{j=1}^n P_j = P_D + P_L \quad (2)$$

where  $P_j$  is the real power generated by unit  $j$ ,  $P_D$  represents the real or direct power demand and  $P_L$  represents the transmission line loss. For generators, the capacity limitation is given by

$$P_{j\min} \leq P_j \leq P_{j\max} \text{ for } j = 1, 2, 3, \dots, n \quad (3)$$

where  $P_{j\min}$  denotes the least active power output of generating unit  $j$  and  $P_{j\max}$  denotes the highest active power output of generating unit  $j$ . The fuel cost function for generating unit  $j$  is denoted as

$$F_{c_j}(P_j) = a_i P_j^2 + b_i P_j + c_i \quad (4)$$

The fuel cost coefficients of generating unit  $j$  are  $a_i$ ,  $b_i$  and  $c_i$ .

## 2.1 Genetic Algorithm

Genetic algorithms (GA) are powerful search engines that are based on population genetics, natural selection and evolution principles [6]. GA searches the solution space of a given problem domain globally. The parameter set of the problem to be addressed is represented by a set of strings in GA. Each encoding corresponds to a unique member of the GA population. During the beginning of such GA cycle, each member of population is given a random value. The GA searches for better individuals in the space of potential chromosomes. Fitness values guide the search. Selection, crossover and mutation are the most commonly used genetic operators. The selection rule determines who will be depicted for next source of GA. The matching guidelines are applied to certain individuals in order to develop new ones in future generation. Physically strong people have a better chance of being chosen and, as a result, depicted at ensuing generation. The GA becomes more efficient by combining mutation and crossover, thereby causing premature loss of genetic information.

## 2.2 Evolutionary Programming

After a number of repetitions, evolutionary programming finds the global optimum solution by generating starting parental vectors that are uniformly distributed in intervals within the boundaries [7]. Initialization, production of child vectors by mutation and competition, and selection of best vectors to evaluate the best fitness solution are the main stages of this technique. One can implement EP technique into practise in the following way.

### Initialization

At the beginning, population would be created subsequently solving Eqs. (2) and (3). The real power outputs of the producing units, which are spread equally between their minimum and maximum limits, are the constituents of parent vectors ( $P_i$ ). For previous generations, the slack bus generator vector is calculated using the Newton–Raphson method.

### Mutation

From each parent vector, an offspring vector  $P_i'$  is created by  $I$  denoted  $N$  is random Gaussian parameter having mean = zero value and SD of unity or one.  $(0, \sigma_1^2)$

$$P_i' = P_i + N(0, \sigma_i^2) \quad \text{for } i = 1, 2, \dots, n - 1 \quad (5)$$

where

$$\sigma_i = \beta * \sum_{i=1}^n \frac{F_i(P_i)}{F_{\min}} (P_{i \max} - P_{i \min}) \tag{6}$$

The offspring vector should achieve the basic and highest generation limits of the units. The optimization with constraints has the ELD problem that can be changed to an unrestricted problem with the penalty factor (PF) given in the following equation. EP technique’s fitness function is expressed as follows.

$$\text{Fitness function} = F_t + \text{PF} * \varphi \tag{7}$$

At the start, the above equation second term becomes zero then later when mutation occurs it becomes nonzero if the generating vectors exceed its boundaries. As a result, the fitness function is reduced to just  $F_t$ , which is evaluated in the same way as it would be for parent vectors similarly it is done for offspring vectors.

**Competition and Selection**

Within the contesting pool, the parent trial vectors  $P_i$  and their progeny  $P_i'$  (a total of 2 k vectors) compete for survival. The cost of parent vectors  $F_t(P_i)$  and offspring vectors  $F_t(P_i')$  determines competition in this population. The best vector with the lowest cost is picked as the new parent for the following generation, whether it is parent vector  $P_i$  (or) child vector  $P_i'$ . The process of initialization and mutation is repeated until the fitness value does not significantly improve.

**2.3 Particle Swarm Optimization**

Eberhart et al. created PSO, a breakthrough optimization method. It employs a large sum of particles grouped in a swarm. Each particle searches for global minimum as it moves throughout the search space. Particles move about in a multidimensional search space in a PSO system. During flight, each particle modifies its position depending on its own and its neighbours’ experiences, choosing the best position that it and its neighbours have encountered. The set of particles that surround a particle, as well as its history, determine its direction. Let  $x_d$  and  $v_d$  represent a particle’s position (position) and velocity (velocity) in a search area, accordingly. A particle’s optimum previous location is preserved and denoted as  $p_{\text{best}}$ . Among all the particles in the group,  $g_{\text{best}}$  is the best particle index [8]. Finally, the formulas below can be used to compute each particle’s velocity and position:

$$v_{d+1} = k * (w * v_d + c_1.\text{rand}() * (p_{\text{best}} - x_d)) + c_2.\text{rand}() * (g_{\text{best}} - x_d) \tag{8}$$

$$x_{d+1} = x_d + v_{d+1} \tag{9}$$

where  $d$  is the generation pointer,  $x_d$  is the particle's current position,  $v_d$  is the particle's velocity, The inertia weight factor is  $w$ , the acceleration constants are  $c_1$  and  $c_2$ , and  $\text{rand}()$  returns a uniform random value between 0 and 1.  $W$  strikes a solid balance between global and local explorations with a good choice of inertia weight. In general, the equation is used to calculate the inertia weight  $w$ .

$$w = w_{\max} - \frac{w_{\max} - w_{\min}}{\text{iter}_{\max}} X_{\text{iter}} \quad (10)$$

$\text{Iter}_{\max}$  is the maximum number of generations, while  $\text{iter}$  is the number of generations currently in use. In the preceding technique, the particle velocity is regulated by a maximum value  $v_{\max}$ . The option  $v_{\max}$  specifies the resolution, or fitness, with which regions between the current and target coordinates should be searched. This constraint promotes local problem space exploration and correctly reflects human learning's modest modifications. If  $v_{\max}$  is too high, particles may fly through good solutions.

## 2.4 Hybrid Particle Swarm Optimization

This employs PSO mechanism as well as a natural selection mechanism, the genetic algorithm (GA). Because PSO's search procedure is heavily reliant on  $p_{\text{best}}$  and  $g_{\text{best}}$ ,  $p_{\text{best}}$  and  $g_{\text{best}}$  restrict the search area. PSO adjusts the current searching points iteratively [9]. HPSO, on the other hand, can use the selection mechanism to jump immediately from the present searching spots to the correct location. Agent positions with low evaluation values are replaced by ones with high evaluation values using the selection. It may take some time for the original PSO to penetrate the solution space's current effective area. HPSO, but from other hand, uses selection process that directs low-valued particles to the present effective area, resulting in a more targeted search, particularly in that area.

## 2.5 Multi-agent System

For numerous years, the science of artificial intelligence has researched agent-based computation [10]. A multi-agent system (MAS) is a computer system in which multiple agents communicate or collaborate to achieve a set of objectives. A real or virtual entity with the traits given below is referred to as an agent.

1. Agents exist in a certain environment and act in that environment.
2. Agents are aware of their immediate surroundings and can communicate with supplementary agents available in that particular area.
3. Agents strive to attain certain goals or fulfil specific tasks.



4. Because of their learning abilities, agents can respond to changes in themselves in a timely manner.

Many complicated problems can be computed and optimized using MAS. MAS agents work together as a society, cooperating (or competing) to attain both their own and common goals. This cooperative and competitive characteristics of a particle in PSO relate to its essential nature. As a result, to construct a new optimal algorithm, here we integrate two methods, namely PSO technology and MAS technology. When using MAS for solving problems, the four important criteria have to be clarified under context:

1. Each agent’s relevance and function within MAS technology;
2. The surroundings in which multi-agents dwell;
3. The proper description of the confined surroundings; and
4. Rubrics framed for the behaviour that regulates the communication among multi-agents. This is agent universe’s law.

## 2.6 Multi-agent Particle Swarm Optimization

The proposed MAPSO approach for addressing economic power dispatch is formed by combining MAS and PSO in this study. An agent in MAPSO is candidate solution to the optimization problem. Firstly build a lattice-like setting in which each agent is fixed to a lattice point. Each agent competes and collaborates with their neighbours to find the best solution as rapidly as possible, and through self-learning, they can apply information in finding first-rate ideal result. It can accelerate information flow among agents using the PSO evolution mechanism, and the proposed MAPSO technique can achieve the purpose of optimizing the objective function’s value.

### Agent Purpose

A PSO particle that represents a potential solution to the optimization problem is referred to as an agent in MAPSO. As a result, the agent in the optimization problem has a fitness value.

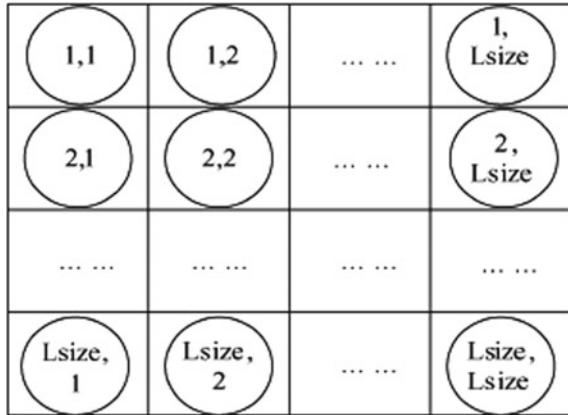
### Definition of an Environment

Each agent can be represented as circle, which will be static on a lattice-point setting as shown in Fig. 1. The information in the circle represents where it is in relation to the rest of the world. Each agent’s database holds two pieces of information: the current velocity of the particle and its location in the search space, according to the way a particle is represented in PSO.  $L$  is  $L_{size} \times L_{size}$ , with  $L_{size}$  being a value. Total magnitude of PSO can be determined by sum of individuals.

### Definition of the Local Environment

The local environment must be defined because each MAS agent can only perceive its immediate surroundings. The following are the neighbours of  $i, j, N_{ij}$ : if the agent at  $(i, j)$  is represented as  $i, j, i, j = 1, 2, \dots, L_{size}$ .  $N_{ij} = \{\alpha_i^1, \alpha_j^1, \alpha_i^2, \alpha_j^2\}$ .

**Fig. 1** Structure of the environment



where

$$\begin{aligned}
 i^1 &= \left\{ \begin{array}{l} i - 1 \quad i \neq 1 \\ L_{size} \quad i = 1 \end{array} \right\}, & j^1 &= \left\{ \begin{array}{l} j - 1 \quad j \neq 1 \\ L_{size} \quad j = 1 \end{array} \right\} \\
 i^2 &= \left\{ \begin{array}{l} i + 1 \quad i \neq L_{size} \\ 1 \quad i = L_{size} \end{array} \right\}, & j^2 &= \left\{ \begin{array}{l} j + 1 \quad j \neq L_{size} \\ 1 \quad j = L_{size} \end{array} \right\}
 \end{aligned} \tag{11}$$

Each agent is surrounded by four neighbours starting at (11). They generate a limited local environment that the agent can only perceive.

**Agent Behavioural Strategies**

Each agent has a set of behaviours that help it achieve its objectives in a timely and accurate manner. Each MAPSO agent contests and collaborates itself with neighbours for disseminating useful data throughout the surroundings and that can take advantage of the PSO’s evolution process and knowledge. Three operators are constructed for the agents based on these behaviours.

**Cooperation and Competition Operator (C&C Operator)**

Assume that the competition and collaboration operators are assigned to the agent precisely on (i, j), and they correspond to  $i, j = (1, 2, \dots, n)$  denotes a genuine vector in the solution space. Agent  $i, j$  is a winner if  $m = \text{Min } N_{i,j} = (m_1, m_2, \dots, M_n)$  is an agent having the lowest fitness value compared with  $N_i$ , neighbours,  $j$ ’s, namely  $m \in N_{i,j}$  and  $N_{i,j}$ ; otherwise, it is a loser.

$$f(\alpha_{i,j}) \leq f(m) \tag{12}$$

if  $\alpha_{i,j}$  wins, it can stay within lattice without repositioning itself in the solution space. Whether it loses, this will expire, and the lattice point will be taken by a new age  $New_{i,j}$ .

## PSO Operator

The proposed MAPSO technique includes the PSO evolution mechanism to achieve the optimal solution rapidly and correctly, based on operators mentioned above. Here PSO is used to rapidly disseminate data across the entire lattice while enhancing evaluation performance. As a result, every unit modifies their position of the particle based on its own history data in addition to the expertise of the best overall agent in the setting, using the best position both this and the best agent have encountered. If a violation occurs, each agent's position is renewed by (8) and (9), respectively, and it is set at the search space's top or bottom limit.

## 2.7 Bacterial Foraging Algorithm (BF)

Engineers can employ animal search and optimal foraging decision-making to solve challenges [11]. Natural selection tends to eliminate species with poor foraging strategies via means of locating, handling and consuming food, while encouraging the spread of genes from animals with strong foraging techniques, according to the bacterium foraging principle. It is inclined to choose the best option depending on the likelihood of reproducing. Poor foraging techniques are either eliminated or moulded into superior ones over several generations. These optimization models have the potential to establish communal foraging surroundings in which many constraints interact to solve engineering challenges like load dispatching. Chemotaxis, swarming, reproduction, and elimination and dispersal are the four mechanisms that drive *Escherichia coli*'s foraging strategy in human intestines.

### Chemotaxis

Swimming and tumbling are used in this process, which is dependent on the rotation of the flagella in each bacterium. Throughout the bacterium's life, it decides whether to move in a predetermined direction (swimming) or in an entirely different direction (tumbling). A unit length random direction, such as ( $j$ ), is created to depict a tumble. After a tumble, this will be utilized to define the direction of travel. A more specific definition is as follows:

$$\theta^i(j + 1, k, l) = \theta^i(j, k, l) + C(i, t)\varphi(j)$$

The position of the  $i$ th bacterium during the  $j$ th chemotactic,  $k$ th reproductive and  $l$ th elimination and dispersal phases is shown by  $i(j, k, l)$ . The magnitude of the step taken in the random direction given by the tumble is determined by  $C(i, t)$ . The "run length unit" is denoted by the letter "C."

### Swarming

Bacterium that discovered optimal food route in a group must seek to lure neighbouring bacteria.

## 2.8 Particle Swirl Algorithm (PSA)

PSA is a brand-new optimization technique for determining a function's lowest and highest values. The particle swirl algorithm (PSA) [12] is a novel method that uses many particles that twirl around a centre. The swirl motion's main point is the vortex, which depends on idea of water spiralling in the direction of ditch. Vortex isn't stationary; it is always moving towards the solution space's highest recorded fitness value. The search is focussed on regions where good results have previously been obtained because the particles are drawn to the vortex. The findings of the particle swirl method were compared to those of a gradient-based optimization strategy and a particle swarm optimization technique on three different functions with ten different peak (goal) locations. For every example, PSA outperformed the other optimization strategies (all three functions and all peak locations).

$$w = w_{\max} - \frac{w_{\max} - w_{\min}}{\text{iter}_{\max}} X_{\text{iter}} \quad (13)$$

$$x_{d+1} = x_d + v_{d+1} \quad (14)$$

## 2.9 Hybrid Multi-Agent-Based Particle Swarm Optimization (HMAPSO)

Novel algorithm called hybrid multi-agent-based particle swarm optimization (HMAPSO) [13] has been developed, which is a hybrid of PSO, GA and MAS. Complex optimization issues have been successfully solved using GA approaches. Recent investigations have found that GA performance has several drawbacks. This efficiency loss is most noticeable in applications with highly epistatic functions or when the parameters being optimized are closely connected. Furthermore, GA's early convergence limits its search capability and decreases its performance, increasing the probability of a local optimum being discovered. The main difficulty in applying GA to engineering problems is its premature convergence, which occurs when the GA degenerates to a local optimum rather than a global one in many cases. Dr. Eberhart and Dr. Kennedy established the concept of PSO so that it can manage huge combinatorial optimization difficulties in power systems to avoid the challenges of premature convergence of GA. PSO has a larger search capacity, near the end of the race, there is a reduction in search capacity. As a result, PSO has more opportunity to discover local optimal value towards concluding the track when addressing issues with more local optima. The particle will migrate away from the global ideal position in PSO if the weight of inertia and the prior velocity are both larger than zero. If all of the particles are near zero, algorithm ends working, resulting in premature convergence, commonly referred as stagnation. To address the stagnation issue, a new

algorithm known as hybrid particle swarm optimization (HPSO) that incorporates the breeding and subpopulation processes of GA in PSO proves that PSO is not the best optimal option. HPSO's strongest feature is that it avoids premature convergence and achieves optimal results in fewer iterations. Complex power system problems can be solved using MAS technology. A multi-agent system (MAS) is a computational system in which multiple agents work together to attain concrete objectives. In MAPSO, the agent is given as a particle to PSO after having competition, cooperation and neighbourhood operators. When compared to traditional optimization methods, PSO and its hybrid algorithms with multi-agent have been shown to converge at better optimal solutions. MAPSO's strong suit is teamwork in a multidimensional search space. Unlike individual particles in PSO, agents work together to achieve a common aim. The advantages of MAPSO and HPSO are combined in HMAPSO. As a result, the suggested method has been demonstrated to have superior properties such as precise and high-quality solutions, reliable convergence and efficient computing. This approach improves global optimization's capacity and efficiency.

### 3 Test System and Discussion

In order to validate the proposed BF approach, it is tested with three test systems having nonlinear characteristics. The bus data and line data of IEEE 14 bus, New England 39 bus system and IEEE 118 bus system are taken from [14]. Based on the actual load flow research utilizing NR, for the IEEE 14 bus first load demand, full input, power losses are as follows:

$$\begin{aligned}
 P_{\text{load}} &= 1.8920 \text{ p.u.}; & Q_{\text{load}} &= 1.0720 \text{ p.u.} \\
 \sum P_G &= 1.9164 \text{ p.u.}; & \sum Q_G &= 1.0041 \text{ p.u.} \\
 P_{\text{loss}} &= 0.0244 \text{ p.u.}; & Q_{\text{loss}} &= 0.0899 \text{ p.u.}
 \end{aligned}$$

Based on the real load flow research utilizing the NR approach. The following are the starting demand, cumulative generations and power dissipation for the New England 39 bus system:

$$\begin{aligned}
 P_{\text{load}} &= 12.5 \text{ p.u.}; & Q_{\text{load}} &= 3.364 \text{ p.u.}; \\
 \sum P_G &= 12.7866 \text{ p.u.}; & \sum Q_G &= 3.2108 \text{ p.u.}; \\
 P_{\text{loss}} &= 0.27864 \text{ p.u.}; & Q_{\text{loss}} &= 1.2167 \text{ p.u.}
 \end{aligned}$$

Based on the real load flow research utilizing the NR approach. The following are the starting demand, cumulative generations and power dissipation for the IEEE 118 bus system

$$P_{\text{load}} = 42.42 \text{ p.u.}; \quad Q_{\text{load}} = 1.439 \text{ p.u.};$$

**Table 1** For IEEE 14 bus, New England 39 Bus/IEEE 118 bus systems, parameter values for GA, HPSO, MAPSO and HMAPSO for optimal cost of production

Parameter	IEEE 14, New England 39 and IEEE 118 bus systems						
	GA	PSO	MAPSO	HMAPSO	HPSO	MAS	BF
Total Particles	6/7/54	6/7/54	6/7/54	6/7/54	6/7/54	6/7/54	6/7/54
Population size	30	20	20	20	20	20	20
No. of iterations	40	250	250	250	250	700	700
$C_1$	–	2	2	2	2	2	2
$C_2$	–	2	2	2	2	2	2
W	–	0.2–0.9	0.2–0.9	0.2–0.9	0.2–0.9	0.2–0.9	0.2–0.9
C&C Operator	–	–	0–1	0–1	–	–	–
Crossover prospect level	0.9	0.05	–	0.05	0.05	–	–
Mutation prospect level	0.004	0.15	–	0.15	0.15	–	–

$$\sum P_G = 43.7486 \text{ p.u.}; \quad \sum Q_G = 7.9568 \text{ p.u.};$$

$$P_{\text{loss}} = 1.32863 \text{ p.u.}; \quad Q_{\text{loss}} = 7.8379 \text{ p.u.}$$

This paper presents the average of 50 trials for the beginning condition, generation cost and execution time. The application was developed in MATPOWER [15, 16] software. Table 1 shows the parameter values for GA, PSO, HPSO, MAPSO and HMAPSO, MAS and BF for the optimal cost of production for IEEE 14 bus, New England 39 Bus/IEEE 118 bus systems.

Solutions obtained by the mentioned approach show its effectiveness plus efficiency of the computation. This algorithm is neither complex nor difficult to implement and solve the nonlinear optimization problems. By this algorithm, better improvement of reduction of cost of generation. The proposed algorithm gives accurate results and converges quickly with less average execution time. The results support the feasibility of the suggested approach, demonstrating its effectiveness and superiority over existing algorithms, as well as its suitability for a wide range of optimization tasks.

The proposed HMAPSO algorithm saves 10.38%, 11.31% and 15.92% for generations in comparison with the Newton–Raphson method, with convergence times of 0.1298 s, 0.2549 s and 0.3252 s for IEEE bus systems such as 14, 39, 118, respectively. Generally, the suggested algorithm HMAPSO (breeding strategy and subpopulation strategy included MAPSO) appears to converge to a superior optimum solution with a lower average execution time.

A novel algorithm HMAPSO is created so that it can acquire the best power system generating cost. According to simulation results shown in Table 2, it is capable of obtaining the lowest cost of power system generation for IEEE 14 bus, New England 39 bus and IEEE 118 bus systems.

**Table 2** Generation cost and computational time using various algorithms for IEEE 14 bus, New England 39 bus/IEEE 118 bus systems

Methods	IEEE 14 bus system		New England bus system		IEEE 118 bus system	
	Cost (Rs/MW)	Time (s)	Cost (Rs/MW)	Time (s)	Cost (Rs/MW)	Time (s)
$\lambda$	854.562	0.3451	45,889	0.4561	250,266	0.5686
GA	853.002	0.3214	46,666	0.4469	250,234	0.4987
EP	858.282	0.2000	47,776	0.4310	257,360	0.6981
PSO	851.752	0.3375	47,606	0.4333	199,851	0.5877
HPSO	864.282	0.1873	43,756	0.3784	198,641	1.8324
MAS	857.202	0.2386	43,987	0.4420	200,566	0.4551
MAPSO	860.062	1.7665	47,618	0.3415	253,217	1.4527
<b>HMAPSO</b>	822.932	<b>0.1598</b>	<b>40,573</b>	<b>0.2849</b>	<b>157,354</b>	<b>0.3652</b>
BFT	852.892	3.4368	45,328	11.632	356,641	0.5784
PSA	854.882	0.2255	47,207	0.4365	269,501	0.6354

Bolded values are the superior values and their respective cost and time are lesser than the other methods results. So in order to show the best result in the comparison table, we bolded those superior values

## 4 Conclusion

This study investigated the use of a novel technique known as hybrid multi-agent particle swarm optimization (HMAPSO) to address economic load dispatch by considering diverse load demand. Simulation observations suggest the projected technology is capable to deliver lower generation costs while also reducing transmission line losses significantly. Furthermore, as compared to the GA, EP, PSO, HPSO, MAPSO, PSA and BF techniques, the execution time for the three test systems under consideration is essentially constant and shorter. As a result, the BF technique looks to be a powerful and efficient tool for addressing highly nonlinear discontinuous cost functions and obtaining a globally improved optimum solution for the ELD issue. We believe that a neuro-fuzzy technique can automate into online decision rules by converting the offline simulation studies.

## References

1. Danila Shirly AR, Sudhilaya M, Priyadarshini Y, Shamni J, Poorani J (2021) Improving efficiency and power loss minimization in landsman DC-DC converter using particle swarm optimization technique (PSO). In: 2021 2nd international conference for emerging technology (INCET). IEEE Xplore, pp 1–6. <https://doi.org/10.1109/INCET51464.2021.9456156>
2. Wang C, Shahidehpour SM (1993) Effects of ramp-rate limits on unit commitment and economic dispatch. IEEE Trans Power Syst 8(3):1341–1349
3. Abido MA (2002) Optimal power flow using particle swarm optimization. Electric Power Energy Syst 24:563–571

4. Chau KW (2007) A split-step particle swarm optimization algorithm in river stage forecasting. *J Hydrol* 346(3–4):131–135
5. Esmim AAA, Lambert-Torres G, Zambroni de souza AC (2005) A hybrid particle swarm optimization applied to loss power minimization. *IEEE Trans Power Syst* 20(2):859–866
6. Zwe-Lee G (2003) Particle swarm optimization to solving the economic dispatch considering the generator constraints. *IEEE Trans Power Syst* 18(3):1187–1195
7. Bakirtzis A, Petridis V, Kazarlis S (1994) A genetic algorithm solution to the economic dispatch problem. *Proc IEE Part C* 141(4):377–382
8. Fogel DB (1992) An analysis of evolutionary programming. In: *Proceedings of the 1st annual conference on evolutionary programming*. Evolutionary Programming Society, pp 43–51
9. Park JB, Lee K-S, Shin J-R, Lee KY (2005) A particle swarm optimization for economic dispatch with non-smooth cost functions. *IEEE Trans Power Syst* 20(1):34–42
10. Khamsawang S, Wannakarn P, Jiriwibhakorn S (2010) Hybrid PSO-DE for solving the economic dispatch problem with generator constraints. *IEEE Trans Power Syst* 978-1-4244-5585
11. Davidson EM, McArthur SDJ, McDonald JR, Cumming T, Watt I (2006) Applying multi-agent system technology in practice: automated management and analysis of SCADA and digital fault recorder data. *IEEE Trans Power Syst* 21(2):559–567
12. Tripathy M, Mishra S (2007) Bacteria foraging-based solution to optimize both real power loss and voltage stability limit. *IEEE Trans Power Syst* 22(1):240–248
13. Menser S, Hereford J (2005) A new optimization technique. In: *Southeast conference, 2006. Proceedings of the IEEE*, 31 Mar 2005–2 Apr 2005, pp 250–255
14. Kumar R, Sharma D, Kumar A (2009) A new hybrid multi-agent-based particle swarm optimisation technique. *Int J Bio-Inspired Comput* 1(4):259–269
15. Matthew S (2005) *An introduction to particle swarm optimization*. University of Idaho, Moscow, pp 1–8
16. Zimmerman R, Gan D (1997) *MATPOWER: A Matlab power system simulation package*. Cornell University Press, Ithaca



# Implementation of Binary Particle Swarm Optimization for Image Thresholding using Memristor Crossbar Array



Priyanka B. Ganganaik, Omkar Mukul Gowaikar, V. Jeffry Louis, Rajesh K. Tripathy, Venkateswaran Rajagopalan, B. V. V. S. N. Prabhakar Rao, and Souvik Kundu

## 1 Introduction

Swarm intelligence-based algorithms are increasingly replacing the traditional search algorithms with high-time complexity. These algorithms are modelled after various natural phenomena such as ant colonies, beehives, bird flocks, fish shoals, and animal herds. [1]. Particle swarm optimization (PSO) [2] is considered as one of the most widely used swarm intelligence algorithms, which makes very few approximations about the function which needs to be optimized, and they can search a large search space for an optimal solution. No information is required about the gradient of the objective function; therefore, even non-differentiable functions can be optimized using PSO. PSO has applications in various domains such as multi-objective function optimization [3], geotechnical engineering [4], solar energy systems [5], improving and optimizing neural network performance [6, 7], and image segmentation [8–10].

On the other hand, image thresholding is one of the fundamental methods for image segmentation based on similarity in features within the regions. Amongst various ways of image thresholding for segmentation, the two most successful

---

P. B. Ganganaik and O. M. Gowaikar are both the authors contributed equally to this work.

---

P. B. Ganganaik (✉) · O. M. Gowaikar · V. Jeffry Louis · R. K. Tripathy · V. Rajagopalan · B. V. V. S. N. Prabhakar Rao · S. Kundu (✉)

Department of Electrical & Electronics Engineering, BITS-Pilani Hyderabad Campus, Hyderabad 500078, India

e-mail: [p20170436@hyderabad.bits-pilani.ac.in](mailto:p20170436@hyderabad.bits-pilani.ac.in)

S. Kundu

e-mail: [souvikelt@gmail.com](mailto:souvikelt@gmail.com)

*Present Address:*

S. Kundu

IMEC, Kapeldreef 75, 3001 Heverlee, Belgium

methods are Otsu's method and Kapur's entropy method. Greyscale image pixel intensities are generally saved as 8 (or integral multiples of 8) bit integers. Thus, discrete optimal thresholds are desired. The thresholds are 8-bit integers as pixel intensities range from 0 to 255. As image thresholding involves finding optimal thresholds from a discrete search space of thresholds, binary particle swarm optimization (BPSO) [11], a variant of PSO, is better suited for this application. Hence, finding them by employing BPSO would be more prudent when compared to any other continuous variant of PSO [12].

Although the swarm intelligence-based algorithms are much faster than the conventional search algorithms, their speed can further be improved significantly. In von Neumann architecture-based computing systems, the memory block and the central processing unit block are separate and are connected via a data bus. With advances in technology, there is an increase in the speed of both the central processing unit (CPU) as well as the random access memory (RAM). However, the rate of increase in the speed of CPUs is much larger than the rate of increase in RAM speed. This increasing disparity in the rates of increase of speed of the CPU and RAM is referred to as the "memory wall". The shared bus between the CPU and memory leads to a limited data transfer rate between them. This conundrum is referred to as the von Neumann bottleneck [13]. With the theoretical conceptualization of memristors by Leon Chua in 1971 [14] and subsequent synthesis of a memristive device by Hewlett Packard in 2008, new avenues have opened for non-von Neumann architecture [15].

Memristors have also demonstrated several unique physical properties like low-energy consumption, nanoscale size, sub-nanosecond switching speed, and extended memory [16–19]. Several architectures of memristors, mainly the memristive crossbar arrays, have been used in neuromorphic circuits like synapses in spiking neural networks [20], applications in chaotic theory [21], and also for implementing machine learning algorithms like the principle component analysis [22]. Recently, very few studies have tried implementing swarm intelligence-based algorithms using memristors. A memristor crossbar-based implementation of ant colony optimization [23], another widely used swarm intelligence-based algorithm, was used in [19, 24], for edge detection. A modified variant of particle swarm optimization, chaotic particle swarm optimization (CPSO), has been used to initialize parameters of a neural network [25].

BPSO has proven to be a very effective algorithm, especially for optimizing functions having their domain defined on discrete search spaces. In this work, BPSO for the image multi-thresholding problem using a novel Pt/Cu:ZnO/Nb:STO memristor crossbar array is implemented. To the best of our knowledge, this is the first study pertaining to the implementation of BPSO on memristor crossbar. The Otsu's function and the Kapur's entropy function are considered as the objective functions or the functions left for BPSO to optimize. It is hypothesized that results obtained from the implementation of BPSO on the memristor crossbar are at par very close to the optimal thresholds. Further, this (Otsu and Kapur's entropy with BPSO) is applied for thresholding 4 T2-weighted transaxial brain magnetic resonance imaging (MRI) scans of the widely used online open-access medical image repository by Harvard Medical School [26].

This manuscript is further organized as follows: In Sect. 2, the Otsu and Kapur's entropy functions are elaborated. In Sect. 3, brief description of the BPSO algorithm is given. In Sect. 4, the memristor model and modelling of the crossbar for implementation are described. In Sect. 5, results for multi-thresholding on standard as well as the brain MR images followed by the results with device variations are explained. Conclusions are given in Sect. 6.

## 2 Objective Functions

In swarm intelligence, the functions optimized by the swarm intelligence algorithm (BPSO in this case) are known as objective functions for the image thresholding problem. Our aim is to divide the pixels in an image  $I$  into  $N$  groups of pixels which has similar grey levels. It is done by finding  $N - 1$  optimal intensity thresholds. It is noted here that  $t_0$  corresponds to the minimum possible intensity (0 intensity),  $t_1$  to  $t_{N-1}$  are the  $N - 1$  optimal thresholds, and  $t_N$  is the maximum possible intensity (255 in case of 8-bit greyscale images). The swarm intelligence algorithm assigns all pixels having pixel intensity in the range between threshold  $t_i$  and threshold  $t_{i+1}$  to the grey-level value of  $G_i$  such that  $i$  can take integer values between 0 and  $N$ . It is mathematically represented as follows:

$$G_i = \{f(x, y) \in I | t_i \leq f(x, y) < t_{i+1}\} \quad (1)$$

Here,  $f(x, y)$  is the integer pixel intensity of image  $I$  at position  $(x, y)$ . So, as seen in Eq. (1), all pixels with intensities in the interval  $[t_i, t_{i+1} - 1]$  are associated with group  $G_i$  where  $i$  is an integer between 0 and  $N - 1$ . To find the optimal thresholds  $t_1$  to  $t_{N-1}$ , the Kapur's entropy function [27] and the Otsu's thresholding function [28] are used. These objective functions take the image greyscale histogram and a set of  $N - 1$  thresholds as input. The image greyscale histogram is a discrete graph of the number of pixels versus pixel intensity. The swarm intelligence algorithm uses the objective functions to adjust the thresholds for optimal results.

### 2.1 Kapur's Entropy Function

Kapur's entropy function is a widely used function for thresholding. It is based on discrete entropy, which is the measure of unpredictability or the degree of randomness. In information theory, entropy (often referred to as Shannon's entropy) is mathematically defined as [29, 30]:

$$H(X) = - \sum_{i=0}^n P(x_i) \log(P(x_i)) \quad (2)$$

where  $X$  is a discrete random variable which can take  $n$  discrete values.  $x_i$  corresponds to the  $i$ th value taken by  $X$ .  $P(x_i)$  is the probability of  $X$  assuming the value  $x_i$ . It is known that for a discrete random variable  $X$ , if the entropy of  $X$  is to be maximized, it has to follow a finite uniform probability mass function (PMF), which has the maximum entropy. The swarm intelligence algorithm tries to maximize Kapur’s entropy function for the image thresholding application so that the modified histogram after thresholding is very close to a discrete uniform PMF. Hence, the contrast is maximized for the thresholded image. As it is known that the Kapur’s entropy function is a histogram-based thresholding function, let  $h(j)$  be the number of pixels in image  $I$  with intensity “ $j$ ”. For a particular pixel intensity  $j$ , the probability that a pixel in image  $I$  has intensity  $j$  is

$$p_j = \frac{h(j)}{\sum_{l=L \min}^{L \max} h(l)} \tag{3}$$

where  $L \min$  and  $L \max$  are the minimum and maximum possible intensities in image  $I$ . For greyscale 8-bit images,  $L \min$  is 0, and  $L \max$  is  $255 (2^8 - 1)$ . The expressions for Kapur’s entropy for multi-thresholding from Eqs. (2) and (3) are given below [31]:

$$H_i = - \sum_{j=t_i}^{t_{i+1}-1} \frac{p_j}{\omega_i} \ln \left( \frac{p_j}{\omega_i} \right); \quad \omega_i = \sum_{j=t_i}^{t_{i+1}-1} p_j \tag{4}$$

It is noticed that Eq. (4) is similar to Eq. (2) where  $\frac{p_j}{\omega_i}$  can be thought of as  $P(x_j)$  where  $l$  is an integer between 0 to  $N - 1$  and  $j$  is a particular pixel intensity. The objective function is then just the summation of all the entropies, which is given below [31]:

$$f_{\text{Kapur}}(I, \mathbf{t}) = \sum_{i=0}^{N-1} H_i \tag{5}$$

where  $\mathbf{t} = [t_1, \dots, t_i, \dots, t_{N-1}]$  and  $I$  are the image concerned. Our goal is to find  $\mathbf{t}^*$  using the swarm intelligence algorithm such that  $\mathbf{t}^*$  maximizes the Kapur’s entropy function or in other words,  $f_{\text{Kapur}}(I, \mathbf{t}^*)$  is the maximum value of Kapur’s entropy function.

### 2.2 Otsu’s Function

Otsu’s function is based on the concept of reducing intra-class variance and increasing inter-class variance. This simply means that in a group or grey level, say  $G_i$ , almost

all pixels are homogenous and almost in the same range; however, for two adjacent groups or grey levels  $G_i$  and  $G_{i+1}$ , the pixels are drastically different. The within-class or intra-class variance is defined as

$$\sigma_i = \omega_i(\mu_i - \mu_T)^2 \quad (6)$$

where  $\omega_i$  is the same as in Eq. (6),  $\mu_i$  is the mean level of each class (9) and  $\mu_T$  is the mean intensity of the Image I.

$$\mu_i = \frac{\sum_{j=i}^{i+1-1} j p_j}{\omega_i} \quad (7)$$

The Otsu function is defined as

$$f_{\text{Otsu}}(I, t) = \sum_{i=0}^{N-1} \sigma_i \quad (8)$$

Again, the goal of this work here is to find  $t^*$  such that  $t^*$  maximizes the Otsu's function. It is observed that if brute-force is used, that is, by manually trying out all possible unordered pairs of  $N$  thresholds, the time complexity would be  $O(n^N)$ . Although this is an effective technique for bi-thresholding, that is, when  $N = 1$ . As  $N$  increases, there is a dire need for faster techniques such as swarm intelligence algorithms to choose the  $N$  thresholds.

### 3 Binary Particle Swarm Optimization (BPSO)

Particle swarm optimization is a swarm intelligence algorithm proposed by Kennedy and Eberhart in 1995 [31]. It is modelled after the social behaviour of a flock of birds or a shoal of fish. It is used for optimizing the objective functions. The domain on which the objective function is defined is referred to as the search space. A swarm of agents or particles is initialized as random points in the search space. Each of these particles can be represented by its current position, its velocity, and the personal best position, and the particle has attained till that point of time. The position and the velocity have dimensions equal to that of the dimension of the domain. Each particle makes its next move by updating velocity followed by position based on the best position attained by any of the particles till that point of time or the global best position and its personal best position. The velocity and position are updated as follows:

$$v_i^{t+1} = W v_i^t + c1 r1 . (x_{gb}^t - x_i^t) + c2 r2 . (x_{i,pb}^t - x_i^t) \quad (9)$$

$$\mathbf{x}_i^{t+1} = \mathbf{x}_i^t + \mathbf{v}_i^{t+1} \quad (10)$$

where  $\mathbf{v}_i$ ,  $\mathbf{x}_i$ , and  $\mathbf{x}_{i,pb}$  are the velocity, position, and personal best position reached so far by the “ $i$ ”th particle, respectively.  $\mathbf{x}_{gb}$  is the global best position of the swarm, and  $W$  is a scalar often referred to as the momentum weight. The constant scalars  $c1$  and  $c2$  are referred to as the acceleration coefficients.  $\mathbf{r1}$  and  $\mathbf{r2}$  are two variable vectors of the same dimension as that of the search space. Each element of these vectors is between 0 and 1, which is randomly updated after every iteration. The personal best and the global best are decided depending on the output of a function which inputs the current position vector of a vector and outputs a scalar. This function is referred to as the fitness function. This fitness function is usually the objective function itself, although they can differ.

For discrete-valued search spaces, Kennedy and Eberhart came up with a discrete binary variant of PSO, the binary particle swarm optimization (BPSO) [31]. Here, the positions of the particles are in discrete binary form. The particle velocities, however, are of continuous nature. The following is how the particle velocity and position updating in BPSO take place:

$$\mathbf{v}_i^{t+1} = W\mathbf{v}_i^t + c1\mathbf{r1} \cdot (\mathbf{x}_{gb}^t - \mathbf{x}_i^t) + c2\mathbf{r2} \cdot (\mathbf{x}_{i,pb}^t - \mathbf{x}_i^t) \quad (11)$$

For every dimension of the position and velocity vectors,

$$x_{i,d} = \begin{cases} 1, & \text{if } \text{sigmoid}(v_{i,d}) > r \\ 0, & \text{otherwise} \end{cases} \quad (12)$$

Here,  $r$  is a random number between 0 and 1 which is updated after every iteration. It is noted that in Eqs. (11) and (12), the current and personal best positions of the particles as well as the global best position are all discrete binary vectors, that is, all the elements in these vectors are either 0 or 1.

So, to essentially formulate the problem of  $N$ -level greyscale image thresholding, the search space can be thought of as an  $8 \times N$  dimensional binary discrete vector space. As stated earlier, each threshold is in the range 0–255. To encode the decimal number in binary space, 8 bits for one threshold are needed. So, for  $N$  thresholds, an  $8 \times N$  dimensional vector space is required. So, one particle corresponds to an  $8 \times N$  dimensional vector or  $N$  thresholds. At the termination of the algorithm, the global best position, which corresponds to the  $N$  thresholds, is considered as the optimal solution, and the  $N$  thresholds corresponding to this global best position are the optimal thresholds. In the next section, BPSO using memristor crossbar is implemented with both the Kapur’s entropy function and the Otsu’s function as the objective function and fitness function.

## 4 Memristor Implementation of BPSO

### 4.1 Memristor Model and Crossbar

Memristors are two-terminal non-linear devices whose resistance can be changed by applying a suitable voltage across it or passing a suitable current through it. Due to this property, they have been used for non-volatile memory applications. In our earlier work, a Pt/Cu:ZnO/Nb:STO memristor was fabricated and for the detailed fabrication procedures [32]. The voltage threshold adapted memristor (VTEAM) model [33] was fitted to this device. At any instant of time, a voltage-controlled memristor can be mathematically represented as

$$\frac{dw}{dt} = f(w, v) \quad (13)$$

$$i(t) = G(w, v).v(t) \quad (14)$$

$w$  is the internal state variable of the device.  $v(t)$  and  $i(t)$  are the voltage across the device and the current passing through the device, respectively.  $G$  is the conductance of the device. As per the VTEAM model,  $f(w, v)$  is

$$\frac{dw}{dt} = \begin{cases} k_{\text{off}} \left( \frac{v(t)}{v_{\text{off}}} - 1 \right)^{a_{\text{off}}} f_{\text{off}}(w), & 0 < v_{\text{off}} < v \\ 0, & v_{\text{on}} < v < v_{\text{off}} \\ k_{\text{on}} \left( \frac{v(t)}{v_{\text{on}}} - 1 \right)^{a_{\text{on}}} f_{\text{on}}(w), & v < v_{\text{on}} < 0 \end{cases} \quad (15)$$

Here,  $v_{\text{on}}$  and  $v_{\text{off}}$  are known as the on and off voltages, respectively. The parameters  $k_{\text{off}}$ ,  $k_{\text{on}}$ ,  $a_{\text{off}}$ , and  $a_{\text{on}}$  are constants and these are device dependent.  $f_{\text{off}}$  and  $f_{\text{on}}$  are known as the window functions, and they are used to ensure that  $w$  is bounded such that  $w \in [w_{\text{on}}, w_{\text{off}}]$ . They are rectangular step-functions defined as follows:

$$f_{\text{off}} = \begin{cases} 1 & \text{if } w < 1 \\ 0 & \text{otherwise} \end{cases} \quad (16)$$

$$f_{\text{on}} = \begin{cases} 1 & \text{if } w > 0 \\ 0 & \text{otherwise} \end{cases} \quad (17)$$

For the model described in [33],

$$w = \frac{x}{D} \quad (18)$$

where  $x$  is the length of the doped/polarized region, and  $D$  is the active material thickness. The current-voltage relationship for the model is

$$v(t) = i(t) * \left( R_{on} + (R_{off} - R_{on}) \left( \frac{w - w_{on}}{w_{off} - w_{on}} \right) \right) \tag{19}$$

As  $x$  can take values between 0 and  $D$ ,  $w_{on} = 0$  and  $w_{off} = 1$ . Substituting this in the above equation, it can be deduced that

$$v(t) = i(t) * (R_{on} + (R_{off} - R_{on})w) \tag{20}$$

The memristor can function as a switch; hence, memristor-aided logic (MAGIC) [15] and the implementation of implication logic (IMPLY) have been proposed in [34]. The low-resistance state (LRS) when memristance is  $R_{on}$ , that is, when  $w = 0$ , can be considered logic 1, and the high-resistance state (HRS) when memristance is  $R_{off}$ , that is, when  $w = 1$ , can be considered logic 0. The logic state of a memristor can be changed by applying sufficient voltage across it. To change state from 0 to 1, a negative voltage,  $v_{set}$ , with magnitude greater than  $v_{on}$  is applied across the memristor. Similarly, to change the logic state from 0 to 1 a positive voltage,  $v_{reset}$ , of magnitude greater than  $v_{off}$  is applied across the memristor. Apart from these binary states, the analogue nature of memristors permits multiple states in between the LRS and HRS.

Memristors can also be employed in a structure called the memristor crossbar array. The memristor crossbar array is a structure comprising of vertical and horizontal metallic wires with memristors at the intersection of each vertical and horizontal wire [35]. The memristance of any memristor can be tuned by applying a suitable voltage,  $v_i$ , across at the ends of the vertical and horizontal wires which have the corresponding memristor at their intersection. To ensure that this does not change the memristance of the other memristors in the row or column, a voltage of  $v_i/2$  is applied across these memristors such that  $|v_i/2| < |v_{off}|$  (or  $|v_{on}| < |v_i|$ ). From Eq. (15), this ensures that  $dw/dt = 0$  so the memristance will not change. One of the most important applications of memristor crossbars is the implementation of the matrix dot product, and this paves the way for various neuromorphic architectures using this structure.

For implementing matrix dot product using memristor crossbar, first, the crossbar with memconductance of all memristors proportional to the elements of the first matrix is initialized. Later, voltages with magnitude in the range  $(v_{on}, v_{off})$  and time pulse width proportional to the corresponding element of the other matrix involved for the dot product are applied. As mentioned above, this ensures that their conductance does not change. From Eq. (14), the output currents are obtained and integrated. Mathematically, this is equivalent to the dot product as

$$i = g_{i,j} * v_m \tag{21}$$



$$Q = \int_0^{t_{i,j}} i dt = g_{i,j} * v_m * t_{i,j} \tag{22}$$

where  $g_{i,j}$  is the memconductance of the memristor located at the intersection of the  $i$ th row and  $j$ th column,  $v_m$  is the magnitude of the voltage pulse,  $t_{i,j}$  is the time pulse voltage width applied across the memristor, and  $Q$  is the charge obtained from integrating the current. In Eq. (22), it is observed that the charge  $Q$  is directly proportional to the product  $g_{i,j} * t_{i,j}$ .

### 4.2 Implementation of BPSO using Memristor Crossbar

As matrix multiplications are the key operations in most of the AI-based algorithms, implementing them using the memristor crossbar tackles the von Neumann bottleneck faced by these algorithms. BPSO involves many matrix-vector dot products, and the use of memristor crossbars can improve the speed of the algorithm.

Discrete binary particles can be modelled using a memristor crossbar array, as shown in Fig. 1a. A particle has a discrete binary position corresponding to 8 N

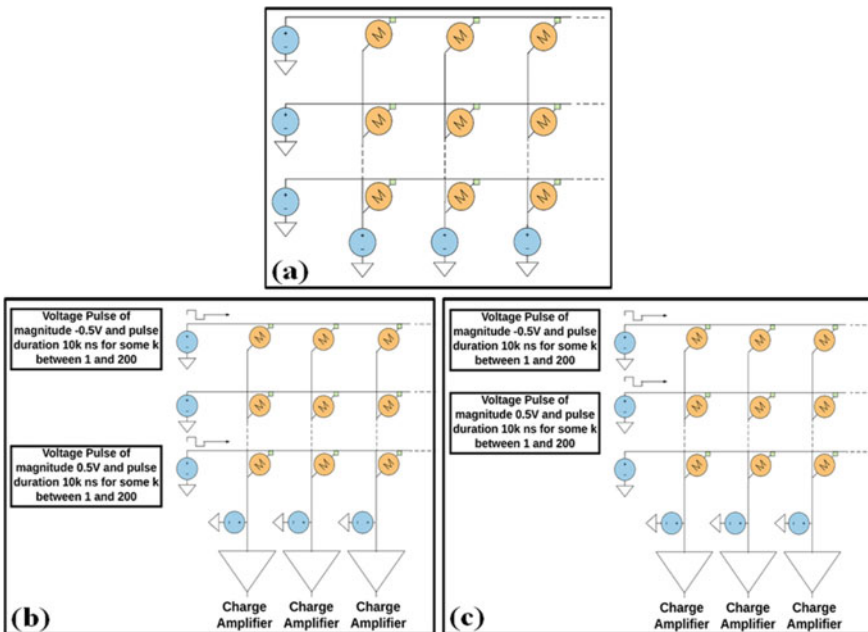


Fig. 1 a Memristor crossbar array, b calculating  $c1r1.(x'_{gb} - x'_i)$  in crossbar, c calculating  $c2r2.(x'_{i.pb} - x'_i)$  in crossbar

bits, where  $N$  corresponds to the number of thresholds. It also has its personal best position, which is also represented using  $8 N$  bits. A design of a memristor crossbar having  $(2N_s + 1)$  rows and  $8 N$  columns is proposed, where  $N_s$  is the number of particles. Here, 2 rows correspond to one particle, which stores the current position and the personal best position of the particle, respectively. The last row is used to store the global best position. Each memristor represents one bit which can be logic state 0 if it is at a high-resistance state or logic state 1 at a low-resistance state. The particle positions are initialized randomly. This is done by applying  $v_{set}$  or  $v_{reset}$  voltages across each memristor individually.

Reading voltages are considered as 0.5 and  $-0.5$  V as both lies between  $v_{on}$  ( $-1.2$  V) and  $v_{off}$  ( $+1.35$  V). From Eq. (15),  $w$  does not change for this voltage. Hence, the memristor behaves as an ideal resistor for these reading voltages. For BPSO, 2 crossbars are simulated for finding 2 thresholds and 3 thresholds. The optimal parameters for BPSO were decided after multiple runs of the algorithm with varied parameters taking into consideration the proximity of the value of the objective function at the output thresholds to the maximum possible value of the objective function. The following Table 1 summarizes the parameters for BPSO and information about the crossbars:

For the memristor model, the read time is 10 ns. As  $c1$  and  $c2$  are 2 each, all the elements of  $c1r1$  and  $c2r2$  will be in the range 0–2. Positions are modelled as conductances, whereas  $c1r1$  and  $c2r2$  are modelled as voltage pulses. A voltage pulse of magnitude equal to 0.5 V (read voltage) and a random time duration of  $10 * k$  ns is applied, where  $k \in [1, 200]$  to the last row or the global best position. This, in turn, ensures that all the elements of  $c1r1$  and  $c2r2$  have a precision of only 2 significant digits after the decimal point, and each element has a least count of 0.01. Ideally, the random numbers should have as much precision as possible, however, here, there is a trade-off as the maximum pulse width increases 10 folds with an increase in the number of significant digits after the decimal point, and this

**Table 1** Parameters for BPSO and information about the crossbars used as suggested in the pseudocode in Sect. 3 and Eq. (11)

Parameter	$N = 2$	$N = 3$
Number of particles ( $N_s$ )	50	100
Number of iterations	50	50
$W$	1.0	1.0
$c1$	2.0	2.0
$c2$	2.0	2.0
Total number of memristors	1616	4824
Dimensions of crossbar	$101 \times 16$	$201 \times 24$
Area of the crossbar	$1.309 \times 10^{-11} \text{ m}^2$	$3.907 \times 10^{-11} \text{ m}^2$

slows down the implementation drastically. Simultaneously, a voltage pulse with magnitude  $-0.5$  V and the same pulse width as the one above to the row storing the current position of the particle is applied. Further, a charge amplifier is used which will give a voltage output proportional to the charge or integral of the current at the input of the amplifier. This voltage vector, which is the output of the charge amplifiers, corresponds to  $c1r1 \cdot (x_{gb}^t - x_i^t)$  in Eq. (11). This is depicted in Fig. 1b. Here, all the random elements of  $\mathbf{r1}$  have been considered equal, which are generated newly for different particles every iteration. Now, a voltage pulse of  $0.5$  V magnitude and pulse width is chosen randomly between  $10$  ns and  $2$   $\mu$ s is sent to the second row (the row storing the personal best position) and simultaneously another pulse of the same width but magnitude equal to  $-0.5$  V to the first row (the row storing the current position). Here, the output of the charge amplifiers corresponds to  $c2r2 \cdot (x_{i,pb}^t - x_i^t)$  in Eq. (11) and again, all elements of  $\mathbf{r2}$  are assumed to be equal, which is shown in Fig. 1(c).

Thus, for the  $j$ th dimension of the  $i$ th particle, from Eq. (22):

$$Q_j = v_{\text{read}} * k * t_{\text{read}} * (g_{gb,j} - g_{i,j}) \quad (23)$$

$$Q_j = v_{\text{read}} * k * t_{\text{read}} * (g_{pb_{i,j}} - g_{i,j}) \quad (24)$$

Here,  $Q_j$  is the output voltage at the  $j$ th column/dimension.  $g_{gb,j}$  is the  $j$ th dimension of the global position, and  $g_{pb_{i,j}}$  and  $g_{i,j}$  are the  $j$ th component of the personal best and current best positions of the  $i$ th particle, respectively. It is noted here that as  $R_{\text{on}}$  (in the order of  $105$   $\Omega$ ) is much lesser than  $R_{\text{off}}$  (in the order of  $108$   $\Omega$ ), conductance corresponding to the logic 1 state ( $g_{\text{on}} = 1/R_{\text{on}}$ ) is in the order of  $1$   $\mu\Omega^{-1}$ , and the conductance corresponding to the logic 0 states ( $g_{\text{off}} = 1/R_{\text{off}}$ ) is in the order of  $1$   $\text{n}\Omega^{-1}$ ; thus, the following approximation is used:

$$|g_{\text{on}} - g_{\text{off}}| \approx |g_{\text{on}}| \quad (25)$$

Further, the net velocity is calculated from Eq. (11). This part is done using an 8 N bit adder circuitry external to the crossbar. The position is updated as shown in Eq. (12) by applying  $V_{\text{set}}/V_{\text{reset}}$  pulses for sufficient duration of time. This is followed by updating the personal best and global best positions by checking the fitness value. As the proposed architecture is a coprocessor, the crossbar is strapped onto the CPU using the PCI bus. It is noted here that the fitness values for different positions are calculated in the CPU (outside the crossbar). The flowchart given in Fig. 2 summarizes the entire process.

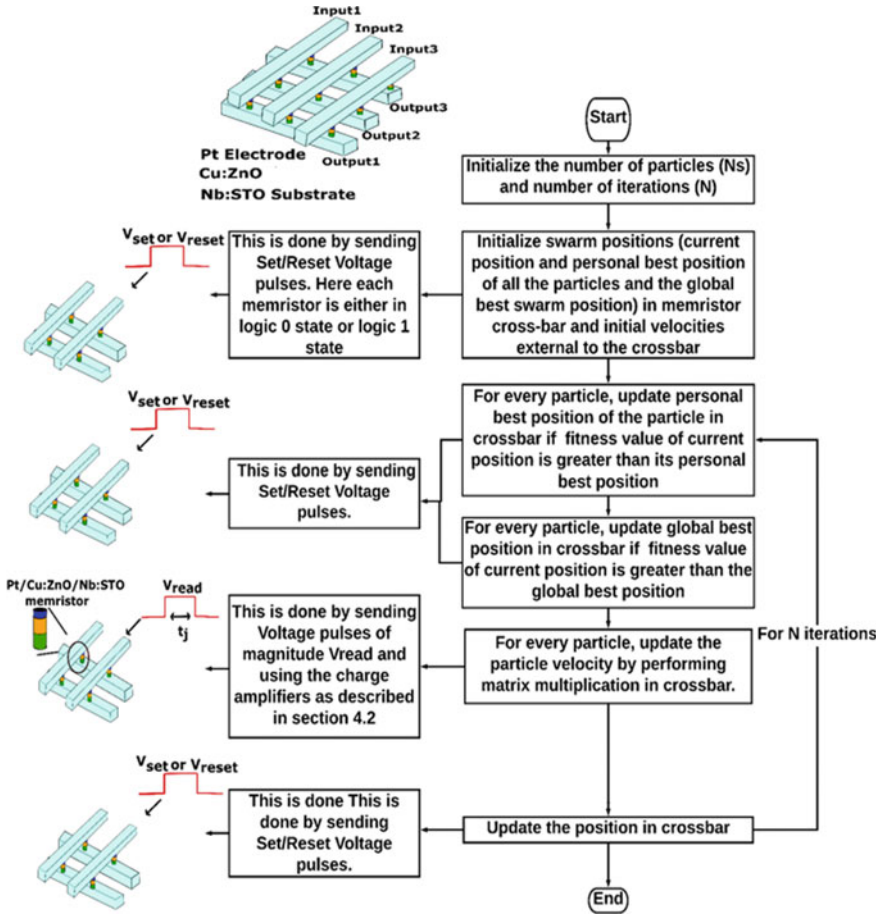
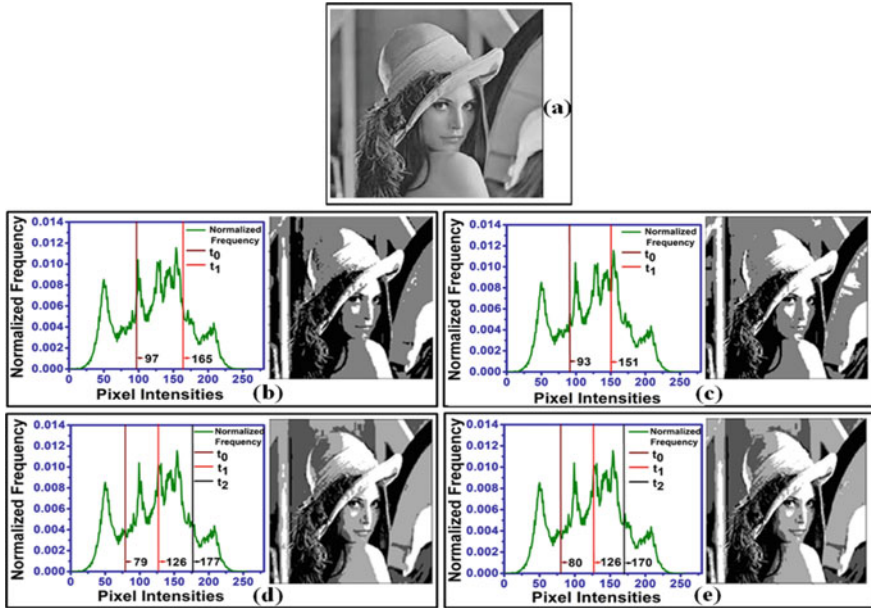


Fig. 2 Flowchart summarizing the entire procedure

## 5 Results

### 5.1 Results and Explanation for Standard Images

As mentioned in Sect. 4.2, two memristor crossbar arrays were simulated for 3 grey-level thresholding and 4 grey-level thresholding, respectively. The histogram of the normalized number of pixels (normalized frequency) vs pixel intensity was obtained from greyscale images. In this study, the results are explained using the ubiquitous “Lena” image (225 × 225) (Fig. 3a) and 4 T2-weighted transaxial brain MRI scans, which were randomly selected from the online open-access medical image repository from Harvard Medical School [26].



**Fig. 3** Brute-force optimization. **a** Greyscale Lena image, **b** histogram and image for 2-thresholds whilst optimizing Kapur’s entropy function, **c** histogram and image for 2-thresholds whilst optimizing Otsu’s function, **d** histogram and image for 3-thresholds whilst optimizing Kapur’s entropy function, **e** histogram and image for 3-thresholds whilst optimizing Otsu’s function

Brute-force optimization is used on the Kapur and Otsu functions, considering unordered pairs of 2 and 3 intensity thresholds ranging from 0 to 255. The results for the same are shown in Fig. 3b–e. For the 2 threshold-based thresholding, a  $101 \times 16$  crossbar is initialized with  $V_{set}$  and  $V_{reset}$  pulses as suggested in Fig. 2. Fifty particles or 50 of the 65,280 possible domain points were arbitrarily considered where the number of iterations was 50. Kapur’s entropy function was considered as the objective function as well as the fitness function. At the termination of the algorithm, the thresholds corresponding to the global best position were considered as the optimal thresholds. The same was repeated considering the Otsu’s function as both the objective function and the fitness function. For 3 threshold-based thresholding, another crossbar was simulated with dimensions  $201 \times 24$ . In this case, the number of particles was taken as 100 or 100 of the 16,581,120 possible domain points were arbitrarily considered. The number of iterations was again taken as 50. A similar process was repeated, taking Kapur’s entropy function and Otsu’s function as both the objective and fitness functions. The results for the same have been depicted in Fig. 4a–d.

To consider device variations, it is assumed that the high-impedance ( $R_{off}$ ) and low-impedance states ( $R_{on}$ ) follow a Gaussian distribution with mean as the designated values and standard deviation equal to 5% of the mean value. For each simulated

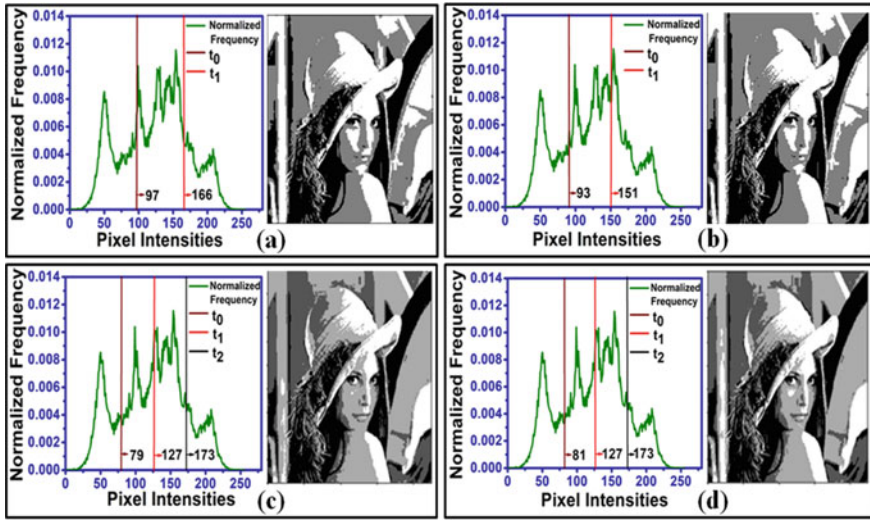


Fig. 4 BPSO implementation using memristor crossbar. **a** Histogram and image for 2-thresholds whilst optimizing Kapur’s entropy function, **b** histogram and image for 2-thresholds whilst optimizing Otsu’s function, **c** histogram and image for 3-thresholds whilst optimizing Kapur’s entropy function, **d** histogram and image for 3-thresholds whilst optimizing Otsu’s function

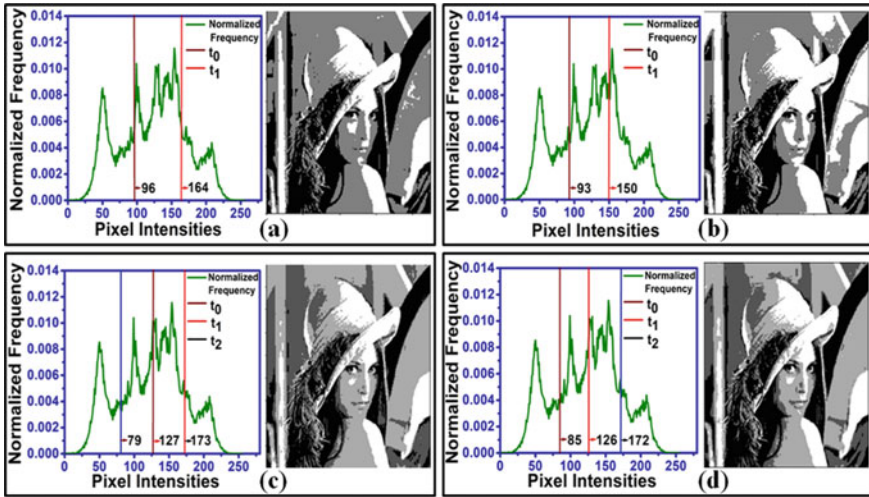
memristor in both the crossbars,  $R_{on}$  and  $R_{off}$  were taken randomly from the corresponding distributions, and the entire process was repeated for both the crossbars. The results for the same are shown in Fig. 5a–d:

From Table 2, it is observed that the memristor crossbar-based BPSO is very efficacious as the maximum error obtained was only 0.1413% when compared to the maximum value of the Kapur and Otsu functions.

### 5.2 Biomedical Image Processing using the Proposed Methodology.

Further, one of the conventional brain MR images is thresholded, which are the essential source of detecting any abnormalities in the brain tissues. Multi-thresholding these images help enhance them by delineating the cerebrospinal fluid (CSF), white matter, and grey matter. This helps unveil some intricacies in the image, thus aiding the diagnosis and leading to better treatment. In this section, the results for 4 randomly selected T2-weighted transaxial brain slices each of size  $256 \times 256$  pixels are discussed, which are shown in Fig. 6a–d.

Also, 4 grey-level thresholding (3 thresholds) for the above 4 images is applied. As described in Sect. 5.1, manual thresholding to see the results for the optimal thresholds maximizing the Kapur’s entropy function and the Otsu’s function is applied,



**Fig. 5** BPSO crossbar implementation with device variations. **a** Histogram and image for 2-thresholds whilst optimizing Kapur’s entropy function, **b** histogram and image for 2-thresholds whilst optimizing Otsu’s function, **c** histogram and image for 3-thresholds whilst optimizing Kapur’s entropy function, **d** histogram and image for 3-thresholds whilst optimizing Otsu’s function

respectively; results are shown in Fig. 6e–h, i–l. Later, the crossbar simulated for 3 thresholds is considered; first, the Kapur’s entropy function is considered as the objective function (Fig. 7(a)-(d)), followed by the Otsu’s function (Fig. 7e–h). Also, the device variations are considered in the same crossbar, which is best depicted in Fig. 8a–d, e–h.

Tables 3 and 4 recapitulate the quantitative results of the crossbar implementation for brain MR image thresholding. Tables 3 and 4 give a summary of results from Kapur’s entropy function as the objective function and Otsu’s function as the objective function.

From these tables below, it is observed that there is a negligible difference between the optimal values and the values from BPSO implemented using the memristor crossbar. The error is significantly less when the crossbar implementation is compared to the brute-force optimization method. It is further noticed that there is no significant difference when the device variations are considered; thus, the system is robust and unsusceptible to the device variations.

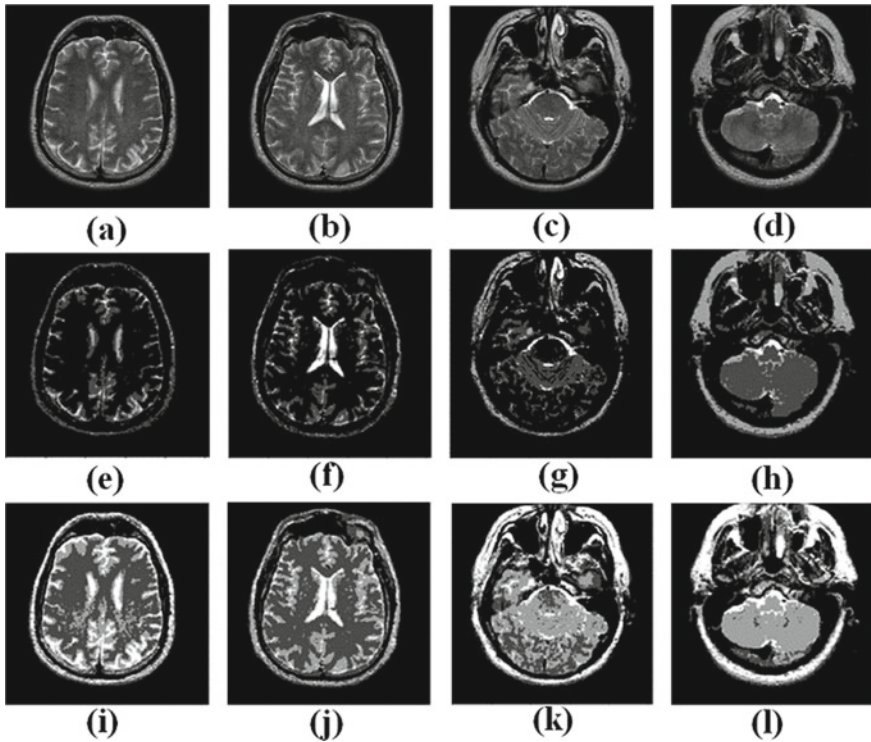
## 6 Conclusion

In this work, a novel memristor crossbar-based implementation of the BPSO algorithm was delineated using the characteristics of the Pt/Cu:ZnO/Nb:STO memristor. Results from memristor crossbar-based implementation agree well with the optimal

**Table 2** Results for the Lena image

Optimizing function	Number of thresholds	The maximum value of the function	The optimum value reached by BPSO	BPSO crossbar implementation with 5% variations	Error for BPSO using the crossbar (%)	Error for BPSO using crossbar considering the device variations (%)	Optimal thresholds	Thresholds from BPSO	Thresholds from crossbar with 5% device variations
Kapur	2	12.426442	12.425018	12.424772	0.011	0.013	97, 165	97, 166	96, 164
Otsu	2	1944.730386	1944.140996	1944.533381	0.030	0.010	93, 151	94, 153	93, 150
Kapur	3	15.422701	15.416080	15.409498	0.043	0.0860	115, 149, 185	115, 149, 183	79, 127, 173
Otsu	3	2108.857182	2108.176481	2105.876604	0.032	0.1413	91, 134, 173	93, 134, 173	85, 126, 172

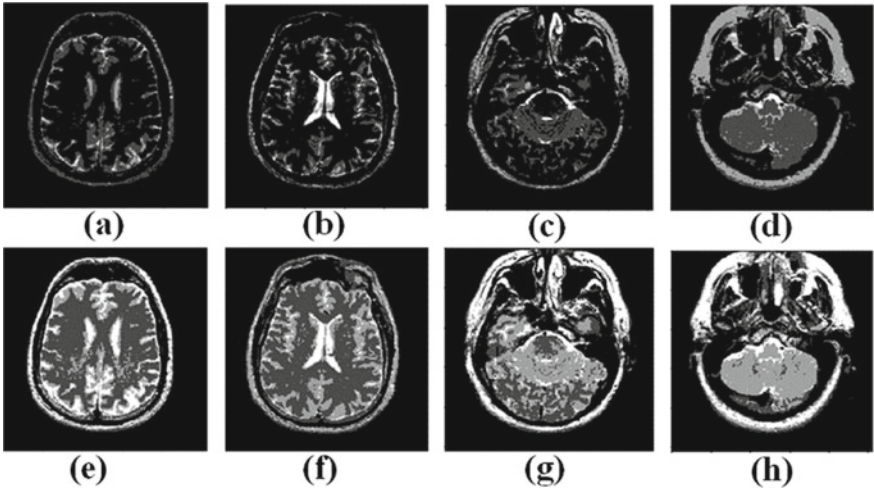




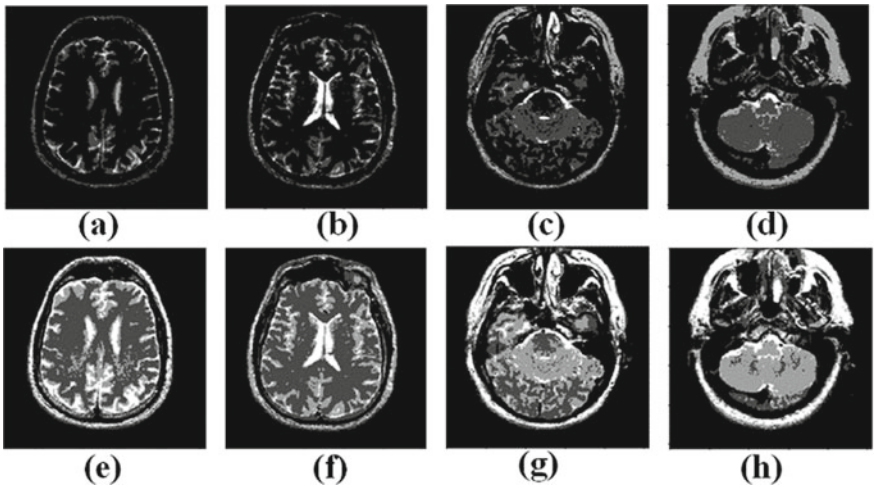
**Fig. 6** a–d Original greyscale brain MR images. e–h show the corresponding 4 grey-level thresholding by optimal values of the Kapur's entropy function for images (a–d). i–l show the corresponding 4 grey-level thresholding by optimal values of the Otsu's function for images (a–d)

values with a maximum error of only 2.5%. The performance of the proposed implementation was validated using a variety of widely used practical greyscale images of different sizes. Our results demonstrate that the memristor crossbar-based implementation of BPSO is a suitable alternative to the conventional CMOS system and lays the foundation for memristive devices to fasten the already fast swarm intelligence algorithms.

Although a lot of work has been proposed for the modelling of neural networks and other machine learning algorithms, very few works on other artificial intelligence-based algorithms have been proposed using memristor circuitry. There is much scope for research in the memristor-based implementation of other AI-based algorithms like genetic algorithms, simulated annealing, and swarm intelligence algorithms. Multi-thresholding for more than 4 greyscale levels can be done at the cost of more memristive devices. The precision of  $c1r1$  and  $c2r2$  can be increased by fabricating memristors with a much lesser read time. Also, other applications of PSO and BPSO like PSO-based neural network performance optimization, optimal filter design, or PSO-based controllers can be designed using memristors. About image processing,



**Fig. 7** a–d show 4 grey-level thresholding by BPSO with Kapur’s entropy function as the objective function. e–h show 4 grey-level thresholding by BPSO with the Otsu’s function as the objective function



**Fig. 8** a–d show 4 grey-level thresholding by BPSO with Kapur’s entropy function as the objective function with 5% device variations. e–h show 4 grey-level thresholding by BPSO with Otsu’s function as the objective function with 5% device variations

various other subtypes of image segmentation like morphological feature extraction can be carried out by changing the objective function. This work can serve as an essential guide to all these future works.

**Table 3** Results for brain images with Kapur's entropy function as the objective function

Figure	Maximum value reached by the Kapur's entropy function	Optimal thresholds (Kapur)	Value reached by crossbar implemented BPSO (Kapur)	Thresholds from crossbar implemented BPSO	Value reached by crossbar implemented BPSO with 5% variations	Thresholds from crossbar implemented BPSO with 5% variations	% Error for crossbar-based BPSO	%Error for crossbar-based BPSO with 5% device variations
Brain image 7 (a)	11.135578	107, 161, 209	11.13170273	112, 162, 209	11.118899	112, 162, 209	0.0348	0.1498
Brain image 7 (b)	11.699239	101, 145, 193	11.40804753	106, 150, 189	11.420763	106, 150, 189	2.4890	2.3803
Brain image 7 (c)	11.515785	97, 143, 187	11.51223486	91, 143, 187	11.501549	91, 143, 187	0.0308	0.1236
Brain image 7 (d)	11.551296	57, 119, 187	11.53687572	61, 119, 188	11.548521	61, 119, 188	0.1248	0.0240

**Table 4** Results for brain images with Otsu's function as the objective function

Figure	Maximum value reached by the Otsu's function	Optimal thresholds (Otsu)	Value reached by crossbar implemented BPSO (Otsu)	Thresholds from crossbar implemented BPSO	Value reached by crossbar implemented BPSO with 5% variations	Thresholds from crossbar implemented BPSO with 5% variations	% Error for crossbar-based BPSO	%Error for crossbar-based BPSO with 5% device variations
Brain image 7 (e)	2796.147836	39, 93, 135	2795.908630	40, 94, 133	2794.171627	44, 94, 131	0.0086	0.0707
Brain image 7 (f)	3041.200405	41, 103, 161	3041.000195	42, 104, 157	3040.677863	41, 103, 166	0.0066	0.0172
Brain image 7 (g)	2954.556065	33, 83, 131	2954.276664	36, 86, 132	2954.276664	36, 86, 131	0.0095	0.0095
Brain image 7 (h)	2824.621006	23, 71, 125	2823.774072	26, 76, 125	2824.424199	23, 72, 123	0.02998	0.0070

## References

1. Beni G, Wang J (1993) Swarm intelligence in cellular robotic systems. *Rob Biol Syst: Towards New Bionics?* (2):703–712. [https://doi.org/10.1007/978-3-642-58069-7\\_38](https://doi.org/10.1007/978-3-642-58069-7_38)
2. Kennedy J, RE (1995) Particle swarm optimization. In: *Proceedings of ICNN'95-international conference on neural networks*, vol 4, pp 1942–1948
3. Mason K, Duggan J, Howley E (2017) Multi-objective dynamic economic emission dispatch using particle swarm optimisation variants. *Neurocomputing* 270:188–197. <https://doi.org/10.1016/j.neucom.2017.03.086>
4. Hajihassani M, Jahed Armaghani D, Kalatehjari R (2018) Applications of particle swarm optimization in geotechnical engineering: a comprehensive review. *Geotech Geol Eng* 36(2):705–722. <https://doi.org/10.1007/s10706-017-0356-z>
5. Elsheikh AH, Abd Elaziz M (2019) Review on applications of particle swarm optimization in solar energy systems. *Int J Environ Sci Technol* 16(2):1159–1170. <https://doi.org/10.1007/s13762-018-1970-x>
6. Han H, Member S, Lu W, Hou Y, Qiao J (2018) An adaptive-PSO-based self-organizing. *IEEE Trans Neural Netw Learn Syst* 29(1):104–117
7. Melo H, Watada J (2016) Gaussian-PSO with fuzzy reasoning based on structural learning for training a neural network. *Neurocomputing* 172:405–412. <https://doi.org/10.1016/j.neucom.2015.03.104>
8. Borjigin S, Sahoo PK (2019) Color image segmentation based on multi-level Tsallis–Havrda–Charvát entropy and 2D histogram using PSO algorithms. *Pattern Recogn* 92:107–118. <https://doi.org/10.1016/j.patcog.2019.03.011>
9. Sivaranjani R, Roomi SMM, Senthilarasi M (2019) Speckle noise removal in SAR images using multi-objective PSO (MOPSO) algorithm. *Appl Soft Comput J* 76:671–681. <https://doi.org/10.1016/j.asoc.2018.12.030>
10. Pham TX, Siarry P, Oulhadj H (2018) Integrating fuzzy entropy clustering with an improved PSO for MRI brain image segmentation. *Appl Soft Comput J* 65:230–242. <https://doi.org/10.1016/j.asoc.2018.01.003>
11. Kennedy J, Eberhart RC (1997) Discrete binary version of the particle swarm algorithm. *Proc IEEE Int Conf Syst Man Cybernet* 5:4104–4108. <https://doi.org/10.1109/icsmc.1997.637339>
12. Djerou L, Khelil N, Dehimi HE, Batouche M (2009) Automatic multilevel thresholding using binary particle swarm optimization for image segmentation. In: *SoCPaR 2009—soft computing and pattern recognition*, pp 66–71. <https://doi.org/10.1109/SoCPaR.2009.25>
13. Wulf WA, McKee SA (1995) Hitting the memory wall. *ACM SIGARCH Comput Arch News* 23(1):20–24. <https://doi.org/10.1145/216585.216588>
14. Leon O (1971) Memristor—the missing circuit element. *IEEE Trans Circuit Theory* c(5):507–519
15. Kvatinsky S, Belousov D, Liman S, Satat G, Wald N, Friedman EG, Kolodny A, Weiser UC (2014) MAGIC—Memristor-aided logic. *IEEE Trans Circuits Syst II: Express Briefs* 61(11):895–899. <https://doi.org/10.1109/TCSII.2014.2357292>
16. Torrezan AC, Strachan JP, Medeiros-Ribeiro G, Williams RS (2011) Sub-nanosecond switching of a tantalum oxide memristor. *Nanotechnology* 22(48):485203. <https://doi.org/10.1088/0957-4484/22/48/485203>
17. Choi BJ, Torrezan AC, Strachan JP, Kotula PG, Lohn AJ, Marinella MJ, Li Z, Williams RS, Yang JJ (2016) High-speed and low-energy nitride memristors. *Adv Funct Mater* 26(29):5290–5296. <https://doi.org/10.1002/adfm.201600680>
18. Zhou J, Cai F, Wang Q, Chen B, Gaba S, Lu WD (2016) Very low-programming-current RRAM with self-rectifying characteristics. *IEEE Electron Device Lett* 37(4):404–407. <https://doi.org/10.1109/LED.2016.2530942>
19. Yu Y, Deng Q, Ren L, Tashi N (2020) Memristor crossbar array based ACO for image edge detection. *Neural Process Lett* 51(2):1891–1905. <https://doi.org/10.1007/s11063-019-10179-6>
20. Jo SH, Chang T, Ebong I, Bhadviya BB, Mazumder P, Lu W (2010) Nanoscale memristor device as synapse in neuromorphic systems. *Nano Lett* 10(4):1297–1301. <https://doi.org/10.1021/nl904092h>

21. Zhang X, Zhang Y, Zhang Z, Mahadevan S, Adamatzky A, Deng Y (2014) Rapid Physarum algorithm for shortest path problem. *Appl Soft Comput J* 23:19–26. <https://doi.org/10.1016/j.asoc.2014.05.032>
22. Choi S, Sheridan P, Lu WD (2015) Data clustering using memristor networks. *Nature Publishing Group*, pp 1–10. <https://doi.org/10.1038/srep10492>
23. Dorigo M, Di Caro G (1999) Ant colony optimization: a new metaheuristic, evolutionary computation. In: *Proceedings of the 1999 congress on CEC 99*, vol 2, pp 1470–1477
24. Pajouhi Z, Roy K (2018) Image edge detection based on swarm intelligence using memristive networks. *IEEE Trans Comput Aided Des Integr Circuits Syst* 37(9):1774–1787. <https://doi.org/10.1109/TCAD.2017.2775227>
25. Geng Y, Duan S, Dong Z, Wang L (2017) A novel PID neural network controller based on memristor. *Chin Control Conf CCC* 2(4):3988–3993. <https://doi.org/10.23919/ChiCC.2017.8027982>
26. Johnson KA, Becker JA, WL (1999) The whole brain atlas
27. Kapur JN, Sahoo PK, AKCW (1985) A new method for gray-level picture thresholding using the entropy of the histogram. *Comput Vis Graph Image Process* 29:273–285
28. Otsu N (1979) Threshold selection method from gray-level histograms. *IEEE Trans Syst Man Cybern*, SMC-9(1):62–66. <https://doi.org/10.1109/TSMC.1979.4310076>
29. GFS (1923) The Bell system technical journal. *J Franklin Inst* 196(4):519–520. [https://doi.org/10.1016/s0016-0032\(23\)90506-5](https://doi.org/10.1016/s0016-0032(23)90506-5)
30. Shannon CE (1948) A mathematical theory of communication. *Bell Syst Tech J* 27(4):623–656. <https://doi.org/10.1002/j.1538-7305.1948.tb00917.x>
31. Portes de Albuquerque M, Esquef IA, Gesualdi Mello AR, Portes de Albuquerque M (2004) Image thresholding using Tsallis entropy. *Pattern Recogn Lett* 25(9):1059–1065. <https://doi.org/10.1016/j.patrec.2004.03.003>
32. Boppidi PKR, Raj PMP, Challagulla S, Gollu SR, Roy S, Banerjee S, Kundu S (2018) Unveiling the dual role of chemically synthesized copper doped zinc oxide for resistive switching applications. *J Appl Phys* 124(21):214901. <https://doi.org/10.1063/1.5052619>
33. Kvatinsky S, Ramadan M, Friedman EG, Kolodny A (2015) VTEAM: a general model for voltage-controlled memristors. *IEEE Trans Circuits Syst II Express Briefs* 62(8):786–790. <https://doi.org/10.1109/TCSII.2015.2433536>
34. Kvatinsky S, Satat G, Wald N, Friedman EG, Kolodny A, Weiser UC (2014) Memristor-based material implication (IMPLY) logic: design principles and methodologies. *IEEE Trans Very Large Scale Integr (VLSI) Syst* 22(10):2054–2066. <https://doi.org/10.1109/TVLSI.2013.2282132>
35. Hu M, Li H, Chen Y, Wu Q, Rose GS, Linderman RW (2014) Memristor crossbar-based neuromorphic computing system: a case study. *IEEE Trans Neural Netw Learn Syst* 25(10):1864–1878. <https://doi.org/10.1109/TNNLS.2013.2296777>

# Performance Analysis of Solar PV-fed BLDC Motor Under Partial Shading Condition Using Various PSO MPPT Algorithms



Pakki Pavan Kumar, V. Hemant Kumar, and R. N. Patel

## 1 Introduction

With increasing concerns on environment and due to limited conventional fuels, usage of solar PV systems has rapidly gained momentum. Solar PV system is the simplest and convenient one which directly deals with photon energy from the sun which is converted into electrical energy. The infiltration of solar PV systems in the world has considerably improved in the previous decade, extending its application from a tiny roof top or building integrated solar PV systems to a giant solar PV power plant, at many types of installation places, from remote rural areas to domestic areas within large cities. At various installation places, particularly in urban locations, while the burden for reducing usage of land, usually leads to compact installations on the roof top and walls, which introduces shading on to the panels due to adjacent panels which are closed placed, frequently leads to functioning of the solar PV system at non-uniform irradiation conditions due to various obstacles adjacent to the system.

Partial shading results in firmly nonlinear outcome on the power output and the total electrical output of the PV system. Based on the extent and levels of shade, many local maximum power points (MPPs) will be developed, impeding the actual tracking of the global optimum operating point, leading to decrease in performance, formation of hotspot, and rapid worsening of the shaded cells. To thoroughly understand this phenomenon and to alleviate its impact on performance of PV system, an appropriate energy model is to be developed, which is accurate and reliable to predict the output of the system under several shading patterns which the system might come across in

---

P. P. Kumar (✉) · V. Hemant Kumar · R. N. Patel  
Department of Electrical Engineering, National Institute of Technology Raipur, Raipur,  
Chhattisgarh, India  
e-mail: [pavan.kumar270489@gmail.com](mailto:pavan.kumar270489@gmail.com)

R. N. Patel  
e-mail: [mpatel.ee@nitrr.ac.in](mailto:mpatel.ee@nitrr.ac.in)

real-time conditions [1, 2]. This kind of model will be suitable for application where calculation of energy generated, for planning of PV installations, optimization of array topologies and techno economic studies. Additionally, one more application of these models is to evaluate the efficiency of MPPT algorithms, when the considered PV system is subjected to partial shading. When there is a mismatch in the qualities of solar cells, the possibility of occurring hotspots in the PV module will be high. The effect of hotspot can be minimized by using bypass diodes, which is proved to be the best possible method. But, the use of bypass diodes poses a serious consequence, if the maximum heat dissipation by the shaded cells is not considered. With the increase in number of solar cells, the shaded cells dissipate power instead of generating power, and at various shading patterns, maximum point of power dissipation occurs which leads to forming of hotspots [3, 4]. Due to formation of hotspot near the shaded cell, the shaded cell will tend to act as a reverse bias diode and starts consuming the power generated by the remaining or healthy cells of the module and even the total module current gets affected.

The shaded cell not only affects the overall output of the system but also has an adverse effect on temperature which might damage/melt the EVA encapsulant, the TPT back sheet and might even lead to fire [5]. Generally, the number of bypass diodes in a solar PV system is decided considering the number of cells in a module; typically for every 12–24 cells, a bypass diode will be used. The bypass diodes do not consider the maximum heat dissipation by the shaded cell and work effectively in preventing the hotspot only in a PV system with shorter strings but not with lengthy strings.

## 2 Partial Shading and Fast Varying Solar Irradiation in PV Systems

The performance of solar PV systems under zero or no shading condition is good and effective, but when there is any shadow cast on the solar PV panels or the solar irradiation is varying rapidly, the performance of whole PV system gets effected [6]. As the name indicates, partial shading is nothing but, shading of even just few cells of a solar PV panel, which potentially compromise the performance of the entire PV system, and fast varying solar irradiation directly indicates the condition where the solar irradiance upon the panel varies at a rapid rate. Generally, the PV panels will be connected both in serial and parallel combinations for producing the required power output [5]. Solar PV arrays are generally split into strings of solar panels. Shade on one of the panels in the string will reduce the output of overall string virtually to zero until the shadow exists. Figure 1 shows real-time partial shading in solar PV panels [7].

One of the latest technologies of minimizing the impact of partial shading in PV systems is by using bypass diodes which minimizes the impact of partial shading by bypassing current flow through the shaded cell or cells. Solar cells made of amorphous



**Fig. 1** Partial shading on roof top panels



silicon are supposed to be better at handling shading effects than the cells made of crystalline silicon, but the amorphous modules have relatively lower efficiency than that of crystalline modules and hence making the crystalline panels a better choice in spite of its poor performance during partial shading conditions. There are few advanced technologies under development for developing solar PV panels, such as super black and transparent solar cells, which offer higher efficiencies even under extreme weather conditions. But, these technologies are very expensive and even not available for commercial use [8]. There are few widely used methodologies to overcome the problem of partial shading, fast varying solar irradiation conditions and its effects [9]. First one is by using maximum power point tracking (MPPT) at individual string inverter, and the second one is by using micro-inverters and power optimizers at panel level. MPPT is the most widely used method, which continuously keeps track of the panel voltage and current, regulates the voltage in relation to the required load voltage and hence maximizing the power output at the load [10]. Micro-inverters are the latest technology in which every panel will have its own inverter. The output of each micro-inverter is synchronized w.r.t. the grid. Power optimizers are the devices which change the voltage or current magnitudes to match with that of unshaded modules and hence maintaining the maximum power output from the string [6]. Both micro-inverters and power optimizers fundamentally allow each and every panel in the system to operate individually, so that overall output of the system is not inexplicably affected due to one of the panels under shade. The system output and the annual yield will be higher using micro-inverters and power optimizers. But, the major drawback of using micro-inverter/power optimizers is that they are expensive when compared to that of system with central string inverters, when the effect of shading is not an issue [11]. The simplest and cost-effective way to mitigate the impact of partial shading is choosing the optimum site for installing the solar PV

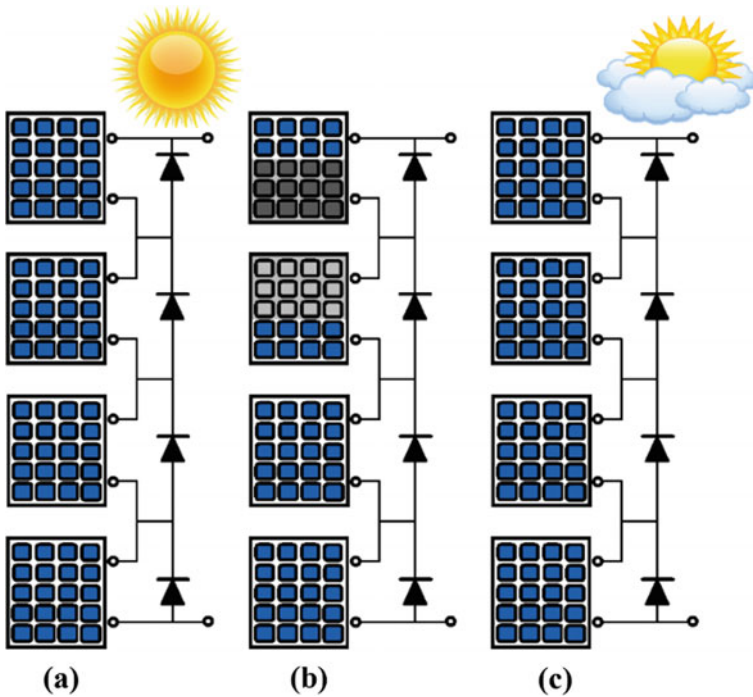


Fig. 2 a Standard test conditions, b Partial shading condition, and c Fast varying solar irradiation

panels where the chance of getting shade over the entire day is practically minimum or ideally zero (Fig. 2).

### 3 Particle Swarm Optimization

Partial shading conditions (PSC) and fast varying solar irradiation (FVSI) are the two most commonly seen effects seen in large solar PV systems. These effects directly affect the characteristic of the solar PV panels by producing more complex and multiple peak characteristics [12]. Consequently, the output of entire PV system gets exaggerated. Generally, the solar PV system gets categorized into groups and subgroups based on the kind of shading effects it is subjected to. Then, the characteristics for each of these groups and subgroups will be developed individually. These individual characteristics are then merged to get global characteristics of the entire solar PV system. The conventional MPPT algorithms like P&O and I.C. MPPT algorithms fail to track the maximum power point from the complex and multiple peak characteristics developed due to these effects [5, 7]; hence, there is a need to implement a better MPPT algorithm which keeps track of multiple peaks and points

out at the global peak. Typical characteristics of solar PV system involving multiple groups with varying partial shading conditions are shown in Figs. 3 and 4.

This optimization algorithm is developed considering the behavior of a flock/swarm of fishes and birds which is in search of food. PSO is the simplest, operative, and search-based meta-heuristic methodology which can solve multi-variable optimization problems involving multiple optimal points [13]. PSO explores the multidimensional search space of the considered fitness function by fine-tuning the paths followed by each agent, basically known as particles, due to the quasi-stochastic style of the paths followed by position vectors. The movement of these particles comprises majorly of two components: One is a probabilistic component, and the other is a predetermined component. Each particle will adjust its trajectory pointing

Fig. 3 P-V characteristics under partial shading

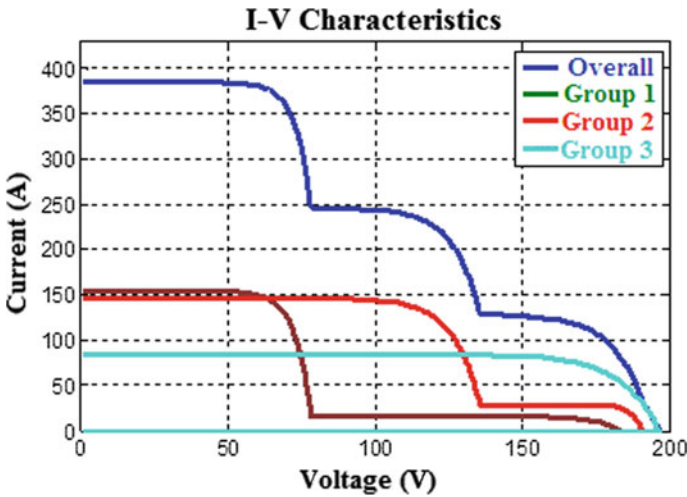
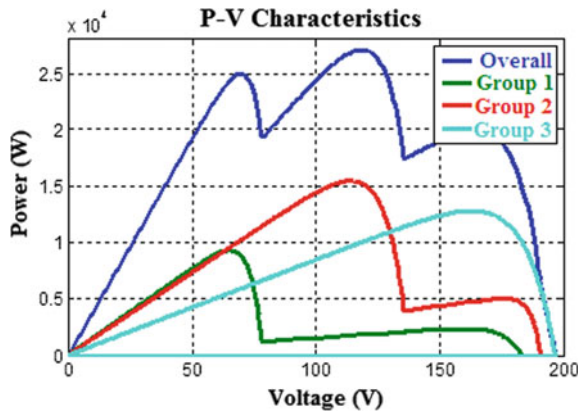


Fig. 4 I-V characteristics under partial shading

at the current best global location  $G_{best}$  and its previous best personal location  $P_{best}$ , and it also has an inclination to travel arbitrarily.

### 3.1 Weighted Particle Swarm Optimization

The PSO uses multiple search agents which are involved in the search process within the search space. These search agents continuously share their information like position and velocity with each other during the search process. The search space is determined based on the decision variables [1, 2]. It considers a set of solutions which is analogous to the particles of a swarm in the considered area of search.

The number of individuals/particles can be decided inherently or randomly within the bounds mentioned. These particles continuously explore for the best solution within the area of search. Within the search space, these particles incessantly update their position relative to the optimal location based upon the respective experience and the information shared by other particles [3, 4].

The particle reaches the optimal location by incrementing or decrementing the velocity relative to the current location. If the value of current location is lesser than optimal location, then the velocity value is increased and decreased when higher [6]. These particles store their personal best location  $P_{best}$  found up to now and also the global best location  $G_{best}$  considering all the particles involved for every iteration. PSO has the capability of carrying out global search by fine-tuning the particle positions [13]. The position of each particle for the subsequent iteration is obtained using the current global best position  $G_{best}$  and the personal best position  $P_{best}$ .

The position ( $x_i$ ) of the particle is adjusted using the expression indicated below.

$$x_i^{n+1} = x_i^n + v_i^{n+1} \tag{1}$$

The velocity required by the particle to reach the optimal location based on their person best and global best locations is indicated below.

$$v_i^{n+1} = v_i^n + \alpha \epsilon_1 (P_{best} - x_i^n) + \beta \epsilon_2 (G_{best} - x_i^n) \tag{2}$$

where  $\alpha$  and  $\beta$  are the acceleration or learning coefficients (mostly equal to two);  $\epsilon_1$  and  $\epsilon_2$  are the random numbers in between 0 and 1;  $v_i^{n+1}$  is the next iteration particle velocity;  $v_i^n$  is the previous velocity of particle;  $x_i^{n+1}$  is the updated particle position;  $x_i^n$  is the previous particle position;  $i = 1, 2, 3, 4, \dots, N$  represents the quantity of particles;  $n$  represents the iteration count. The conventional PSO algorithm can be modified to get various variants which can be used in multiple applications demanding faster convergence, simple algorithm, quality output, etc. The most evident development is perhaps using an inertia function  $\delta(t)$  so that the velocity vector  $v_i^n$  will be replaced by  $\delta v_i^n$ .

$$v_i^{n+1} = \delta v_i^n + \alpha \epsilon_1 (P_{best} - x_i^n) + \beta \epsilon_2 (G_{best} - x_i^n) \tag{3}$$

In most of the cases, the inertia function can be realized as a fixed value, whose value is normally  $\delta \approx (0.5 - 0.9)$ . This is analogous to involving a phantom mass, which alleviates the particle’s trajectory and hence improvising convergence criterion of the algorithm compared to that of conventional PSO algorithm [14].

The following flowchart is considered to obtain the maximum power point of the considered PV system with various shading pattern. To implement this, particle position in the PSO algorithm is declared as the required duty ratio, which basically is required as output from MPPT algorithm. The overall objective of this flowchart is to track the maximum output power of the considered PV array (Fig. 5).

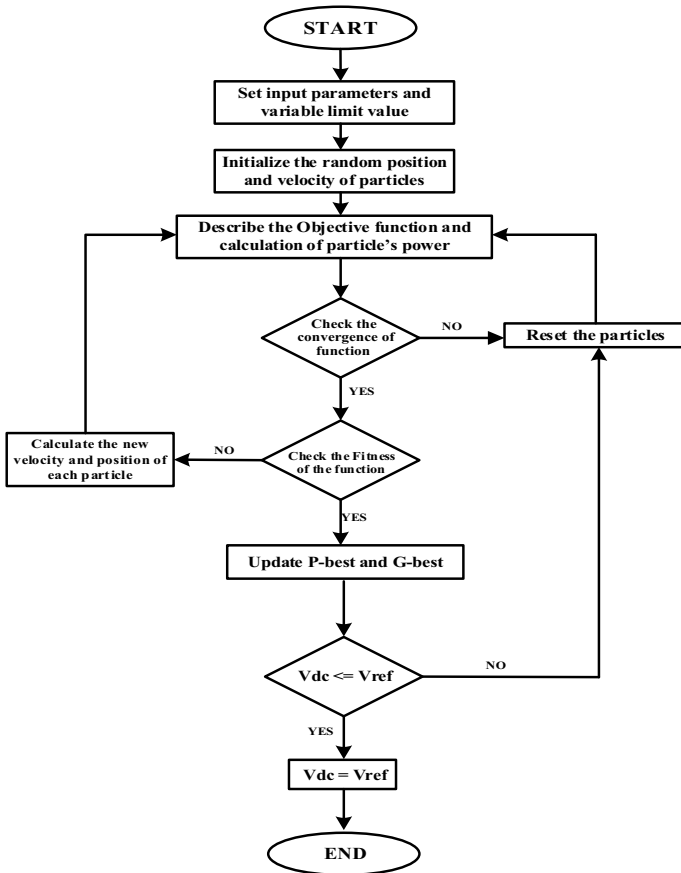


Fig. 5 Flowchart for PSO MPPT

### 3.2 Accelerated PSO Algorithm (APSO)

The classical PSO uses both the personal best  $P_{\text{best}}$  and global best  $G_{\text{best}}$  positions of the considered particles. Accelerated particle swarm optimization (APSO) algorithm is one of the best and easier variants of the PSO algorithm in which only the global best position will be used [15]. Thus, in APSO, the velocity vector is developed by considering the simple formula,

$$v_i^{n+1} = \delta v_i^n + \alpha \epsilon_1 + \beta \epsilon_2 (G_{\text{best}} - x_i^n) \tag{4}$$

The particle’s position will be updated based on Eqs. (4) and (5). For further increasing the algorithm’s convergence criterion, the update of location can be done using a single equation, as shown below, without actually involving the updated velocity vector [16].

$$x_i^{n+1} = \delta v_i^n + \alpha \epsilon + \beta G_{\text{best}} + (1 - \beta)x_i^n \tag{5}$$

Therefore, APSO is the simplest and leads to quick convergence, relative to other variants of PSO, as it consists of two parameters only. A further enhancement to the accelerated PSO can be done by eliminating the random number which introduces the randomness in every iteration. Instead of random number, a monotonically decreasing function as mentioned below can be used.

### 3.3 Constriction Coefficient PSO (CPSO)

In order to ensure an optimum balance between exploration and exploitation, Clerc [17] has used constriction coefficient  $\chi$ . The constriction coefficient ensures faster and stable convergence, reduction in the amplitude of particle oscillations and also avoids the collapse if the right social conditions are in place [10]. The expression for updated velocity vector can be written as

$$v_i^{n+1} = \chi [\delta v_i^n + \alpha \epsilon_1 (P_{\text{best}} - x_i^n) + \beta \epsilon_2 (G_{\text{best}} - x_i^n)] \tag{6}$$

where  $\chi$  can be expressed as

$$\chi = \frac{2k}{|2 - \varphi - \sqrt{\varphi(\varphi - 4)}|}$$

$$\varphi = \varphi_1 + \varphi_2$$

$\varphi_1$  and  $\varphi_2$  can be written as product of social and cognitive acceleration coefficients  $k_1$  and  $k_2$  times random noise  $r_1$  and  $r_2$ . The convergence of the algorithm is assured

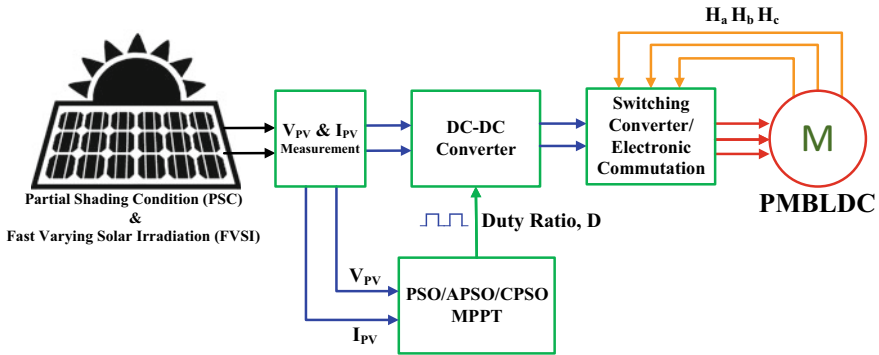


Fig. 6 Proposed topology

with deceleration of particles as the iteration count increases, under the constraints,  $\varphi \geq 4$  and  $0 \leq k \leq 1$ . The value of constant  $k$  decides the balance between exploration and exploitation. If  $k$  is chosen nearer to 1, fast convergence can be observed, and if  $k$  is chosen nearer to 0, slow convergence with a quality output can be observed [18, 19].

### 4 Proposed Topology

In the proposed topology shown in Fig. 6, a DC-DC boost converter is used to boost the panel voltage output to the essential voltage of BLDC motor [20]. The solar PV system is subject to partial shading and fast varying solar irradiation conditions as mentioned earlier which lead to variations in the output voltage magnitude. The DC-DC converter needs to output the required voltage by making use of the variable duty ratio developed by PSO/APSO/CPSO MPPT algorithms, in spite of these conditions. The performance of BLDC motor under partial shading and fast varying solar irradiation conditions using PSO, APSO, and CPSO MPPT algorithms was analyzed in this work.

### 5 Results and Analysis

For the solar PV system described in the earlier section, the output of the overall system under partial shading and fast varying solar irradiation conditions is obtained, from which the performance of the system as well as the BLDC motor drive can be analyzed.

### 5.1 Results—Partial Shading Condition (PSC)

As explained in the previous section regarding partial shading and its effects on the overall solar PV system, the overall output efficiency of the system decreases, due to partial or complete shade on either one or more than one of the panels in the system. When one of the panels in a solar PV system is subjected to partial shading, then the irradiance level w.r.t. that respective panel will be low relative to other panels. And if a panel is under complete shading, then its irradiance should be near to zero. In order to simulate partial shading condition for the system considered in this paper, two panels of the system were subjected to lower irradiance level, indicating partial shading w.r.t. these two panels and the remaining panels was subjected to maximum irradiance level as per the standard testing conditions. One of the panels is subjected to irradiance level of  $500 \text{ W/m}^2$ , the other panel is subjected to irradiance level of  $800 \text{ W/m}^2$ , and all the remaining panels are subjected to irradiance level of  $1000 \text{ W/m}^2$  over the entire simulation time as shown in Fig. 6.

Figures 8 and 9 shown are the voltage output and current output of the boost converter, which are directly used for the BLDC drive. The output voltage obtained using APSO MPPT is much smoother without any ripples, but the steady-state value is  $148.4 \text{ V}$  which is relatively smaller than that of other MPPTs. The classical/weighted PSO has produced steady-state output voltage of  $172.2 \text{ V}$  with low ripples, and the CPSO MPPT has produced a steady-state output voltage of  $169.7 \text{ V}$  with highest ripples among the three MPPTs considered, as it ensures an optimum balance between exploration and exploitation. The output current using APSO MPPT has a smoother current ripple of all other MPPTs but has a lower steady-state value. The classical/weighted PSO has improved steady-state current and moderate distortions in the current ripples.

Figures 10 and 11 shown are the speed and electromagnetic torque output of the BLDC motor drive, which are highly dependent on the output of the boost converter. The rise time of speed response using APSO MPPT is higher than that of other MPPTs, but the steady-state value is relatively low, which is  $1515 \text{ RPM}$ .

The steady-state values of speed response for PSO and CPSO MPPTs are  $1771 \text{ RPM}$  and  $1750 \text{ RPM}$ , respectively, but the speed response of classical/weighted PSO MPPT is much smoother than that of CPSO MPPT. The electromagnetic torque response using APSO MPPT has a smoother torque ripple of all other MPPTs. The classical/weighted PSO MPPT has moderate distortions in the torque ripples, and CPSO MPPT has relatively higher distortions in the torque ripple.

### 5.2 Results—Fast Varying Solar Irradiation (FVSI)

In the above section of results in partial shading condition, the performance of BLDC motor drive under partial shading using PSO, APSO, and CPSO MPPTs is analyzed, but in this section, the performance of BLDC motor drive under fast varying solar



irradiation will be analyzed. Figure 7 shows the variation of solar irradiation to which the solar PV system is subjected to, and hence, the performance is analyzed. Figures 12 and 13 shown are the output voltage and output current from the boost converter under fast varying solar irradiation as shown in Fig. 6, which are directly fed to the BLDC drive. With the decrease in irradiation level in between 2 and 3.5 s of the run time, the output voltage of the boost converter is also changed. The change in output voltage is different for the three variants of PSO MPPTs considered, which can be observed in Fig. 12. As the variations in output voltage of boost converter directly shows impact on the speed response of the BLDC motor drive, and hence, the speed response has changed in between 2 and 3.5 s of the run time. The variations in output voltage with change in irradiation are summarized in Table.1.

Figures 14 and 15 shown are the speed and torque output of the considered BLDCM drive, which are highly dependent on the boost converter output. The variations in speed response with change in irradiation are summarized in Table 2.

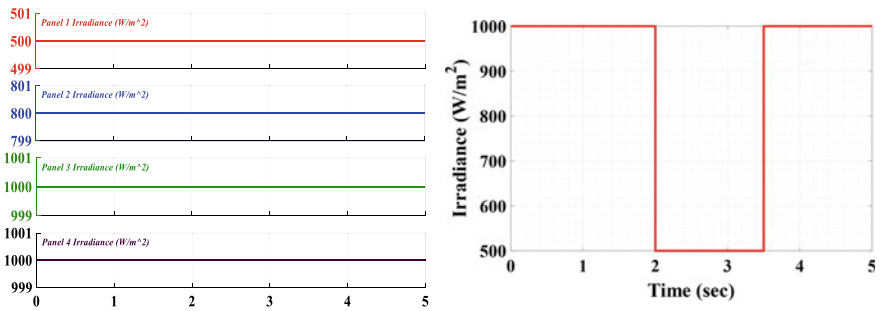
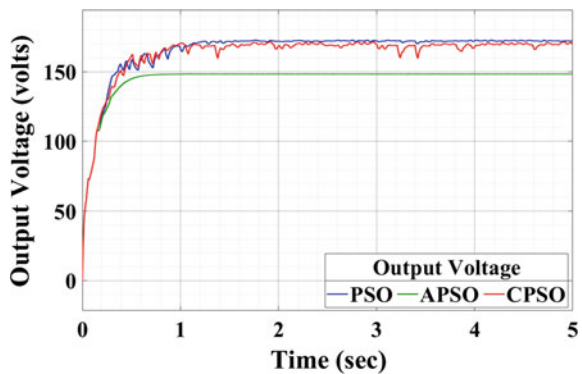
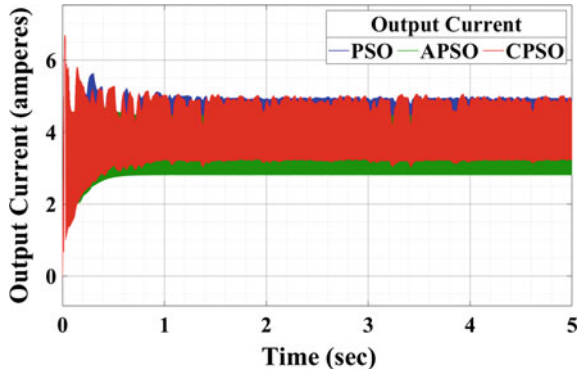


Fig. 7 Solar irradiance under PSC and FVSI conditions

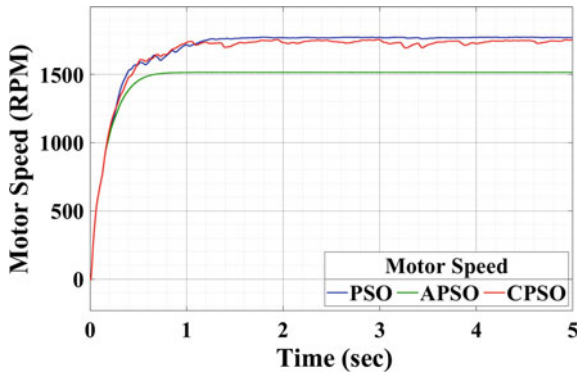
Fig. 8 Output voltage of boost converter-PSC



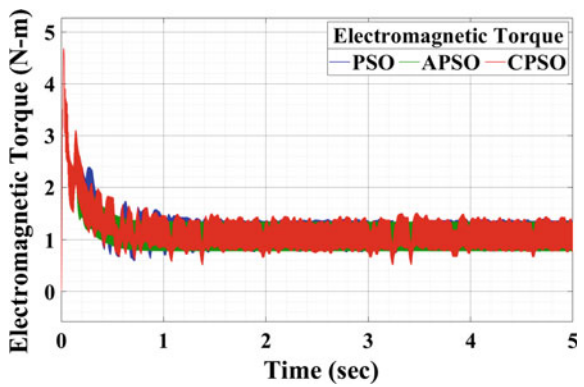
**Fig. 9** Output current of boost converter-PSC



**Fig. 10** Speed response of BLDC motor-PSC



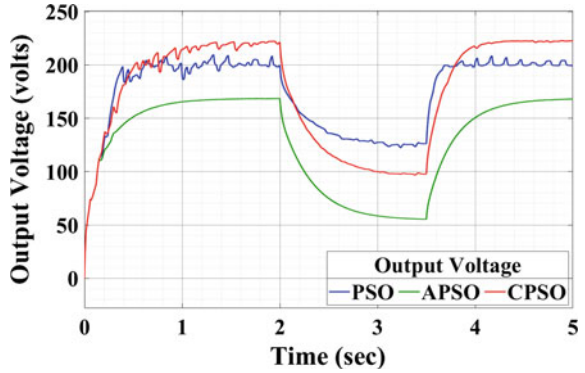
**Fig. 11** Torque response of BLDC motor-PSC



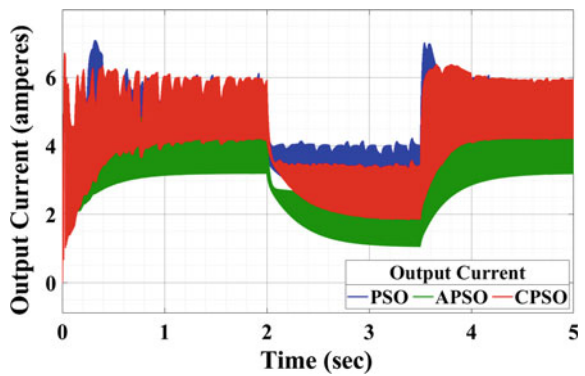
## 6 Conclusion

From the above discussions, it can be concluded that the performance of classical/weighted PSO MPPT is relatively better in the case of partial shading conditions

**Fig. 12** Output voltage of boost converter-FVSI



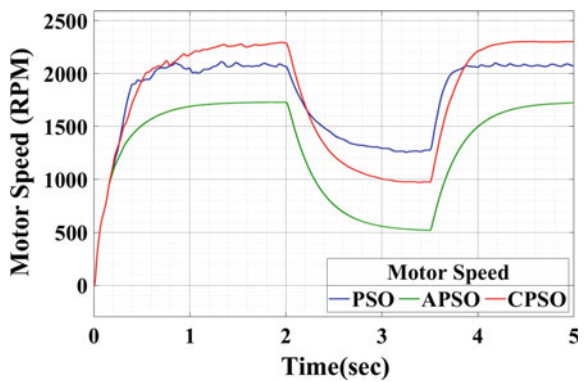
**Fig. 13** Output current of boost converter-FVSI



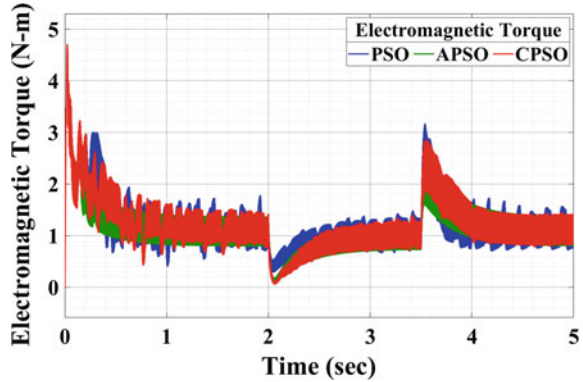
**Table 1** Comparison of voltage response for FVSI

	Voltage at 2 s (V)	Voltage at 3.5 s (V)	Change in voltage (V)	Steady-state value
PSO	199.5	126.4	73.1	199.6 V at 3.88 s
APSO	168.8	55.55	113.25	168.2 V at 5 s
CPSO	221.2	98.01	123.19	222.2 V at 4.36 s

**Fig. 14** Speed response of BLDC motor -FVSI



**Fig. 15** Torque response of BLDC motor -FVSI



**Table 2** Comparison of speed response for FVSI

	Speed at 2 s	Speed at 3.5 s	Change in speed	Steady-state value
PSO	2068 RPM	1278 RPM	790 RPM	2072 RPM at 3.88 s
APSO	1731 RPM	519 RPM	1212 RPM	1724 RPM at 5 s
CPSO	2290 RPM	975 RPM	1315 RPM	2298 RPM at 4.36 s

as its steady-state response is better than the remaining types of PSO MPPT considered. But, the only noticeable drawback is the rise time of response in the case of classical/weighted PSO MPPT which is a bit lower when compared to that of others. The steady-state response of APSO MPPT does not have any ripple, but the steady-state value is very low when compared to that of others. In the response for fast varying solar irradiation, the performance of CPSO MPPT is relatively better than that of the remaining, as the steady-state value is high. But, either the rise time or the time taken to reach that steady-state value is slightly lower than that of PSO MPPT. And even the change in the steady-state values with change in solar irradiation is higher in the case of CPSO MPPT than that of PSO MPPT, which is not acceptable in few of the motor drive applications.

## References

1. de Oliveira FM, da Silva SAO, Durand FR, Sampaio LP, Bacon VD, Campanhol LBG (2016) Grid-tied photovoltaic system based on PSO MPPT technique with active power line conditioning. *IET Power Electron* 1–12
2. Renaudineau H, Donatantonio F, Fontchastagner J, Petrone G, Spagnuolo G, Martin JP, Pierfederici S (2014) PSO-based global MPPT technique for distributed PV power generation. *IEEE Trans Ind Appl*
3. Saptarshi S, Sanchita B, Richard AP (2018) Particle swarm optimization—a survey of historical and recent developments with hybridization perspectives. *Mach Learn Knowl Extract-MDPI*
4. Mostefa K, Berkouk El M (2015) Particle swarm optimization—a survey of historical and recent

- developments with hybridization perspectives. In: 4th International conference on electrical engineering (ICEE)
5. Djaalab A, Rezaoui MM, Merzouk I, Bessous N (2019) Study of the effects of partial shading on PV array. In: International conference on communications and electrical engineering (ICCEE)
  6. Liu G, Zhu J, Tao W, Wang H, Blaabjerg F (2019) MPPT algorithm based on PSO for PV array under partially shaded condition. In: 22nd International conference on electrical machines and systems (ICEMS)
  7. Amardeep C, Shriya G, Dhriti P, Mahfooz F, Varshney G (2015) Effect of partial shading on characteristics of PV panel using Simscape. *Int J Eng Res Appl* 5(10):85–89
  8. Kermadi M, Berkouk EM (2015) A maximum power point tracker based on particle swarm optimization for PV-battery energy system under partial shading conditions. In: 3rd International conference on control, engineering & information technology (CEIT), pp 1–6
  9. Suryavanshi R, Joshi DR, Jangamshetti SH (2012) PSO and P&O based MPPT technique for SPV panel under varying atmospheric conditions. In: International conference on power, signals, controls and computation, pp 1–6
  10. Imran M, Hashim R, Khalid NEA (2013) An overview of particle swarm optimization variants. *Procedia Eng* 53:491–496. ISSN: 1877-7058
  11. Subha R, Himavathi S (2016) Simplified accelerated particle swarm optimization algorithm for efficient maximum power point tracking in partially shaded photovoltaic systems. *IET Renew Power Gener* 10(9):1340–1347
  12. Diqing Z, Wu C, Li Z (2012) Modeling and simulation of partial shaded PV modules. In: International computer science conference ICSC 2012. Communications in computer and information science book series, vol 327. Springer, Berlin, pp 124–134
  13. Liu Y, Xia D, He Z (2011) MPPT of a PV system based on the particle swarm optimization. In: 4th International conference on electric utility deregulation and restructuring and power technologies (DRPT), pp 1094–1096
  14. Cheng Z, Zhou H, Yang H (2010) Research on MPPT control of PV system based on PSO algorithm. In: Chinese control and decision conference, pp 887–892
  15. Subha R, Himavathi S (2016) Accelerated particle swarm optimization algorithm for maximum power point tracking in partially shaded PV systems. In: IEEE international conference on electrical energy systems (ICEES)
  16. Yang X-S, Suash D, Simon F (2012) Accelerated particle swarm optimization and support vector machine for business optimization and applications. In: Conference paper in communications in computer and information science, pp 53–66
  17. Clerc M, Kennedy J (2002) The particle swarm—explosion, stability and convergence in a multidimensional complex space. *IEEE Trans Evol Comput* 6:58–73
  18. Pranava G, Prasad PV (2013) Constriction coefficient particle swarm optimization for economic load dispatch with valve point loading effects. In: International conference on power, energy and control (ICPEC), pp 350–354
  19. Tripathi JP, Ghoshal S (2017) Combining inertia and constriction technique in the PSO applied to fault identification in a hydraulic system. *Proc Inst Mech Eng C J Mech Eng Sci* 231(14):2730–2740
  20. Kumar R, Singh B (2017) Single stage solar PV fed brushless DC motor driven water pump. *IEEE J Emerg Sel Top Power Electron* 5(3):1377–1385

# Design and Analysis of Fuzzy-Based Hybrid PV-Wind Power Quality Improvement for Local Nonlinear Loads Using MLMS



Pasala Gopi, M. Padma Lalitha, and S. Ayisha Jabeen Bhanu

## 1 Introduction

As a significant power supplier for fossil fuels, everyday energy demand is increasingly expanding as a result of population breakouts. In most cases, fossil fuel is utilized after power is created. The growth in fossil fuel energy capacity increases the depletion of fuel supplies renewable energy sources that do not deteriorate after electricity is the only option to meet future needs for energy. On the other side, fossil fuel releases animal gases into the environment and pollutes the biosphere when utilized to generate electricity. As a consequence, renewable energy sources have the primary objective to provide energy to sustain our environment and the nation's long-term growth. The energy supply to a rural area that does not link conventional loads is via an independent distributed system. The battery guarantees that the provided power is dependable because there is no-load connection. The battery stores renewable electricity and supplies it to the user whenever required. With the integration of various renewable energy sources, you may minimize your dependence on batteries and increase your total capacity without expanding the power storage size.

The current study investigates innovative technology that combines solar and wind power generation in a free-standing distributed producing system, providing consumers with optimum dependability. The effectiveness of the fuzzy logic stand-alone hybrid PV-wind MPPT technique is demonstrated in this study. Thus, the fuzzy logic control is available as an analytical model which may be calculated with fractional parameters, decreasing testing time. The results of the simulation are shown to confirm fuzzy logic's PV-wind control capabilities. The structure of the paper is as follows: The second portion gives an overview of the system configuration. The third portion of the study provides an overview of PV system modeling. The

---

P. Gopi (✉) · M. Padma Lalitha · S. Ayisha Jabeen Bhanu  
Department of EEE, Annamacharya Institute of Technology and Sciences, Rajampeta, India  
e-mail: [pgp@aitsrajampet.ac.in](mailto:pgp@aitsrajampet.ac.in)

fourth part of the research provides an overview of wind turbine modeling. Section V explores the fuzzy logic control design. Sections VI and VII provide the results of the MATLAB-based simulation.

## 2 Modeling and Control of the Proposed Wind Photo-Voltaic Cogeneration System

As illustrated in Fig. 1, the proposed system includes the VRS to the interface of the wind generator and the VSI to link the cogeneration system to the grid. A DC wire connects the PV generator to the DC-link condenser [1]. VSR and VSI are two-tier converters, each with a parallel gate polar diode consisting of six cells (IGBT). The next portions include comprehensive modeling and control of the proposed system.

### 2.1 MPPT and P&O Method

A comparison is provided under different wind conditions of multiple control techniques for the removal of maximum power from a wind conversion system (WECS) utilizing a permanent synchronous generator (PMSG). WECS consists of a wind turbine, a PMSG, and a DC/DC transducer connected to a DC load. Maximum power point tracking (MPPT) control system compares the PI, P&O, and fuzzy logic controller (FLC). DC output and load power are the measurements for MPPT control efficiency. Controllers exhibit overall efficiency using steady-state voltage and dynamic system reactivity at varying wind speeds.

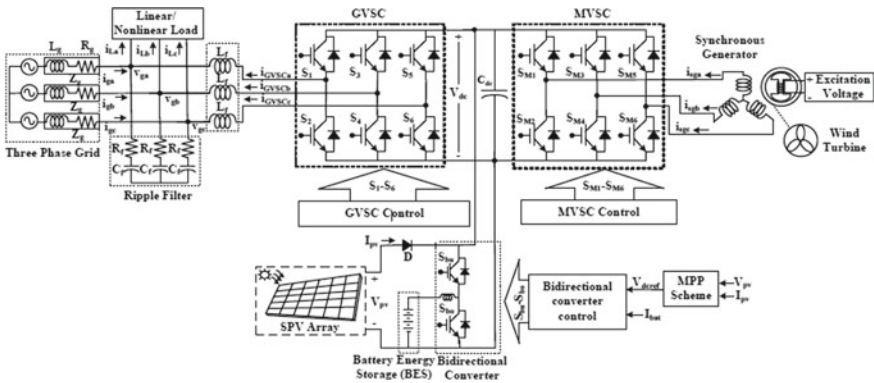


Fig. 1 Proposed wind-PV cogeneration system

### 2.2 Wind Generator

A full wind turbine (FSWT) comprising a permanent magnet synchronous generator (PMSG) is utilized for simple maintenance and low running costs [2]. The following is a model for a wind turbine:

$$P_m = \frac{1}{2} C_p(\beta, \lambda) \rho \pi R^2 v_{wind}^3, \lambda = \frac{R\omega_r}{v_{wind}} \tag{1}$$

Where  $P_m$  is the power developed and comprises rotor coefficient of the blade angle ( $\beta$ ) as nonlinear function,  $R$  is the range of the turbine blade,  $V_{wind}$  is the wind speed, and  $\omega_r$  is the speed of the rotor. In this study, the wind output is set to zero under the normal operating conditions [3].

$$\tilde{v}_s = R_s \tilde{i}_s + L_s \frac{d\tilde{i}_s}{dt} + jp\omega_r (\psi + L_s \tilde{i}_s) \tag{2}$$

$$J \frac{d}{dt} \omega_r + \beta \omega_r = \frac{3}{2} p \psi I_{sq} - T_m \tag{3}$$

The stator voltage and current are included in the complex vectors (Eq. 2). Here the complex vector is given as  $X_{art} = X_4 + jX_N$ , where  $X_4$  and  $X_N$  are direct ( $d$ ) and sqadrate ( $q$ ) axis components. ‘ $j$ ’ is the imaginary unit number,  $P$  is the number of poles.  $R_s$  and  $L_s$  are stator winding inductance and resistances. In (Eq. 3)  $T$  is the mechanical torque, motor inertia and viscous friction is given by  $J$  and  $\beta$ .

### 2.3 Adjusting Machine Voltage (VSR)

Figure 2 illustrates the relationship between the mechanical speed of the rotor and the wind turbine power at different wind speeds. The mechanical rotor speed is equal to the maximum wind power production at any wind speed. The VSR is utilized in Fig. 1 for maximum wind power. With wind speed, the MPPT wind generator algorithm (MPPT1) determines the rotor speed optimal for the scenario. There are two instances [4, 5]. As indicated in Eq. 4, PI speed controller ( $G's(s) = g'a + g2a/s$ ) is used to regulate (more alone) rotor speeds with a maximum (alone) value, and to set the q component of a current stator reference ( $I_{sq}^*$ ), whistled component of the current stator references ( $I_{sd}^*$ ) is set to nil for the maximum torque [5].

$$I_{sq}^* = (\omega_r^* - \omega_r) G_s(s), I_{sd}^* = 0 \tag{4}$$

The distinguishing operator is this instance. The closed-loop transfer function of the speed controller is adjusted to represent about 10% of the bandwidth of the internal current controller utilizing resolution (Eqs. 3 and 4) and setting  $g'A + g2a/s$



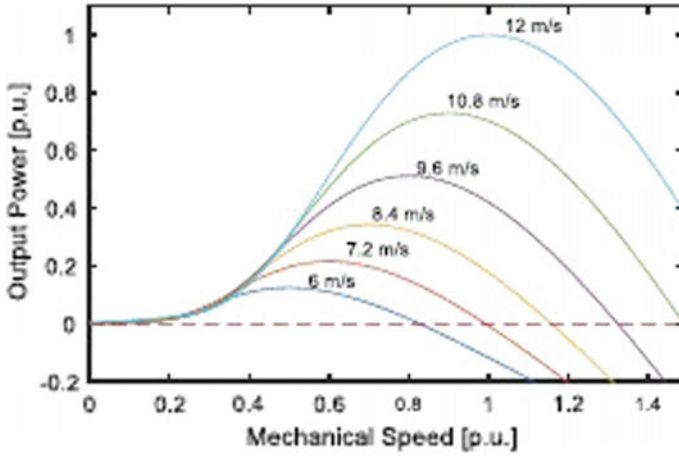


Fig. 2 Mechanical characteristics of the wind turbine at different wind speeds

=  $J$ , assume  $I_{AN} \approx I_{AN}$  within the speed controller’s (as stated in the paragraph below).

$$\omega_r/\omega_r^* = \left(\frac{3}{2}p\psi g_{ps}/j\right) / \left(s + \left(\frac{3}{2}p\psi g_{ps}/j\right)\right) \tag{5}$$

The  $g'A$  and  $g2a/s$  settings of the speed controller can be changed according to your requirements. As illustrated in (Figs. 1 and 5), the PI controller ( $Gs(s) = g'2 + g23$ ) is used to follow the needed references to the PMSG-generated stator streams (Eq. 4).

$$\bar{v}_s = \left(\vec{l}_s - \bar{l}_s\right)G_i(s) + jp\omega_r^o L_s \bar{l}_s + jp\psi H\omega_r \tag{6}$$

The loops are  $Jp$ ; the gain is denoted by  $H$ , and the super script “ $o$ ” indicates the variable’s stable value. A resolution and a delay are the current controller (Eqs. 4, 5, 6).

The filter and the utility-grid impedance are modeled in below equations,

$$\bar{v}_c = \bar{v}_f + R_f \bar{l}_c + l_f \frac{d\bar{l}_c}{dt} + j\omega l_f \bar{l}_c \tag{7}$$

$$\bar{v}_f = \bar{v}_g + R_g \bar{l}_g + l_g \frac{d\bar{l}_g}{dt} + j\omega l_g \bar{l}_g \tag{8}$$

$$\bar{l}_c = c_f \frac{d\bar{v}_f}{dt} + \bar{l}_g + j\omega c_f \bar{v}_f \tag{9}$$

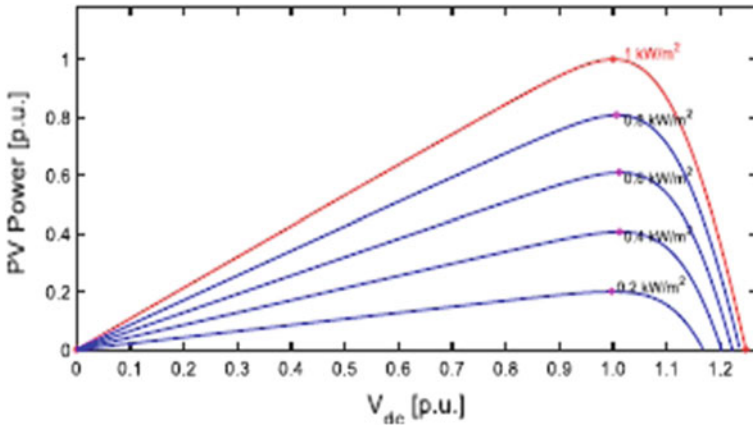


Fig. 3 Characteristics of the PV array at different solar irradiance levels

To minimize over modulation in the VSI, a PV array design should take into account the coordination between the MPPT voltage and PCC RMS voltage, i.e.,  $V_{pcc}$ , of the PV array. For example, the ratio between AC and DC voltage influences the pulse width modulation (PWM) and power converter switching pattern (Fig. 3).

$$\omega = \omega^o + \frac{V_{fq}^c}{V_{fd}^o} K_{\delta}(s) \tag{10}$$

### 3 Design of PV System

The schematic design of a free-standing three-phase hybrid PV-wind system is given in Fig. 1. The work is an important theme. This PV system has a PV array, C-connecting DC, three-phase inverter,  $V_a$ ,  $V_b$ , and  $V_c$  filter. The main purpose of this paper is to display the happy control signals via-inverter control switches in order to control the  $V_{DC}$  voltage across condenser  $C$  and to produce the load voltage input current for the power factor unit. The following statement shows the system’s mathematical model. PV cells and array modeling a PV cell is a simple P-N cross-over diode that converts light into energy. Figure 4 shows a PV cell architecture with an  $I_L$  current source a parallel diode,  $R_{sh}$  shunting resistance, and the  $R_s$  series resistance. In Fig. 4,  $I_{ON}$  is the following diode current.

$$I_{ON} = I_s [\exp[\alpha(V_{pv} + R_s i_{pv})] - 1] \tag{11}$$

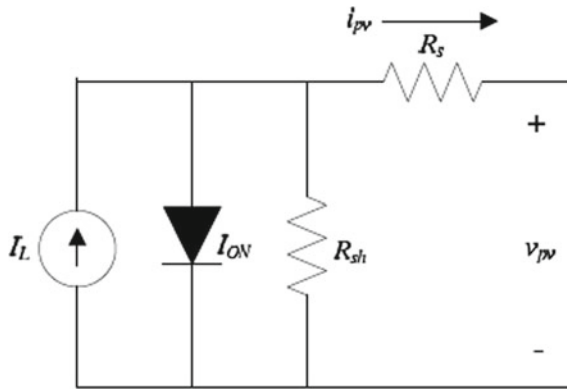


Fig. 4 Equivalent circuit diagram of PV cell

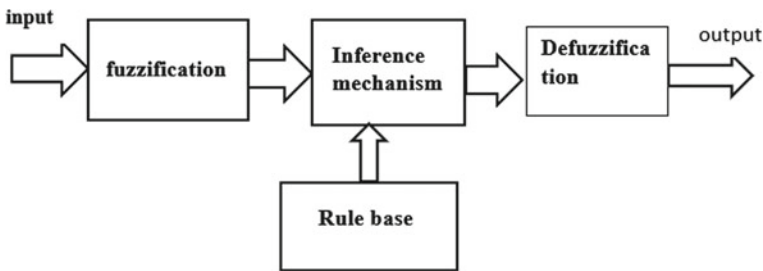


Fig. 5 Block diagram of a fuzzy logic controller

where the Boltzmann constant  $k = 1.3807 \times 10^{-23} \text{ J/K}$  is,  $q = 1.6022$ . The power of the PV array is  $C$ , i.e.,  $V_{DC}$ ; the charge by an electron is  $10^{-19} \text{ C}$ ;  $T_C$  is the absolute working temperature of a Kelvin;  $\alpha$  is ideal for the p-n connection between 1 and 5. The current Kirchhoff (KCL) law in Fig. 1 may now be used to represent the current PV cell output current ( $I_{pv}$ ).

$$i_{pv} = I_L - I_s \left[ \exp\left[\alpha(V_{pv} + R_s i_{pv})\right] - 1 \right] - \frac{v_{pv} + R_s i_{pv}}{R_{sh}} \tag{12}$$

The present light  $I_L$  relies on solar irradiation that may be connected to Eq. 13:

$$I_L = [I_{sc} + K_i(T_c - T_{ref})] \frac{s}{1000} \tag{13}$$

If  $I_{sc}$  is the short-circuit current,  $s$  is solar irradiance,  $K_i$  is the current short-circuit cell coefficient, and  $T_{ref}$  is the cell temperature benchmark. The temperature changes the saturation current of the cell according to Eq. 14,

$$I_s = I_{RS} \left[ \frac{T_C}{T_{ref}} \right]^3 \exp \left[ \frac{qE_g}{Ak} \left( \frac{1}{T_{ref}} - \frac{1}{T_C} \right) \right] \tag{14}$$

where Eq. 14 is the band-length semiconductor energy used in the cell and  $I_{RS}$ , the saturation reverse current in the cell is the radioactive energy reference time and sun. Due to the low tension of the PV cell, many PV cells in series are combined with high voltages. A number of ‘‘PV cells are built and encapsulated under astringent circumstances for the construction of the PV module with glass, plastic, and other transparent material.’’ Parallel to a PV array, a number of modules are linked to produce the required voltage and power. Figure 3 shows the design of the PV array electrical equipollent circuit.

where  $N_s$  is the number of cells in series and  $N_p$  is the number of modules in parallel. In this case, the array  $I_{PV}$  can be written as in below Eq. 15.

$$i_{pv} = N_p I_L - N_p I_s \left[ \exp \left[ \alpha \left( \frac{v_{pv}}{N_s} - \frac{R_s i_{pv}}{N_p} \right) \right] - 1 \right] - \frac{N_p}{R_{sh}} \left( \frac{v_{pv}}{N_s} - \frac{R_s i_{pv}}{N_p} \right) \tag{15}$$

### 4 Wind Modeling

The wind turbine can produce the mechanical power  $P_m$  and the associated torque  $T_m$  sent to the turbine in accordance with the following models:

$$P_m = \frac{1}{2} \rho \pi R^2 V_w^3 C_p(\lambda, \beta) \tag{16}$$

$$T_m = \frac{P_m}{\omega_m} \tag{17}$$

where  $C_p$  is a coefficient of TSR and pitch power and air density ( $\text{kg/m}^3$ ). Aerodynamic design concepts for rotor blades approximate the  $C_p$  coefficient [6].

$$C_p(\lambda, \beta) = C_1 \left[ \frac{C_2}{\lambda_i} - C_3 \beta - C_4 \right] e^{-C_5/\lambda_i} + C_6 \lambda \tag{18}$$

where the empirical constant  $C_1 = 0.5176$ ,  $C_2 = 116$ ;  $C_3 = 0.4$ ;  $C_4 = 5$ ;  $C_5 = 21$ ;  $C_6 = 0.0068$ . This work is utilized in the laboratory to build a DC wind turbine engine simulator (Figs. 2 and 3). The PMSG-diode model corrector is a kind of PMSG-diode model corrector. If the PMSG’s stator bind (allowed) has continuously fluctuated, when bent with them (rad/s) speed, induced EMF, the specified EMF is

$$(V)es = k(m) = ke/P, \text{ it is } (V)es = k(m) = ke/P \tag{19}$$

where  $K$  (Vs/rad) is the continuous induced voltage of the machine,  $P$  is the total number of rotor poles, and  $e$  is the electrical angular rate of the infused voltage stator of the PMSG provides constant terminal phase voltage,  $V_s$ , and output power  $P_g$ .

$$V_s^2 = E_s^2 - (\omega_e L_s I_s)^2 \quad P_g = 3V_s I_s = 3\sqrt{E_s^2 I_s^2 - (\omega_e L_s)^2 I_s^4} \quad (20)$$

With  $E_s$ ,  $I_s$ , and  $L_s$ , all PMSG stators, stator current, and inductance are induced. If the switching angles and as well as switching angles are small, the connections between the diode rectifier voltage and the  $V_{DC}$  can be linked along with the line voltage on the PMSG  $V_t$  terminal [7].

$$V_{DC} = \frac{\sqrt[3]{2}}{\pi} V_t = \frac{\sqrt[3]{6}}{\pi} V_s \quad (21)$$

where  $V_t$  is the RMS value of line-to-line PMSG voltage. When power loss is committed during the correction of diode circuits, the output power of WECS  $P_g$  might be similar.

$$P_g = P_{DC} = 3V_s I_s = V_{DC} I_{DC} \quad (22)$$

The current output diode corrector can be expressed as [8–11] PMSG power output  $P_g$  and EM torque  $T_g$ .

$$\begin{aligned} P_g &= \frac{\sqrt[3]{6}}{\pi} \omega_g I_{DC} \sqrt{k^2 - \frac{6}{\pi^2} (PL_s)^2 I_{DC}^2}, T_g \\ &= \frac{\sqrt[3]{6}}{\pi} I_{DC} \sqrt{k^2 - \frac{6}{\pi^2} (PL_s)^2 I_{DC}^2} \end{aligned} \quad (23)$$

The torque of the generator may be adjusted using the wind turbine rotor speed following,

$$\omega_m = \frac{T_m - T_g}{B_t} \quad (24)$$

where  $B_t$  (Nms/rad) is the rotor turbine friction coefficient. The torque wind turbine and ultimate turbine speed may be changed by changing the output current of the diode rectifier (Eqs. 10 and 11). This method is used to maximize power at different WECS wind speeds.

**Table 1** Control rule table

<i>E</i>	$\Delta E$						
	NB	NM	NS	Z	PS	PM	PB
NB	NB	NB	NB	NB	NM	NS	Z
NM	NB	NB	NB	NM	NS	Z	PS
NS	NB	NB	NM	NS	Z	PS	PM
Z	NB	NM	NS	Z	PS	PM	PB
PS	NM	NS	Z	PS	PM	PB	PB
PM	NS	Z	PS	PM	PB	PB	PB
PB	Z	PS	PM	PB	PB	PB	PB

## 5 Fuzzy Logic Controller

Fuzzy logic uses the fuzzy set theory in which members consist of one or more sets of variables with a certain membership level. Fuzzy logic helps us to emulate human thinking on computers, to quantify inaccurate data, to evaluate nebulously, and finally while still obtaining unambiguous results using defense approach. The FLC consists mostly of three blocks.

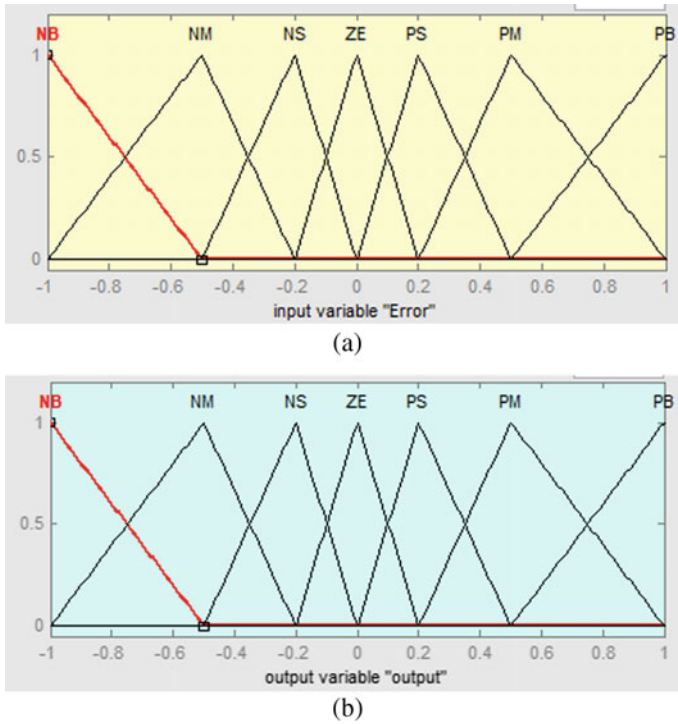
- Fuzzification
- Inference
- Defuzzification (Fig. 5)

The first is the use of a fleeting logic control system for maximum point tracking a hybrid wind-solar system in fixed atmospheric conditions. The model and control technique is a helpful tool for improving the performance of a hybrid power system for smart home applications; the wind turbine subsystem pitch angle is synchronized using a floating logic control system based on the obtained wind speed data (Table 1).

The solar cell output impedance is equal to the load impedance values in the solar subsystems fuzzy logic controller. Figure 6(a) and (b) show the membership function of input error and output gain respectively.

## 6 Results and Discussion

The MATLAB hybrid wind-solar system simulation is verified using the well-known Simulink methodology. The DC engine and SG combo simulate a wind turbine. By stacking related VSC's on top of each other, the feature of changing wind speed is accomplished. Solar PV array characteristics are obtained with a solar PV simulator. SG and grid currents are used for the three-phase VI quantification. Grid voltage,  $V_{DC}$ , and  $V_t$  are all monitored during phase VI quantification. The control algorithm



**Fig. 6** Membership functions of the input error and output gain. (a) input error, (b) output gain

determines these feedbacks settings. Table 2 shows the parameters and its values of the proposed system.

**Table 2** Proposed system specifications

S. No.	Parameters	Values
1	SG parameters	415 V, 5hp
2	Rs	3.1 Ω
3	Poles	4
4	Field excitation voltage	220 V
5	DC motor	220 V, 19 A, 5hp
6	Wind speed range	7–12 m/s
7	DC link voltage: Vdc	400 V
8	Solar PV array	2.4 kW
9	Parameters of BES	240 V, 56 Ah
10	Ripple filter parameters	Rf = 5 Ω, Cf = 10 μF
11	Grid parameters: Vgab	220 V, 50 Hz

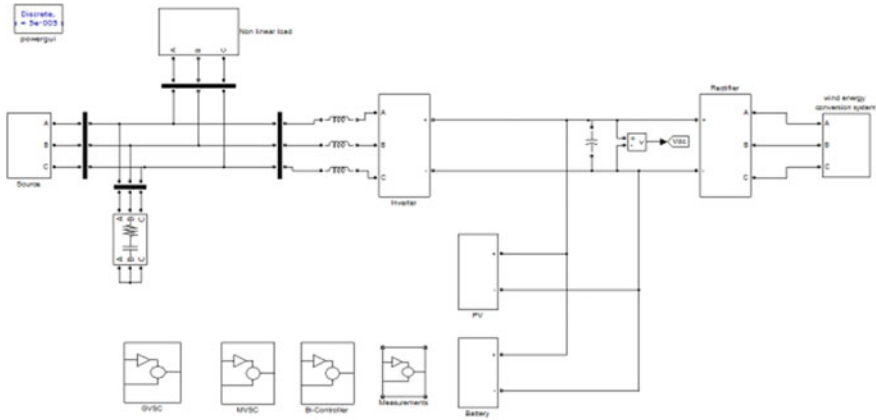


Fig. 7 Simulink model of the proposed hybrid wind-solar cogeneration

**Case I: Steady-State Behavior of the Micro-grid**

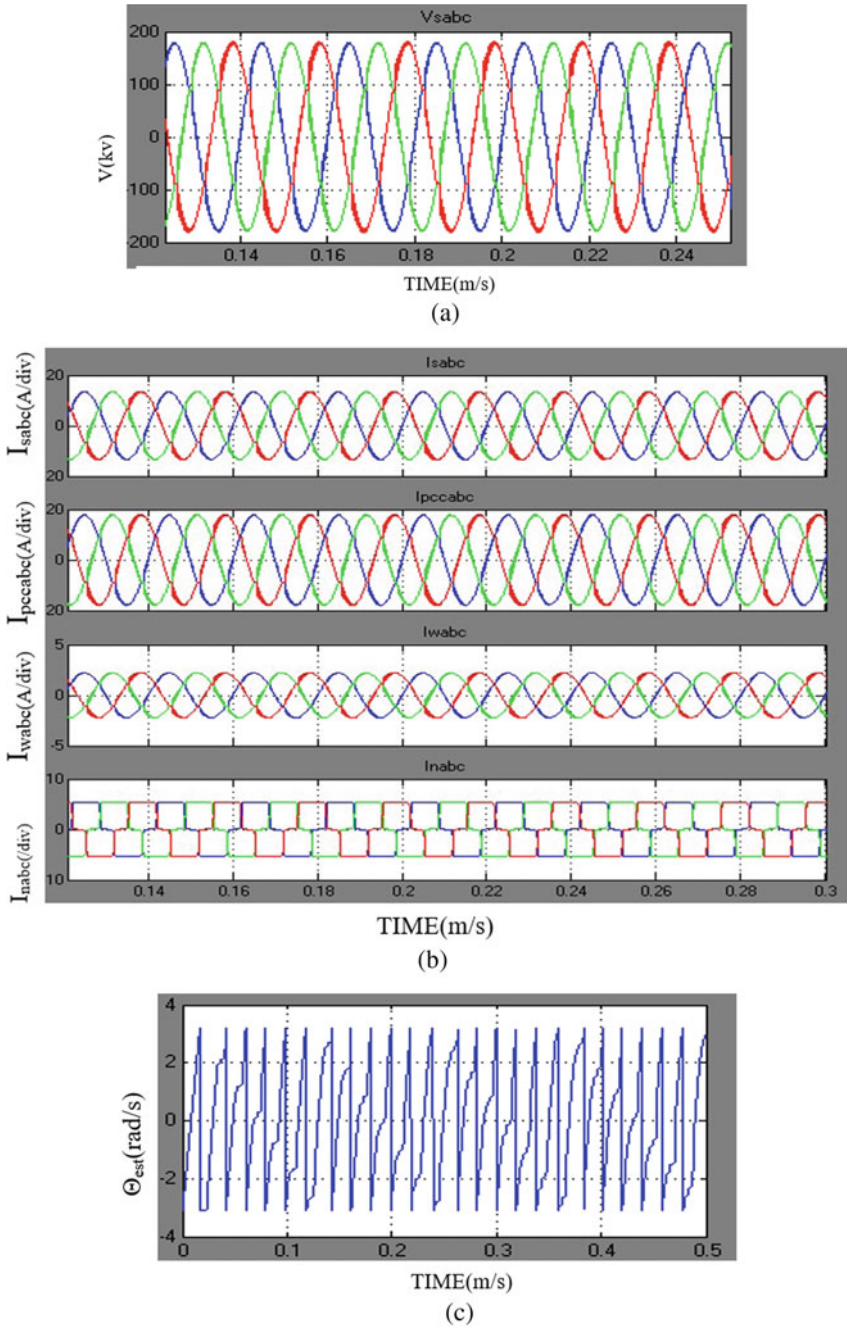
The performance of a constant micro-grid is shown in Fig. 7. Solar and wind power at the rated levels are evaluated, and a bidirectional converter maintains the DC connection tension. The generation of SG is minimal at low wind speeds like 7.2 m/s; the amplitude of the produced currents is tiny. The greatest generation from SG is achieved with the rapid increase in wind speed, as illustrated in Fig. 7.

The current load ( $I_{nabc}$ ), composed of harmonics, is composed of balanced, sinusoidal grid currents. In Fig. 8, you find the GVSC ( $I_{pccabc}$ ) currents, loading current, and grid currents. GVSC offers off-setting currents. The grid streams ( $I_{sabc}$ ) are sinusoidal, phased-inversed, and played in Fig. 8.

**Case II: Response of Micro-grid under Wind Speed Change**

Inherently, unpredictable wind speeds are thus needed to assess their impacts in order to operate the micro-grid correctly. As the wind speeds rise and decrease together with a one-phase grid voltage in Fig. 9, the DC connection voltage and AC network terminal voltage are depicted. The two-way converter overcomes the transitory voltage of the DC connection, and the two-way converter keeps the DC connecting voltage at a constant value. The impact of wind speed change and SG, the direct axis ( $I_d$ ), and the square axis ( $I_q$ ) current are presented in Fig. 9. The wind speed is rising from 7.2 to 12 m/s, with the  $I_d$  and  $I_q$  SG components simultaneously increasing. Internal wind speed signal fluctuations and no wind situations are observed in Fig. 9  $w_{wff}$  is rising size with wind-driven SG production growing at wind speed. Changes in the weight and DC loss component ( $W_{loss}$ ) occur at the same time ( $W_{net}$ ). The  $w_{wff}$  term is 0 when the wind velocity is zero and the wind velocity is less than the down. Figure 10 demonstrates system power variations dependent on wind speed differences. Figure 10 demonstrates changes in power with increasing and decreasing demand for the load. If the output of renewable energy is



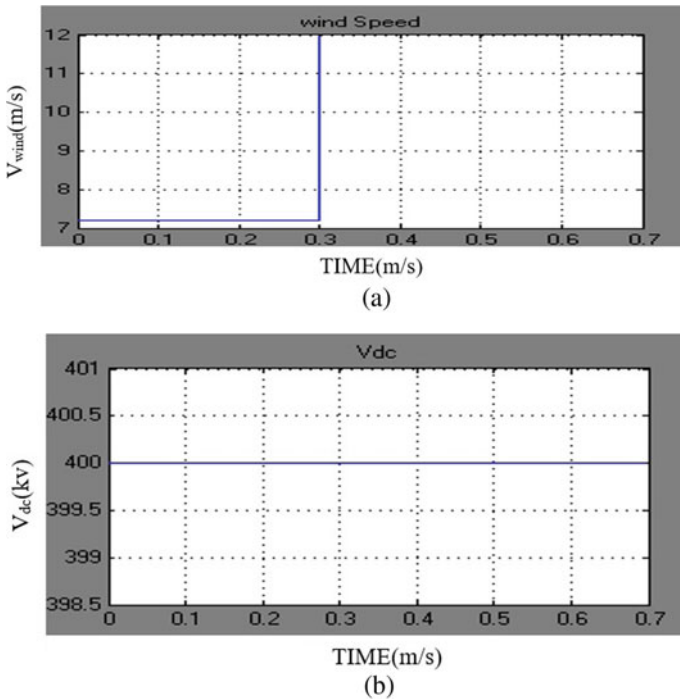


**Fig. 8** Response of the micro-grid at steady-state condition, **a** source voltage ( $V_{sabc}$ ), **b**  $I_{sabc}$ ,  $I_{pccabc}$ ,  $I_{wabc}$ , and  $I_{nabc}$ , **c** phase angle ( $\theta_{est}$ )

rated demand for load increases, the excess energy generated by the system will be lowered and vice versa.

**Case III: Response of Micro-grid under Solar Insolation Change**

Sun irradiation fluctuates during the Dayan dissent at night. In Fig. 10, the micro-grids performance is optically shielded from solar radiation issues. The solar panel’s electricity increases or decreases when the sunshine falls, and the photovoltaic current rises or falls. The two-way converter controls the DC connection’s transient voltage. The AC grid voltage is maintained sinewed. Figure 10 illustrates changes to the internal signal. The photovoltaic term ( $W_{pvff}$ ) is reduced to 0 at zero irradiation. Since wind, BES, and grid generation absorb the charge demand, the load weight component remains constant. The variance of  $W_{loss}$  and  $W_{net}$  has simultaneously enhanced optics. The solar, wind, load, and grid power changes as a function of solar irradiation transmutations are shown in Fig. 10. Micro-grid production consumes load demand.



**Fig. 9** Response of the micro-grid at change in wind speed, **a** wind speed (Ws), **b** DC link voltage ( $V_{dc}$ ), **c**  $I_{sabc}$ ,  $I_{pccabc}$ ,  $I_{wabc}$  and  $I_{nabc}$ , **d** phase angle ( $\theta_{est}$ )

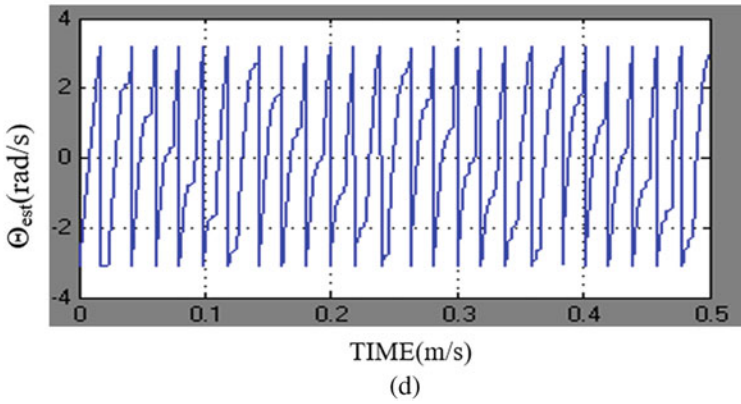
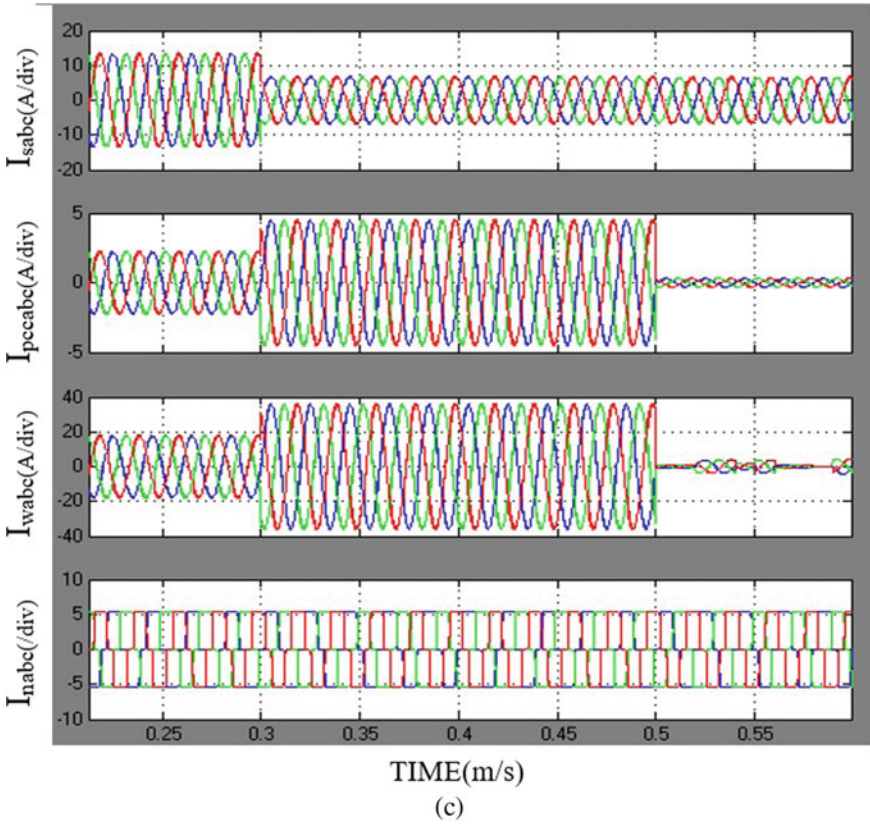
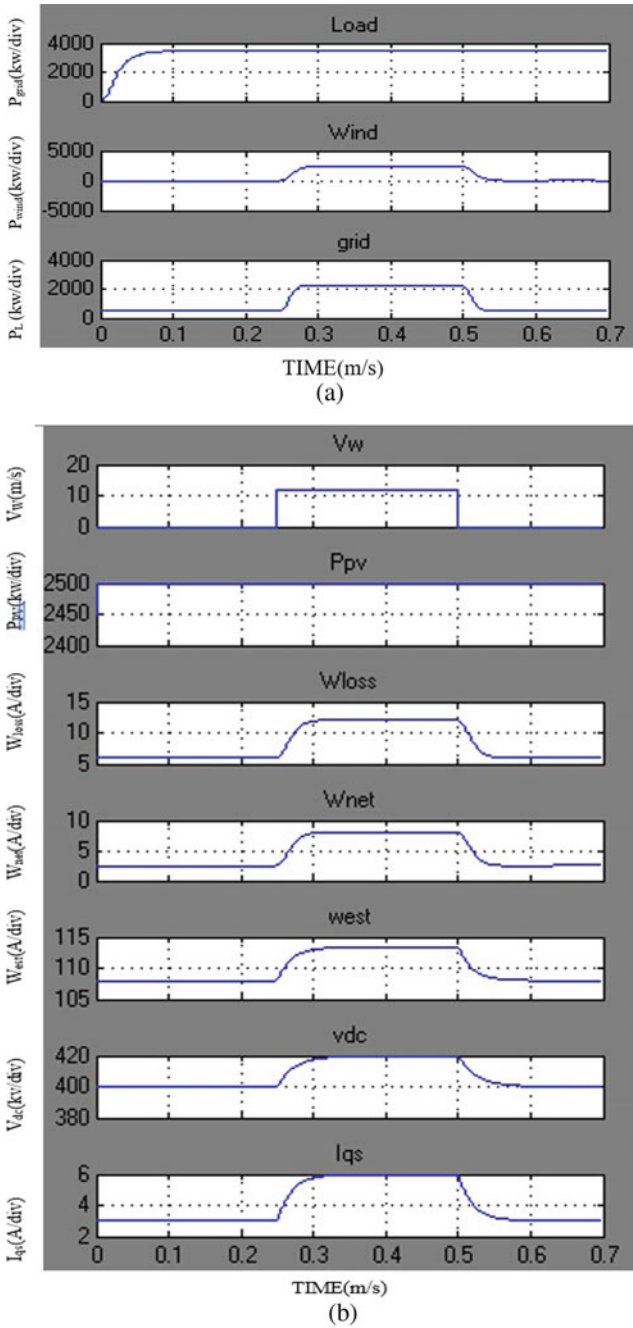


Fig. 9 (continued)



**Fig. 10** Response of the micro-grid at solar irradiations, **a** active power of load, wind, and grid, **b**  $V_w$ ,  $P_{pv}$ ,  $W_{loss}$ ,  $W_{net}$ ,  $W_{est}$ ,  $V_{dc}$ ,  $I_{qs}$

## 7 Conclusion

Vector-controlled back-to-back grid devices are used in this article for the proof of wind-PV cogeneration systems. After wind speed swings, the MVSC catches the highest wind power. The GVSC function on the grid consists of extracting the maximum PV power from the PV generator, balancing the power output via the DC-link condenser, and maintaining the PCC voltage unit in various operating modes. The proposed system is of the fuzzy logic controller of the system improves below points, and a small stability survey was carried out for the entire system. Some of the benefits of the system area follow: (1) Wind and solar accumulators' reliability and efficiency have risen. (2) The extraction of wind and PV electricity is entirely based on an independent VSR and VSI MPPT extraction. (3) The VSI maintains DC-link voltage control in all operating circumstances, therefore improving its damping efficiency. (4) Basic system structure and controller architecture. (5) Fault tying may be performed with existing auspice systems. In the MATLAB/Simulink system, the results of time-domain simulations were dampened.

## References

1. Singh B, Solanki J (2009) A comparison of control algorithms for DSTATCOM. *IEEE Trans Ind Electron* 56(7):2738–2745
2. Chauhan PJ, Reddy BD, Bhandari S, Panda SKB (2019) Battery energy storage for seamless transitions of wind generator in standalone microgrid. *IEEE Trans Ind Appl* 55(1):69–77
3. Shan Y, Hu J, Chan KW, Fu Q, Guerrero JM (2018) Model predictive control of bi-directional DC-DC converters and AC/DC interlinking converters—a new control method for PV-wind-battery microgrids. *IEEE Trans Sustain Energy* Early Access
4. Nguyen C, Lee H (2017) Power management approach to minimize battery capacity in wind energy conversion systems. *IEEE Trans Ind Appl* 53(5):4843–4854
5. Chen X, Sun H, Wen J, Lee WJ, Yuan X, Li N, Yao L (2011) Integrating wind farm to the grid using hybrid multi terminal HVDC technology. *IEEE Trans Ind Appl* 47(2):965–972
6. Malysz P, Sirouspour S, Emadi A (2014) An optimal energy storage control strategy for grid-connected microgrids. *IEEE Trans Smart Grid* 5(4):1785–1796
7. Kshirsagar P, Krishnan R (2012) High-efficiency current excitation strategy for variable-speed non-sinusoidal back-EMF PMS Mmachines. *IEEE Trans Ind Appl* 48(6):1875–1889
8. Hou X, Sun Y, Lu J, Zhnag X, Koh LH, Su M, Guerrero JM (2018) Distributed hierarchical control of AC microgrid operating in grid-connected, islanded and their transition modes. *IEEE Access* 6:77388–77401
9. Boudoudouh S, Maaroufi M (2019) Renewable energy sources integration and control in railway microgrid. *IEEE Trans Ind Appl* Early Access
10. Millner AR, Smith C, Jaddivada R, Ilic M (2018) Component standards for stable microgrids. *IEEE Trans Power Syst* Early Access
11. Radwan M, Mohamed Y (2019) Grid-connected wind-solar cogeneration using back-to-back voltage source converters. *IEEE Trans Sustain Energy* Early Access

# Small-Signal Analysis of Inverter to Maximise Power from Two PV Subarrays Under Different Environmental Conditions



S. Sneha and P. B. Savitha

## 1 Introduction

The grid-connected PV power system is one of the promising arrangements for the utilisation of photovoltaic energy. The ever-increasing demand for energy or the technological advancement in the integration of renewables resources, especially photovoltaic power, has brought a lot of advancement and continuous improvement in extracting maximum power and its utilisation. There are various arrangements of grid-connected photovoltaic systems. The transformerless grid-connected PV is one of the arrangements. In every arrangement, the aim is to deliver maximum power to the grid with minimum distortion in spite of variation in environmental conditions. The design or the arrangement of the photovoltaic arrays/modules plays an important role in delivering desired power. The PV arrays are reconfigured or grouped into subarrays in such a way to increase the DC link voltage. In case of mismatched environmental conditions, the array output voltages are varied and the DC link voltage is reduced thereby the desired power is not delivered [1, 2]. One of the biggest reliability issues of PV systems is the difference between their expected and actual power outputs. When large number of modules are connected in series, the power production from the array gets reduced when the modules are exposed to varying environmental conditions such as shade [3, 4]. Topologies created from neutral point clamped (NPC)-based inverters [5] suffer greatly from mismatched operation problems. To address the aforementioned issue, efforts have been made to extract the maximum possible power from each of the PV modules while they are subjected to

---

S. Sneha (✉) · P. B. Savitha

Department of Electrical and Electronics Engineering, Dayananda Sagar College of Engineering, Bangalore 560078, India

e-mail: [snehashankar82@gmail.com](mailto:snehashankar82@gmail.com)

P. B. Savitha

e-mail: [savithapb-eee@dayanandsagar.edu](mailto:savithapb-eee@dayanandsagar.edu)

mismatched operating circumstances. Schemes based on the control circuit [3, 4] had sought to do this. However, they accomplished this through a two-stage arrangement, resulting in a complicated control algorithm and a low operational efficiency. Power extraction during mismatched environment condition is obtained by selecting suitable interconnections of PV modules [6]. However, the method described above is inadequate for low-power single-phase grid-connected transformerless PV systems. In order to simplify the control configuration and to reduce the number of components, schemes reported in [7] combine all the PV modules into two subarrays, and then each of the subarray is made to operate at their respective MPP. However, the reported overall efficiency of both schemes is poor. The insertion of a buck and boost stage in single-phase grid-connected transformerless PV inverters improves power extraction during MEC. The small-signal analysis made to understand the behaviour of the circuit helps in understanding the behaviour of the system, especially with nonlinear features. It is a typical analytical approach in electronics engineering that uses linear equations to simulate the behaviour of electronic circuits including nonlinear components. By using the well-known small-signal modelling [8–12], the equivalent small-signal equations and the new equivalent small-signal circuit can be obtained.

In this paper, solar PV arrays are divided into two serially connected subarrays. Control of each subarray is carried out using a buck and boost based inverter. During MEC, the maximum power evacuated from the subarrays is determined. For maximum power tracking, perturb and observe mppt algorithm is used. A detailed Small-Signal model is carried out. The inverter's topological configuration and its control strategy ensure that the magnitude of leakage current affiliated with PV arrays remains within the allowable limit.

## 2 Inverter Modelling and Its Operation

The schematic diagram of the circuit is shown in Fig. 1. The working and modelling is discussed in this section stage by stage. Converter 1 consists of self-commutated switches,  $S_1$  with its anti-parallel diode  $D_1$  and switch  $S_3$  with  $D_3$ , as well as freewheeling diodes,  $D_{fw1}$ ,  $D_{fw3}$ , respectively. It also consists of filter inductor and capacitors,  $L_1$ ,  $C_3$  and  $C_1$ , respectively. Similarly, converter 2 is made up of  $S_2$  with  $D_2$  and switch  $S_4$  with  $D_4$ , and freewheeling diodes such as  $D_{fw2}$ ,  $D_{fw4}$ , respectively. It also consists of filter inductor and capacitors  $L_2$ ,  $C_4$  and  $C_2$ , respectively. In order to track the maximum power from the panels, perturb and observe ( $P$  and  $O$ ) maximum power point tracking (mppt) is used. Inverter stage has four self-commutated switches namely  $S_a$ ,  $S_b$ ,  $S_c$  and  $S_d$  along with their anti-parallel diodes  $D_a$ ,  $D_b$ ,  $D_c$  and  $D_d$ , respectively. This stage is connected to the grid through the filter inductor and resistor  $L_g$  and  $R_g$ , respectively. Converter 1 operates in buck mode when  $V_{pva} \geq v_{C1}$  and it operates in boost mode when  $V_{pva} < v_{C1}$ . Converter 2 operates in buck mode when  $V_{pvb} \geq v_{C2}$  and it operates in boost mode when  $V_{pvb} < v_{C2}$  where  $V_{pva}$ ,  $V_{pvb}$  are the

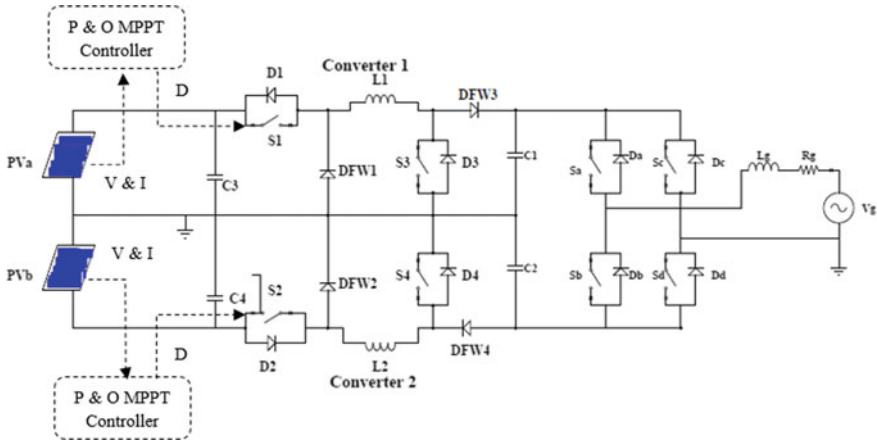


Fig. 1 Buck-boost based grid-tied inverter [13]

maximum power point (MPP) voltages of  $P_{va}$  and  $P_{vb}$  and  $v_{C1}$ ,  $v_{C2}$  are the output voltages of converter 1 and converter 2, respectively.

### 2.1 Modelling of PV

Photovoltaic model with series resistance ( $R_S$ -model) is achieved with the inclusion of series resistance  $R_S$  and is shown in Fig. 2, hence the output current can be derived as:

$$I = I_{ph} - I_s \left[ \exp\left(\frac{V + IR_s \frac{N_s}{N_p}}{N_s V_T}\right) - 1 \right] \tag{1}$$

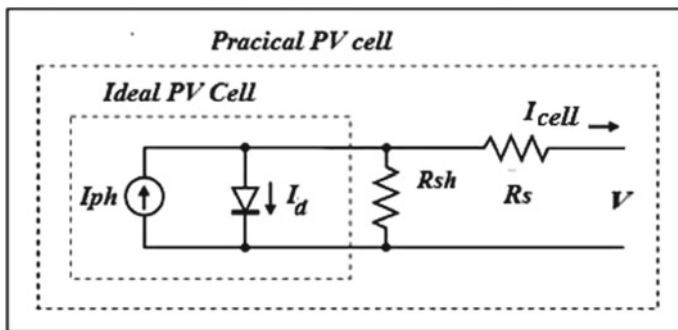


Fig. 2 Single diode equivalent circuit [8]



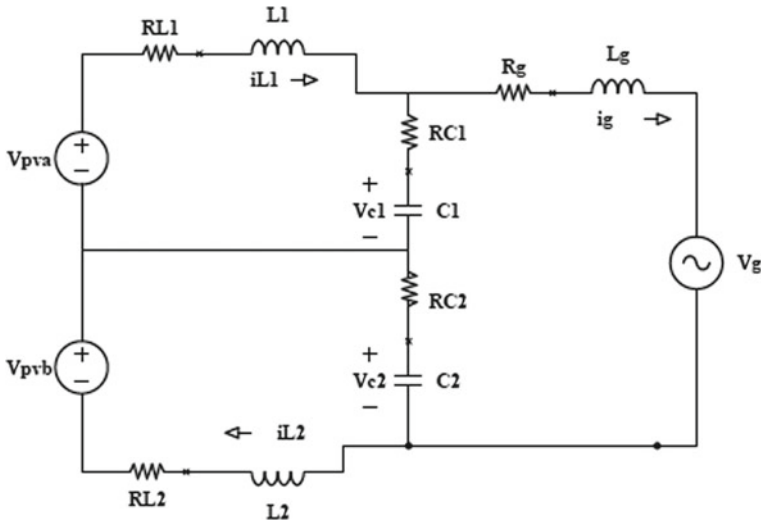


Fig. 3 Equivalent circuit when switches  $S_1$  and  $S_2$  are turned on

Equation (1) does not adequately represent the behaviour of the cell when subjected to environmental variations, especially at low voltages. A more practical model can be seen in Fig. 3 where series  $R_s$ , and parallel resistances  $R_{sh}$ , are introduced. Applying Kirchhoff’s current law to the node where  $I_{ph}$ , diode,  $R_{sh}$  and  $R_s$  meet,

$$I = I_{ph} - I_d - I_{sh}$$

$$I = I_{ph}N_P - N_P \left\{ I_s \left[ \exp \left( \frac{V + IR_s \frac{N_S}{N_P}}{N_S V_T} \right) - 1 \right] \right\} - \left[ \frac{V + IR_s \frac{N_S}{N_P}}{\frac{N_S}{N_P} R_{sh}} \right] \tag{2}$$

where:  $N_S$  = number of series connected cells,  $N_P$  = number of parallel connected modules,  $I_s$  = cell saturation of dark current,  $I_d$  = diode current,  $I_{sh}$  = shunt current,  $V_T$  (thermal voltage) =  $kT_C/q$ ,  $k$  = Boltzmann’s constant, =  $1.38 \times 10^{-23}$  J/K,  $q$  = electron charge ( $1.6 \times 10^{-19}$  C).

### 2.2 Small-Signal Analysis of Buck-Boost Inverter

The buck-boost based inverter small-signal model is carried out and it is tested for four different conditions. The assumptions made for the state variables are current through inductors and voltage across capacitors.

$$\begin{aligned} i_{L1}(t) &= x_1(t); i_{L2}(t) = x_2(t); V_{C1}(t) = x_3(t); V_{C2}(t) \\ &= x_4(t); i_g(t) = x_5(t); \end{aligned} \tag{3}$$

*Input and Output variables are considered as*

$$V_{pva}(t) = u_1(t), V_{pvb}(t) = u_2(t) \quad \text{and} \quad V_g(t), = y(t)$$

**Buck mode**

In buck mode, switches  $S_1$  and  $S_2$  are turned on and off sinusoidally, while switches  $S_3$  and  $S_4$  are kept off.

*By applying KVL to the Fig. 3,*

$$\dot{i}_{L1}(t) = -\frac{R_{L1}}{L_1}i_{L1}(t) - \frac{R_{C1}}{L_1}i_{L1}(t) - \frac{1}{L_1}V_{C1}(t) + \frac{R_{C1}}{L_1}i_g(t) + \frac{1}{L_1}V_{pva}(t) \tag{4}$$

$$\dot{i}_{L2}(t) = -\frac{R_{L2}}{L_2}i_{L2}(t) - \frac{R_{C2}}{L_2}i_{L1}(t) - \frac{1}{L_2}V_{C2}(t) + \frac{R_{C2}}{L_2}i_g(t) + \frac{1}{L_2}V_{pvb}(t) \tag{5}$$

$$\begin{aligned} \dot{i}_g(t) &= \frac{R_{C1}}{L_g}i_{L2}(t) + \frac{R_{C2}}{L_g}i_{L2}(t) + \frac{1}{L_g}V_{C1}(t) + \frac{1}{L_g}V_{C2}(t) \\ &\quad - \frac{R_{C1}}{L_g}i_g(t) - \frac{R_{C2}}{L_g}i_g(t) - \frac{R_g}{L_g}i_g(t) \end{aligned} \tag{6}$$

Voltage across capacitors  $C_1$  and  $C_2$  are

$$v_{C1}(t) = \frac{1}{C_1} \int [i_{L1}(t) - i_g(t)] \quad \text{and} \quad v_{C2}(t) = \frac{1}{C_2} \int [i_{L2}(t) - i_g(t)]$$

Differentiating on both the sides,

$$\dot{v}_{C1}(t) = \frac{i_{L1}(t)}{C_1} - \frac{i_g(t)}{C_1} \quad \text{and} \quad \dot{v}_{C2}(t) = \frac{i_{L2}(t)}{C_2} - \frac{i_g(t)}{C_2} \tag{7}$$

The state model can be obtained by combining Eqs. (4), (5), (6) and (7):

$$\begin{bmatrix} \dot{i}_{L1}(t) \\ \dot{i}_{L2}(t) \\ \dot{v}_{C1}(t) \\ \dot{v}_{C2}(t) \\ \dot{i}_g(t) \end{bmatrix} = \begin{bmatrix} -\frac{R_{L1}+R_{C1}}{L_1} & 0 & -\frac{1}{L_1} & 0 & \frac{R_{C1}}{L_1} \\ 0 & -\frac{R_{L2}+R_{C2}}{L_2} & 0 & -\frac{1}{L_2} & \frac{R_{C2}}{L_2} \\ \frac{1}{C_1} & 0 & 0 & 0 & -\frac{1}{C_1} \\ 0 & \frac{1}{C_2} & 0 & 0 & -\frac{1}{C_2} \\ \frac{R_{C1}}{L_g} & \frac{R_{C2}}{L_g} & \frac{1}{L_g} & \frac{1}{L_g} & -\frac{R_{C1}+R_{C2}+R_g}{L_g} \end{bmatrix} \begin{bmatrix} i_{L1}(t) \\ i_{L2}(t) \\ v_{C1}(t) \\ v_{C2}(t) \\ i_g(t) \end{bmatrix}$$

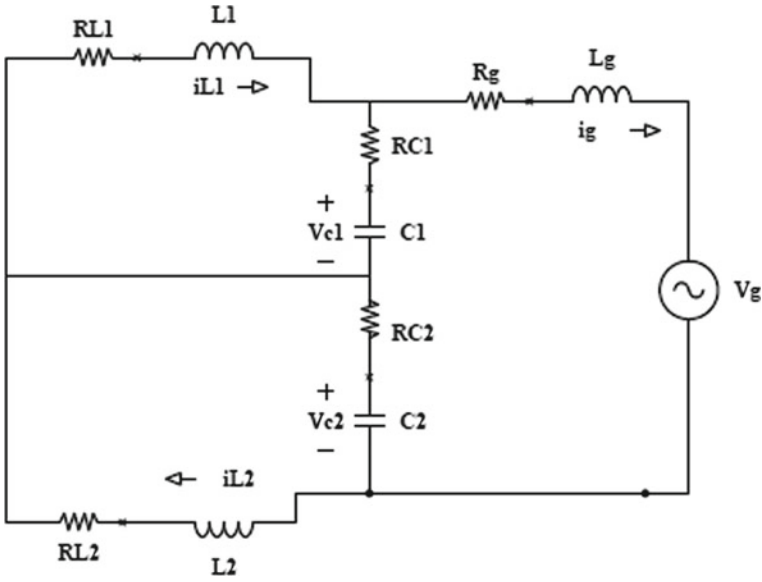


Fig. 4 Equivalent circuit when switches  $S_1$  and  $S_2$  are turned off

$$+ \begin{bmatrix} \frac{1}{L_1} & 0 & 0 \\ 0 & \frac{1}{L_2} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -\frac{1}{L_g} \end{bmatrix} \begin{bmatrix} v_{pva}(t) \\ v_{pvb}(t) \\ v_g(t) \end{bmatrix} \tag{8}$$

Figure 4 represents the equivalent circuit of buck mode operation when switches  $S_1$  and  $S_2$  are turned off which is same as equivalent circuit of buck mode operation when switches  $S_1$  and  $S_2$  are turned on but the only difference is that input source  $V_{pva}$  and  $V_{pvb}$  are absent here so,  $u_1(t)$  and  $u_2(t)$  is equal to zero.

Hence, the state model can be obtained as:

$$\begin{bmatrix} \dot{i}_{L1}(t) \\ \dot{i}_{L2}(t) \\ \dot{v}_{C1}(t) \\ \dot{v}_{C2}(t) \\ \dot{i}_g(t) \end{bmatrix} = \begin{bmatrix} -\frac{R_{L1}+R_{C1}}{L_1} & 0 & -\frac{1}{L_1} & 0 & \frac{R_{C1}}{L_1} \\ 0 & -\frac{R_{L2}+R_{C2}}{L_2} & 0 & -\frac{1}{L_2} & \frac{R_{C2}}{L_2} \\ \frac{1}{C_1} & 0 & 0 & 0 & -\frac{1}{C_1} \\ 0 & \frac{1}{C_2} & 0 & 0 & -\frac{1}{C_2} \\ \frac{R_{C1}}{L_g} & \frac{R_{C2}}{L_g} & \frac{1}{L_g} & \frac{1}{L_g} & -\frac{R_{C1}+R_{C2}+R_g}{L_g} \end{bmatrix} \begin{bmatrix} i_{L1}(t) \\ i_{L2}(t) \\ v_{C1}(t) \\ v_{C2}(t) \\ i_g(t) \end{bmatrix}$$

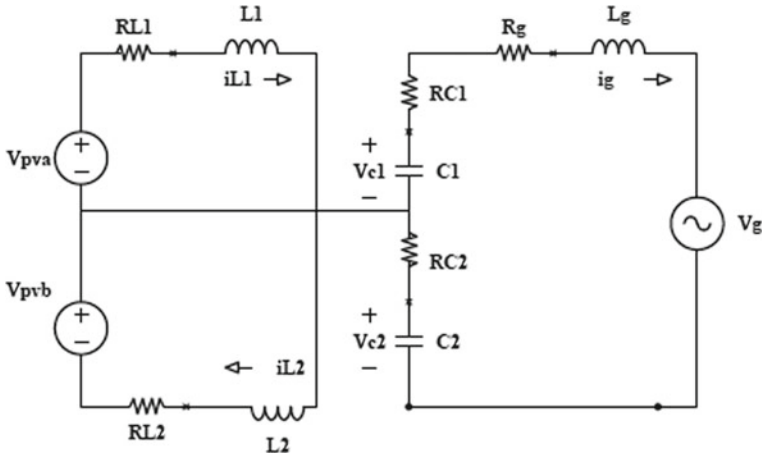


Fig. 5 Equivalent Circuit when switches  $S_3$  and  $S_4$  are turned on

$$+ \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -\frac{1}{L_g} \end{bmatrix} \begin{bmatrix} v_{pva}(t) \\ v_{pvb}(t) \\ v_g(t) \end{bmatrix} \tag{9}$$

**Boost mode**

In boost mode, switches  $S_3$  and  $S_4$  are turned on and off sinusoidally, and switches  $S_1$  and  $S_2$  remain on throughout the operation mode.

Figure 5 represents the equivalent circuit of boost mode when switches  $S_3$  and  $S_4$  are turned on and switches  $S_1$  and  $S_2$  are also kept on throughout this operation.

By applying KVL to Fig. 5,

$$i_{L1}'(t) = -\frac{R_{L1}}{L_1}i_{L1}(t) + \frac{1}{L_1}V_{pva}(t) \tag{10}$$

$$i_{L2}'(t) = -\frac{R_{L2}}{L_2}i_{L2}(t) + \frac{1}{L_2}V_{pvb}(t) \tag{11}$$

$$i_g'(t) = \frac{1}{L_g}v_{C1}(t) + \frac{1}{L_g}v_{C2}(t) - \frac{R_{C1}}{L_g}i_g(t) - \frac{R_{C2}}{L_g}i_g(t) + \frac{R_g}{L_g}i_g(t) - \frac{v_g}{L_g} \tag{12}$$

Voltage across capacitors  $C_1$  and  $C_2$  are

$$v_{C1}(t) = \frac{1}{C_1} \int [-i_g(t)] \text{ and } v_{C2}(t) = \frac{1}{C_2} \int [i_{L2}(t) - i_g(t)]$$

Differentiating on both the sides,

$$v_{c1}(t) = -\frac{i_g(t)}{C_1} \text{ and } v_{c2}(t) = -\frac{i_g(t)}{C_2} \tag{13}$$

The state model can be obtained by combining Eqs. (10), (11), (12) and (13)

$$\begin{bmatrix} \dot{i}_{L1}(t) \\ \dot{i}_{L2}(t) \\ \dot{v}_{C1}(t) \\ \dot{v}_{C2}(t) \\ \dot{i}_g(t) \end{bmatrix} = \begin{bmatrix} -\frac{R_{L1}}{L_1} & 0 & 0 & 0 & 0 \\ 0 & -\frac{R_{L2}}{L_2} & 0 & 0 & 0 \\ \frac{1}{C_1} & 0 & 0 & 0 & -\frac{1}{C_1} \\ 0 & \frac{1}{C_2} & 0 & 0 & -\frac{1}{C_2} \\ 0 & 0 & \frac{1}{L_g} & \frac{1}{L_g} & -\frac{R_{C1}+R_{C2}+R_g}{L_g} \end{bmatrix} \begin{bmatrix} i_{L1}(t) \\ i_{L2}(t) \\ v_{C1}(t) \\ v_{C2}(t) \\ i_g(t) \end{bmatrix} + \begin{bmatrix} \frac{1}{L_1} & 0 & 0 \\ 0 & \frac{1}{L_2} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -\frac{1}{L_g} \end{bmatrix} \begin{bmatrix} v_{pva}(t) \\ v_{pvb}(t) \\ v_g(t) \end{bmatrix} \tag{14}$$

Figure 6 represents the equivalent circuit of boost mode when switches  $S_3$  and  $S_4$  turned off and switches  $S_1$  and  $S_2$  are made on throughout. This circuit is same as buck mode operation when switches  $S_1$  and  $S_2$  are made on. Hence, the state model remains the same.

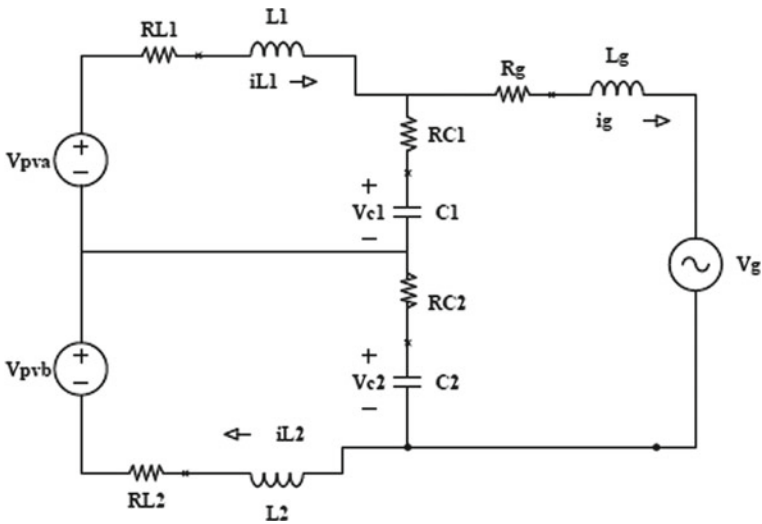


Fig. 6 Equivalent Circuit when switches  $S_3$  and  $S_4$  are turned off

$$\begin{bmatrix} i_{L1}(t) \\ i_{L2}(t) \\ v_{C1}(t) \\ v_{C2}(t) \\ i_g(t) \end{bmatrix} = \begin{bmatrix} -\frac{R_{L1}+R_{C1}}{L_1} & 0 & -\frac{1}{L_1} & 0 & \frac{R_{C1}}{L_1} \\ 0 & -\frac{R_{L2}+R_{C2}}{L_2} & 0 & -\frac{1}{L_2} & \frac{R_{C2}}{L_2} \\ \frac{1}{C_1} & 0 & 0 & 0 & -\frac{1}{C_1} \\ 0 & \frac{1}{C_2} & 0 & 0 & -\frac{1}{C_2} \\ \frac{R_{C1}}{L_g} & \frac{R_{C2}}{L_g} & \frac{1}{L_g} & \frac{1}{L_g} & -\frac{R_{C1}+R_{C2}+R_g}{L_g} \end{bmatrix} \begin{bmatrix} i_{L1}(t) \\ i_{L2}(t) \\ v_{C1}(t) \\ v_{C2}(t) \\ i_g(t) \end{bmatrix} + \begin{bmatrix} \frac{1}{L_1} & 0 & 0 \\ 0 & \frac{1}{L_2} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -\frac{1}{L_g} \end{bmatrix} \begin{bmatrix} v_{pva}(t) \\ v_{pvb}(t) \\ v_g(t) \end{bmatrix} \tag{15}$$

Here, the parasitic resistances of  $L_1$ ,  $L_2$ ,  $L_g$ ,  $C_1$  and  $C_2$  are  $R_{L1}$ ,  $R_{L2}$ ,  $R_g$ ,  $R_{C1}$  and  $R_{C2}$ , respectively. Because the grid frequency  $f_g$  is sufficiently lower than the switching frequency  $f_s$ , the state space averaging approach is used. To simplify the study  $V_{pva}$ ,  $V_{pvb}$  and  $v_g$  are assumed to be stiff voltage sources, with the impact of input filter capacitors ignored. The system parameters values are considered to be as follows:  $R_{L1} = R_{L2} = 0.12$ ,  $R_g = 0.04$ ,  $R_{C1} = R_{C2} = 0.26$ ,  $V_{pva}$ ,  $V_{pvb}$ , = 130 V.

### 2.3 Working of the Inverter

Figure 7 shows the working of the buck mode operation of the inverter for positive and negative half cycles, respectively. For both buck and boost modes of operation, during the positive half cycle, switches  $S_a$  and  $S_d$  are kept on while switches  $S_b$  and  $S_c$  are kept off permanently. During the negative half cycle, switches  $S_b$  and  $S_c$  are kept on while switches  $S_a$  and  $S_d$  are kept off permanently.

During buck mode, duty ratios of the switches  $S_1$  and  $S_2$  are varied sinusoidally to maintain sinusoidal grid current ( $i_g$ ), while switches  $S_3$  and  $S_4$  are kept off. During boost mode, duty ratios of the switches  $S_3$  and switches  $S_4$  are varied sinusoidally to maintain sinusoidal grid current ( $i_g$ ), while switches  $S_1$  and  $S_2$  are kept on throughout this mode.

If the irradiation level and operating temperature of panel 1 ( $P_{va}$ ) and panel 2 ( $P_{vb}$ ) are different, then the voltages at maximum power point ( $V_{pva}$  and  $V_{pvb}$ ), current at MPP ( $I_{pva}$  and  $I_{pvb}$ ) and power at MPP ( $P_{pva}$  and  $P_{pvb}$ ) will differ corresponding to  $P_{va}$  and  $P_{vb}$ , respectively.

If all the subarrays are running at their respective MPPs and the losses occurred in power processing stages are ignored, then the total power involved with  $C_1$  and  $C_2$  are:

$$P_{c1} = P_{pva} \text{ and } P_{c2} = P_{pvb} \tag{16}$$

$P_g$  is the average power injected into the grid over a half cycle,

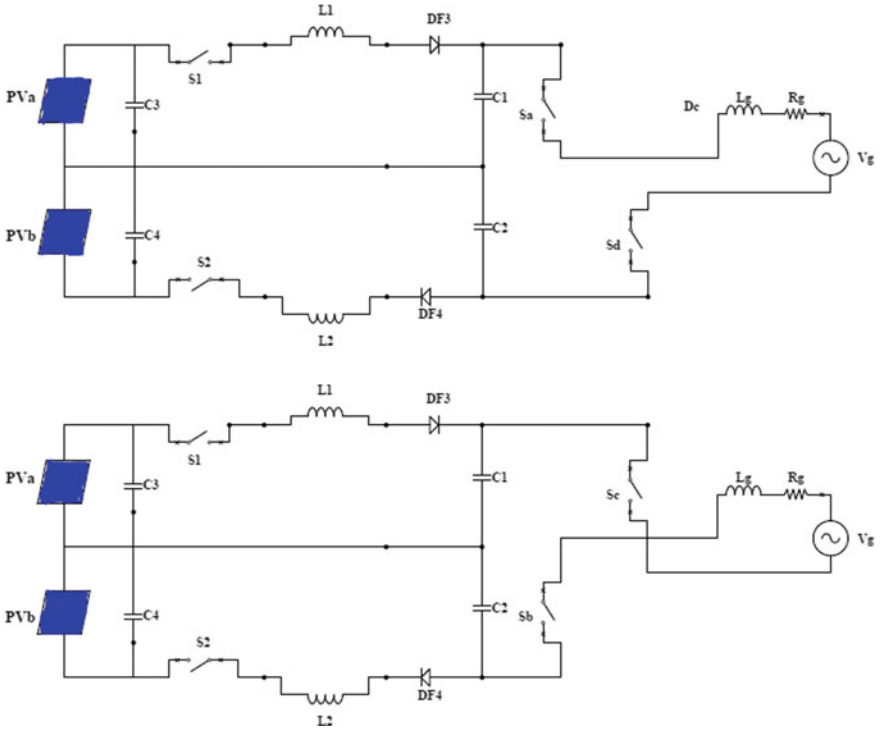


Fig. 7 Active operating states of buck mode during positive and negative half cycles, respectively

$$P_g = P_{pva} + P_{pvb} \tag{17}$$

At any half cycle

$$v_g = v_{c1} + v_{c2} \tag{18}$$

Therefore, instantaneous injected power is

$$(P_g) \text{ is } P_g = v_g i_g = (v_{c1} + v_{c2}) i_g \tag{19}$$

Here,  $v_{c1}$  and  $v_{c2}$  are the instantaneous quantities of  $V_{c1}$  and  $V_{c2}$ , respectively.

As  $i_g$  is in phase with  $v_g$ ,

$$I_g = \frac{P_g}{V_g} \tag{20}$$

Here,  $V_g$  and  $I_g$  represents rms values of  $v_g$  and  $i_g$ .

The power injected to the grid is

$$P_g = \frac{1}{\pi} \int_0^{\pi} p_g d(\omega t) = P_{c1} + P_{c2} \quad (21)$$

Hence,

$$P_{c1} = \int_0^{\pi} V_{c1m} \sin(\omega t) I_{gm} \sin(\omega t) d(\omega t)$$

$$P_{c1} = \frac{V_{c1m} I_{gm}}{2} \quad \text{and} \quad P_{c2} = \frac{V_{c2m} I_{gm}}{2} \quad (22)$$

where  $V_{c1m}$ ,  $V_{c2m}$  and  $I_{gm}$  are the amplitudes of  $v_{c1}$ ,  $v_{c2}$  and  $i_g$ .

By combining (16) and (22):

$$V_{c1m} = \frac{\sqrt{2} * P_{PVa}}{P_g / V_g} \quad \text{and} \quad V_{c2m} = \frac{\sqrt{2} * P_{PVb}}{P_g / V_g} \quad [\text{from(20)}] \quad (23)$$

By combining (17) and (23):

$$V_{c1m} = \frac{V_m * P_{pva}}{P_{pva} + P_{pvb}} \quad \text{and} \quad V_{c2m} = \frac{V_m * P_{pvb}}{P_{pva} + P_{pvb}} \quad (24)$$

The voltage patterns  $v_{c1}$  and  $v_{c2}$  appear as full wave rectified sinusoidal waveforms with amplitudes  $V_{c1m}$  and  $V_{c2m}$ . The amplitude of  $v_g$  is denoted by  $V_m$ . The magnitudes of  $V_{c1m}$  and  $V_{c2m}$  are determined by the power extracted from each subarray, as shown in (24).

If the power obtained from  $PV_a$  is less than the power extracted from  $PV_b$ , then  $V_{c1m} < V_{c2m}$ , whereas if the power extracted from  $PV_b$  is less than  $PV_a$ , then  $V_{c2m} < V_{c1m}$ .

The controller is intended to achieve the goals such as both subarrays operate simultaneously at their respective MPP and  $i_g$  (grid current) is sinusoidal and in-phase with  $V_g$  (grid voltage) throughout the operating range. The controller used for the circuit is the PI controller and it is used to connect the inverter to the grid. Details are explained in the simulation.

## 2.4 Design of $L_1$ , $L_2$ , $L_g$ , $C_1$ and $C_2$

To design the values of  $L_1$  and  $L_2$ , the inverter's buck mode operation is taken into consideration and it is calculated using Eq. (25).



**Table 1** The designed values of inductors and capacitors are

Elements	Values
Inductors $L_1, L_2$	10 mH, 10 mH
Capacitors $C_1, C_2$	30 mF, 30 mF
Grid frequency	50 Hz

$$\text{i.e., } L_1 = \frac{V_{C12} * \Delta I_{\text{factor}}}{\Delta I_{L1} * f_{sw}} \text{ and } L_2 = \frac{V_{C12} * \Delta I_{\text{factor}}}{\Delta I_{L2} * f_{sw}} \tag{25}$$

which is derived from [14] where  $V_{C12}$  is 320 V, the maximum ripple instant factor,  $\Delta I_{\text{factor}} = 0.12$ , which is chosen from the ripple factor characteristic [15] for modulation index = 0.87. A higher ripple current value reduces the inductor’s size and endogenous losses. As a result, a value of not more than 20% is suggested, i.e. inductor ripple currents,  $\Delta L_1 = \Delta L_2 = 0.19$  A, and the switching frequency,  $f_{sw} = 20$  kHz. To create a high stability margin and a wide control bandwidth, a value smaller than  $L_1$  or  $L_2$  is used for  $L_g$ .

Capacitor value is determined by the reactive power absorbed in the rated state. The design theory provided in [16] is used to select the value of the filter capacitors  $C_1$  and  $C_2$ . Taking  $x$  as a proportion of reactive power absorbed under rated conditions,

$$C_r = xC(x \text{ value must be lesser than 1, here it is assumed as 0.5})$$

The condition limits the upper capacitor value as,

$$C = \frac{0.05(P_{\text{rated}})}{V_g^2 * \omega_o} \tag{26}$$

Here  $P_{\text{rated}}$  is the rated power, i.e. 780 W while  $V_g$  is the fundamental RMS voltage of the grid, i.e. 220 V,  $\omega_o$  is the fundamental frequency in radians per second, i.e. ( $\omega_o = 2 * \pi * f_g$ ) where  $f_g$  is the grid frequency = 50 Hz (Table 1).

### 3 Simulation Model and Analysis

To validate the model, the MATLAB/SIMULINK simulation is carried out for various conditions which contribute to mismatched power. The variation in insolation is taken from 500 to 1000 W/m<sup>2</sup>. To simulate, the CANADIAN PV module of capacity 165 W is selected [17]. Its detailed parameters and its specifications are listed in Table 2.

The simulation is carried out considering mismatched environmental conditions for subarrays, and it is simulated for 7 s. At standard test condition (STC), the MPP parameters of each subarray are as follows:  $P_{pva} = P_{pvb} = 583$  W,  $V_{pva} = V_{pvb} = 102$  V and  $I_{pva} = I_{pvb} = 5.72$  A. With reference to Table 2, the simulation of the

**Table 2** The final designed values for the converter

Elements	Values
PV array details	Canadian solar polycrystalline modules CS6P-165PE. Each subarray has four series connected PV modules [18]
$V_g$ and $f_g$	220 V and 50 Hz
MPPT algorithm	Perturb and observe
Inductors $L_1, L_2$ and $L_g$	10 mH, 10 mH, 0.08 mH
Capacitors $C_1, C_2$	30 mF, 30 mF
Switching frequency of switches $S_1$ to $S_4$	20 kHz
Carrier frequency of switches $S_5$ to $S_8$	7500 Hz
Load $R$	57.6 $\Omega$

circuit is done and it summarises the variation in irradiation level and the temperature with respect to time for the two subarrays.

## 4 Results

The variation of irradiance and other parameters are tabulated in Table 3. For the power changing from 250 to 700 W, the maximum power is obtained during 5–6 s because the inverter is analysed for different environmental conditions and its efficiency is tested. The simulation is carried out with the help of data given in Tables 1, 2 and 3. Figure 8a represents the output of PV arrays from 0 to 7 s for different variable irradiance conditions. From the waveforms, it is clear that the maximum power is obtained during 5–6 s since PV modules are divided into two serially connected subarrays and buck and boost based inverter is used in order to control each of the subarray, and Fig. 8b represents the waveforms of inductor currents and Fig. 8c represents the waveforms of capacitors voltages from 0 to 7 s for different variable irradiance conditions. Figure 9 represents the waveforms of inductors currents and Fig. 10 represents the waveforms of voltages across capacitors from 0 to 7 s for different variable irradiance conditions.

Table 4 represents the efficiency of the buck-boost inverter for the time duration variations from 0 to 7 s. From the table, it is clear that the peak efficiency achieved is about 96.47%.

**Table 3** Simulated output volatges and power during dynamic variations in temperature and in isolation

Time in (s)	Irrad. in $PV_a$ ( $W/m^2$ )	Irrad. in $PV_b$ ( $W/m^2$ )	Tempin $PV_a$ ( $^{\circ}C$ )	Tempin $PV_b$ ( $^{\circ}C$ )	$P_{pv_a}$ (W)	$P_{pv_b}$ (W)	$I_{gm}$ (A)	$V_{C1m}$ (V)	$V_{C2m}$ (V)	$I_{L1m}$ (A)	$I_{L2m}$ (A)
0-1	500	800	25	25	200	450	5.4	150	185	5.3	8.2
1-2	600	800	25	25	230	475	5.4	142	183	6.3	8.2
2-3	700	800	25	25	270	460	5.4	145	180	7.2	8.2
3-4	800	800	25	25	383	383	5.4	164	164	7.9	8.2
4-5	900	800	25	25	578	280	5.4	192	135	9.3	8.2
5-6	1000	800	25	25	649	280	5.4	198	135	10.2	8.2
6-7	1000	800	30	25	640	278	5.4	195	135	10.2	8.2

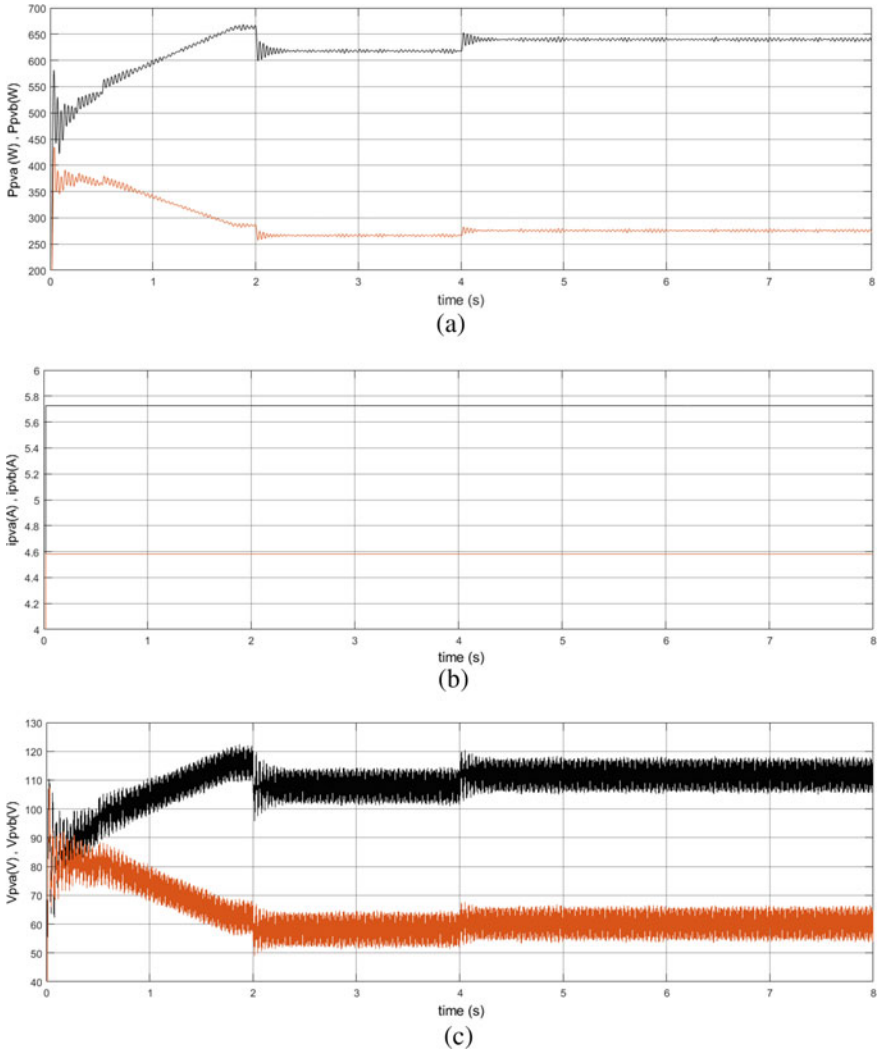
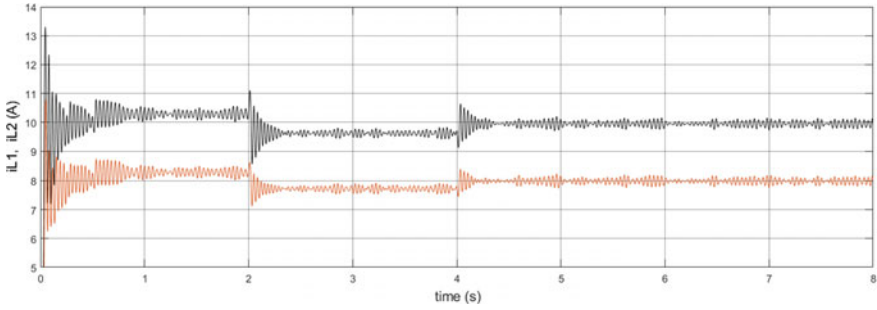
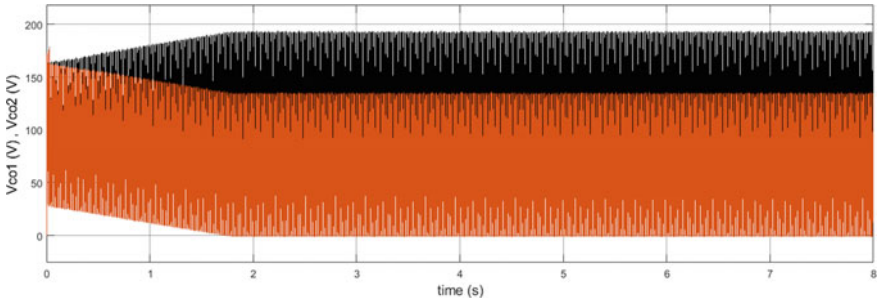


Fig. 8 Simulated waveform, a power versus time, b current versus time, c voltage versus time



**Fig. 9** Simulated waveform: inductors currents versus time



**Fig. 10** Simulated waveform: capacitors voltage versus time

**Table 4** Efficiency values obtained for the inverter

Time in (s)	AC output power of inverter in watts	DC input power of inverter in watts	Efficiency in %
0–1	1680	1758	95.56
1–2	1680	1747.3	96.14
2–3	1680	1741.35	96.47
3–4	1680	1759.72	95.46
4–5	1680	1765.06	95.18
5–6	1680	1771.2	94.85
6–7	1680	1777.05	94.53

## 5 Conclusion

In this work, an effort is made to limit the number of series connected PV subarrays so as to meet the desired DC link voltage in the case of the transformerless grid-connected system during mismatched environment conditions is discussed. To overcome this, an effort is made by rearranging/grouping the PV array into two subarrays and the output voltage is fed to the buck-boost based inverter. It is analysed by the detailed small-signal of buck-boost based modelling. And the detailed simulation of the inverter is carried out for mismatched environment condition using MALAB/SIMULINK.

## References

1. Amoiralis EI, Tsili MA, Kladas AG (2012) Power transformer economic evaluation in decentralized electricity markets. *IEEE Trans Ind Electron* 59(5):2329–2341
2. Bowtell L, Ahfock A (2010) Direct current offset controller for transformerless single-phase photovoltaic grid-connected inverters. *IET Renew Power Gener*
3. Shimizu T, Hirakata M, Kamezawa T, Watanabe H (2001) Generation control circuit for photovoltaic modules. *IEEE Trans Power Electron* 16(3)
4. Zhang L, Sun K, Feng L, Wu H, Xing Y (2013) A Family of neutral point clamped full-bridge topologies for transformerless gridtied inverters. *IEEE Trans Power Electron* 28(2):730–739
5. Bidram A, Davoudi A, Balog RS (2012) Control and circuit techniques to mitigate partial shading effects in photo-voltaic arrays. *IEEE J Photovolt* 2(4):532–546
6. Debnath D, Chatterjee K (2016) Maximising power yield in a transformerless single phase grid connected inverter servicing two separate photovoltaic panels. *IET Renew Power Gener* 10(8):1087–1095
7. Tian H, Mancilla-David F, Ellis K, Jenkins P, Muljadi E (2012) A detailed performance model for photovoltaic systems. National Renewable Energy Laboratory, USA NREL/JA-5500–54601, July 2012
8. Spier DW, Oggier GG, da Silva SAO, from Federal University of Technology—UTFPR-CP (2017) Modeling and analysis of a dc-dc boost-buck converter for renewable energy applications. *IEEE*
9. Zhao Z, Xu M, Chen Q, Lai J, Cho Y (2012) Derivation, analysis, and implementation of a Boost-Buck converter-based high-efficiency PV inverter. *IEEE Trans Power Electron* 27(3):1304–1313
10. Raveendhra D, Thakur P, Narasimha Raju BL (2013) Design and small signal analysis of solar PV fed FPGA based closed loop control Bi-directional DC-DC converter. In: International conference on circuits, power and computing technologies
11. Katti P, Pradeepa S, Arjun M (2020) Small signal model for PV fed buck converter in continuous and discontinuous conduction modes. *Int J Eng Appl Sci Technol* 5(3). ISSN No. 2455-2143
12. Spiera DW, Oggier GG, da Silva SAO (2019) Dynamic modeling and analysis of the bidirectional DC-DC boost-buck converter for renewable energy applications. *Sustain Energy Technol Assessm* 34:133–145
13. Dutta S, Chatterjee K (2017) A buck & boost based grid connected PV inverter maximizing power yield from two PV arrays in mismatched environmental conditions. *IEEE Trans Ind Electron*
14. Dutta S, Debnath D, Chatterjee K (2017) A grid connected single phase transformerless inverter controlling two solar PV arrays operating under different atmospheric conditions. *IEEE Trans Ind Electron*

15. Araujo SV, Zacharias P, Mallwitz R (2010) Highly efficient single phase transformerless inverters for grid-connected photovoltaic systems. *IEEE Trans Ind Electron*
16. Wu W, He Y, Blaabjerg F (2012) An LLCL power filter for single phase grid-tied inverter. *IEEE Trans Power Electron* 27(2):782–789
17. Information on Canadian solar module CS6P-165PE, available: [www.solarhub.com/product-catalog/pv-modules/124](http://www.solarhub.com/product-catalog/pv-modules/124)
18. Savitha PB, Shashikala MS, PuttaBuddhi KL (2016) Modeling of photovoltaic array and control of grid connected photovoltaic system to provide quality power to grid. In: International conference on electrical, electronics, communication, computer and optimization techniques (ICECCOT)

# Some Properties of Silicene Nanoribbons When Doped with Two Lead Atoms in an Electric Field



Hoang Van Ngoc

## 1 Introduction

Science and technology are developing more and more, creating products of great value. The continuous development of science and technology requires the creation of new materials with many new features. In recent years, scientists have been constantly working to create new materials, especially low-dimensional materials. Low-dimensional materials are an important step forward in miniaturizing electronic devices, reducing product costs, and saving power consumption. Low-dimensional materials have properties, principles, and many effects that are different from bulk materials. Silicene is a two-dimensional material, and it has a band gap [1–3]. The doping of atoms in silicene helps to adjust the band gap of silicene [4]. The external electric field is also one of the causes to change the band gap of silicone [5]. Changing the band gap width of the material helps to create applications with the appropriate band gap. Silicon nanoribbons (SNRs) are one-dimensional materials whose edges are modified by hydrogen or other atoms [4]. The change of atoms at the two edges of the SNRs causes its energy band structure to change as well. In this paper, we study the simulation of doping two Pb atoms in a unit cell of SNRs. There is a constant external field acting on the system, which has a magnitude of  $0.1 \text{ V/\AA}$ . Thanks to the change of doping positions, we study three configurations to find the most optimal configuration. These configurations will be optimized for simulation, energy of formation calculation, energy band drawing, and density of states (dos) diagram. The DFT theory is used for the calculation, which is a proven theory of reliability. The subject configurations are stable in the electric field. The combination

---

H. Van Ngoc (✉)

Institute of Applied Technology, Thu Dau Mot University, Thu Dau Mot, Binh Duong Province, Vietnam

e-mail: [ngochv@tdmu.edu.vn](mailto:ngochv@tdmu.edu.vn)



of lead doping and the external electric field produces a system with many distinct electronic properties.

## 2 Configurations

Figure 1 shows the initial structure of silicene nanoribbons (SNRs), and Fig. 1 is the structure of a unit lattice cell. This is a buckling structure, and its buckling has a magnitude of about 0.538 Å. The distance between adjacent Si atoms is from about 2.29 to 2.21 Å (Figs. 2, 3, and 4).

The energy of formation is calculated according to the formula [6].

$$E_f = E_t - E_p + 2 * E_{Si} - 2 * E_{Pb} \tag{1}$$

Here,  $E_f$  is the energy of formation;  $E_t$  and  $E_p$  are the total energy of doping system and pristine system;  $E_{Si}$  and  $E_{Pb}$  are the energies of free atom Si and Pb, respectively.

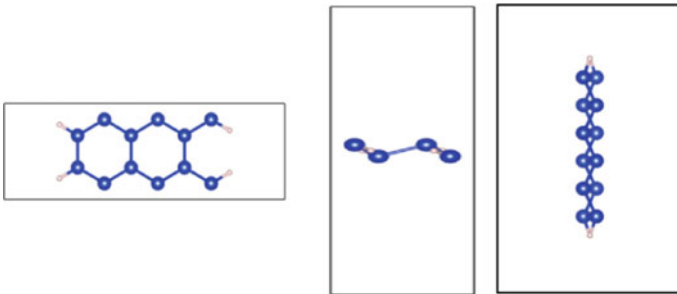


Fig. 1 Pristine system

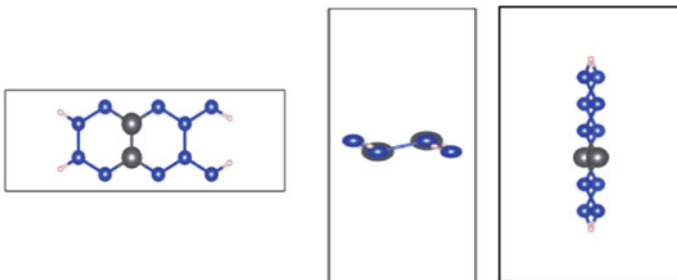


Fig. 2 Ortho-structure

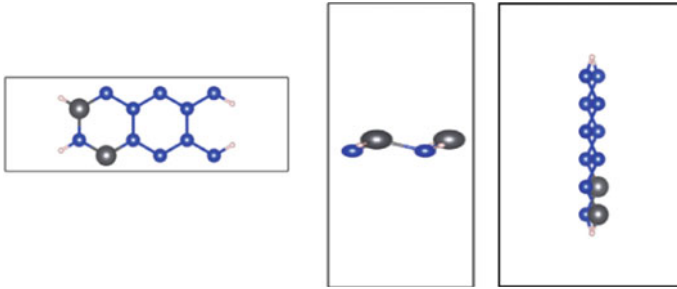


Fig. 3 Meta-structure

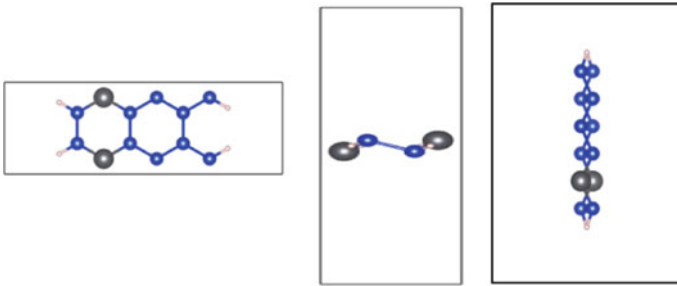


Fig. 4 Para-structure

Table 1 Energy of formation

Doped configuration	$E_t$ (eV)	$E_p$ (eV)	$E_{Si}$ (eV)	$E_{Pb}$ (eV)	$E_f$ (eV)
Ortho	-70.01	-69.41	-0.135	-0.132	-0.606
Meta	-70.10	-69.41	-0.135	-0.132	-0.696
Para	-69.65	-69.41	-0.135	-0.132	-0.246

Table 1 shows formation energy of doped systems; looking at the data table, we see that the meta-doped system gives the lowest formation energy, and we can see that this is the most optimal configuration (Fig. 5).

### 3 Discussions

Figure 1 is the original undoped structure, this is a semiconductor, and its band gap is 0.325 eV. Figures 6, 7, and 8 show the ortho-, meta-, and para-structures, respectively; the post-doped configurations are also semiconductors with band gaps of 0.4, 0.26, and 0.41 eV, respectively. Thus, look back at the original structure, the ortho- and

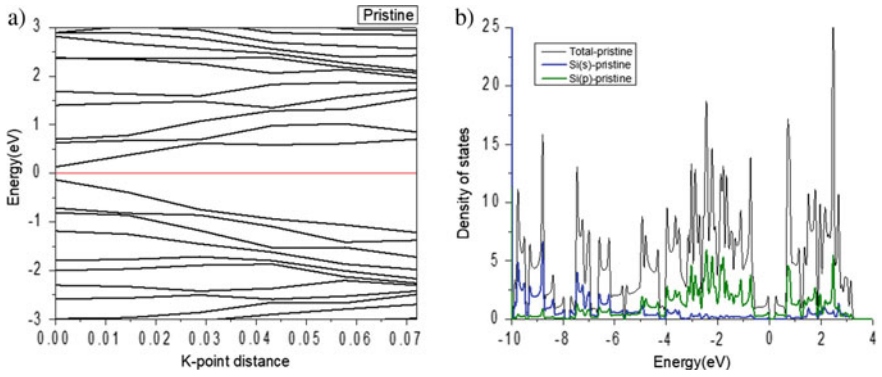


Fig. 5 Energy band (a) and Dos (b) of the original structure

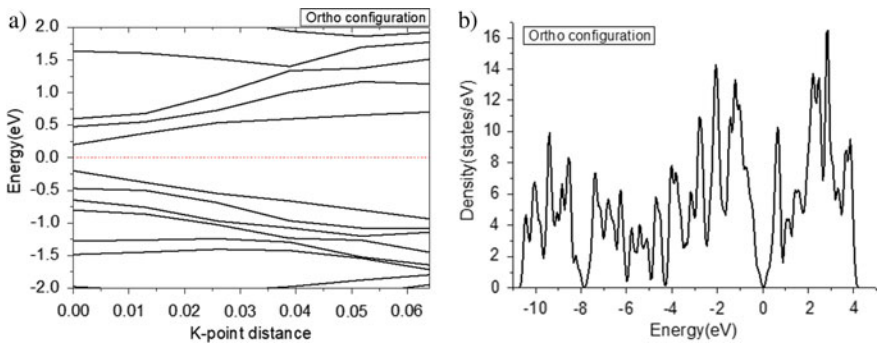


Fig. 6 Energy band (a) and Dos (b) of the ortho-structure

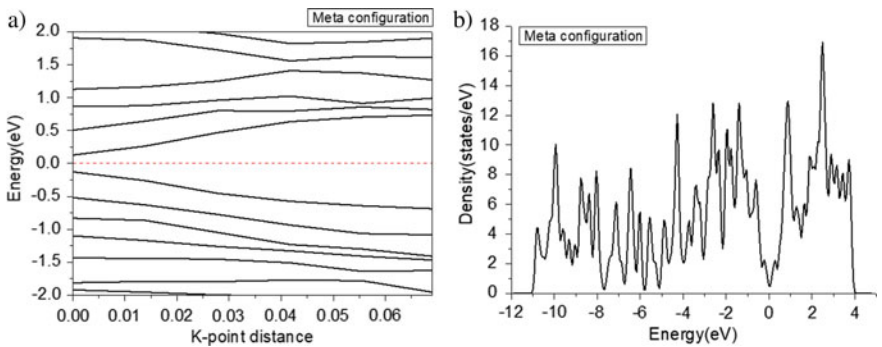
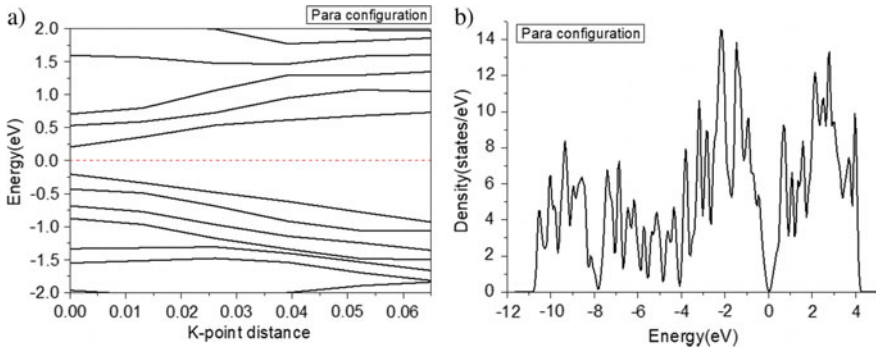


Fig. 7 Energy band (a) and Dos (b) of the meta-structure

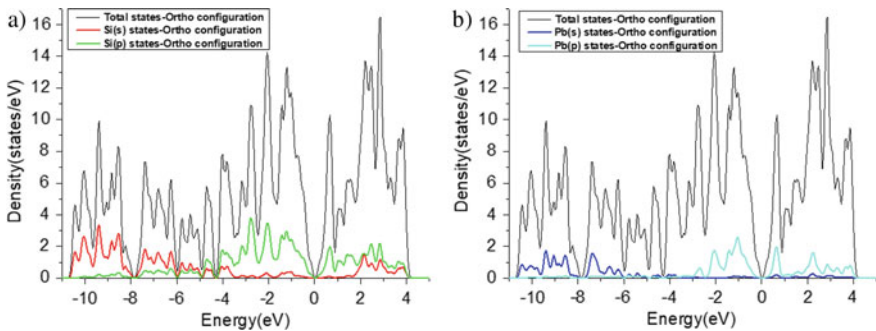


**Fig. 8** Energy band (a) and Dos (b) of the para-structure

para-structures have a wider band gap, while the meta-structure has a narrower band gap, and the reason for this difference is because the doping position of Pb is changed.

The concentration of states is relatively similar for the three doped configurations, but the state peaks of the three configurations are different. The change of states, as well as band gap, is partly due to lead doping and partly due to the influence of electrical field.

Figures 9, 10, and 11 show the contributions of partial states to the formation of conduction and covalent regions. The contributions of silicon states are large due to the large density of silicon particles in the structure. Although the Pb density is small, the contribution of Pb(s) and Pb(p) states is relatively large, which indicates the mobility of the outer shell electrons of lead. Lead and silicon both belong to group IV, so the semiconductor system will conduct electricity with the same density of electrons and holes. There are overlapping peaks representing the hybridization between lead and silicon which gives stability in the doping configuration.



**Fig. 9** States of Si (a) and Pb (b) of the ortho-structure

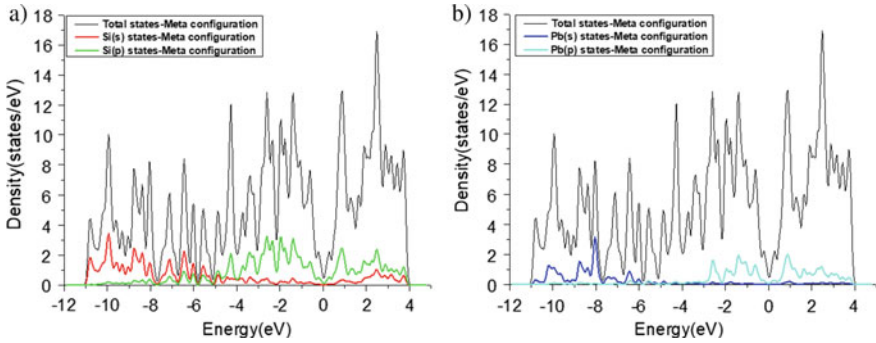


Fig. 10 States of Si (a) and Pb (b) of the meta-structure

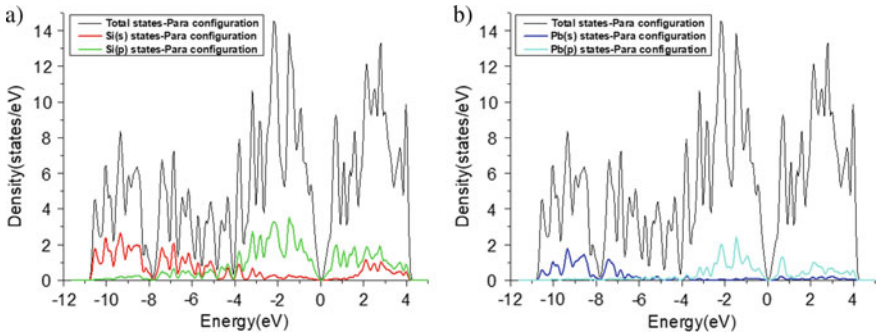


Fig. 11 States of Si (a) and Pb (b) of the para-structure

### 4 Conclusions

The article has studied Pb doping in SNRs, and there are two Pb atoms in an unit cell. The external field effect on the system has a constant magnitude of  $0.1 \text{ V/\AA}$ . The doping site change produces the three studied configurations as ortho-, meta-, and para-structures. DFT theory is used to determine the energy of formation of the system. The simulation systems are stable in the electric field. The contributions of partial states are also studied here. All three doped configurations are semiconductors, electron, and hole conductors with the same density. We can conclude that the presence of Pb and the electrical field are the causes for the change in the energy band, which leads to more diverse applications in electronic technology.

**Acknowledgements** This research used resources of the high-performance computer cluster (HPCC) at Thu Dau Mot University (TDMU), Binh Duong Province, Vietnam.

## References

1. Cahangirov S, Topsakal M, Aktürk E, Şahin H, Ciraci S (2009) Two- and one-dimensional honeycomb structures of silicon and germanium. *Phys Rev Lett* 102:236804
2. Aufray B, Kara A, Vizzini SB, Oughaddou H, LéAndri C, Ealet B, Le Lay G (2010) Graphene-like silicon nanoribbons on Ag(110): a possible formation of silicene. *Appl Phys Lett* 96:183102
3. Guzmán-Verri G, Lew Yan Voon L (2007) Electronic structure of silicon-based nanostructures. *Phys Rev B* 76:075131
4. Aghaei SM, Monshi MM, Calizo I (2016) A theoretical study of gas adsorption on silicene nanoribbons and its application in a highly sensitive molecule sensor. *RSC Adv* 6:94417–94428
5. Kamal C (2012) Controlling band gap in silicene monolayer using external electric field. *Mesoscale Nanoscale Phys* 13:1–10
6. Pang Q, Zhang Y, Zhang J-M, Jib V, Xuc K-W (2011) Electronic and magnetic properties of pristine and chemically functionalized germanene nanoribbons. *Nanoscale* 3:4330–4338

# Optimal Solution of Economic Emission Load Dispatch Using a Hybrid Technique



Gutta Prem Kumar, G. V. Nagesh Kumar, and Vempalle Rafi

## 1 Introduction

In a huge integrated network of power systems, the operation of generating units has become pivotal. Deciding on real-time problems multi-objective is essential. The multi-objective model is significantly different compared with the mono-objective model. Economic load dispatch concern with a single objective problem related to a diminishing of fuel value. The goal of economic dispatch (ED) is to scrutinize the production of generation from power units to encounter the load demand fulfilling curtailment of inequality and equality. Enhancement of awareness in the public about environmental issues and safety in 1990 passage amendment concern to clean air act [1] pressurized utilities to concentrate on the strategies of operation of the thermal power unit with the reduction in atmospheric emissions. Minimization of emissions like  $\text{SO}_2$  and  $\text{NO}_x$  leads to consideration of another objective which opens on to the development of multi-objective economic emission dispatch (MEED) [2].

A power system has collaborated with many optimization problems. Among them, MEED is crucial and desirable for cost benefits as well as the reduction in emission to avoid the contamination of the atmosphere. From the aspect of economic, valve point loading [3] creates inequality in the cost functions, in input and production aspects as well as ripples in the heat rate curve. MEED is a highly nonlinear and nonconvex optimization problem. Over the past decade, many techniques have been applied for the optimal value of the MEED optimization problem. In classical techniques, MEED transfigured into a mono-objective problem [4, 5], but the drawback fails in providing particulars related to the trade-off curve. Addressing two objectives simultaneously has fascinated many researchers due to sophisticated evolutionary

---

G. Prem Kumar (✉) · G. V. Nagesh Kumar · V. Rafi  
Department of Electrical Engineering, JNTUA College of Engineering, Pulivendula,  
Andhra Pradesh 516390, India  
e-mail: [guttapremkumar247@gmail.com](mailto:guttapremkumar247@gmail.com)

methods. Some of them are non-dominant sorting genetic algorithm-II [6], strength Pareto evolutionary algorithm (SPEA) [7], niched Pareto genetic algorithm [8], Tabu search [9], and multi-objective particle swarm optimization [10]. In recent periods, from the PSO point of view, many approaches handle the multi-objective problems in achieving the global solution.

The conventional PSO suffers from struck of local optima in the search space. Many optimization techniques as their dos and don'ts. To enhance the performance as well as more computational efficiency in searching space, hybridization was focused. The hybrid method concerns the concept of integration of two or more algorithms. Combined algorithms are modified evolutionary programming and sequential quadratic programming (EP/SQP) [11], augmented Lagrange multiple with self-adaptive differential evolution (ALM/SADE) [12], fuzzy cluster-based swarm optimization (FCPSO) [13], genetic algorithm and evolutionary programming (GA/EP) [14], non-dominant sorting genetic algorithm-II, and population variant differential evolution (NSGA-II/PVDE) [15]. In this paper, a novel hybrid method is applied which is the combination of PSO and TSA techniques. To achieve better performance, PSO focuses on inspecting the total search space globally and Tabu search exploiting the below space nothing but a subspace.

The rest of the paper is ordered as follows. Section 2 deals with the problem formulation, Sect. 3 presents the methodology applied in solving optimization problem which is a multi-objective, results and its discussion are provided in Sect. 4. Section 5 draws out a conclusion.

## 2 Formulation of Multi-objective Optimization Problem

### 2.1 Cost Function

The cost function is indicated in quadratic form with its cost coefficients. With the incorporation of the valve point loading effect, linear function is transfigured into nonlinear and non-smooth convex function. The second term  $\sin$  represents the inclusion of the valve point effect.

$$F_c = F_o(P_{Go}) = a_o + b_o P_{Go} + c_o P_{Go}^2 + |e_o(\sin(f_o(P_{Gomin} - P_{Go})))| \quad (1)$$

where  $a_o$ ,  $b_o$ ,  $c_o$ , are the power unit 'o' coefficients of cost, and  $e_o, f_o$  are constants of the cost function of generating unit 'o' reflecting the effects of valve point.  $P_{Gomin}$  represents the low power generation of unit 'o'.

SO<sub>2</sub> and NO<sub>x</sub> are the atmospheric pollutants liberated from the combustion of fossil fuels during the operation of thermal units which is mathematically modeled as the combination of quadratic form with an exponential term which is shown in Eq. (2).



$$E_c = E(P_{Go}) = \sum_{o=1}^n (\alpha_o + \beta_o P_{Go} + \gamma_o P_{Go}^2) + \xi_o \exp(\lambda_o P_{Go}) \tag{2}$$

where  $\alpha_o, \beta_o, \gamma_o, \xi_o$  and  $\lambda_o$  are the ‘o’ generator emission coefficients,  $P_{Go}$  is the real power of the generating unit ‘o’.  $E$  represents the total emission of the thermal plant. Multi-objective dispatch problem converted into a single-objective problem using price penalty factor.

$$T_c = a_o + b_o P_{Go} + c_o P_{Go}^2 + |e_o(\sin(f_o(P_{Gomin} - P_{Go})))| + h_k [(\alpha_o + \beta_o P_{Go} + \gamma_o P_{Go}^2) + \xi_o \exp(\lambda_o P_{Go})] \tag{3}$$

Four different combinations of price penalty factor of power unit ‘o’ presented. Min/min, min/max, max/max, and max/min are presented in Eqs. (4)–(7).

$$h_{o,min / max} = \frac{(a_o P_{o,min}^2 + b_o P_{o,min} + c_o)}{(d_o P_{o,max}^2 + e_o P_{o,max} + f_o)} \tag{4}$$

$$h_{o,max / min} = \frac{(a_o P_{o,max}^2 + b_o P_{o,max} + c_o)}{(d_o P_{o,min}^2 + e_o P_{o,min} + f_o)} \tag{5}$$

$$h_{o,max / max} = \frac{(a_o P_{o,max}^2 + b_o P_{o,max} + c_o)}{(d_o P_{o,max}^2 + e_o P_{o,max} + f_o)} \tag{6}$$

$$h_{o,min / min} = \frac{(a_o P_{o,min}^2 + b_o P_{o,min} + c_o)}{(d_o P_{o,min}^2 + e_o P_{o,min} + f_o)} \tag{7}$$

### 2.2 Constraint of Equality

In this equality constraint, total real power generation of all thermal units in the power plant has to meet the load demand. The summation of active power has to equate to the load demand.

$$\sum_{o=1}^n P_{Go} = P_D \tag{8}$$

### 2.3 Constraint of Inequality

The generation of active power has to be maintained within the limits of low and high values of power generation which described the inequality constraint.

$$P_{Gomin} \leq P_{Go} \leq P_{Gomax} \tag{9}$$

where  $P_{Gomin}$  and  $P_{Gomax}$  are the lower power and largest power generation of thermal unit ‘o’.

## 3 Methodology

### 3.1 Particle Swarm Optimization

In 1995, Eberhart and Kennedy introduced an optimization technique, particle swarm optimization which belongs to the category of swarm intelligence. This stochastic technique is a population-based one and is stimulated by the behavior of birds swarming or fish schooling. PSO was independent of genetic operators like mutation and crossover. PSO algorithm is efficient for multi-dimensional spaces of search and painless implementation with effective computational efficiency and good characteristics of convergence. In this method, an optimal solution is predicted in a search space with the prescribed number of particles with modernized generations.

Particle insists a bird in a bird flock which prescribes every single solution. Every particle is associated with a certain velocity and is affected by social behavior and cognitive at the period of exploration. PSO is a high-potential operative for multi-objective, nonlinear, multi-model, as well as for complex problems. Based on the performance, alteration of positions of the particles takes place and no peculiar tuning of the parameter is required. In evaluating the local optima exchange of flying information exist among all the particles. Based on the pertinent information tribes explore auspicious areas in finding the global optima.

Modernization of generations at each iteration is implemented by two parameters, namely pbest and gbest. Initially, each particle’s position and velocity are initialized. In a search space having n-dimensional the velocity and position of a ‘w’ particle are transfigured as a vector of  $P_w = \{P_{w1}, P_{w2}, P_{w3} \dots P_{wm}\}$  and respective velocity is  $V_w = \{V_{w1}, V_{w2}, V_{w3} \dots V_{wm}\}$ . Let  $gbest = \{P_{1best}, P_{2best}, P_{3best} \dots P_{nbest}\}$  and  $pbest_w = \{P_{w1best}, P_{w2best}, P_{w3best} \dots P_{wnbest}\}$  be the neighbor best and particle ‘w’ best positions, respectively.

$$\begin{aligned} V_w^{o+1} &= K(\omega * v_w^o + C_{a1} * r_1 * (P_{bestw} - P_w^o) + C_{a2} * r_2 * (G_{bestw} - P_w^o)) \\ P_w^{o+1} &= P_w^o + V_w^{o+1} \end{aligned} \tag{10}$$

where  $P_w^o$  and  $v_w^o$  present position and velocity of 'w' particle at wth iteration, ' $\omega$ ' indicates the weight of inertial,  $C_{a1}$  and  $C_{a2}$  represents the coefficients of acceleration,  $r1$  and  $r2$  is a random number. Inertia weight is expressed as

$$\omega = \omega_{\max} - (\omega_{\max} - \omega_{\min}) * \text{iter}/\text{maxiter};$$

maxiter is a large iteration value, iter is the current iteration value.  $\omega_{\max}$ ,  $\omega_{\min}$  are the maximum, minimal inertial weight, acceleration coefficients are predicted as follows

$$\begin{aligned} C_{a1} &= C_{1\max} - (C_{1\max} - C_{1\min}) * \text{iter}/\text{maxiter}; \\ C_{a2} &= C_{2\max} - (C_{2\max} - C_{2\min}) * \text{iter}/\text{maxiter}; \\ \theta &= (C_{a1} + C_{a2}); \end{aligned} \tag{11}$$

$$K = \frac{2}{|2 - \theta\sqrt{\theta^2 - 4\theta}|} \tag{12}$$

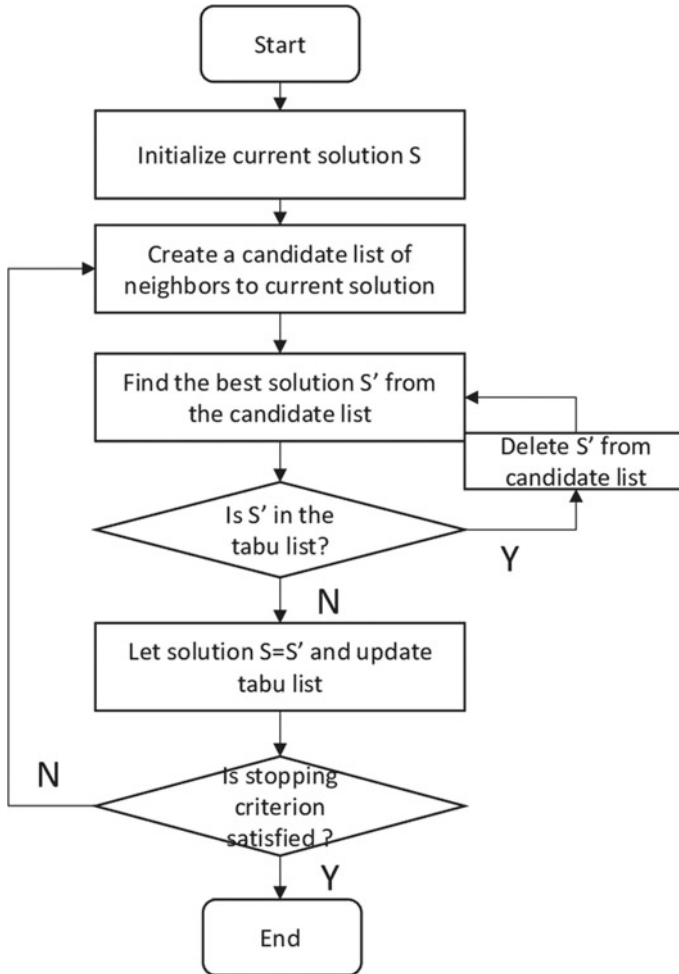
### 3.2 Tabu Search

In 1989, Glover introduced the Tabu search approach in finding the optimal value of optimization problems. Hill climbing is the root of the Tabu search algorithm (TSA) and the most satisfactory and powerful technique of optimization concerned with the combinatorial. The positive sign of TSA is to migrate from the poor solutions and desires in evaluating good solutions. Aspiration criteria and Tabu list are two variables associated with TSA. The function of the Tabu list is to create the temporary storage of fixed attributes to keep away from the newly generated solution until the existence in the Tabu list. Entry of new attributes make free existence attribute by controlling such directions leads to steer clear from the local optima.

Aspiration criteria is another pivotal component that evaluates which attribute makes free and which attribute to make allowable if the move is more economical as well as to allow banned attributes by satisfying the aspiration criteria. Figure 1 shows the flowchart which illustrates the process of Tabu search algorithm.

### 3.3 Hybrid Method

Hybridization is the process of combining two or more algorithms to enhance the performance as well as to enhance the computational efficiency with the improved convergence characteristics. In the proposed hybrid method, combination of PSO and TS methods are applied. The procedure of hybrid method is as follows



**Fig. 1** Flow chart of Tabu search method

1. Initial parameters are initialized, i.e., population length, the maximum number of iterations, minimum and maximum values of power generation, cost and emission coefficients, and minimum and maximum velocities, Tabu list length, Tabu list, restriction period, and maximum count.
2. Generate the random population and obtain the fitness value of each population using Eq. (3). The randomly generated population has to satisfy the constraints (Eqs. 8 and 9). Half the population is subjected to PSO with random velocities, and another half of the population is concerned with TSA.
3. The velocity of the particle is modernized using Eq. (10).

4. New positions of the particles are to be evaluated and updated the pbest and gbest values are upgraded, go to step 3 until reaching stopping criteria from which a minimum best value is predicted.
5. Let the initial solution is considered by the rest of the population for TSA and evaluate the cost value. The current best solution is initialized as the best solution.
6. Create the neighbor solutions and predict the fitness values and evaluate the best neighbor solution. If the best neighbor solution is greater than the best solution, then update the Tabu list and solution go to step 7 else update with the initial solution.
7. Reset the Tabu status if recency period is greater than the restriction period and frequency greater than maximum frequency set the Tabu status direction.
8. If the count reaches the maximum value stop process, else go to step 6 with an increment in the count. Finally, the largest values eliminated from the result of PSO and TSA.

## 4 Results and Discussion

The proposed hybrid technique is tested on a test system which is a nonlinear and nonconvex optimization problem. In the MEED problem, price penalty factor and valve point, the loading effect is contemplated. Price penalty factor is considered in four ways namely low/high, high/low, max/max, and min/min and is evaluated using Eqs. (4)–(7). The parameters that are initialized, size of population 40, maximum number of iteration are 100, maximum acceleration value 0.9, and minimum acceleration value 0.4,  $C_{1\max}$  value 2.075 and  $C_{1\min}$  value 2.025,  $C_{2\max}$  value 2.025, and  $C_{2\min}$  value is 2.125.

Case (i) In this study, to predict the efficiency of the proposed technique, it is applied to a five-unit system. The data relating to emission and cost coefficients with valve point effect, generation limits, are taken [16]. Three different loads are considered 450, 650, and 850 MW, and the price penalty factor with four ways as well as valve point loading effect is implemented for the three loads. The corresponding scheduling of thermal units, as well as total cost values with 450, 650, 850 (MW) with low/high, high/low, high/high, and min/min price penalty factor, is shown in Tables 1, 2 and 3.

Table 4 shows the comparison between the hybrid method and BSA and conventional method PSO. With the hybrid method, an improved cost value is achieved for three different demand loads. Among different price penalty factors, minimum values of cost and emission are achieved with the low/high penalty factor, and the maximum values of cost and emission are obtained with high/low penalty factor. Figures 2 and 3 show the decay in the total cost value with an increase in the number of iteration. With low/high penalty factor achieved the fuel price value is 1565956.923 (\$/h) and

**Table 1** Comparison of cost values with different price penalty factors for a load of 450 MW

Load	Load demand PD = 450 MW			
	Min/Max	Min/Min	Max/Max	Max/Min
PG1 (MW)	75	75	75	75
PG2 (MW)	104.162	104.162	104.162	103.995
PG3 (MW)	62.737	62.737	62.737	62.969
PG4 (MW)	52.105	52.042	52.042	48.900
PG5 (MW)	155.994	156.057	156.057	159.134
Total cost (\$/h)	1,607,654.413	3,133,804.052	3,131,326.415	61,091,229.740

**Table 2** Comparison of cost values with different price penalty factors for a load of 650 MW

Load	Load demand PD = 650 MW			
	Min/Max	Min/Min	Max/Max	Max/Min
PG1 (MW)	74.999	75	75	75
PG2 (MW)	124.999	125	125	125
PG3 (MW)	104.388	104.387	104.387	104.562
PG4 (MW)	86.539	86.382	86.382	82.161
PG5 (MW)	259.072	259.230	259.230	263.275
Total cost (\$/h)	3,722,858.022	7,252,969.814	7,248,282.808	138,014,913.268

**Table 3** Comparison of cost values with different price penalty factors for a load of 850 MW

Load	Load demand PD = 850 MW			
	Min/Max	Min/Min	Max/Max	Max/Min
PG1 (MW)	75	75	75	75
PG2 (MW)	125	125	125	125
PG3 (MW)	175	175	175	175
PG4 (MW)	170.463	170.577	170.414	170.666
PG5 (MW)	304.536	304.422	304.585	304.333
Total cost (\$/h)	7,338,560.787	14,297,805.311	14,286,505.776	271,044,474.883

**Table 4** Comparison of cost values of the hybrid method with other methods for different loads

Load	Min/max cost penalty factor		
	BSA	PSO	PSO-TSA
450 (MW)	2,075,700.00	2,075,900.00	1,607,654.413
650 (MW)	4,325,900.00	4,326,300.00	3,722,858.022
850 (MW)	7,396,900.00	7,397,100.00	7,338,560.787

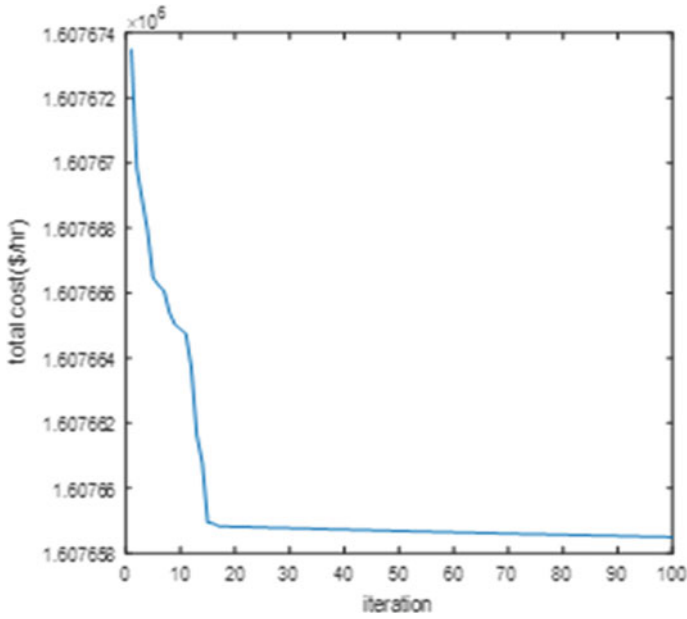


Fig. 2 Cost versus iteration with 450 MW

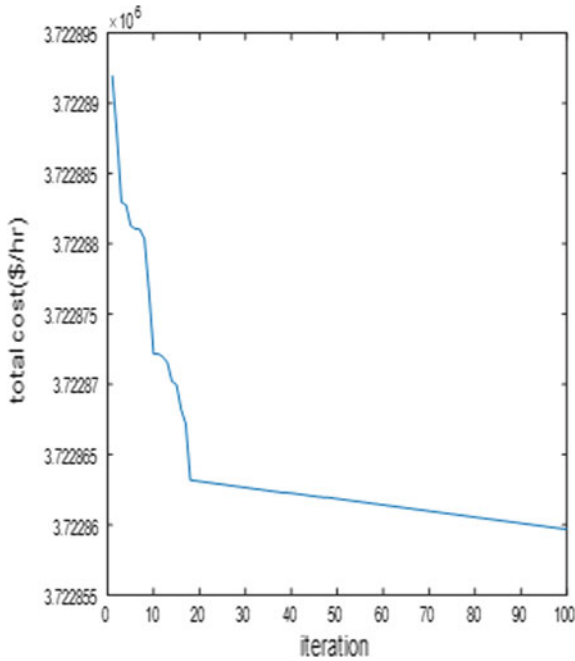


Fig. 3 Cost versus iteration with 650 MW

the emission value 1,310,755.196 (\$/h) at 450 MW load. Respective cost and emission values for a load demand of 650 MW and 850 MW are 3,624,414.063 (\$/h), 2,668,235.407 (kg/h) and 7,143,845.158 (\$/h), 4,509,745.300 (kg/h).

The potential of the hybrid method illustrates the minimization in the running cost, emission values as well as in total price for different cost penalty factors at different load demands. Scheduling of thermal units in load sharing within the generation limits is maintained and equating the load demand resembling the inequality and equality constraints. The hybrid method enhances the performance as well as the efficiency of computing the multi-objective optimization problem which is contradictory.

## 5 Conclusion

The hybrid method is proposed to predict the optimal solution of the multi-objective economic emission load dispatch optimization problem. Nonlinear and nonconvex objective function with cost and emission as objectives solved by min/max, max/min, max/max, and min/min price penalty factor using PSO and TSA. To illustrate the potential of the hybrid method, test on a test system with five thermal units, and results are evaluated for different penalty factors. Comparison with different penalty factors illustrates cost, emission, and total cost value are minimum with min/max penalty factor rather than compared with other penalty factors.

## References

1. El-Keib AA, Ma H, Hart JL (1994) Economic dispatch in view of the Clean Air Act of 1990. *IEEE Trans Power Syst* 9(2):972–978
2. Vempalle R, Dhal PK (2021) Optimal location of distributed generators for maximum savings using PSO-SSA optimization algorithm. In: 2021 11th international conference on cloud computing, data science & engineering (confluence), pp 624–629
3. Singh NJ, Dhillon JS, Kothari DP (2016) Synergic predator-prey optimization for economic thermal power dispatch problem. *Appl Soft Comput* 43:298–311
4. Palanichamy C, Babu NS (2008) Analytical solution for combined economic and emissions dispatch. *Electr Power Syst Res* 78(7):1129–1139
5. Dhillon JS, Parti SC, Kothari DP (1993) Stochastic economic emission load dispatch. *Electr Power Syst Res* 26(3):179–188
6. Navaneetha C, Nagesh Kumar GV, Rafi V (2021) Reactive power sharing in Islanded microgrid by using voltage recovery operation of droop control strategy. In: 2021 international conference on intelligent technologies (CONIT), pp 1–7
7. Rajesh K, Visali N (2020) Tradeoff curve of an economic emission load dispatch using NSGA-II and PVDE. *ARPN J EngAppl Sci* 15(1)
8. Abido MA (2003) A niched Pareto genetic algorithm for multiobjective environmental/economic dispatch. *Electric Power Energy Syst* 25(2):97–99
9. Li F, Morgan R, Williams D (1997) Hybrid genetic approaches to ramping rate constrained dynamic economic dispatch. *Electr Power Syst Res* 43:97–103
10. Victoire TAA, Jeyakumar AE (2005) A modified hybrid EP-SQP approach for dynamic dispatch with valve-point effect. *Int J Elect Power Energy Syst* 27(8):594–601



11. Thitithamrongchai C, Eua-Arpon B (2006) Self-adaptive differential evolution method with augmented lagrange multiplier for power economic dispatch of units with valve-point effects and multiple fuels. IEEE
12. Padma R, Vali SH, Rafi V (2021) Reducing fault current by adaptive stabilizer in distribution system. In: 2021 international conference on intelligent technologies (CONIT), pp 1–9
13. Patel R, Arora K, Katiyar S (2016) Control analysis of PMSG based wind energy conversion system using buck-boost converter. Second Int Conf Comput Intel Commun Technol (CICT) 2016:395–402
14. Li RM, Williams D (1997) Hybrid genetic approaches to ramping rate constrained dynamic economic dispatch. *Electr. Power Syst Res* 43: 97–103
15. Rajesh K, Visali N (2020) Tradeoff Curve of an economic emission load dispatch using NSGA-II and PVDE. *ARNP J Eng Appl Sci* 15(1)
16. Krishnamurthy S, Tzoneva R (2011) Comparative analyses of Min-Max and Max-Max price penalty factor approach for multi-criteria power system dispatch problem with valve point effect loading using Lagrange's method. In: 2011 International conference on power and energy systems, pp 1-7

# Solution of Unit Commitment Using Genetic Algorithm with Population Refreshment



Gundra Bhargavi, G. V. Nagesh Kumar, and Vempalle Rafi

## 1 Introduction

In the operation of a power system, the common problem often concerns the scheduling of thermal units. According to energy management, the generation of real power has to encounter the growing load demand with small-scale utilization of fuel cost [1]. The pivotal, first, and foremost thing is evaluating the scheduling operation of each thermal unit in the power system. The task of optimization enhanced significantly because of the insufficiency of energy. There are many courses of action to perform the planning operation of generators to equate the load demand. Among them, unit commitment (UC) is one such course of action that provides more flexibility in the solvent operation of the power network. Predicting the on/off commitments of power units in an order of sequence over a period of time is termed unit commitment. Solving optimization problem concern to unit commitment enumerates the scheduling of thermal unit over a time period and also minimizes the generation fuel cost. Incorporation of the security factors can be termed as security constraint unit commitment [2].

The complexity of the problem enhances equivalently with the enlargement in the size of the system. Consequently, there is an increase in the importance of unit commitment. UC entails predicting the shutdown and start-up schedule of committed units to satisfy load demand over a time period [3]. UC is associated with two performances, namely scheduling of committed unit as well as economic dispatch performance. Solvent dispatch is a nonlinear, non-convex, and non-smooth optimization problem. The optimal solution of the UC optimization problem evaluates the generation of thermal units for each hour and start-up and shut down constraints of running

---

G. Bhargavi (✉) · G. V. Nagesh Kumar · V. Rafi  
Department of Electrical Engineering, JNTUA College of Engineering, Pulivendula, Andhra Pradesh 516390, India  
e-mail: [gundra.bhargavilatha@gmail.com](mailto:gundra.bhargavilatha@gmail.com)

units. UC predicts the scheduling of power units optimally with the aim of reduction in cost generation by perpetuating the system constraints.

UC is a mixed-integer combinational problem of optimization, high-dimensional, and nonlinear. Over a decade, many optimization techniques were implemented for achieving an optimal value of a UC optimization problem. Based on the availability of literature, classical and stochastic are classified. Priority list method [4], Lagrangian relaxation [5], Branch bound method [6], and Dynamic Programming method [7] concern the classical optimization techniques. The drawback associated with the classical techniques is miniature; UC problems are easy to solve, as the complexity of the problem enhances the rate of convergences gets deteriorates which is unsuitable. Stochastic techniques are heuristic in nature. Evolutionary programming [8], particle swarm optimization [9], differential evolution [10], and bacterial foraging algorithm [11] are related to the stochastic techniques.

These heuristic techniques are applicable for the large and medium networks of the power system. In this paper, heuristic method, a genetic algorithm, is applied with the refreshment of parent population and tested on a test system IEEE 39 bus system with ten thermal units. To explicate the accuracy of the proposed method, results are contrasted with the existing literature methods. Accuracy and efficacy with refreshment population are better in terms of fuel cost. The genetic algorithm with population refreshment affirms the reliability and efficiency of complex UC optimization problems. The rest of the papers are as follows: Sect. 2 deals with problem formulation of the unit commitment optimization problem. Section 3 presents the methodology in solving the UC problem, Sect. 4 discusses the results that are obtained with the heuristic method, and finally, Sect. 5 presents the conclusion.

## 2 Problem Formulation

The total cost is incorporated with fuel cost and start-up cost, and the objective function of total cost minimization is represented in Eq. (1)

$$\min \sum_{w=1}^M \sum_{z=1}^T [F_w(\text{Pg}_w(z)) + \text{SUC}_w(1 - U_w(z-1))]U_w(z) \quad (1)$$

where  $\text{Pg}_w$  represents the active power of 'w' unit,  $U_w$  indicates the 'w' power unit on/off status,  $T$  is time in hours equivalent to 24, and the total number of generators is represented as 'M'. The quadratic form of the cost function is represented as

$$F_w(\text{Pg}_w(z)) = a_w + b_w \text{Pg}_w(z) + c_w \text{Pg}_w^2(z) \quad (2)$$

$F_w$  is the cost function,  $a_w$ ,  $b_w$ ,  $c_w$  are the cost coefficients of thermal unit 'w'.  $\text{Pg}_w(z)$  is the real power of unit 'w' at hour 'z.' The respective start-up cost of unit 'w' is

expressed as

$$SUC(w) = \begin{cases} HC(w), & \text{if } MD(w, z) \leq TC(w) \leq MD_{zo} \\ CSC(w), & \text{if } TC(w) > MD_{zo} \end{cases} \quad (3)$$

where  $MD_{zo} = MD(w, z) + CST(w)$ .

$CSC(w)$  is the cold start-up cost of thermal unit ‘w,’  $HC(w)$  is the hot start-up cost,  $MD(w)$  is the minimum downtime,  $TC(w)$  indicates the off duration of power unit ‘w.’  $CST(w)$  is the cold start-up time of unit ‘w.’ The UC optimization problem is associated with inequality and equality constraints.

### 2.1 Equality Constraint

The total power generation from the thermal units must meet the requirement of load demand which is termed as equality constraint. At any instant of time, the total load must be equated with the real power generation of power units. The mathematical representation is represented as

$$\sum_{w=1}^M Pg_w(z) \cdot U_w(z) = d_o(z) \quad (4)$$

where  $d_o$  is the initial load demand in megawatt at hour.

### 2.2 Spinning Reserve Constraint

The term spinning reserve indicates additional supply has to provide from the generating units to avoid the shortages that are concern with the loss of units [12]. The maintenance of reserve capacity is required. The mathematical equation is represented as

$$\sum_{w=1}^M Pg_w(z) \cdot U_w(z) \geq d_o(z) + SR(z) \quad (5)$$

$SR(z)$  is the spinning reserve at hour ‘z.’

### 2.3 Power Generation Limits Constraints

The active power generated from each thermal unit has to maintain within the maximum and minimum power generation limits. This constraint is related to the inequality constraint.

$$Pg_{w,\min} \leq Pg_w \leq Pg_{w,\max} \quad (6)$$

where  $Pg_{w,\min}$  is the minimum power limit generation of thermal unit 'w,' and  $Pg_{w,\max}$  indicates the maximum power generation limit of 'w' thermal unit.

### 2.4 Minimum Down and Uptime

The minimum time that is required for the commitment and decommitment and vice versa is indicated as a minimum up and downtime.

$$X_{w,\text{off}} \geq MDT_w \quad (7)$$

$$X_{w,\text{on}} \geq MUT_w \quad (8)$$

where  $X_{w,\text{off}}, X_{w,\text{on}}$  represents the off/on status of thermal unit 'w,' and  $MDT_w, MUT_w$  are the minimum downtime and minimum uptime of thermal unit 'w'.

## 3 Methodology

### 3.1 Genetic Algorithm

In 1975, genetic algorithm was developed by John Holland and a publicized by Goldberg. It is based on the natural selection and mechanism of genetics. To obtain a superior optimal value of an optimization problem, GA explores the search space in predicting the fittest individuals based on the genetic operators. They are three main genetic operators are associated with GA, namely crossover, selection, and mutation. Based on the analysis of the optimization problem, the GA is chosen. The downside associated with binary GA, enhancement in the problem dimension delays the precise and convergence characteristics.

GA is population-based which is incorporated with chromosomes. Each chromosome is the combination of bits 0's and 1's, and each bit is termed as 'gene.' GA is based on the fundamental concept of a healthy continuance solution in a specified

population. The algorithm commences from an approximate and finally reaches the near solution. Based on the strength of chromosomes, best solutions are repositied during every generation of newly population. Genetic operators play a pivotal role in evaluating the optimal solution to an optimization problem.

### **Selection Operator**

The initial operator concerned to the GA is selection operator. It is a procedure in predicting the fittest population for the process of reproduction. The possibility for the participation of reproduction for the next generation will be eventually high for the fittest population. Fitness is termed as strength of the individual population with the available sources. Populations are selected for the next generation based on the performance for a given optimization problem. The main criteria of the operator is to eliminate the unfit population and to continue with high strength population with the maintenance of population size.

### **Crossover**

Crossover is the swapping of information between two parent populations for the production offspring. Crossover is the pivotal among the three genetic operators of GA. New chromosomes are generated with both properties of two parent population. To enhance the diversity locally for each individual binomial, mutual crossover is implemented. The set value of the crossover is chosen between 0 and 1, i.e., a random number for achieving an optimal value of an optimization problem.

### **Mutation**

The most essential operator of genetic operators is mutation. It is the procedure, exchange of gene randomly in chromosome. The set value of the mutation operator is selected a very low value. Mutation operator helps in forming a massive change in a newly generated population, i.e., a modified structure for each population. The range of mutation value is selected between 0 and 1. Mutation produces analogous structure with a little specified probability by changing the portion of individual structure randomly.

## ***3.2 Procedure of Applying GA with Refreshed Population***

The step involved in solving unit commitment optimization problem with population refreshment is as follows. With reference to [13], the population variant is implemented in the initial stage.

1. Explicate the fitness functions with incorporation of fuel cost function and start-up cost function.
2. Initialize the population size satisfying the constraints, crossover rate, mutation rate, and encode the objective functions.

3. Predict the strength of each individual population and refresh the population. The refreshment of population is as follows
  - (a)  $B_{\min} = 0.5 * (Pg_{w,\max} - Pg_{w,\min})$
  - (b) Evaluate the median of  $Pg_{w,\max}$  and  $Pg_{w,\min}$
  - (c) Predict the average value of median  $Pg_{w,\min}$  and median of  $Pg_{w,\max}$
  - (d) If average value  $< B_{\min}$ 
    - Lower =  $Pg_{w,\min} + \beta * B_{\min}$ ;
    - Higher =  $Pg_{w,\max} - \beta * B_{\min}$ ;
  - (e) Generate the new parent population with modified generation limits in a search space.
4. Evaluate the fitness of the new parent population and perform the genetic operators for the generation of offspring's.
5. Combine the parent and child population and select the best fitness population of initial size.
6. Continue the process until the final iteration of criteria is achieved and obtain the optimal value of cost minimization optimization problem.

## 4 Results and Discussion

The potential of the GA with population refreshment in the initial stage is applied and tested on IEEE 39 bus system with ten thermal units. The initial population size is considered as 40, crossover rate 0.9, and mutation rate 0.05. Maximum and minimum power generation limits, emission coefficients, cost coefficients as well as load demand are considered [14] and maximum number of generation are 100. Based on this parameters, the optimal value of minimization of cost function is obtained for over 30 trails to confirm the strength GA with refreshed population. The value of beta is selected such that the variation in limits results better optimal value. The load demand with time horizon of one hour for period of 24 h is shown in Table 1.

### Case study

In this case, the heuristic algorithm is applied on IEEE 39 bus system with 10 thermal units. The tentative on/off status of the generating units are obtained based

**Table 1** Load demand on hourly base

Hour	1	2	3	4	5	6	7	8	9	10	11	12
Demand (MW)	700	750	850	950	1000	1100	1150	1200	1300	1400	1450	1500
Hour	13	14	15	16	17	18	19	20	21	22	23	24
Demand (MW)	1400	1300	1200	1050	1000	1100	1200	1400	1300	1100	900	800

on proposed method GA with population refreshment which is implemented to evaluate the economic dispatch. The simulation is performed on MATLAB software. The committed thermal units are scheduled for different 24 loads with time horizon of one hour is shown in Table 2. The generated power of all thermal units are maintained within the generation limits, i.e., inequality constraint. The start-up cost and the running cost value of all thermal units for different load demand is shown in Table 3.

The highest and least running cost are obtained at loads of 900 and 1500 MW. In comparison with existing methods concerned to GA, the proposed method shows a better reduction in the total cost value which is shown in Table 4. Compared with binary coded genetic algorithm, the reduction in the total cost value with the proposed method is 2.22% and with modified genetic algorithm is 0.64%. During the refreshment of the population, the generation limits get modified as a consequence;

**Table 2** Unit commitment of thermal power units for different load demands

Pg1	Pg2	Pg3	Pg4	Pg5	Pg6	Pg7	Pg8	Pg9	Pg10
1	1	0	0	0	0	0	0	0	0
1	1	0	0	0	0	0	0	0	0
1	1	0	0	0	0	0	0	0	0
1	1	0	0	1	0	0	0	0	0
1	1	0	0	1	0	0	0	0	0
1	1	1	0	1	0	0	0	0	0
1	1	1	0	1	0	0	0	0	0
1	1	1	1	1	0	0	0	0	0
1	1	1	1	1	0	1	0	0	0
1	1	1	1	1	1	1	0	0	0
1	1	1	1	1	1	1	0	1	0
1	1	1	1	1	1	1	0	0	0
1	1	1	1	1	0	0	0	0	0
1	1	1	0	1	0	0	0	1	0
1	1	1	0	1	0	0	0	0	0
1	1	1	0	1	0	0	0	0	0
1	1	1	0	1	0	1	0	0	0
1	1	1	0	1	0	1	0	0	0
1	1	1	1	1	0	1	0	0	0
1	1	1	1	1	0	0	0	0	0
1	1	1	1	0	0	0	0	0	0
1	1	0	1	0	0	0	0	0	0
1	1	0	1	0	0	0	0	0	0



**Table 3.** Scheduling of thermal units for different load demands

Pg1 (MW)	Pg2 (MW)	Pg3 (MW)	Pg4 (MW)	Pg5 (MW)	Pg6 (MW)	Pg7 (MW)	Pg8 (MW)	Pg9 (MW)	Pg10 (MW)	SUC	Running cost
455	245	0	0	0	0	0	0	0	0	0	13,683.129
455	295	0	0	0	0	0	0	0	0	0	14,554.499
454.999	395	0	0	0	0	0	0	0	0	0	16,301.889
455.000	454.999	0	0	40	0	0	0	0	0	900	18,597.667
454.999	454.999	0	0	90	0	0	0	0	0	0	19,608.537
455	455	129.450	0	60.549	0	0	0	0	0	1100	21,893.111
454.999	454.999	129.449	0	110.550	0	0	0	0	0	0	22,912.160
454.999	454.999	129.449	129.449	31.100	0	0	0	0	0	1120	24,153.486
455	455	129.450	129.450	131.100	0	0	0	0	0	0	26,188.042
455	455	129.450	129.450	162	0	69.100	0	0	0	520	29,233.424
454.999	454.999	129.449	129.449	161.999	79.699	39.400	0	0	0	340	30,596.349
455	455	129.450	129.450	162.000	79.700	34.625	0	54.774	0	60	32,628.986
454.999	454.999	129.449	129.449	161.999	43.800	25.300	0	0	0	0	29,373.793
455	455	129.450	129.450	131.100	0	0	0	0	0	0	26,188.042
455	455	129.450	0	150.325	0	0	0	10.225	0	60	24,681.093
455	440.550	129.449	0	25	0	0	0	0	0	0	20,927.253
455	390.550	129.450	0	25	0	0	0	0	0	0	20,051.371
454.999	454.999	129.449	0	35.249	0	25.300	0	0	01	260	22,567.382
454.999	455	129.449	0	135.249	0	25.300	0	0	0	0	24,605.241
455	455	129.450	129.450	162	0	69.100	0	0	0	560	29,233.424
455	455	129.450	129.450	131.100	0	0	0	0	0	0	26,188.042

(continued)

**Table 3** (continued)

Pg1 (MW)	Pg2 (MW)	Pg3 (MW)	Pg4 (MW)	Pg5 (MW)	Pg6 (MW)	Pg7 (MW)	Pg8 (MW)	Pg9 (MW)	Pg10 (MW)	SUC	Running cost
455	386.100	129.450	129.450	0	0	0	0	0	0	0	21,879,788
455	315.550	0	129.450	0	0	0	0	0	0	0	17,764.365
454.999	215.550	0	129.449	0	0	0	0	0	0	0	16,021.901
										<b>4920</b>	<b>549,832,984</b>

**Table 4** Comparison of total cost value with different GA methods

Method	Binary coded GA	ICGA	IICGA	IGA	DMHGA	MGA	Proposed method
Total cost	5,67,367	5,66,404	5,55,421.7	5,63,931	5,56,760	5,58,341.43	5,54,752.9846

there is reduction in the search space. Over 30 trials are made, and at each trial is commenced with different initial population.

## 5 Conclusion

The population refreshment with GA is applied for solving UC optimization problem which is concerned with ten thermal units considering cost minimization as main criteria. Variation of generation limits results a modification of search space which is more essential in achieving a better optimal value of cost minimization. The potential of the proposed method evaluated in achieving an optimal value for reduction in total cost and made a comparison with the existing related GA methods. For the production of optimal value, equilibrium constraints and inequality constraints are taken into account. The deviation in the total cost value with the proposed method is 0.642%.

## References

1. Victoire TA, Jeyakumar AE (2005) Unit commitment by a tabu-searchbased hybrid-optimization technique, *IEE Proc. Gener Transm Distrib* 152(4):563–574
2. Shahidepour M, Yamin H, Li Z (2002) *Market operations in electric power systems*. Wiley, New York
3. Wood AJ, Wollenberg BF (1996) *Power generation, operation, and control*, 2nd edn. Wiley, New York
4. Senjyu T, Shimabukuro K, Uezato K, Funabashi T (2003) A fast technique for unit commitment problem by extended priority list. *IEEE Trans Power Syst* 18(2):882–888
5. Singhal (2011) Generation scheduling methodology for thermal units using lagrangian relaxation. In: *Proceedings of 2nd IEEE international conference current trends in technology*, pp 1–6
6. Chen CL, Wang SC (1993) Branch-and-bound scheduling for thermal generating units. *IEEE Trans Energy Convers* 8(2):184–189
7. Singhal, Sh N (2011) Dynamic programming approach for solving power generating unit commitment problem. In: *Proceedings of IEEE international conference on computer and communication technology*, pp 298–303
8. Tanaka, Hasegawa (1999) An evolutionary programming solution to the unit commitment problem. *IEEE Trans Power Syst* 14(4):1452–1459
9. Wang, Yuan (2009) An improved binary particle swarm optimization for unit commitment problem. *Exp Syst Appl* 36:8049–8055

10. Vempalle R, Dhal PK, (2020) Loss minimization by reconfiguration along with distributed generator placement at radial distribution system with hybrid optimization techniques. *Technol Econ Smart Grids Sustain Energy* 5
11. Rafi V, Dhal PK (2020) Maximization savings in distribution networks with optimal location of type-I distributed generator along with reconfiguration using PSO-DA optimization techniques. *Mater Today: Proc* 33:4094
12. Swain R, Sarkar P, Meher KC, Chanda CK (2017) Population variant differential evolution based multiobjective economic emission load dispatch. *Int Trans Electr Energy Syst* 27(10):1–25
13. Selva K, Vijayakumar K, Boopathi CS (2017) Demand response unit commitment problem solution for maximizing generating companies profit. *Energies* 10:1465
14. Nageswara Rao P, Siva Krishna Rao GV, Nagesh Kumar GV (2018) A novel technique for controlling speed and position of bearingless switched reluctance motor employing sensorless sliding mode observer. *Arab J Sci Eng* 43(8):4327–4346

# Modelling Series RLC Circuit with Discrete Fractional Operator



George E. Chatzarakis, A. George Maria Selvam, R. Janagaraj,  
and D. Vignesh

## 1 Introduction

In applied mathematics, developing a model setup is an important task and which is vital in understanding the process in wide range of fields like physics, chemistry, evolutionary models, viscoelasticity, electrical and electronic communications and so on. Setting up an electrical circuit model is beneficial to researchers of almost all fields as it can be used to realize any systems of engineering like hydraulic thermal, prediction of weather, traffic flow, nuclear and mechanical systems and so on. An electrical circuit can be referred to as an electronic circuit if at least active elements are present in the circuits.

Fractional calculus proves to be exceptional medium to describe the hereditary properties of different processes. Although the field of non-integer order calculus came into existence back in 1695, it was in recent times with development of computers and simulation softwares the importance in modelling more complex and physical problem is realized [1, 2]. Some recent applications of modelling conduction of heat through a semi-infinite solid [3] and impedance of transmission lines are calculated numerically by [4]. The memory and hereditary effects of fractional order equations are one of the most important aspects of modelling any processes [5].

---

G. E. Chatzarakis

Department of Electrical and Electronic Engineering Educators, School of Pedagogical and Technological Education (ASPETE), Marousi, 15122 Athens, Greece

A. George Maria Selvam (✉) · D. Vignesh

PG & Research Department of Mathematics, Sacred Heart College (Autonomous), Tirupattur, Tamil Nadu 635601, India  
e-mail: [agmshc@gmail.com](mailto:agmshc@gmail.com)

R. Janagaraj

Department of Mathematics, Faculty of Engineering, Karpagam Academy of Higher Education, Coimbatore, Tamil Nadu 641021, India

The discrete time fractional calculus has been of recent interest to researchers, and several works on application to inverted pendulum, pantograph and vibrating ear drum model are carried out in [6–8]. The term fractional difference calculus was introduced by Miller and Ross in 1989 [9], but it was in last few decades the theoretical framework of the forward and backward discrete fractional operators was developed by works of [10–13]. Most common difference operators of fractional order that are employed in modelling are Riemann–Liouville and Caputo type operators [14]. The work of [15] illustrates the numerical methods that can be employed for fractional difference equation. The analysis of qualitative behaviour of the fractional difference equations like oscillation and non-oscillations, stability and asymptotic behaviour has attracted a wide range of researchers as in [16–22]. Inspired from the above literature, this article develops a model of series RLC circuit using discrete fractional operator and is validated with the simulations obtained from simulation software PSPICE.

The article is structured with some basic definitions and lemmas, which are essential throughout this work in Sect. 2. The solution of fractional equations of damped and undamped series RLC circuit systems is obtained in Sect. 3. Some examples are illustrated based on the theoretical results in Sect. 4 followed with a brief conclusion in the last section.

## 2 Preliminaries and Basic Lemmas

In this section, some preliminary results of discrete fractional calculus are introduced, which will be used throughout the paper.

**Definition 1** [10] The falling factorial is defined as  $\ell^{(\alpha)} = \frac{\Gamma(\ell+1)}{\Gamma(\ell+1-\alpha)}$ , for any real number and denotes the Gamma function. Also  $\Delta\ell^{(\alpha)} = \alpha\ell^{(\alpha-1)}$ , where  $\Delta$  is the forward difference operator with  $\Delta u(\ell) = u(\ell + 1) - u(\ell)$ .

**Definition 2** [10] Let  $\alpha > 0$ . The  $\alpha$ -th fractional sum  $\Delta^{-\alpha} : N_a \rightarrow N_{a+\alpha}$  of  $f$  is defined by  $\Delta^{-\alpha} f(\ell) = \frac{1}{\Gamma(\alpha)} \sum_{m=0}^{\ell-\alpha} (\ell - m - 1)^{(\alpha-1)} f(m)$ , where  $f$  is defined for  $m \equiv a \pmod{1}$  and  $\Delta^{-\alpha} f$  is defined for  $\ell \equiv (a + \alpha) \pmod{1}$ .

**Definition 3** [10] Let  $\alpha > 0$  and  $\beta - 1 < \alpha < \beta$ , where  $\beta$  denotes a positive integer  $\beta = \lceil \alpha \rceil$ . Set  $\nu = \beta - \alpha$ . The  $\alpha$ -th fractional difference is defined as

$$\Delta^\alpha f(\ell) = \Delta^{\beta-\nu} f(\ell) = \Delta^\beta \Delta^{-\nu} f(\ell).$$

**Definition 4** [10] The shift operator  $\mathbb{E}$  is defined as  $\mathbb{E}^\nu f(\ell) = f(\ell + \nu), \nu \in \mathbb{N}$ .

**Lemma 5** [23] Let  $f, g : \mathbb{N}_0 \rightarrow \mathbb{R}, \alpha, \beta > 0$  and  $c, d$  be scalar. Then the following equalities hold:

- $\Delta^{-\alpha} \Delta^{-\beta} f(\ell) = \Delta^{-(\alpha+\beta)} f(\ell) = \Delta^{-\beta} \Delta^{-\alpha} f(\ell)$

2.  $\Delta^\alpha [cf(\ell) + dg(\ell)] = c\Delta^\alpha f(\ell) + d\Delta^\alpha g(\ell)$
3.  $\Delta \Delta^{-\alpha} f(\ell) = \Delta^{-(\alpha-1)} f(\ell)$

**Lemma 6** [10] For any positive real number  $\mu$ , the power rule is

$$\Delta^{-\alpha} \ell^{(\mu)} = \frac{\Gamma(\mu + 1)}{\Gamma(\mu + 1 + \alpha)} \ell^{(\mu+\alpha)}.$$

**Lemma 7** [10] For any  $\alpha > 0$ , the following equality holds:

$$\Delta^{-\alpha} \Delta f(\ell) = \Delta \Delta^{-\alpha} f(\ell) - \frac{(\ell - a)^{(\alpha-1)}}{\Gamma(\alpha)} f(a).$$

**Lemma 8** [9] Let  $f, g$  be real-valued function defined on  $\mathbb{N}_0$  and  $\mathbb{N}_\alpha \cup \mathbb{N}_0, \alpha \in (0, 1]$ . Then

$$\Delta^{-\alpha} (fg)(\ell) = \sum_{q=0}^{\infty} \binom{-\alpha}{q} [\Delta^q g(\ell)] [\Delta^{-(\alpha+q)} f(\ell + q)],$$

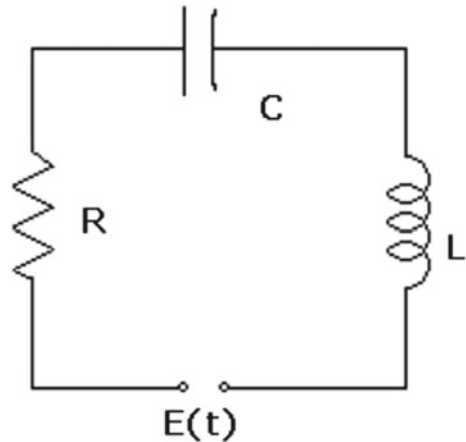
where  $\ell \equiv \alpha \text{ mod}(1)$  and  $\binom{-\alpha}{q} = \frac{\Gamma(1-\alpha)}{\Gamma(1+q)\Gamma(1-\alpha-q)}.$

*Result 9* [23] Let  $f, \ell$  be defined on  $\mathbb{N}_0$  and on  $\mathbb{N}_\alpha \cup \mathbb{N}_0$ , respectively, and  $0 < \alpha \leq 1$ , then the  $\alpha$ -th order difference of the product  $\ell f(\ell)$  is

$$\Delta^\alpha [\ell f(\ell)] = \alpha \Delta^{-(1-\alpha)} f(\ell + 1) + \ell \Delta^\alpha f(\ell).$$

### 3 LRC Series Circuit

The possibility of developing an electric circuit was made realistic in 1800, when battery with continuously flowing current was invented by Alessandro Volta (see [24, 25]). The electric lighting is the first and foremost application of the electrical circuits. RLC circuits are useful in tuning narrow frequency range in radio receivers and television. One of the most common and basic RLC circuits is used in construction of large electrical networks employed in communication systems. Mathematical models of different physical systems are represented by similar type differential equations as in differential equations representing the mass damper system and RLC circuits. It can be justified with the fact that an electrical circuit is developed for a mechanical system, and the current in the circuit will be equivalent to the displacement of the considered mechanical system with introduction of some suitable factors. In order to increase the decay of oscillations in a RLC circuit, a resistor can be introduced. Thus, constructing electrical circuits for a mechanical system is cost and time efficient due

**Fig. 1** LRC series circuit

to the fact that the circuit elements can be assembled with ease. Immediate and accurate measurements on electrical circuits explain its practical importance and go-to option than the mechanical system.

A series RLC circuit is the series connection of three basic passive circuit elements: resistor with resistance ( $R$ ), capacitor with capacitance ( $C$ ) and inductor of inductance ( $L$ ) with possible connection to external source  $E(t)$ ; see Fig. 1. Let us now consider a differential equation of second order with forcing term and damping given by

$$m \frac{d^2x}{dt^2} + \beta \frac{dx}{dt} + kx(t) = f(t). \quad (1)$$

Let the current through the series RLC circuit in Fig. 1 be  $i(\ell)$ . Employing *Kirchhoff's second law* with the voltage  $E(\ell)$  introduced to the circuit yields

$$L \frac{di}{d\ell} + Ri + \frac{1}{C}w = E(\ell). \quad (2)$$

From the relation between the charge  $w(\ell)$  and current  $i(\ell)$  given by  $i = \frac{dw}{d\ell}$ , the Eq. (2) is transformed into the following

$$L \frac{d^2w}{d\ell^2} + R \frac{dw}{d\ell} + \frac{1}{C}w(\ell) = E(\ell). \quad (3)$$



### 3.1 Undamped Series Circuit Analogue System

Consider a RLC circuit in the absence of external electromotive force  $E(\ell) = 0$  and resistance  $R = 0$ , with non-decaying electrical vibrations for increasing  $\ell$ , and then the circuit is undamped circuit and *simple harmonic*. So that

$$L \frac{d^2 w}{d\ell^2} + \frac{1}{C} w(\ell) = 0. \tag{4}$$

Suppose that  $\sigma = \frac{1}{\sqrt{LC}}$  in Eq. (4),

$$\frac{d^2 w}{d\ell^2} + \sigma^2 w(\ell) = 0. \tag{5}$$

Form the auxiliary equation, the above equation can be solved. It is not easy to apply the discrete fractional order operator to (5). So apply the existence of a different method to prove the result. By applying the discrete fractional order operator, first transform the Eq. (5) into a first-order differential equation [26].

Let  $w(\ell) = \ell^{2+\alpha} u(\ell)$  for  $\ell \neq 0$ , (5) lead to

$$\frac{d^2}{d\ell^2} \ell^{2+\alpha} u(\ell) + \sigma^2 \ell^{2+\alpha} u(\ell) = 0,$$

i.e.

$$\ell u_2(\ell) + 2(2 + \alpha)u_1(\ell) + [(2 + \alpha)(1 + \alpha)\ell^{-1} + \sigma^2 \ell]u(\ell) = 0, \tag{6}$$

here  $u_n(\ell) = \frac{d^n u}{d\ell^n}, n = 0, 1, 2, \dots$

When  $(2 + \alpha)(1 + \alpha) = 0$ , Eq. (6) has two values of  $\alpha$ 's,  $\alpha = -2$  and  $\alpha = -1$ . For the case  $\alpha = -2$ , again obtain Eq. (5). Let  $\alpha = -1$  and  $u(\ell) = e^{\eta\ell} v(\ell)$  for  $\ell \neq 0$ . Then  $u_1(\ell) = e^{\eta\ell} v_1(\ell) + \eta e^{\eta\ell} v(\ell)$  and  $u_2(\ell) = e^{\eta\ell} v_2(\ell) + 2\eta e^{\eta\ell} v_1(\ell) + \eta^2 e^{\eta\ell} v(\ell)$ . From (6) leads to

$$\ell v_2(\ell) + 2(1 + \eta\ell)v_1(\ell) + [(\eta^2 + \sigma^2)\ell + 2\eta]v(\ell) = 0. \tag{7}$$

If  $\eta^2 + \sigma^2 = 0$ , then  $\eta = \pm i\sigma$ . Therefore, (7) becomes

$$\ell v_2(\ell) + 2(1 \pm i\sigma\ell)v_1(\ell) \pm 2i\sigma v(\ell) = 0. \tag{8}$$

Applying forward  $\Delta^\gamma$  discrete fractional order operator to (8) leads to

$$\Delta^\gamma[\ell v_2(\ell)] + 2\Delta^\gamma[v_1(\ell)] \pm 2i\sigma[\ell v_1(\ell)] \pm 2i\sigma\Delta^\gamma[v(\ell)] = 0. \tag{9}$$

Using Result (9) and Definition (4) leads to

$$\begin{aligned}\Delta^\gamma[\ell v_2(\ell)] &= \ell v_{2+\gamma}(\ell) + \gamma \mathbb{E} v_{1+\gamma}(\ell) \\ \Delta^\gamma[\ell v_1(\ell)] &= \ell v_{1+\gamma}(\ell) + \gamma \mathbb{E} v_\gamma(\ell).\end{aligned}$$

Using above equations in (9) yields

$$\ell v_{2+\gamma}(\ell) + [\gamma \mathbb{E} + 2(1 + \pm i \sigma \ell)] v_{1+\gamma}(\ell) \pm i 2 \sigma (1 + \gamma \mathbb{E}) v_\gamma(\ell) = 0. \quad (10)$$

If  $1 + \gamma \mathbb{E} = 0$ , then  $\gamma = -\mathbb{E}^{-1} = -1$ . So (10) implies that

$$\ell v_1(\ell) + [1 + \pm 2i \sigma \ell] v(\ell) = 0,$$

or

$$v_1(\ell) + (\ell^{-1} \pm 2i \sigma) v(\ell) = 0, \quad (11)$$

which is a homogeneous first-order linear differential equation. Solving Eq. (11), yields  $v^I = A \ell^{-1} e^{2i \sigma \ell}$  and  $v^{II} = B \ell^{-1} e^{-2i \sigma \ell}$ . Hence, the general solution is

$$w(\ell) = C \cos \sigma \ell + D \sin \sigma \ell, \quad (12)$$

where  $C = A + B$  and  $D = B - A$  are arbitrary constants. For (5), the initial conditions are

$$w(0) = w_0 \quad \text{and} \quad w_1(0) = i_0. \quad (13)$$

Using Eq. (13) in (12) yields

$$C = w_0 \quad \text{and} \quad D = \frac{i_0}{\sigma}.$$

Hence,

$$w(\ell) = w_0 \cos \sigma \ell + \frac{i_0}{\sigma} \sin \sigma \ell, \quad (14)$$

where the motion is a reciprocation action. This motion is called *simple harmonic motion*.

Now, suppose that

$$w(\ell) = F \sin(\sigma \ell + \theta).$$

From (14)

$$w_0 = F \sin \theta \quad \text{and} \quad \frac{i_0}{\sigma} = F \cos \theta.$$

Now  $\tan \theta = \frac{\sigma w_0}{i_0}$  and  $F = \sqrt{\left(\frac{i_0}{\sigma}\right)^2 + w_0^2}$ , where  $\theta$  and  $F$  are phase angle and vibration amplitude, respectively. The reciprocation action  $T = \frac{2\pi}{\sigma} = 2\pi\sqrt{LC}$ .

### 3.2 Damped Series Circuit Analogue System

If  $E(\ell) = 0$  in Eq. (3), the electrical vibration of the circuit is said to be free. Then

$$L \frac{d^2 w}{d\ell^2} + R \frac{dw}{d\ell} + \frac{1}{C} w(\ell) = 0,$$

or

$$\frac{d^2 w}{d\ell^2} + k_1 \frac{dw}{d\ell} + k_2 w(\ell) = 0, \quad (15)$$

where  $k_1 = \frac{R}{L}$  and  $k_2 = \frac{1}{LC}$ .

Taking  $w(\ell) = \ell^{2+\delta} u(\ell)$  for  $\ell \neq 0$  (15) yields

$$\frac{d^2}{d\ell^2} [\ell^{2+\delta} u(\ell)] + k_1 \frac{d}{d\ell} [\ell^{2+\delta} u(\ell)] + k_2 \ell^{2+\delta} u(\ell) = 0,$$

or

$$\ell u_2(\ell) + [k_1 \ell + 2(2 + \delta)] u_1(\ell) + [(\delta^2 + 3\delta + 2)\ell^{-1} + k_1(2 + \delta) + k_2 \ell] u(\ell) = 0. \quad (16)$$

Suppose that  $\delta = -1$  and  $u(\ell) = e^{\beta\ell} v(\ell)$  for  $\ell \neq 0$ , which leads  $u_1(\ell) = e^{\beta\ell} v_1(\ell) + e^{\beta\ell} v(\ell)$  and  $u_2(\ell) = e^{\beta\ell} v_2(\ell) + 2\beta e^{\beta\ell} v_1(\ell) + \beta^2 e^{\beta\ell} v(\ell)$ . From (16),

$$\ell v_2(\ell) + [2 + (2\beta + k_1)\ell] v_1(\ell) + [(\beta^2 + k_1\beta + k_2)\ell + k_1 + 2\beta] v(\ell) = 0. \quad (17)$$

For  $\beta^2 + k_1\beta + k_2 = 0$

$$\beta = -\frac{k_1}{2} \pm \frac{1}{2} \sqrt{k_1^2 - 4k_2},$$

or

$$\beta = \varepsilon \pm \tau,$$

where  $\varepsilon = -\frac{k_1}{2}$  and  $\tau = \frac{1}{2} \sqrt{k_1^2 - 4k_2}$ . From (17), it leads to

$$\ell v_2(\ell) + (\rho + 2)v_1(\ell) + \rho v(\ell) = 0, \quad (18)$$

where  $\rho = k_1 + 2\beta$ . Apply  $\Delta$ -discrete fractional order operator on both sides of Eq. (18)

$$\ell v_{2+\gamma}(\ell) + (\gamma \mathbb{E} + \rho \ell + 2)v_{1+\gamma}(\ell) + \rho(\gamma \mathbb{E} + 1)v_\gamma(\ell) = 0.$$

For  $\gamma = -\mathbb{E}^{-1} = -1$  yields

$$\ell v_1(\ell) + (\rho \ell + 1)v(\ell) = 0.$$

Solving the above equation,

$$v(\ell) = A\ell^{-1}e^{-\rho\ell},$$

or

$$u(\ell) = e^{\beta\ell}v(\ell) = e^{\beta\ell}A\ell^{-1}e^{-\rho\ell}$$

or

$$u(\ell) = A\ell^{-1}e^{\frac{\rho}{2}\ell - \frac{k_1}{2}\ell - \rho\ell}.$$

Now,

$$w(\ell) = Ae^{-\left(\frac{\rho+k_1}{2}\right)\ell}.$$

Here, three cases are possible,

1. **Overdamped:**  $k_1^2 > 4k_2$  or the damping coefficient is greater than the angular frequency, and the result of  $k_1^2 > 4k_2$  inside the radical is positive. Therefore, the solution is

$$w(\ell) = Be^{-\left(\frac{k_1 - \sqrt{k_1^2 - 4k_2}}{2}\right)\ell} + Ce^{-\left(\frac{k_1 + \sqrt{k_1^2 - 4k_2}}{2}\right)\ell},$$

where  $B, C$  are arbitrary constants.

2. **Critically Damped:**  $k_1 = 4k_2$  or the damping factor is equal to the angular frequency, and the result of  $k_1 = 4k_2$  inside the radical is zero. In this case, one root of the equation and solution is

$$w(\ell) = (B + C\ell)e^{-\frac{k_1\ell}{2}}.$$

3. **Underdamped:**  $k_1 < 4k_2$  or the damping coefficient is smaller than the angular frequency, and the result of  $k_1 < 4k_2$  inside the radical is negative. Therefore, the roots are complex and solution is

$$w(\ell) = e^{-\frac{k_1 \ell}{2}} (B \cos \mu \ell + C \sin \mu \ell),$$

where  $\mu = \frac{1}{2} \sqrt{4k_2 - k_1^2}$ .

Here, the critically damped solution and the overdamped solutions will not exhibit an oscillatory behaviour. Because the equation has two distinct solutions. Let  $w(\ell) = D e^{-\frac{k_1 \ell}{2}} \cos(\mu \ell - \rho) = e^{-\frac{k_1 \ell}{2}} (D \cos \mu \ell \cos \rho + D \sin \mu \ell \sin \rho)$ . Then

$$B^2 + C^2 = D^2 \quad \text{and} \quad \tan \rho = \frac{C}{B}.$$

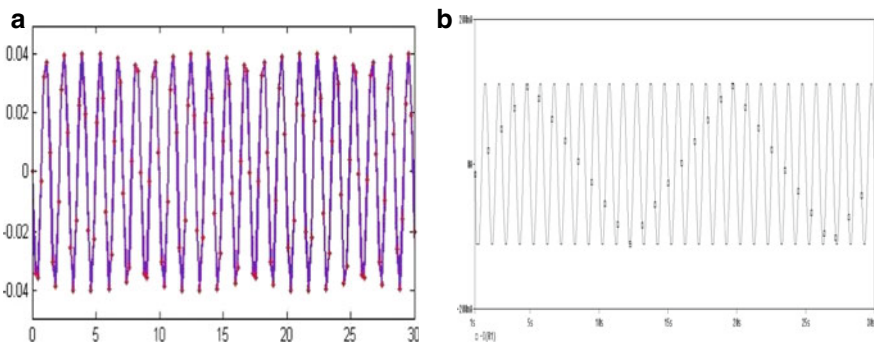
where  $e^{-\frac{k_1 \ell}{2}}$  is the damping factor and the motion is a vibration motion.

### 4 Analytical and Experimental Results

This section presents some numerical examples to show the validity of the theoretical results. Also using the electrical circuit, the experimental results analysed to verify the solutions obtained by the discrete RL fractional differential equation.

**Example 1** Let  $L = 1h, C = 0.0004f, E(\ell) = 0$  and the initial conditions are  $w(0) = 0c$  and  $i(0) = 2A$ . Now,  $\sigma = \frac{1}{\sqrt{LC}} = \frac{1}{\sqrt{0.0004}} = 50, T = 2\pi \sqrt{LC} = 0.1257$  and  $\frac{i_0}{\sigma} = \frac{2}{50} = \frac{1}{25}$ . Therefore,  $w(\ell) = \frac{1}{25} \sin(50\ell)$ , which is undamped series circuit analogue system. The theoretical and experimental graphs are displayed in Fig. 2.

**Example 2** Let  $L = \frac{1}{4}h, R = 30 \Omega, C = \frac{1}{300}f, E(\ell) = 0$  and the initial conditions are  $w(0) = 4c$  and  $i(0) = 0A$ . Now,  $k_1 = \frac{R}{L} = 120, k_2 = \frac{1}{LC} = 1200$ . Hence,  $k_1^2 > 4k_2$ , which is overdamped. Therefore,  $w(\ell) = \frac{3}{2} \left(1 + \frac{3}{\sqrt{6}}\right) e^{(20\sqrt{6}-60)\ell} +$

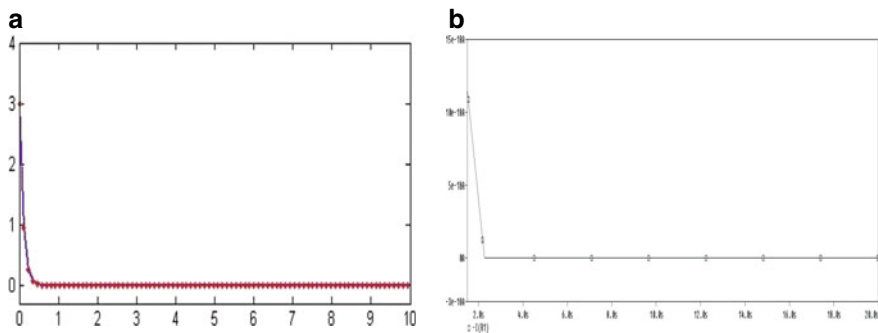


**Fig. 2** Comparison of undamped series circuit analogue system with **a** theoretical and **b** experimental

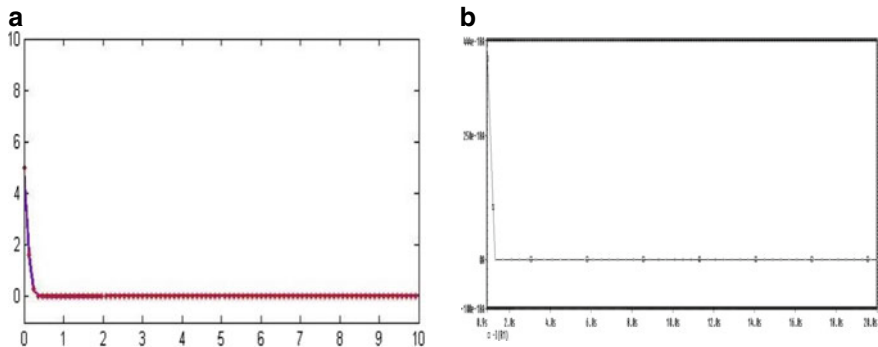
$\frac{3}{2}\left(1 - \frac{3}{\sqrt{6}}\right)e^{-(20\sqrt{6}+60)\ell}$ . The theoretical and experimental results of overdamped series circuit are exhibited; see Fig. 3.

**Example 3** Let  $L = \frac{1}{20}h, R = 2\Omega, C = \frac{1}{20}f, E(\ell) = 0$ , and the initial conditions are  $w(0) = 5c$  and  $i(0) = 0A$ . Now,  $k_1 = \frac{R}{L} = 40, k_2 = \frac{1}{LC} = 400$ . Hence,  $k_1^2 = 4k_2$ , which is critically damped. Therefore,  $w(\ell) = (5 + 100\ell)e^{-20\ell}$ . The theoretical and experimental results of critically damped series circuit are exhibited; see Fig. 4.

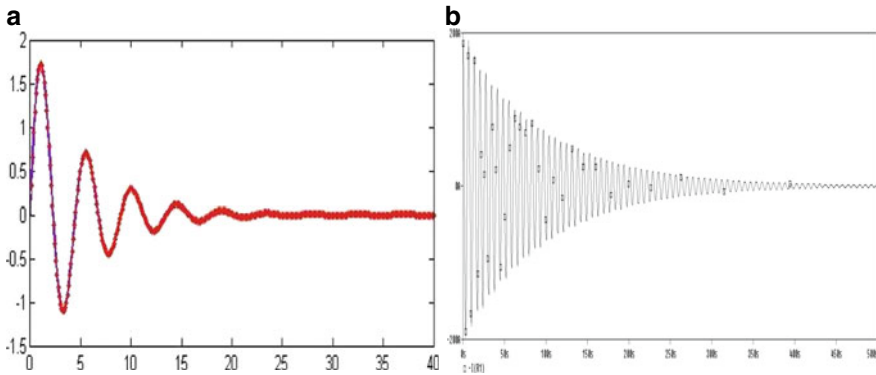
**Example 4** Let  $L = 1h, R = 0.4\Omega, C = \frac{1}{2}f, E(\ell) = 0$ , and the initial conditions are  $w(0) = 0c$  and  $i(0) = 3A$ . Now,  $k_1 = \frac{R}{L} = 0.4, k_2 = \frac{1}{LC} = 2$ . Here  $k_1^2 < 4k_2$ , Therefore,  $w(\ell) = \frac{15}{7}e^{-\frac{\ell}{5}}\sin\left(\frac{7}{5}\ell\right)$ . which is underdamped, the theoretical and experimental results are presented in Fig. 5.



**Fig. 3** Comparison of overdamped series circuit analogue system with **a** theoretical and **b** experimental



**Fig. 4** Comparison of critically damped series circuit analogue system with **a** theoretical and **b** experimental



**Fig. 5** Comparison of underdamped series circuit analogue system with **a** theoretical and **b** experimental

### 4.1 Discussion

Numerical examples presented in this section with suitably supported simulations illustrate the following behaviour of the fractional order RLC circuit. Figure 2 illustrates that an undamped RLC circuit exhibits a uniform electrical vibration for increasing time which is supported by the uniform oscillation obtained for numerical values given in Example 1. Impact of damping is presented in following Examples 2, 3 and 4 describing the overdamped, critically damped and underdamped cases. It can be observed from Figs. 3 and 4 for the case of overdamped and critically damped, where there are no electrical vibrations, and with increasing time, the damping effect causes the system to attain stability at origin. In case of underdamped cases in Example 4, the circuit undergoes initial electrical vibrations which over period time decays and approaches origin.

### 4.2 Experimental Implementation of Fractional Derivatives in Series RLC Circuit

The introduction of the fractional order element in the RLC circuit is analysed, and its dynamics for different fractional orders are presented in Fig. 6. The amplitude of the oscillations for different fractional orders states the behavioural changes taking place in the RLC circuit with respect to fractional orders.

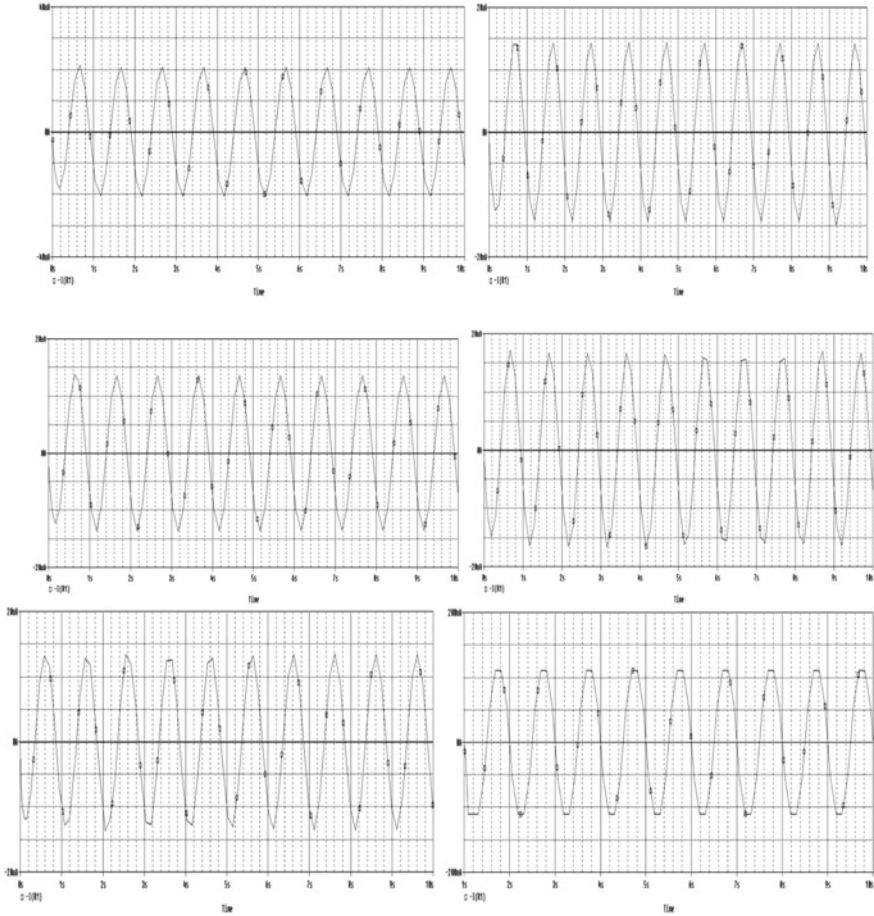


Fig. 6 Experimental analysis of fractional derivatives in series circuit

## 5 Conclusion

This article examined the second-order differential equation representing damped and undamped series RLC circuit system, and using forward discrete fractional operator, the solutions are obtained. Also the theoretical results are verified from the illustrated examples and compared with simulations from PSpice for continuous fractional order derivatives. In the near future work, this method can be applied to the nonlinear equations such as Duffing equation, pendulum equation and van der Pol's equation.

**Acknowledgements** The authors are thankful to the referee for useful remarks that developed the paper.



## References

1. Hilfer R (2000) Applications of fractional calculus in physics. Applications of fractional calculus in physics. World Scientific
2. Oldham K, Spanier J (1974) The fractional calculus, mathematics in science and engineering. Academic Press
3. Battaglia JL, Cois O, Puigsegur L, Oustaloup A (2001) Solving an inverse heat conduction problem using a non-integer identified model. *Int J Heat Mass Transf* 44:2671–2680
4. Wang JC (1987) Realizations of generalized Warburg impedance with RC ladder networks and transmission lines. *J Electrochem Soc* 134(8):1915–1920
5. Podlubny I (1999) Fractional differential equations, mathematics in science and engineering. Academic Press, San Diego, vol 198
6. Selvam AGM, Baleanu D, Alzabut J, Vignesh D, Abbas S (2020) On Hyers–Ulam Mittag-Leffler stability of discrete fractional Duffing equation with application on inverted pendulum. *Adv Differ Equ* 2020(456)
7. Alzabut J, Maria Selvam G, El-Nabulsi RA, Vignesh D, Samei ME (2021) Asymptotic stability of nonlinear discrete fractional pantograph equations with non-local initial conditions. *Symmetry* 13(473)
8. Maria Selvam G, Alzabut J, Dhakshinamoorthy V, Jonnalagadda JM, Abodayeh K (2021) Existence and stability of nonlinear discrete fractional initial value problems with application to vibrating eardrum. *Math Biosci Eng* 18(4):3907–3921
9. Miller KS, Ross B (1989) Fractional difference calculus. In: Proceedings of the international symposium on Univalent Func., *Frac. Calc. Appl.*, Ellis Horwood Ser. Math., pp 139–152
10. Atici FM, Eloe PW (2008) Initial value problems in discrete fractional calculus. *Proc Am Math Soc* 137(3):981–989. <https://doi.org/10.1090/S0002-9939-08-09626-3>
11. Atici FM, Eloe PW (2009) Discrete fractional calculus with the nabla operator. *Electron J Qual Theory Differ Equ* 3:1–12
12. Goodrich CS, Peterson AC (2015) Discrete fractional calculus. Springer International Publishing, Switzerland. <https://doi.org/10.1007/978-3-319-25562-0>
13. Holm M (2011) The theory of discrete fractional calculus: development and application. PhD thesis. University of Nebraska-Lincoln, Lincoln, Nebraska
14. Abdeljawad T (2011) On Riemann and Caputo fractional differences. *Comput Math Appl* 62(3):1602–1611
15. Anastassiou GA (2010) Principles of delta fractional calculus on time scales and inequalities. *Math Comput Model* 52(3–4):556–566
16. Adiguzel H (2018) Oscillatory behavior of solutions of certain fractional difference equations. *Adv Differ Equ* 2018(445):1–13
17. Chatzarakis GE, Gokulraj P, Kalaimani T (2018) Oscillation test for fractional difference equations. *Tatra Mt Math Publ* 71(1):53–64. <https://doi.org/10.2478/tmmp-2018-0005>
18. Maria Selvam G, Janagaraj R (2019) Oscillation criteria of a class of fractional order damped difference equations. *Int J Appl Math* 32(3):433–441. <https://doi.org/10.12732/ijam.v32i3.5>
19. Secer A, Adiguzel H (2016) Oscillation of solutions for a class of nonlinear fractional difference equations. *J Nonlinear Sci Appl* 9:5862–5869
20. Chatzarakis GE, George Maria Selvam A, Janagaraj R, Miliaras GN (2022) Oscillation Behaviour of Solutions for a Class of a Discrete Nonlinear Fractional-Order Derivatives. *Tatra Mountains Math Publ* 79(2):101–118 <https://doi.org/10.2478/tmmp-2021-0022>
21. Chatzarakis GE, George Maria Selvam A, Janagaraj R, Miliaras GN (2020) Oscillation criteria for a class of nonlinear discrete fractional order equations with damping term. *Math Slovaca* 70(5):1165–1182. <https://doi.org/10.1515/ms-2017-0422>
22. Selvam A, Alzabut J, Janagaraj R, Adiguzel H (2020) Oscillation analysis for nonlinear discrete fractional order delay and neutral equations with forcing term. *Dyn Syst Appl* 29(2). <https://doi.org/10.46719/dsa20202929>
23. Sengul S (2010) Discrete fractional calculus and its applications to tumor growth. Masters Theses & Specialist Projects, pp 1–161

24. Mehta VK, Mehta R (2013) Principles of electronics, 11th edn. S. Chand & Company Pvt. Limited
25. Sedha RS (2012) A textbook of applied electronics, 3rd edn. S. Chand & Company Pvt. Limited
26. Ozturk O (2016) A study on nabla discrete fractional operator in mass-spring-damper system. *New Trends Math Sci* 4:137–144. <https://doi.org/10.20852/ntmsci.2016422559>

# Design of Cascaded Multilevel Inverter-based STATCOM for Reactive Power Control with Different Novel PWM Algorithms



Ch. Lokeshwar Reddy and G. Sree Lakshmi

## 1 Introduction

The importance of electric power quality has been drastically increasing year to year. With the use of latest equipment plenty of research is going on to increase the power quality, these could improve the transmission capability and the superiority of the voltage waveform. This equipment's are termed as Flexible AC Transmission System (FACTS) devices. The FACTS devices are made with power electronic components. As per literature several FACTS devices are available, out of which Static Synchronous Compensator (STATCOM) is the most important device in power systems to increase the power quality by supplying the reactive power components required by the load [1]. STATCOM is gradually exploited to increase the power transmission capability and stabilize the voltage profile.

The STATCOM reduces the burden on source and smoothen the voltage profile of the electric lines, by supplying or absorbing the reactive powers. STATCOM is primarily one of the important shunt types of FACTS device; it comprises of an inverter with capacitors as energy storage elements on the DC side. The STATCOM is controlled with a control device to meet the load requirements [2]. With the use of latest power semiconductor devices and novel topologies of inverters, the voltage and power capacities of the power systems have increased. The output reactive power of the converter is controlled based on variations in connected load. The inverter is directly attached to medium level of voltages without the interface of

---

Ch. Lokeshwar Reddy (✉) · G. Sree Lakshmi  
Department of Electrical and Electronics Engineering, CVR College of Engineering, Mangalpalli,  
Hyderabad, India  
e-mail: [lokeshwar.reddy@cvr.ac.in](mailto:lokeshwar.reddy@cvr.ac.in)

G. Sree Lakshmi  
e-mail: [g.sreelakshmi@cvr.ac.in](mailto:g.sreelakshmi@cvr.ac.in)

coupling transformers. So, the cost and the losses associated with in transformer can be reduced.

## 2 Cascaded Multilevel STATCOM

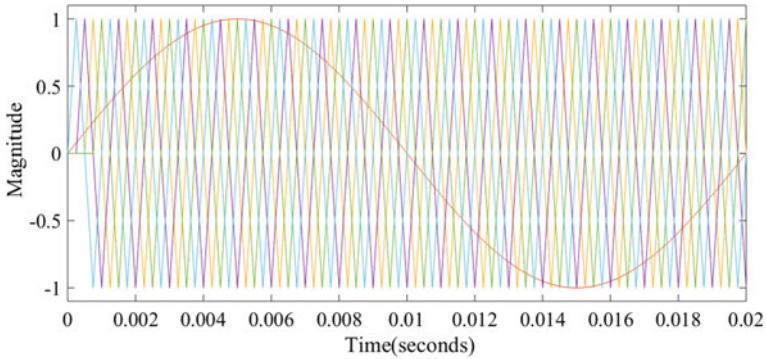
Multilevel converters have been chosen significantly for STATCOM application to boost the compensator capability for medium or high voltage and large power appliances. Mainly multilevel inverters are categorized into three categories for STATCOM application. The cascaded multilevel inverter has many merits when compared to other multilevel inverters and techniques.

Several advantages of cascaded H-bridge topology are:

It reduces harmonics in the output. It removes the need of transformer to deliver the required voltage levels. Modularized layout and flexibility of circuit for packing which is simple and ease of fabrication. Very easy to deliver high voltage levels by connecting H-bridges in series for each phase. No extra circuit is needed for stabilizing of DC link voltages and decreases the capacitance of DC link by employing a proper modulation method.

Shunt compensator like STATCOM which receives or delivers the reactive power components can be done by the implementation of closed loop control to the inverter. Unbalance of voltages among DC capacitors, the switching losses of the devices due to different conduction time periods, implementation of control circuit for real time environment [3, 4] are the major challenges in cascaded H-bridge STATCOM. Balancing of DC link voltage and implementation of control algorithm is to be accomplished by using cascaded H-bridge inverter as a STATCOM.

Balancing of DC link voltages in the inverters can be achieved by implementing a suitable PWM technique. The commonly implemented PWMs for cascaded H-bridge multilevel STATCOM are multi carrier PWM techniques. Mostly, these are Phase Shifted Carrier Pulse Width Modulation (PSCPWM), where carriers are shifted horizontally and the Space Vector Pulse Width Modulation (SVPWM) techniques [5]. The frequency of switching devices can be improved by using latest developed power electronic devices. With the use of different control techniques along with high-frequency PWM techniques the harmonics generated by the STATCOM can be estimated. The working of STATCOM can be explained by using a couple of differential equations and by using switching functions the switching device can be modeled.



**Fig. 1** Reference and carrier signal arrangement for SPSCPWM

### 3 Modulation Methods for Cascaded Multilevel STATCOM

The balancing and reduction of ripples in the voltages across DC link, minimizing the harmonics in the output of STATCOM, providing reactive components of power needed by the load can be achieved by implementing a suitable multi carrier type PWM method. Several PWM techniques are preferred which are sinusoidal, space vector and third harmonic injected.

#### 3.1 Sinusoidal Pulse Width Modulation

For multilevel inverters the modulation techniques are developed based on the alignment carrier signals. The carrier signals moved horizontally with sinusoidal as reference is Sinusoidal Phase Shifted Carrier PWM (SPSCPWM). The reference and carrier signal arrangement for SPSCPWM is shown in Fig. 1. The SPSCPWM technique is generally preferred for the cascaded multilevel STATCOM, due to this method the power sharing among all switches is same and the execution is very easy [6]. In all the Phase Shifted Carrier PWM techniques no harmonics will present up to  $2N$ th carrier group, where  $N$  is the sum of H-bridge in each phase [7].

#### 3.2 Modified Space Vector Pulse Width Modulation

In common SVPWM technique the switching time periods for multilevel inverters is to be determined by mapping of the outer sector to the inner sector where the space vector is located [8]. Switching times can determined for the present sector based on relative switching vectors. During real time applications the computation time is improved due to the existence of higher sectors [9].

Suitable offset voltages can be added to the sinusoidal reference voltages during the comparison with carrier signals to generate gate signals. Then the performance of multi carrier-based SPWM is almost same as SVPWM technique [10]. The evaluation of voltage offset is reliant on modulus function and is depends on levels in the phase voltage, amount of DC link and phase voltages. The subsequent equations are utilized to determine the offset time  $T_{\text{offset}}$ .

$$T_a = \frac{V_a * T_s}{V_{dc}} \tag{1}$$

$$T_b = \frac{V_b * T_s}{V_{dc}} \tag{2}$$

$$T_c = \frac{V_c * T_s}{V_{dc}} \tag{3}$$

$T_a$ ,  $T_b$  and  $T_c$  are the switching times for  $a$ ,  $b$  and  $c$  phases,  $V_a$ ,  $V_b$  and  $V_c$  are the amounts of phase voltages and  $T_s$  is the sampling time period.

$$T_{\text{offset}} = \left[ \frac{T_0}{2} - T_{\text{min}} \right] \tag{4}$$

$$T_o = [T_s - T_{\text{effect}}] \tag{5}$$

$$T_{\text{effect}} = T_{\text{max}} - T_{\text{min}} \tag{6}$$

$T_{\text{max}}$  and  $T_{\text{min}}$  highest and lowest values of phase voltages in a sampling period.

In a sampling period an offset voltage is added to reference voltages of each phase, so the performance with multi carrier SPWM method is almost same as the performance of SVPWM method. The generated modified SVPWM reference signals

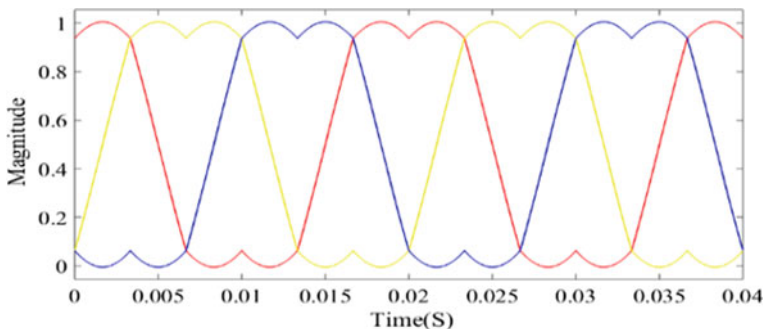
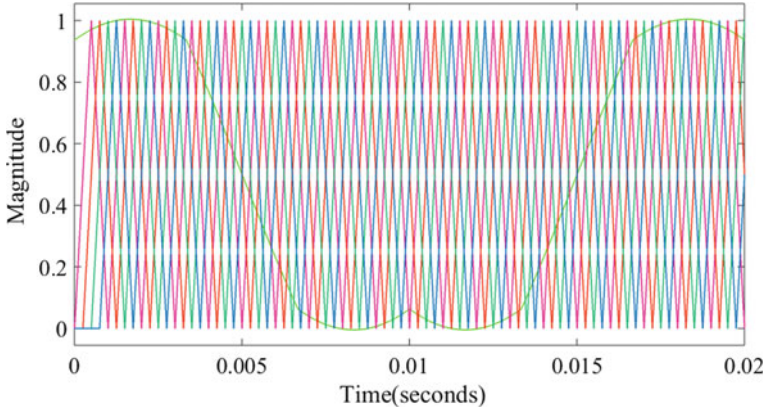


Fig. 2 Modified SVPWM reference signals



**Fig. 3** Reference and carrier waveform arrangements SVPSPWM

shown in Fig. 2. The switching signals can be delivered by the comparison of reference signal with shifted carrier signals. Figure 3 shows the arrangement of reference and carrier signals for the modified Space Vector PWM (MSVPWM) technique.

In a sampling period an offset voltage is added to reference voltages of each phase, so the performance with multi carrier SPWM method is almost same as the performance of SVPWM method.

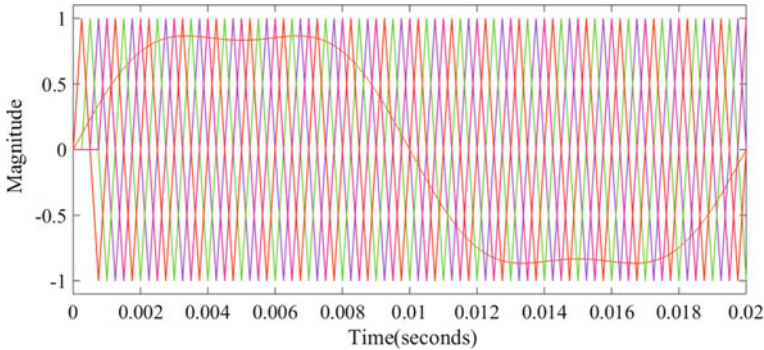
### 3.3 *Third Harmonic Injected Pulse Width Modulation*

Modulation strength of the three-phase inverter is to be enhanced by the addition of third harmonic wave to the reference wave of every phase. The fundamental component of line voltages will not be affected since the common mode voltage components balance among the phase voltages only. Based on the modulation index considered to exploit the reference voltage, the third harmonic magnitude of signal inserted is 1/6th to rise the output of fundamental signal; the degree of reference p.u is 1.15. The subsequent equations are considered to generate the reference signal in Third Harmonic Injected Pulse Width Modulation (THIPWM).

$$V_a(t) = 1.15 \sin(\omega t) + (1/6) \sin(3\omega t) \tag{7}$$

$$V_b(t) = 1.15 \sin(\omega t - 2 * \pi / 3) + (1/6) \sin(3\omega t) \tag{8}$$

$$V_c(t) = 1.15 \sin(\omega t - 4 * \pi / 3) + (1/6) \sin(3\omega t) \tag{9}$$



**Fig. 4** Reference and carrier waveform arrangements for THIPWM Technique

The Third Harmonic Injected Phase Shifted Carrier (THIPSC) reference and carrier waveform arrangement is shown in Fig. 4.

### 4 Control Strategy for Cascaded Multilevel Inverter-Based STATCOM

The cascaded H-bridge five-level STATCOM as shown in Fig. 5. The cascaded five-level inverter built STATCOM is to be coupled with the grid. Each H-bridge inverter produces 3 levels of output, to produce 5 levels two H-bridges are coupled in series in a phase. Generally,  $2m + 1$  levels voltages in phase will be produced, where  $m$  is total H-bridges connected in series in each phase.

In the STATCOM model the three-phase values like  $i_{al}, i_{bl}, i_{cl}, v_a, v_b, v_c, i_{as}, i_{bs}$  and  $i_{cs}$  as currents of the load, voltages of the source and currents of the inverter, all the quantities are converted in to  $i_{dl}^*, i_q^*, v_d, v_q, i_d$  and  $i_q$  in the reference frame of synchronously rotating. The complete control technique block diagram is shown in Fig. 6. The reference values of  $d$ - $q$  axes of the inverter output voltages  $e_d$  and  $e_q$  can be controlled as

$$e_d = V_d + X_1 - \omega L i_q \tag{10}$$

$$e_q = \omega L i_d + X_2 \tag{11}$$

The source voltage value of direct axis is  $V_d$  and  $i_{dl}^*, i_q^* i_d$ , and  $i_q$ , are the  $d$ ,  $q$  axes load currents and inverter currents considered. The parameters of control  $X_1$  and  $X_2$  are controlled as

$$x_1 = (i_d^* - i_d) \left( k_{p2} + \frac{k_{i2}}{s} \right) \tag{12}$$



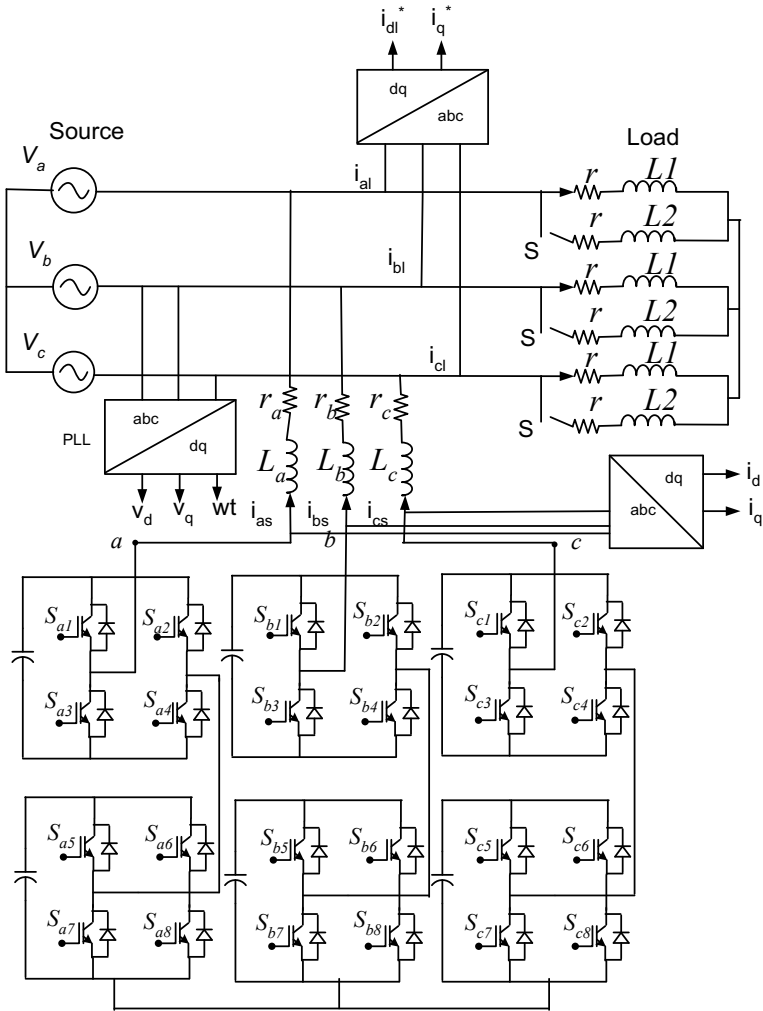


Fig. 5 STATCOM cascaded five-level

$$x_2 = (i_q^* - i_q) \left( k_{p3} + \frac{k_{i3}}{s} \right) \tag{13}$$

The reference current component of *d*-axis  $i_d^*$  is

$$i_d^* = \left[ (V_{dc}^*) - (V_{dc1} + V_{dc2} + V_{dc3} + V_{dc4} + V_{dc5}) \left( k_{p1} + \frac{k_{i1}}{s} \right) \right] \tag{14}$$

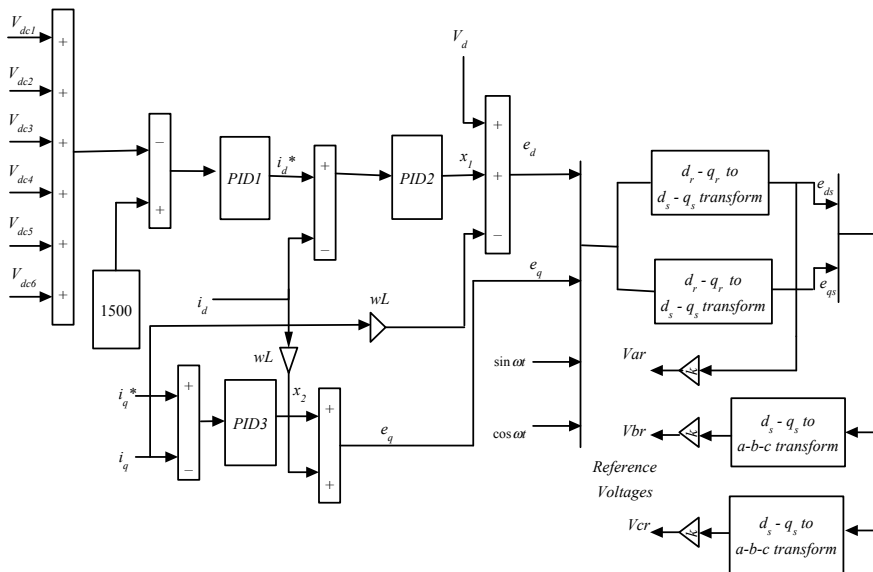


Fig. 6 PWM generation and abc to dqo and vice-versa control block diagram

The magnitude of reference voltage of DC link is  $V_{dc}^*$  and the voltages of the capacitors of DC link in every H-bridge are  $V_{dc1}$  to  $V_{dc6}$ . By considering the stationary frame as reference and  $e_d, e_q$  are transformed into synchronous rotating frame as:

$$e_{ds} = (\sin \omega t)e_q + (\cos \omega t)e_d \tag{15}$$

$$e_{qs} = (\cos \omega t)e_q - (\sin \omega t)e_d \tag{16}$$

By the consideration of different voltages of the inverter by taking reference signals are considered as

$$v_{ar} = e_{ds} \tag{17}$$

$$v_{br} = -\frac{1}{2}e_{ds} + \frac{\sqrt{3}}{2}e_{qs} \tag{18}$$

$$v_{cr} = -\frac{1}{2}e_{ds} - \frac{\sqrt{3}}{2}e_{qs} \tag{19}$$

From  $V_{dc}^*$  and  $i_q^*$  inner and outer loops produce direct and quadrature axes voltages reference,  $e_d$  and  $e_q$  for the cascade multilevel inverter which is considered with STATCOM for reactive power compensation for different load changes. By considering the reference voltages that are to be controlled which is interfaced with

supply, the reactive currents which are essential to the load draws active currents to regulate the DC link voltage  $V_{dc}^*$ .

## 5 Simulation Results and Analysis

A cascaded five-level inverter is taken as STATCOM for simulation. The STATCOM simulation is done in MATLAB/SIMULINK for various step variations in the load. The inverter which can be controlled by adopting different PWM techniques like Sinusoidal Phase Disposition (SPD) PWM, Modified Space Vector Phase Shifted Carrier (MSVPSC) PWM, Sinusoidal Phase Shifted Carrier (SPSC) PWM, Third Harmonic Injected Phase Shifted Carrier (THIPSC) PWM and Modified Space Vector Phase Disposition (MSVPD) PWM techniques. The selected PI controller values for voltage and current control loops and the simulation parameters are shown in Tables 1 and 2 correspondingly.

### 5.1 STATCOM Performance for Sinusoidal Phase Shifted Carrier PWM

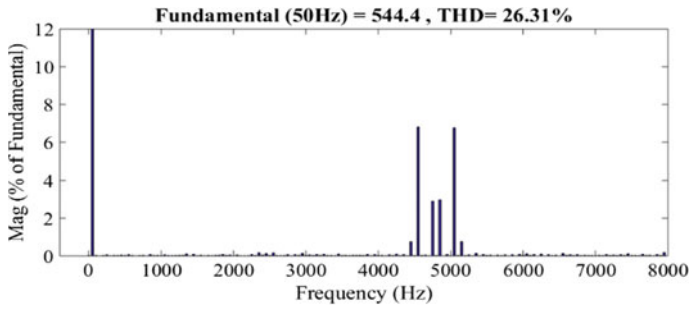
The various STATCOM performances by applying SPSCPWM method is shown in Fig. 7, from the results it is seen that reference voltage of 1500 V is spreads

**Table 1** PID controller tuning of parameters

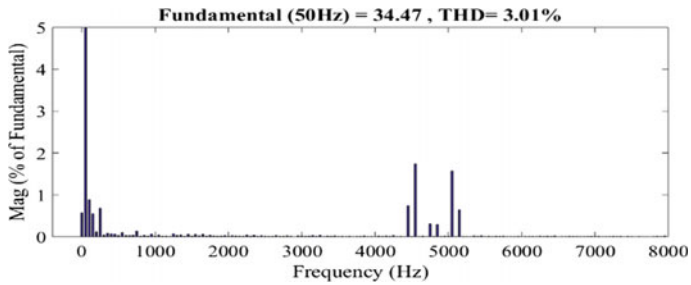
Controlled variable	PID controller	Proportional controller gain	Integral controller gain
Voltage	PID1	0.08	0.2
Voltage	PID2	0.07	0.2
Current	PID3	4	6

**Table 2** System parameters for simulation

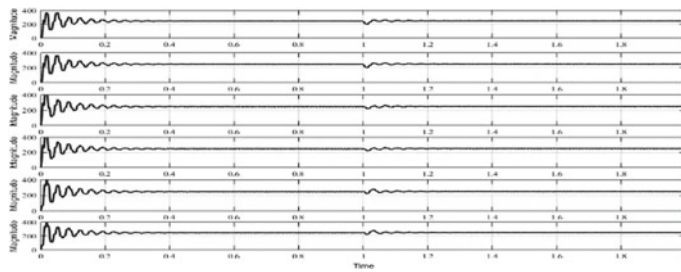
Voltage at supply side	(Phase-phase) RMS 415 V
Voltage across DC link	250 V constant for all PWM methods
Load values	$C = 1.5 \text{ mF}$ , $R = 4.6 \text{ } \Omega$ (for capacitive load) $L = 12.4 \text{ mH}$ , $R = 4.6 \text{ } \Omega$ (for inductive load)
Fundamental frequency	50 Hz
DC link capacitance	10 mF
STATCOM Interfacing resistance	0.13 $\Omega$
Frequency for switching	1200 Hz
STATCOM Interfacing Inductance	2.19e-3 H



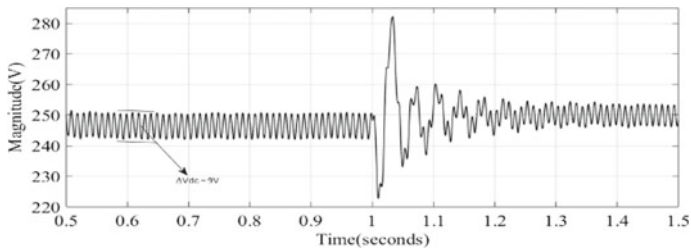
a. Output Voltage Harmonic Spectrum



b. Harmonics Spectrum of Output current

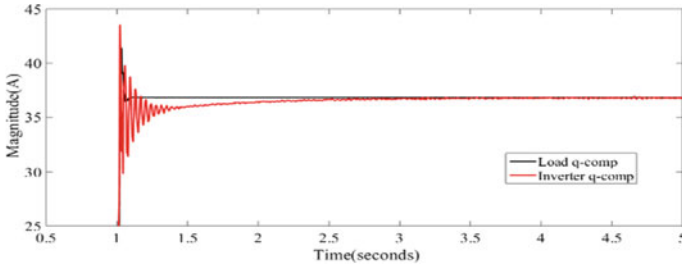
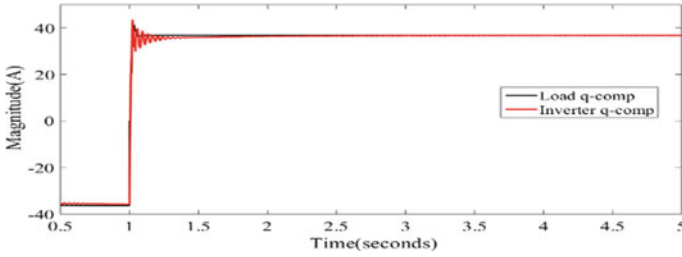


c. Voltages across DC-Link capacitor

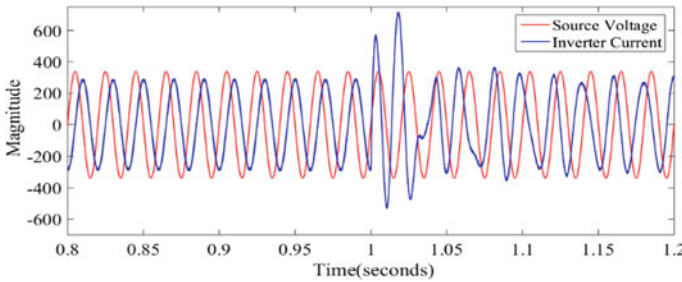


d. Ripples in voltage of DC-link capacitor

**Fig. 7** Responses of STATCOM for sudden change of load from inductive to capacitive at 1 s with SPSCPWM technique



**e. Reactive components of currents-Load and inverter**



**f. Phasor relations of source voltage and inverter current**

**Fig. 7** (continued)

similarly across every capacitor of the DC link in the H-bridges. The ripples in the voltage across DC link capacitors are about 9 V and it is very low value. The output current harmonics are of 3% and in the voltage, it is 26.31%. The harmonics are not present below the frequency 4.8 kHz, because of the SPSCPWM technique. If the load changes quickly also the STATCOM in a position to compensate reactive power components, in this method it takes around 1.5 s to reach stability. The source voltage and inverter currents exactly follow phasor relations before and after changing the load.

## ***5.2 Performance of STATCOM for Sinusoidal Phase Disposition Pulse Width Modulation***

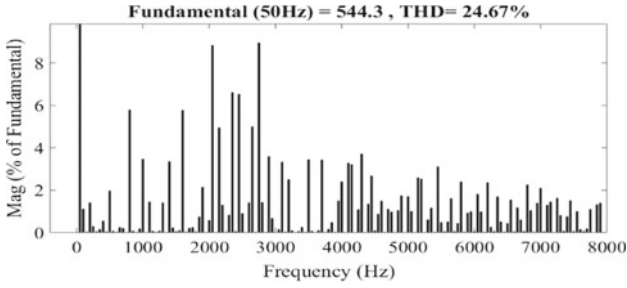
The various STATCOM performances by applying SPDPWM method is shown in Fig. 8, from the results it is seen that the ripples in the voltage among the capacitors of DC link is near to 30 V, which is very high value. The harmonics in output current are 12.09% and in the voltage is 24.67%. The harmonics in output currents are very high. If the load changes suddenly also the STATCOM in a position to compensate reactive power components, but it takes around 2.5 s to reach stability. The inverter currents lag with respect to source voltage for inductive load and it is leading after 1 s by changing the load to capacitive. The ripples in DC link, harmonics in output current, settling time for reactive power components, the ripples in reactive power components during load changes are very high in this method with respect to SPSCPWM technique.

## ***5.3 STATCOM Performance for Third Harmonic Phase Shifted Carrier PWM***

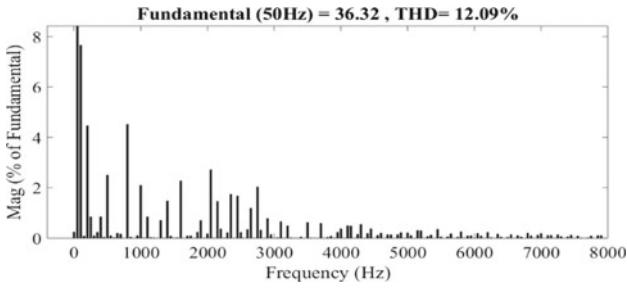
The performances of STATCOM for the application of THIPSCPWM technique is shown in Fig. 9, it is seen that the use of THIPSCPWM technique the ripple content across DC link capacitor voltage is around 10 V and which is low similar to SPSCPWM. The output current harmonics are 3.43% and, in the voltage, it is 22.51%. If the load changes suddenly also the STATCOM in a position to compensate reactive power components, it takes around 1.5 s to reach stability. The inverter currents change from lagging to leading with respect to source voltage subsequently shifting the load from RL to RC. Like in SPSCPWM technique here also no harmonics detected below 4.8 kHz frequency. Over all this technique performance is even better than the performance of SPSCPWM technique.

## ***5.4 STATCOM Performances for Modified Space Vector Phase Shifted Carrier PWM***

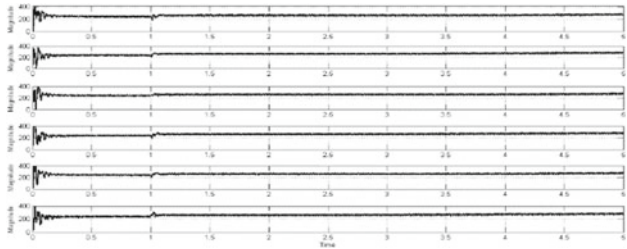
The simulation results of the STATCOM for the MSVPSCPWM method is shown in Fig. 10, it is seen that the ripples in voltage across the DC link capacitors is around 9 V and which is low similar to SPSCPWM and THIPSCPWM. The output current harmonics are of 2.92% and in the voltage is 19.44%. If the load changes suddenly also the STATCOM in a position to compensate reactive power components, it takes around 1.5 s to reach stability. Like in SPSCPWM and THIPSCPWM techniques here also no harmonics present up to 4.8 kHz frequency. This technique performance is better than the performance of SPSCPWM and THIPSCPWM techniques in terms all the performances.



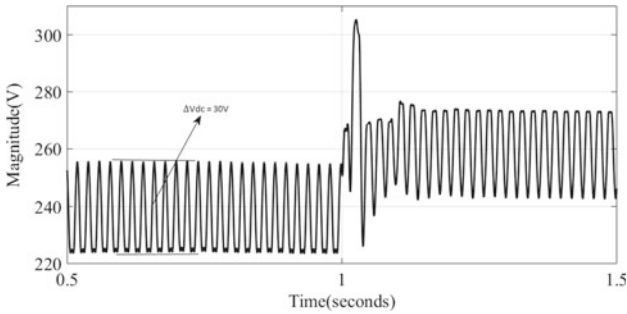
a. Output Voltage Harmonic Spectrum



b. Harmonics Spectrum of Output current

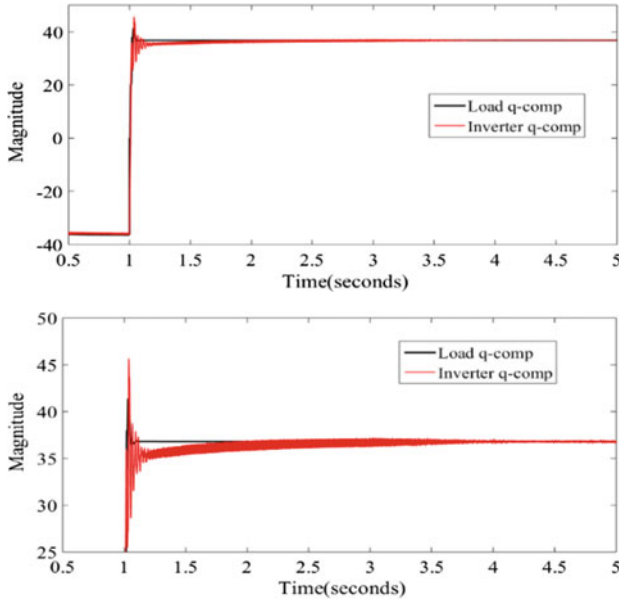


c. Voltages across DC-Link capacitor

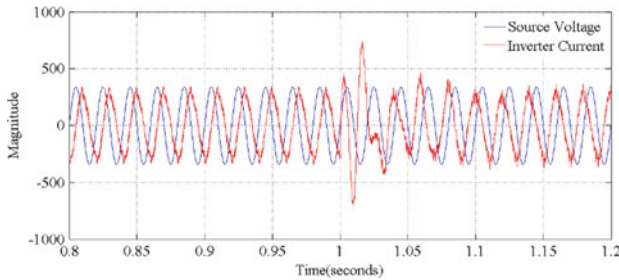


d. Ripples in voltage of DC-link capacitor

**Fig. 8** Responses of STATCOM for sudden change of load from inductive to capacitive at 1 s with SPDPWM technique



e. Reactive components of currents-Load and inverter



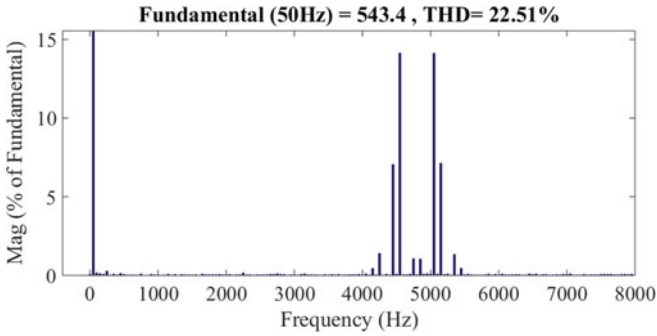
f. Phasor relations of source voltage and inverter current

Fig. 8 (continued)

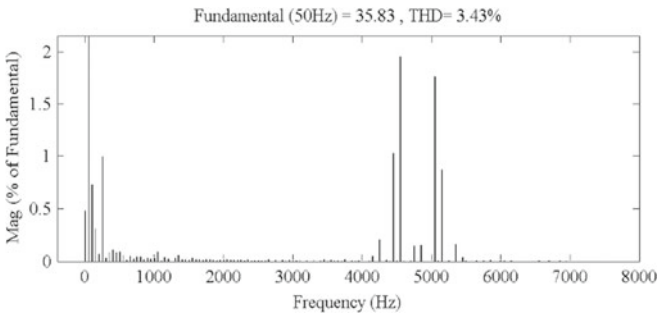
### 5.5 STATCOM Performance for Modified Space Vector Phase Disposition PWM

The simulation results of the STATCOM for the MSVPDPWM technique is shown in Fig. 11, it is seen that ripples in the voltage across the DC link capacitor is around 20 V and which is high with respect to SPSCPWM and THIPSCPWM. The output current harmonics are 13.10% and in the voltage is 27.22%. If the load changes suddenly also the STATCOM in a position to compensate reactive power components, it takes around 1.25 s to reach stability. The performance assessment of the STATCOM for various PWM methods is shown in Table 3 and Fig. 12.

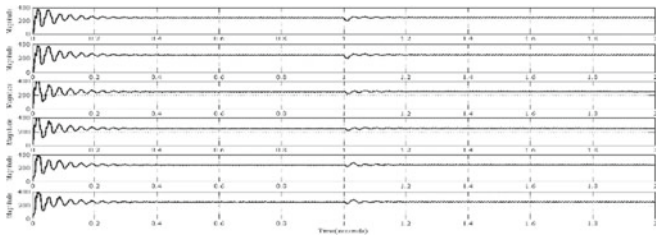




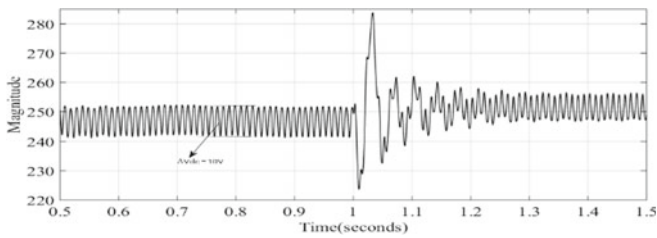
**a. Output Voltage Harmonic Spectrum**



**b. Harmonics Spectrum of Output current**

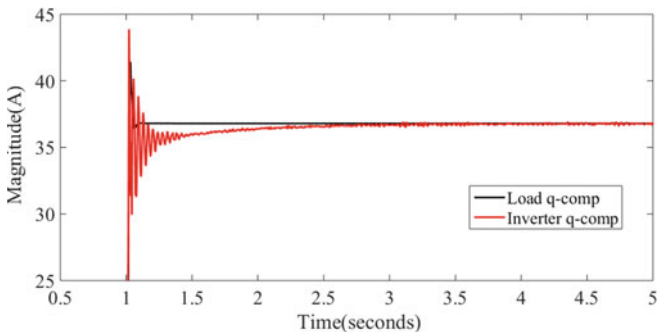
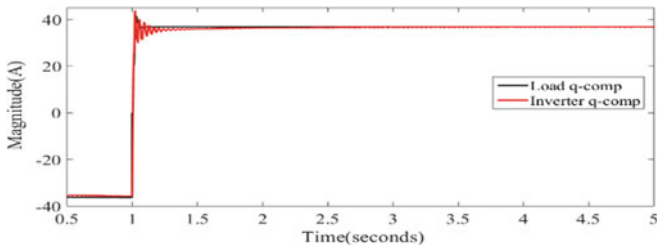


**c. Voltages across DC-Link capacitor**

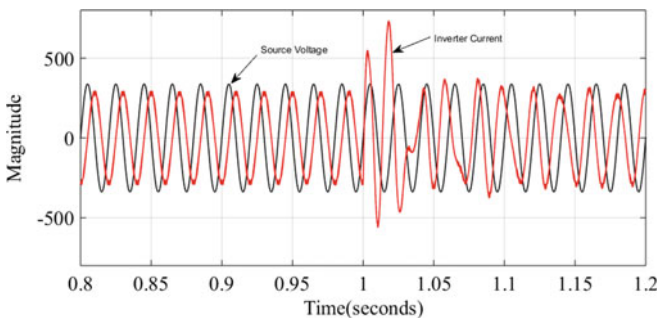


**d. Ripples in voltage of DC-link capacitor**

**Fig. 9** Responses of STATCOM for sudden change of load from inductive to capacitive at 1 s with THIPSCPWM



**e. Reactive components of currents-Load and inverter**

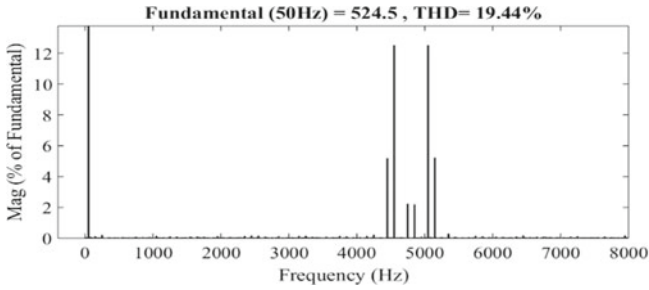


**f. Phasor relations of source voltage and inverter current**

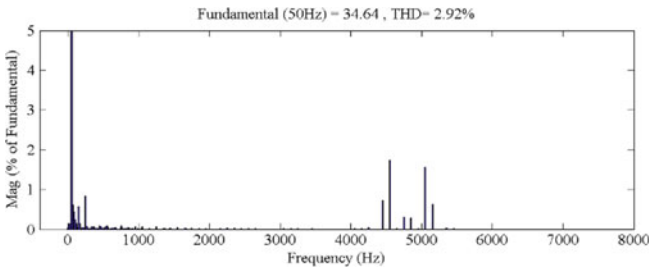
**Fig. 9** (continued)

## 6 Conclusion

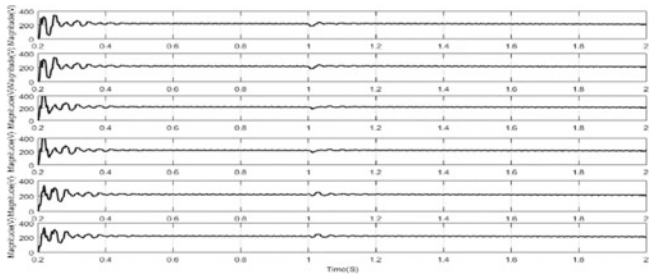
The main objective of the cascaded multilevel inverter-based STATCOM is to compensate reactive powers under different load conditions. The DC link voltage balancing and supplying of reactive powers required by sudden variations of load is done by implementing various PWM methods. The MSVPSPWM strategy obtained better output voltages and currents with THD in output voltage are 19.44% and in output current 2.92%. The ripple in DC link voltage is minimum in SPSC and MSVPSC techniques. Settling time obtained for reactive power components supplied by STATCOM during sudden change of load is low for all the PWM techniques except



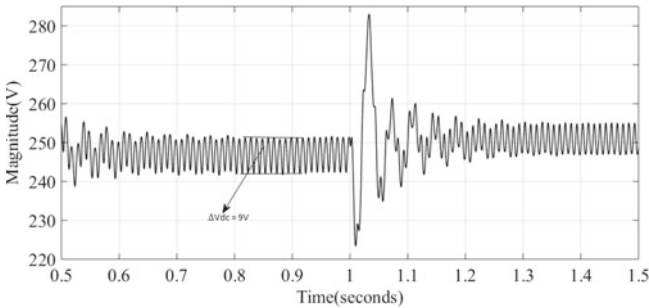
**a. Output Voltage Harmonic Spectrum**



**b. Harmonics Spectrum of Output current**

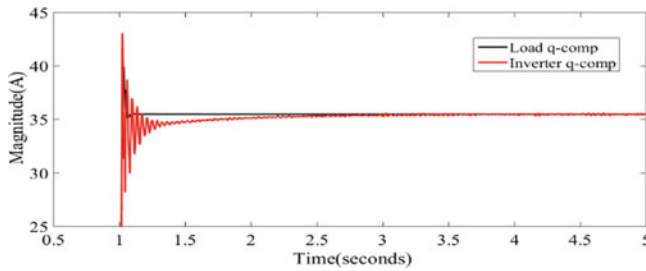
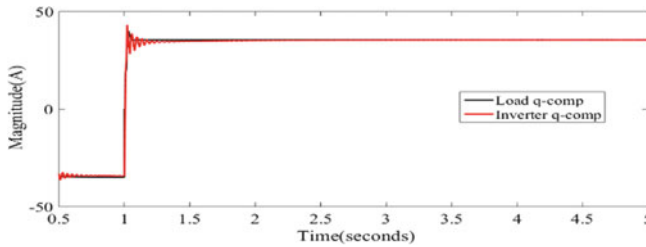


**c. Voltages across DC-Link capacitor**

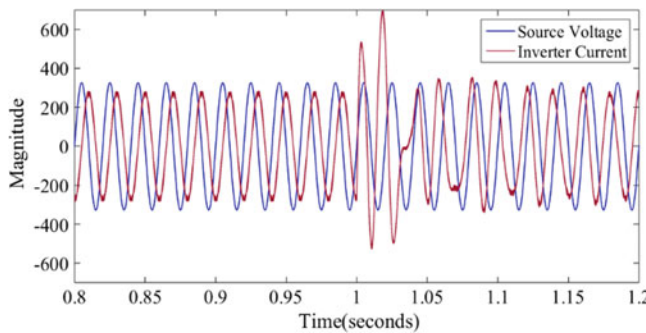


**d. Ripples in voltage of DC-link capacitor**

**Fig. 10** Responses of STATCOM for sudden change of load from inductive to capacitive at 1 s with MSVPSCPWM



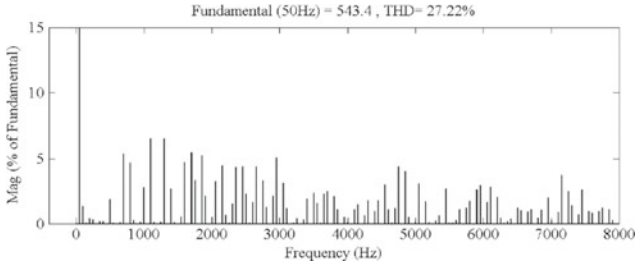
**e. Reactive components of currents-Load and inverter**



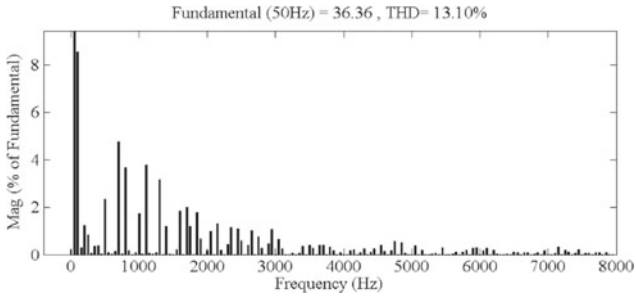
**f. Phasor relations of source voltage and inverter current**

**Fig. 10** (continued)

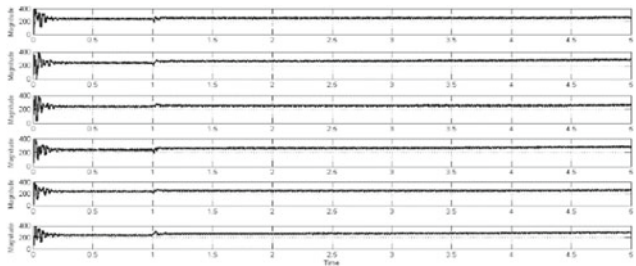
SPD. In all the Phase Shifted Carrier PWM techniques no harmonics will be observed before 4.8 kHz frequency. From the comparison of PWM techniques, it is observed that MSVPSCPWM technique performance is better and can be implemented easily for all levels of cascaded multilevel STATCOM.



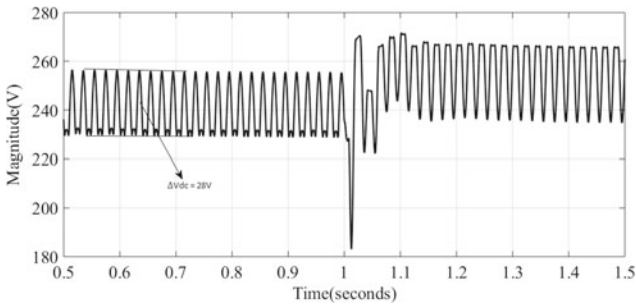
a. Output Voltage Harmonic Spectrum



b. Harmonics Spectrum of Output current

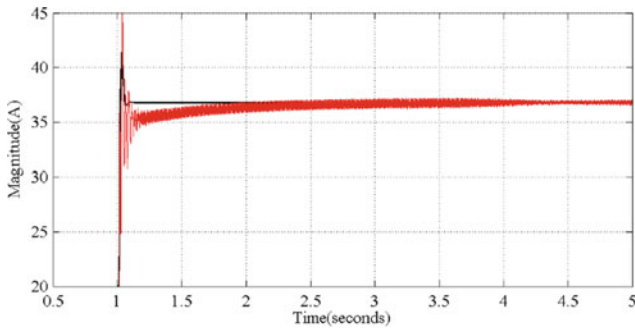
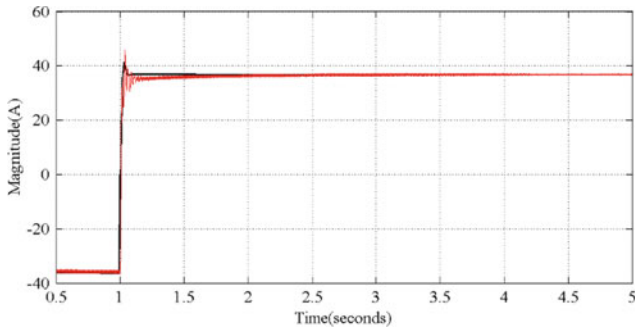


c. Voltages across DC-Link capacitor

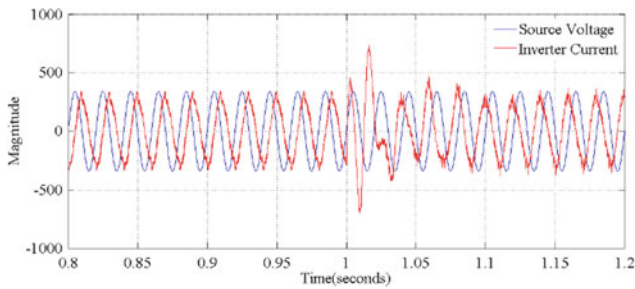


d. Ripples in voltage of DC-link capacitor

Fig. 11 Responses of STATCOM for sudden change of load from inductive to capacitive at 1 s with MSVPDPWM



**e. Reactive components of currents-Load and inverter**



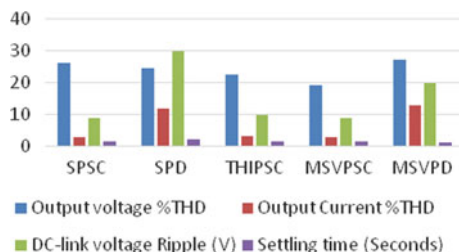
**f. Phasor relations of source voltage and inverter current**

**Fig. 11** (continued)

**Table 3** STATCOM performances for various PWM methods

S. No	PWM technique(s)	Output voltage %THD	Output Current %THD	DC link voltage ripple (V)	Settling time (Seconds)
1	SPSC	26.31	3.01	9	1.5
2	SPD	24.67	12.09	30	2.5
3	THIPSC	22.51	3.43	10	1.5
4	MSVPSC	19.44	2.92	9	1.5
5	MSVPD	27.22	13.10	20	1.25

**Fig. 12** Comparison of different PWM technique performances of STATCOM



## References

- Xu R et al (2015) A novel control method for transformerless H-bridge cascaded statcom with star configuration. *IEEE Trans Power Electron* 30(3):1189–1202. <https://doi.org/10.1109/TPEL.2014.2320251>
- Celanovic N, Boroyevic D (2001) A fast space vector modulation algorithm for multilevel three-phase converters. *IEEE Trans Ind Appl* 37(2):637–641. <https://doi.org/10.1109/28.913731>
- Rodriguez J, Lai JS, Peng FZ (2002) Multilevel inverters: a survey of topologies, controls, and applications. *IEEE Trans Ind Electron* 49(4):724–738. <https://doi.org/10.1109/TIE.2002.801052>
- Malinowski M, Gopakumar K, Rodriguez J, Pérez MA (2010) A survey on cascaded multilevel inverters. *IEEE Trans Ind Electron* 57(7):2197–2206. <https://doi.org/10.1109/TIE.2009.2030767>
- Surendra Babu NNV, Al Hosani K (2015) A novel DC voltage control for a cascade H-bridge multilevel STATCOM. In: *IECON 2015—41st annual conference of the IEEE industrial electronics society*, 003330–003335. <https://doi.org/10.1109/IECON.2015.7392613>
- Naderi R, Rahmati A (2008) Phase-shifted carrier PWM technique for general cascaded inverters. *IEEE Trans Power Electron* 23(3):1257–1269. <https://doi.org/10.1109/tpe.2008.921186>
- Reddy CL, Kumar PS, Sushama M (2015) Babu NNVS a five level cascaded H-bridge multilevel STATCOM. In: *IEEE Asia Pacific conference on postgraduate research in microelectronics and electronics (Primeasia)*, pp 36–41. <https://doi.org/10.1109/PrimeAsia.2015.7450466>
- Lokeshwar Reddy C, Satish Kumar P, Sushama M (2017) Design and performance analysis of 7-level diode clamped multilevel inverter using modified space vector pulse width modulation techniques. *Int J Eng* 30(11):1762–1770. <https://doi.org/10.5829/ije.2017.30.11b.18>
- Reddy CL, Satish Kumar P, Sushama M (2017) Performance comparison of star connected cascaded statcom using different PWM techniques. *Int J P Electron Drive Syst* 8(3):1303–1319. <https://doi.org/10.11591/ijpeds.v8i3>
- Chintala LR, Satish Kumar P, Malaji S (2016) Improvement in performance of cascaded multilevel inverter using triangular and trapezoidal triangular multi carrier SVPWM. *Adv Electr Electron Eng* 14(5):562–570. <https://doi.org/10.15598/aeee.v14i5.1767>

# Outlier Detection Using Linear Regression in Wind and Solar Integrated Power Systems



Priyanka Khirwadkar Shukla and R. Mahalakshmi

## 1 Introduction

The traditional power system has four components—generation, transmission, distribution and utilization. Recent years has seen unbundling of the components and led to restructured power system. The present day restructured power system is a complex mix of renewable and non-renewable or traditional energy sources.

Restructured power system provides a competitive market for buying and selling electricity. Maintaining the power quality is prime importance in recent years. The faults occurring in the power system need to be monitored continuously and mitigated as soon as possible to avoid the chances of cascaded faults leading to blackout of the entire power system. To overcome this problem, digitalization of power system comes into view and hence, introducing the concept of Smart Grid (SG). SG uses low cost and high-quality sensors to gather data from vast areas over a long duration of time. This data monitors the regulation and operational parameters of the power system and hence, enabling the engineer to make quick decision.

This integration of renewable energy sources in traditional power systems has led to more emphasis on system reliability, system stability, condition monitoring (CM), fault diagnosis, unit commitment (UC) and economic dispatch (ED) which are interdependent [1]. The rapidly changing power system parameters require an equally fast paced method for computation over traditional methods and hence, require an automated process. The machine learning (ML) algorithms provide such a platform. ML is defined as a subsection of artificial intelligence. It helps to design and develop

---

P. Khirwadkar Shukla (✉) · R. Mahalakshmi  
Department of Electrical and Electronics Engineering, Amrita School of Engineering, Amrita  
Vishwa Vidyapeetham, Bengaluru, India  
e-mail: [pri.gecb@gmail.com](mailto:pri.gecb@gmail.com); [k\\_priyanka@blr.amrita.edu](mailto:k_priyanka@blr.amrita.edu)

R. Mahalakshmi  
e-mail: [d\\_mahalakshmi@blr.amrita.edu](mailto:d_mahalakshmi@blr.amrita.edu)



algorithms to evaluate behavior built on empirical data. It programs the computers to enhance output by using example or experience. However, outlier detection is required before executing any ML algorithms to obtain proper results.

Outlier detection is the process of discovery of dataset points with different behaviors than as expected. Such objects are called outliers or anomalies [2]. Outliers are abnormal values in a dataset that do not follow the regular distribution and can undoubtedly distort any regression model. Therefore, outliers must be carefully handled to get the correct information from the dataset. Statistical approach method for outlier detection has two methods: parametric and non-parametric methods. In parametric method, the models increase with model complication and not with data size. In non-parametric method, the model is learned from input dataset rather than assumptions. These methods are only applicable for numerical data [2].

Various literature has explored the traditional fault diagnosis methods, but they cannot be applied to the dynamic power system. It is further unsuitable for complex systems [3–9]. Haerberlin and Beutler [3] introduced novel method in which power was standardized. The method led to dividing losses into its components which led to thorough on-line and off-line scrutiny of irregular system malfunctions. A procedure was developed in [4] where the satellite gave information regarding solar irradiance which was swapped with measurements obtained from the site.

New analysis techniques where fault was categorized in four ways and the related loss of energy was calculated in [5]. The real time monitoring and detection of fault using inferential process for connected and disconnected mode was introduced in [6].

In [7], proposed system constantly checked of the observed capture losses for a PV system. A new monitoring was proposed in [8] where irradiation level data was not acquired from PV module specification or the sensors. [9] used decision tree for a fault detection and on PV array.

Khalil et al. [10] used Principal Analysis Component (PCA) and [11] used predictive models for ML algorithms in hardware and semiconductor industries, respectively. Random forest supervised model was used in [12] to find the correlation between different impact factors and target outputs to detect chemical properties of an element. Bishnu and Bhattacharjee [13–16] used ML algorithms—k-means clustering, quad-tree-based algorithm, ensemble learning, ANN and SVM to predict software faults. Though, [10–16] have used ML algorithms for fault detection they have not mentioned outlier detection.

ML algorithms were used in [17–19] for wind and solar connected systems. Chinmay Krishnan and Sindhu Thampatty [20] used ML for predicting sub-synchronous resonance oscillations. However, the outlier detection was not considered. Hence, in this paper, outlier detection using linear regression has been implemented on time—series dataset using Python.

This paper proceeds as follows: Sect. 2 shows the flow chart for the paper. Section 3 explains the data preprocessing of the dataset followed by Sect. 4 for linear regression. The results are discussed in Sect. 5 and the conclusion for this paper is given in Sect. 6.

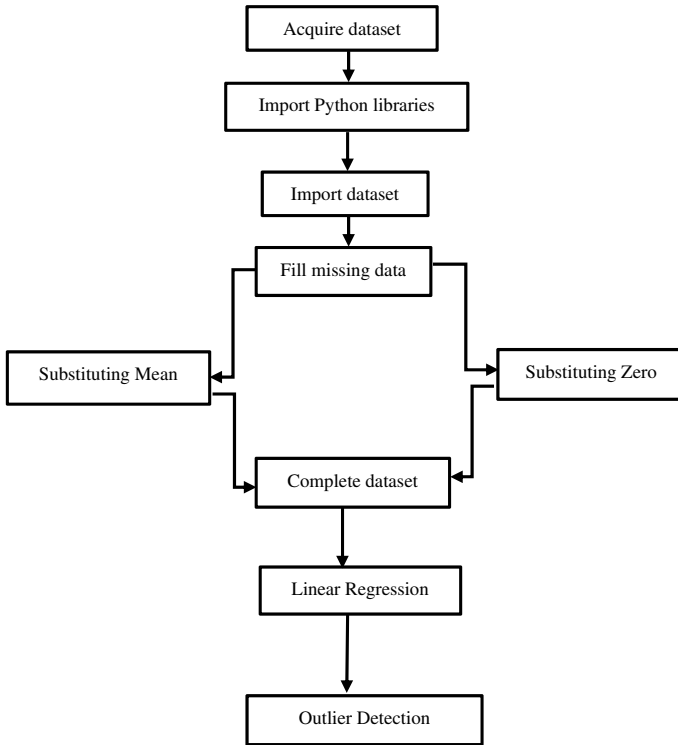


Fig. 1 Schematic diagram

## 2 Flow Chart

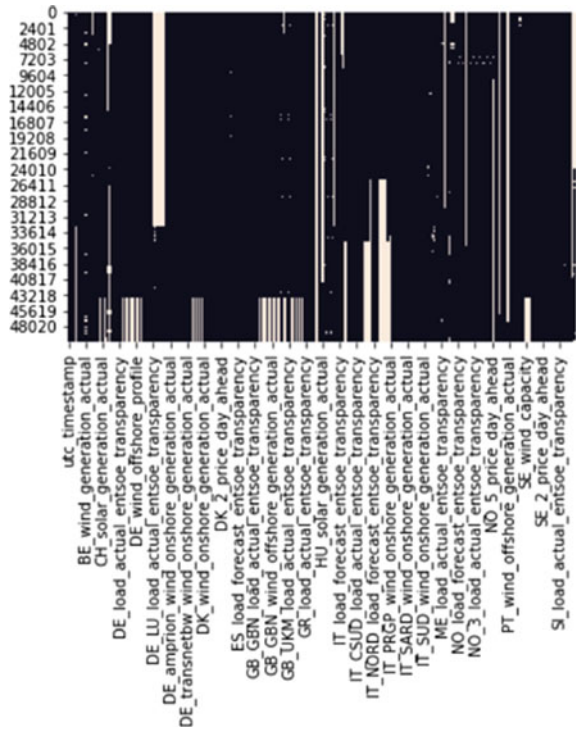
This section shows the step-by-step approach followed for outlier detection (Fig. 1). Once the data is complete and structured, the ML algorithms for outlier detection can be applied. In this paper, linear regression is applied for outlier detection.

## 3 Data Preprocessing

Data preprocessing is taking the raw data which is unstructured, incomplete and inconsistent and structuring it such that the ML algorithms can understand it and the relevant outputs can be obtained. Data preprocessing is an exhaustive process; however, utmost caution should be taken as it is the most important step in developing the ML models.

The dataset used in this paper was time series 60 min single index data for Great Britain which included Northern Ireland and United Kingdom also. The dataset had various features day ahead price, actual load, load forecast, solar capacity, actual

Fig. 2 Incomplete dataset



solar generation, solar profile, wind capacity, actual wind generation, wind profile, wind offshore capacity, actual offshore wind generation, wind onshore capacity and actual onshore wind generation.

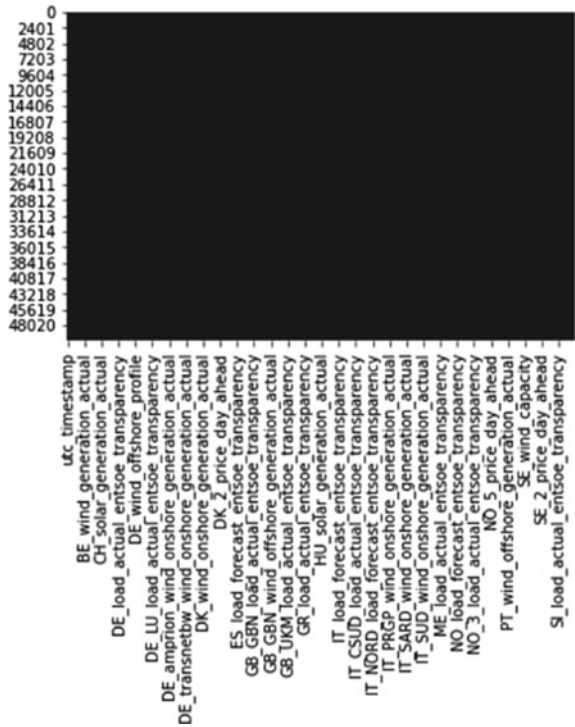
The dataset was having missing values (Fig. 2). This missing data was filled by using the mean value of the column or zero value, depending upon the feature being filled (Fig. 3).

For preprocessing, Python software was used. The dataset was uploaded in the Python software and relevant libraries were imported. The missing data was then filled by mean value or zero value and hence, a complete and structured dataset was obtained.

## 4 Linear Regression

Linear regression is linear process for modeling the correlation between an independent variable and its response. As linear regression is simple, interpretable, scientifically acceptable and widespread available, it is the first choice for analysis of any given dataset. Hence, this paper also uses linear regression for outlier detection.

Fig. 3 Complete dataset



Let,

$$x = (x_1, x_2, \dots, x_r)$$

where

$x$  independent features (actual load, load forecast, solar capacity, actual solar generation, solar profile, wind capacity, actual wind generation, wind profile, wind offshore capacity, actual offshore wind generation, wind onshore capacity and actual onshore wind generation).

$r$  no. of predictors.

$Y$  dependent feature (day ahead price).

Assume  $y$  and  $x$  are linearly related.

$$Y = \beta_0 + \beta_1 x_1 + \dots + \beta_r x_r + \varepsilon \tag{1}$$

where  $\beta_0, \beta_1, \dots, \beta_r$ —regression coefficients.

$\varepsilon$ —random error.

LR evaluates the  $\beta$  values.

For every  $i = 1$  to  $n$  values,  $f(x_i)$  and  $Y_i$  should have minimum difference and this difference is called residual. The weights of smallest residuals are the most appropriate for LR. The weights are calculated by the summing up of all the squared residuals (SSR) and selecting the minimum value for  $i = 1$  to  $n$ .

$$SSR = \sum_i (Y_i - f(x_i))^2 \tag{2}$$

This method is called the ordinary least squares (OLS) regression. The  $r$ -squared value indicates the correlation between  $Y$  and  $x$ . The  $r$ -squared value ranges from 0 to 1 or 0 to 100%. The best fitted LR model has larger  $r$ -squared for better explanation of the relationship of the output with different inputs.

## 5 Results and Discussion

The time series single index 60 min dataset was selected for the analysis. The dataset had 30 features. Of these 30 features, day ahead price was the dependent feature and the remaining 29 features were independent. The dataset was initially incomplete and was filled by using mean value of the feature or zero value. The linear regression algorithm was applied to the preprocessed dataset to get the results. OLS regression method was used. Tables 1, 2 and 3 show the OLS regression results which are obtained by running the ordinary least squares regression algorithm on Python.

**Table 1** Ordinary least squares regression results

Dep. variable	y	R-squared	0.319
Model	OLS	Adj. R-squared	0.319
Method:	Least squares	F-statistic	635.9
No. observations	35,280	Prob (F-statistic)	0.00
Df residuals	35,253	Log-likelihood	-1.485e + 05
Df model	26	AIC	2.971e + 05
Covariance type	Non-robust	BIC	2.973e + 05

**Table 2** Ordinary least squares regression results continued

Omnibus	86,864.639	Durbin-Watson	2.007
Prob (Omnibus)	0.000	Jarque-Bera (JB)	2,174,192,916.368
Skew	26.153	Prob (JB)	0.00
Kurtosis	1218.034	Cond. No	1.02e +16

**Table 3** Ordinary least squares regression results continued

Const	Coef	Std. err	Const	Coef	Std. err
	43.1465	3.189	×15	0.0062	0.001
×1	0.0003	0.000	×16	-0.0033	0.001
×2	0.0011	0.000	×17	0.0129	0.006
×3	-0.1220	0.010	×18	-3.5475	0.249
×4	-0.0013	5.23e-05	×19	-0.0034	0.002
×5	-0.0629	0.852	×20	-0.0002	0.000
×6	-5.0687	0.260	×21	0.0001	0.000
×7	0.1379	0.166	×22	0.1270	0.010
×8	-168.531	27.466	×23	-0.0013	5.23e-05
×9	0.7440	0.129	×24	3.5816	0.271
×10	-0.0670	0.083	×25	-0.0065	0.002
×11	64.9987	10.067	×26	0.7440	0.129
×12	1.4828	0.279	×27	-0.0670	0.083
×13	-0.1394	0.166	×28	-0.0074	0.242
×14	105.4003	17.645	×29	0.0074	0.003

From Table 1, the following conclusions are drawn—R—squared value: percent variation in dependent feature corresponding to independent features.

Adj. R-squared: modified r-squared value for regression.

Prob (F-Statistic): Significance of regression.

AIC/BIC: Akaike’s Information Criteria/Bayesian information criteria used for model selection.

From Table 2, the following conclusions are drawn

Prob (Omnibus): Should be close to 1 to fulfill OLS. Here, the value is 0 that means the OLS is not satisfied.

Durbin-Watson: Should be between 1 and 2. Here, the value is 2.007 that means regression results are unreliable.

Prob (Jarque–Bera): Large value indicates errors are not normally distributed.

The *r*-squared value is calculated using the weights obtained from Table 3.

The standard errors assume that the covariance matrix of the errors is correctly specified. Table 1 shows the coefficient of determination and standard error. The range of *t*-value is from -23.891 to 13.531 and the probability of *t*-value is from 0 to 1. The smallest eigenvalue is  $2.05 \times 10^{-18}$ . This indicates that the matrix is either highly multicollinear or singular.

$$r\_squared\ value = 0.43912243123252714$$

Figure 4 shows the error terms density of predicted and actual *y* variable. Figure 5 shows the residual values with respect to the *x* variable data used for training of

Fig. 4 Error

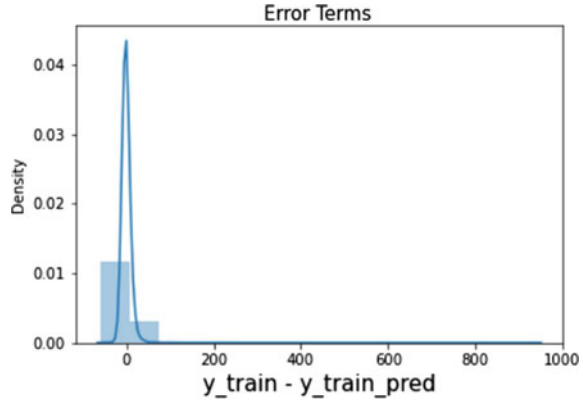
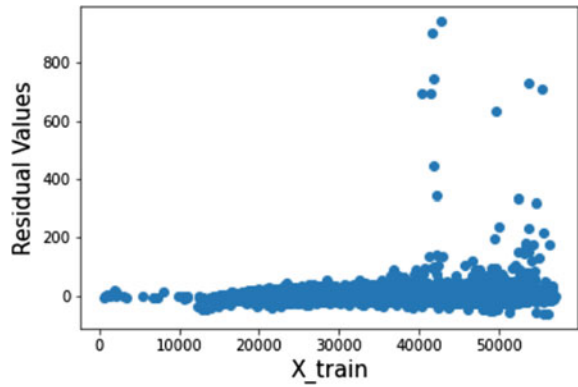


Fig. 5 Residual values



the algorithm. The residual values are error terms calculated based on predicted and actual  $y$  variables.

The  $r$ -squared value obtained is 43.91%. The greater the  $r$ -squared value the stronger the strength of correlation among the various features being analyzed. The coefficient of determination or the  $r$ -squared value with 70% is considered good and with 50% is considered a moderate fit. Hence, the  $r$ -squared value obtained indicates that the ML algorithm used does not explain the variation in day ahead price with respect to various features and hence, it does not explain the regression model properly. So, it is not a suitable model for outlier detection.

## 6 Conclusion

The utilization of renewable energy has increased tremendously over the last two decades for most of the countries. The integration of renewable energy sources to

traditional power systems has led to power system stability issues. Hence, it is of prime importance to have real time fast monitoring and control actions. The traditional methods are not fast paced to match the rapidly changing power system. Hence, a fast and automatic process is required.

ML provides such a platform where the algorithms are trained using examples and experiences. However, the dataset used for training and testing purposes should not have any misleading/erroneous data. Hence, outlier detection comes as a precursor of ML algorithms and ensures correct information is obtained from the dataset. So, outlier detection was explored in this paper.

The outlier detection was done on a real time series 60 min single index dataset. Initially, the dataset was incomplete and hence, it was preprocessed to obtain a complete and structured dataset. The missing values of the dataset were filled by using either the mean value or 0 value depending upon the feature being filled.

The dataset had various features: day ahead price, actual load, load forecast, solar capacity, actual solar generation, solar profile, wind capacity, actual wind generation, wind profile, wind offshore capacity, actual offshore wind generation, wind onshore capacity and actual onshore wind generation which were correlated.

Once the dataset was preprocessed, linear regression using the ordinary least square method was implemented. The coefficient of determination or  $r$ -squared value was calculated. However, the results obtained showed that this method was not suitable for outlier detection for the chosen dataset. Hence, as a future scope, the other ML methods are being suggested.

## References

1. Bhattacharya M, Paramati S, Ozturk I, Bhattacharya S (2016) The effect of renewable energy consumption on economic growth: evidence from top 38 countries. *Appl Energy* 162:733–741
2. Modi K, Oza B (2016) Outlier analysis approaches in data mining. *IJIRT* 3(7)
3. Haeblerlin H, Beutler C (1995) Normalized representation of energy and power for analysis of performance and on-line error detection in PV-systems. In: 13th EU PV conference on photovoltaic solar energy conversion, Nice France
4. Drews A, de Keizer AC, Beyer HG, Lorenz E, Betcke J, van Sark WGJHM, Heydenreich W, Wiemken E, Stettler S, Toggweiler P, Bofinger S, Schneider M, Heilscher G, Heinemann D (2007) Monitoring and remote failure detection of grid connected PV systems based on satellite observations. *Sol Energy* 81:548–564
5. Firth SK, Lomas KJ, Rees SJ (2010) A simple model of PV system performance and its use in fault detection. *Sol Energy* 84:624–635
6. Vergura S, Acciani G, Amoroso V, Patrono GE, Vacca F (2009) Descriptive and inferential statistics for supervising and monitoring the operation of PV plants. *IEEE Trans Ind Electron* 56:4456–4464
7. Chouder A, Silvestre S (2010) Automatic supervision and fault detection of PV systems based on power losses analysis. *Energy Convers Manage* 51:1929–1937
8. Kang BK, Kim ST, Bae SH, Park JW (2012) Diagnosis of output power lowering in a pv array by using the kalman-filter algorithm. *IEEE Trans Energy Convers* 1–10
9. Zhao Y, Yang Y, Lehman B, de Palma JF, Mosesian J, Lyons R (2012) Decision tree—based fault detection and classification in solar photovoltaic arrays. In: *Proceedings of 27th IEEE APEC*, 93–99



10. Khalil K, Eldash O, Kumar A, Bayoumi M (2020) Machine learning-based approach for hardware faults prediction. *IEEE Trans Circ Syst I Regul Pap* 67(11):3880–3892
11. Munirathinam S, Ramadoss B (2016) Predictive models for equipment fault detection in the semiconductor manufacturing process. *IACSIT Int J Eng Technol* 8(4):273–285
12. Wang X, Han D, Hong Y, Sun H, Zhang J, Zhang J (2019) Machine learning enabled prediction of mechanical properties of tungsten disulfide monolayer. *ACS Omega* 4(6):10121–10128
13. Bishnu P, Bhattacharjee V (2011) Software fault prediction using quad tree-based k-means clustering algorithm. *IEEE Trans Knowl Data Eng* 24(6):1146–1150
14. Wójcicki B, Dabrowski R (2018) Applying machine learning to software fault prediction. *e-Informatica Softw Eng J* 12(1)
15. Khuat T, Le M (2019) Ensemble learning for software fault prediction problem with imbalanced data. *Int J Electr Comput Eng* 9(4):3241
16. Gondra I (2008) Applying machine learning to software fault-proneness prediction. *J Syst Softw* 81(2):186–195
17. Gautam M, Raviteja S, Mahalakshmi R (2019) Household energy management model to maximize solar power utilization using machine learning. *Procedia Comput Sci* 165:90–96
18. Gautam M, Raviteja S, Mahalakshmi R (2019) Energy management in electrical power system employing machine learning. In: *International conference on smart systems and inventive technology (ICSSIT)*. Tirunelveli, India, 915–920 2019
19. Mahalakshmi R, Reddy KS, Gautam M (2020) Control of Standalone DFIG based wind turbine generator using machine learning algorithm. In: *4th international conference on electronics, communication and aerospace technology (ICECA)*. Coimbatore, India, pp 204–209
20. Chinmay Krishnan KC, Sindhu Thampatty KC (2020) Prediction of sub-synchronous resonance oscillations-a machine learning approach. In: *IEEE students conference on engineering systems*

# Photovoltaic Systems Incorporated with Energy Storage System for Agricultural Implementation



V. Hemant Kumar, Pakki Pavan Kumar, R. N. Patel, and Vivek Bargate

## 1 Introduction

Renewable energy generation is friendly power option and which is quick creating and sunlight based photovoltaic (PV) frameworks have outperformed record establishments lately. Writing on solar PV associated themes covers procedures for most maximum point following (MPPT), which were as of late surveyed for instance in, power quality upgrades and explicit regulators and explicit power electronic circuitries for DC to DC converters, DC to AC converters, and single level power converters. Since the solar radiance disparity between PV boards associated in arrangement may restrict the force yield, the focal point of a portion of the distributed exploration is on exhibit reconfiguration to adjust impacts. Different papers propose segments incorporated converters, to such an extent that the MPP for every segment can be trailed, yielding most of the extreme extraction of energy. PV Module coordinated DC-DC converters linked with a main DC/AC converter were projected in many configurations.

These days, sustainable power has been increasingly more alluring because of the extreme ecological assurance guidelines and the deficiency of regular fuel sources. Photovoltaic era is the approach which utilizes photo-voltaic cell to change sun-based

---

V. H. Kumar (✉) · P. P. Kumar · R. N. Patel  
Department of Electrical Engineering, National Institute of Technology Raipur, Raipur, C.G.,  
India  
e-mail: [hemant249v@gmail.com](mailto:hemant249v@gmail.com)

R. N. Patel  
e-mail: [rmpatel.ee@nitrr.ac.in](mailto:rmpatel.ee@nitrr.ac.in)

V. Bargate  
Department of Electrical and Electronics Engineering, Chhatrapati Shivaji Institute of  
Technology, Durg, C.G., India

energy over to electrical energy. Photovoltaic power is expecting progressively significant as an environmentally friendly power source due for its unmistakable potential benefits, like basic arrangement, simple assignment, liberated from contamination, low support cost, and so forth. Nonetheless, the detriment is that photovoltaic age is discontinuous, contingent on climate conditions [1]. Hence, energy stockpiling component is important to help get steady and dependable force from PV framework for burdens or utility lattice, and along these lines improve both consistent and dynamic practices of the entire age framework. On account of it develop innovation, minimal expense and high productivity, battery energy stockpiling framework (BESS) is utilized broadly in conveyance age innovation. BESS can be coordinated into PV age framework to shape a half breed PV/Battery age framework, which can be steadier and more dependable [2]. A basic network associated PV/Battery age framework is made out of PV exhibit, battery, power electronic converters, channels, regulators, nearby loads and utility lattice.

The current paper talks about a topology of a PV systems to drive agricultural loads using BLDC motor drive associated with synchronized energy stockpiling arrangement. The PV framework under examination is isolated into a few areas, and each part has its own DC-DC converters for conversion and allowing steady power flow between the sources to preserve constant DC voltage across the link. In the projected plan, the converters are made to acquire a voltage waveform practically identical with that would be accessible from a staggered converter. The framework and its strategy to control are recreated on a two individual PV power framework utilizing the MATLAB/Simulink programming. A battery energy stockpiling framework (BESS) is utilized to give capacity to the lattice if there should arise an occurrence of fractional or complete concealing after a proposed control plot as per which if the battery's condition of charge is lower than the PV systems. The results of motor drive to validates that if there are dynamics in load variations there is a constant voltage across it for its operation under varying irradiance levels.

## 2 Proposed Multi-PV System and Control

In PV applications, a numeral number of different methods have been established and used over the last few decades. An intense review of such systems available in many of the review papers. Only the some of the common approaches used in smaller residential scale installations.

To use in high level of power-based instruments and technology, it is usually seen that the PV plant is divided into many smaller units for optimal working. Each smaller divided sections of the PV array is linked to boost converter, as a result the generated PV voltage is tracked at its mpp region [3]. It has been that a simple boost converter has many advantages over other dc to dc converters when there are multiple PV linked in the array. The output from the boost dc/dc converter is connected to a bi-directional converter, which generally used to charge and discharge the BESS. Figure 1 shows block diagram of our proposed topology. The battery for energy-

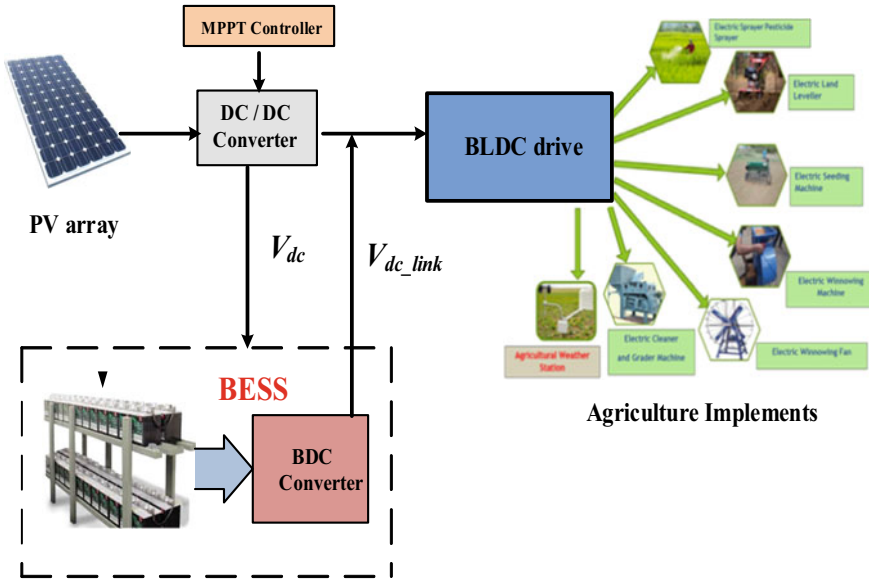


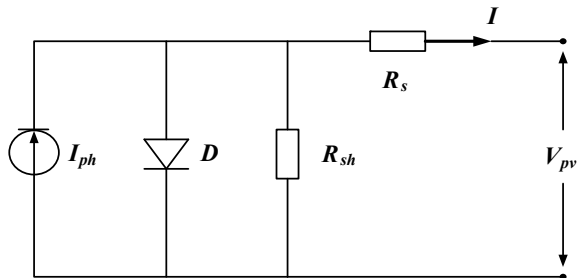
Fig. 1 Proposed solar PV system for driving Agri-loads

storage system (BESS) is connected with the system such that it able to supply power to the DC link done its own BDC whenever there is less power generation due to partial/fully shading on the PV-panels.

### 2.1 Single Model of Photo-Voltaic Cell

A Photo-voltaic cell is the utmost essential part in any PV framework [4–6]. A single diode PV mathematic model is pertinent to reproduce silicon photovoltaic cells, which comprises of a photo-current source  $I_{ph}$ , a non-linear diode, internal series and parallel resistances  $R_s$  and  $R_{sh}$  respectively, as shown in below Fig. 2.

Fig. 2 Single-diode mathematical PV model



The mathematic equation for the voltage-current for a single-diode equivalent circuit can be described as:

$$I = I_{ph} - I_s \left( e^{\frac{q(V+IR_s)}{AkT}} - 1 \right) - \frac{V + IR_s}{R_{sh}} \quad (1)$$

where, “ $I_{ph}$ ” is photo-current; “ $I_s$ ” is diode saturation current; “ $q$ ” is coulomb constant ( $1.602 \times 10^{-19}$  °C); “ $k$ ” is Boltzmann’s constant ( $1.38 \times 10^{23}$  J/K); “ $T$ ” is cell temperature (Kelvin); “ $A$ ” is Ideality factor of P–N junction; “ $R$  and  $R_{sh}$ ” are intrinsic series and shunt resistances [7].

Radiations from sun and temperature at PN junction of PV cell, which can be depicted in below as:

$$I_{ph} = \left( \frac{S}{S_{ref}} \right) [I_{ph,ref} + C_T (T - T_{ref})] \quad (2)$$

where, “ $S$ ” is the radiations coming from sun ( $W/m^2$ ); “ $S_{ref}$ ,  $T_{ref}$  and  $I_{ph,ref}$ ” is the reference values of solar radiations, absolute temperature of cell, tested standard conditions photo current respectively;  $C_T$  is the temperature coefficient (A/K) [8].

Variation of diode saturation current with respect to the temperature of cell is given as:

$$I_s = I_{s\_ref} \left( \frac{T}{T_{ref}} \right)^3 e^{\left[ \frac{qE_g}{Ak} \left( \frac{1}{T_{ref}} - \frac{1}{T} \right) \right]} \quad (3)$$

where  $I_{s\_ref}$  is the reference saturation current of diode in tested standard conditions;  $E_g$  is the band-gap energy of the semiconductor cell (eV), which depends on the material chosen for making cell.

## 2.2 MPPT Algorithm—P&O

Irradiance, panel temperature, and the  $I$ – $V$  characteristic of a PV array are all accounted for variations in MPP (as they are the dependent on the irradiance levels, ambient temperature, heat exchange process efficacy, and panels’ operating point). It’s vital to keep track of the MPP in order to enhance the generating capacity of a PV system under a certain set of functional parameters.

The conversion of incident solar radiation into electrical energy is 30 to 40% efficient in a standard solar PV panel. As a result, MPPT techniques are utilized to increase the efficiency of PV panels [9]. According to the Maximal Power Transfer Theorem, a circuit’s power output is maximum when the circuit’s source impedance equals the load impedance, hence the maximum power point is reduced to an impedance matching problem. As a result of their ease of implementation, MPPT

algorithms are employed in PV systems to maximize power extraction under all climatic circumstances.

The P&O MPPT algorithm approach is utilized to generate output power, PV computation, and power change by taking the samples of both parameters such as current and voltage of PV Array. The variations in voltage (V) are detected by the MPP tracker, and the PV panel's voltage is raised or altered on a routine basis. The next perturbation will be in the same (opposite) direction if the perturbation causes the differential change in PV power (P) to climb (drop). The duty cycle feed to the DC chopper is adjusted until it achieves maximum power. The system oscillates as a result of this variance, which can be decreased by reducing the perturbation step size [10]. The PV array exhibits several typical PV curves with varying maximum power points depending on the irradiance and cell temperatures. The DC-DC converter is powered by the voltage at the point on the curve where maximum power corresponds to maximum voltage. The P&O MPPT algorithm is depicted in the diagram (Fig. 3).

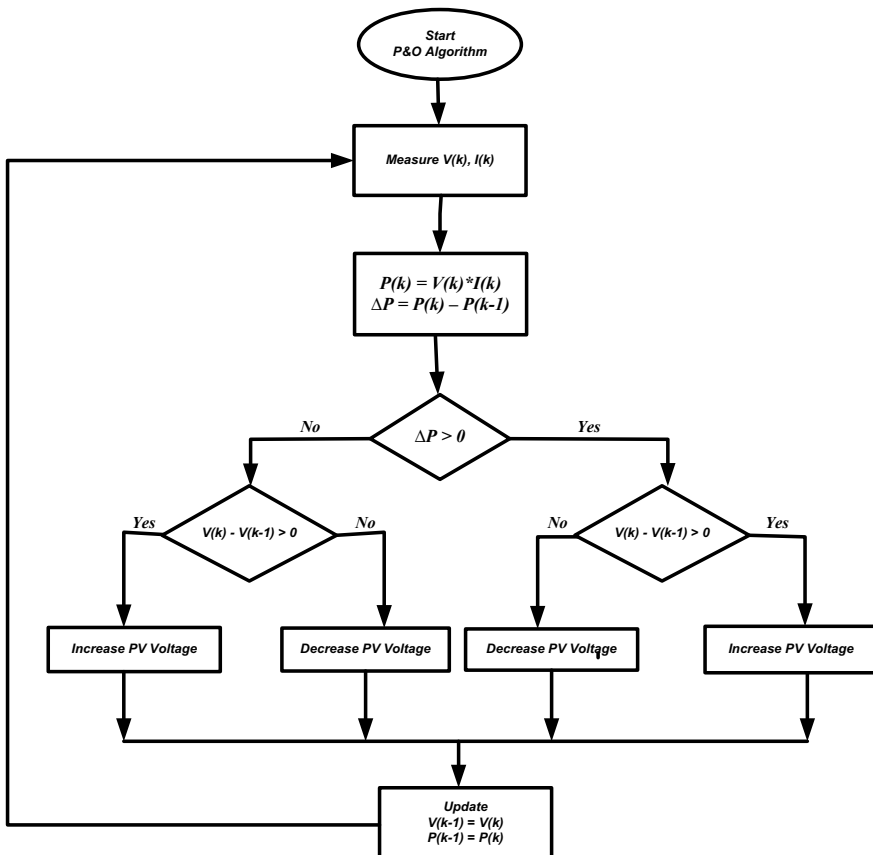


Fig. 3 Flowchart of P&O MPPT algorithm

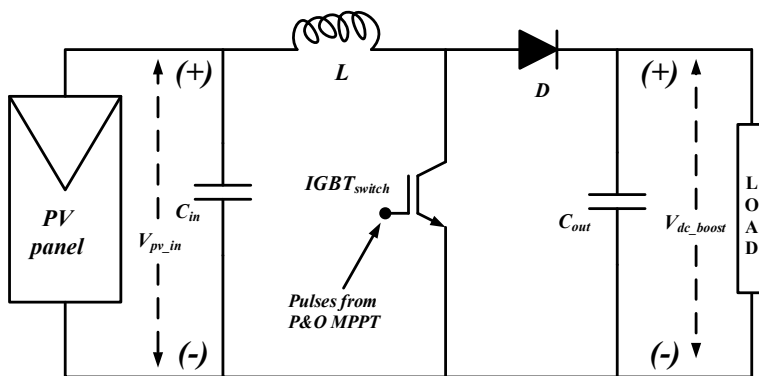


Fig. 4 Circuit diagram of Boost converter

### 2.3 DC/DC Boost Converter

A boost converter constitutes the first stage of any photovoltaic power plant in conjunction with MPPT system. The power output of photovoltaic module is at low voltage level. For load operation the dc link voltage should be at least 1.5 times the PV voltage as a result boost converter comes into existence. The circuit diagram and basic principle of operation of a simple boost converter has been depicted in Fig. 4.

The output voltage from the boost converter is made higher than input voltage by controlling the on time of switch ( $T_{on}$ ). The relation between input and output voltage has been evaluated below. In the following derivation the ripple in the source current has been neglected.

During the time when switch is on the whole source current flows through the inductor ( $L$ ) hence energy is store in the inductor. The store energy in the inductor is expressed as:

$$E_{on} = V_{dc} \times T_{on} \times I \tag{4}$$

where “ $V_{dc}$ ” is the input supply voltage, “ $I$ ” is the current drawn from the source.

When switch is turned off the source current flows through the load. During this period stored energy from the inductor get release and delivers to the load. This released energy can be expressed as:

$$E_{off} = (V_o - V_{dc}) \times T_{off} \times I \tag{5}$$

According to energy conservation theorem the energy stored during  $T_{on}$  in inductor is equal to the amount of energy delivered to the load by the inductor during  $T_{off}$ , where  $T_{off}$  is the turn off period of the switch.

Equating the equations, we have

$$\begin{aligned}
V_{dc} \times T_{on} \times I &= (V_o - V_{dc}) \times T_{off} \times I \\
\Rightarrow \frac{[V_{dc} \times (T_{on} + T_{off})]}{T_{off}} &= V_o \\
\Rightarrow \frac{[V_{dc} \times (T)]}{1 - T_{on}} &= V_o \\
\Rightarrow \frac{V_{dc}}{1 - \delta} &= V_o
\end{aligned} \tag{6}$$

where  $I$  the source is current,  $\delta = \frac{T_{on}}{T}$  is the duty cycle. The output voltage can be varied by varying duty cycle.

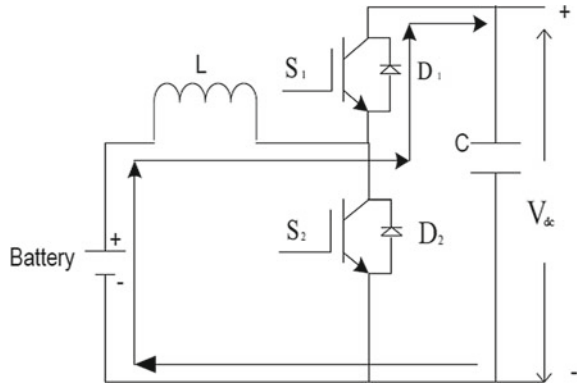
### 3 Bi-Directional Converters

The bi-directional DC to DC converter is basically the extension to conventional DC-DC converter, it is an optimal technique which offers solution to the problem of bi-directional flow of energy [11–13]. With the advancements in technology, few domains will comprise of bi-directional flow and management of energy. Generally, DC-DC converters are unidirectional, the flow of energy is always from source to load. In order to implement the bi-directional flow of energy, typically two conventional unidirectional DC-DC converters will be connected in parallel, one of the DC-DC converters control the energy flow from input to output, while the other converter control the energy flow from output to input. But using two unidirectional DC-DC converters in parallel will cause increased volume and overall cost of the converter [14]. A bi-directional DC-DC converter can also control the magnitude of energy output along with the direction of energy flow. The dc-dc bi-directional converter are fundamentally used to attain the necessary voltage and current magnitudes of the considered electrical machines and battery for both modes of operation. The direction of current in electrical machine will be reversed based on the mode of operation. The acceleration process of electrical machines depends on current flow from the batteries in motoring operation. And alternatively, in generating mode the converter need to flow the energy generated by the machines for charging the batteries [15].

A dc-dc bi-directional converter may interconnect a variety of available sources, allowing the system to be more flexible. The switches in dc-dc bi-directional devices allow electrical energy to flow between two sources that are used to move the force from one direction or the other. These power devices are often used to charge and discharge battery as a system not only to control the charging and discharging current, but also to direct the discharger's yield voltage level to an ideal worthwhile level so that any extra energy is transferred to the battery. It has two modes of operation, buck mode and boost mode. They are widely employed in application's demanding the flow of energy in both directions. These converters are intended for charging a low voltage battery (12 V) during step down mode and for charging or boosting the



**Fig. 5 A**  
DC/DC—bi-directional  
converter



high-voltage battery (400 V/600 V) during emergencies in Hybrid Electric Vehicles (HEV) and Electric Vehicles (EV) [16]. This converter system is comprised of one full-bridge power stage on the HV (high voltage) side that’s also isolated from a full-bridge or push–pull stage on the LV side.

The operating Modes of a battery using BDC:

- Buck Mode: When  $V_{dc} > V_{dc\_ref}$ , battery works in charging mode.
- Boost Mode: When  $V_{dc} < V_{dc\_ref}$ , battery works in discharging mode.

For safe operating modes of a battery, it’s State of Charge (SOC) is kept in between 20 to 80% (Fig. 5).

Generally, pulse-width modulation (PWM) technique is used for controlling the functionality of converter on either side of the converter. In this scenario, if the converter’s input is adjusted or has minor distortions, it has a substantial impact on the output. When the flow energy is from source to load, the converter operates as a step-up (boost) converter and while the flow of energy is from load to source, the converter operates as a step-down (buck) converter.

## 4 Energy Storage Systems

As illustrated in Fig. 1, the secondary source is the battery, which is coupled to the PV through its own converter (i.e. BDC). The battery’s energy is extracted in such a manner that it only drains whenever the PV system is unable to assist the energy requirement. If its state of charge drops below the determined value during other periods, it obtains charges from the grid. The reference active current ( $I^*$ ) is obtained by adding the components viz., reference charging current ( $I^*_{q\_charge}$ ) and reference discharging current ( $I^*_{q\_discharge}$ ) as mentioned in Eq. 9.

$$I^*_{q\_discharge} = [P_{ref} - (P_{pv} + P_{battery})] * \left( K_p + \frac{1}{K_i} \right) \tag{7}$$

when  $V_b > V_{b, \text{minimum}}$ ;

$$I_{q\_charge}^* = [P_{ref} - (V_{b,max} + V_{battery})] * \left( K_p + \frac{1}{K_i} \right) \tag{8}$$

when  $V_b < V_{b, \text{minimum}}$ ;

$$I^* = I_{q\_discharge}^* + I_{q\_charge}^* \tag{9}$$

From Fig. 6, it can be observed that the PI controller for charging the battery is basically used to check whether the aggregate power produced by the battery storage and PV system combinedly, coincide with the reference value, so that the battery is primarily used to supply the total real power which is lost due to shading effect in PV panels. The battery pack power as long as the battery’s terminal voltage is greater than the voltage of the battery that addresses the base minimum state-of-charge. The charging current gets regulated to deliver a current component that indicates the magnitude of the force to raise the battery terminal voltage to a level that correlates to the most extreme state-of-charge. The quantity of discharging and charging flows is being used to compute the overall battery current. This plan’s active part of  $I^*$  was set to zero so that the battery would still not provide responsive force; this part can also fluctuate.

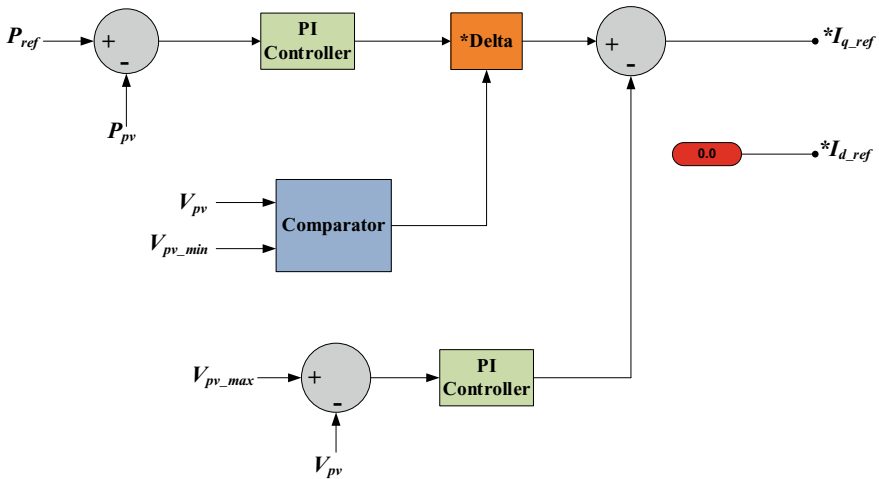


Fig. 6 Control Strategy for the battery charging and discharging

### 5 Simulation Result and Analysis

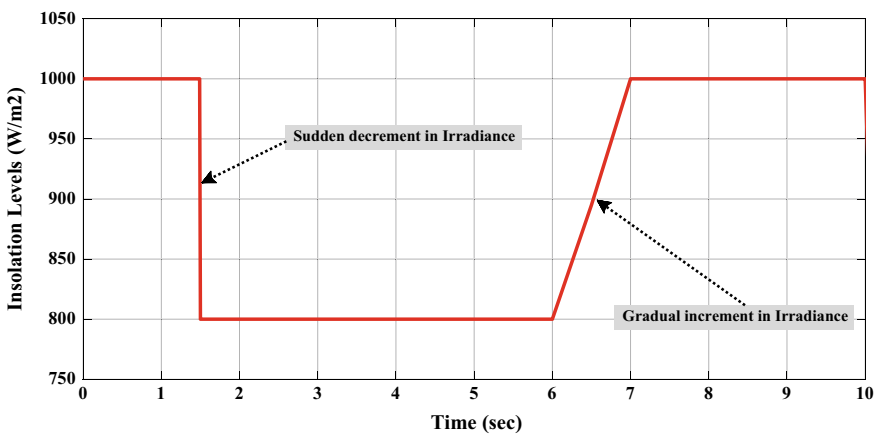
Thorough simulation studies were accomplished on MATLAB/Simulink climate, and the outcomes got for different working conditions are introduced in this segment. The upsides of boundaries utilized in the model for reenactment are recorded in Table 1. The powerful contribution to the proposed framework during the activity is displayed in underneath figures. In Fig. 7 shows unexpected decrement and slow expansion in sun powered radiations levels as a contribution to framework.

For assessing the exhibition of the proposed framework created, the framework is reenacted to testing situations. Figure 8 shows outrageous changes happening while at the same time changing over the sun based light levels to accept uniform insolation at the primary stretch when  $t < 1.5$  s and  $t > 6$  s by maintaining it around 36 V. The temperature is taken as constant value of 25 °C (Fig. 8).

Figure 9 shows the recreation aftereffects of the PV exhibit power utilizing the P&O MPPT. It tends to be seen that the comparing yield force of the PV exhibit moreover gives a clearer understanding during the finding system, a different figure

**Table 1** Parameters of a single PV model

Parameters	Value
Solar Irradiance; $G_{ref}$	1000 W/m <sup>2</sup>
Cell temperature; $T_{ref}$	25 °C
Open circuit voltage; $V_{oc}$	38.8 V
Short circuit current; $I_{sc}$	5.96 Amps
Maximum voltage; $V_m$	36.7 V
Maximum current; $I_m$	5.58 Amps
Temperature co-efficient	6.175%



**Fig. 7** Varying solar radiations input to the PV system

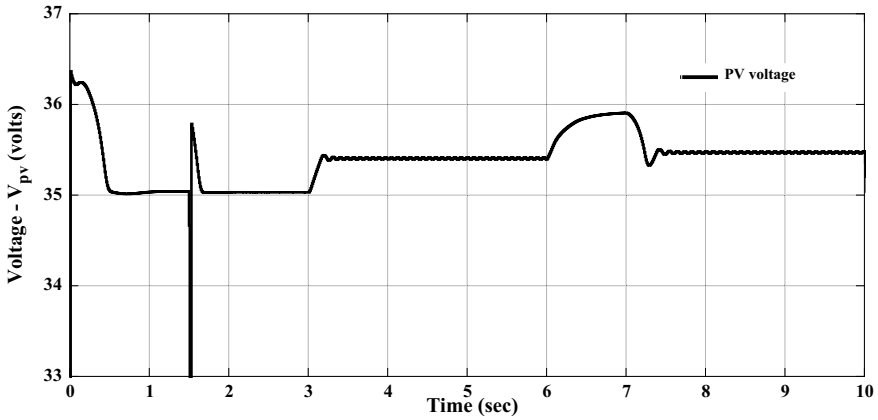


Fig. 8 PV panel output voltage ( $V_m$ )

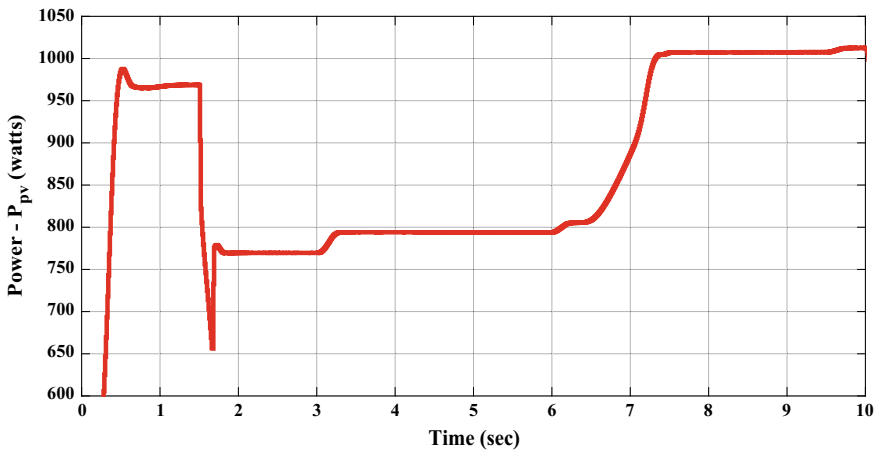


Fig. 9 Power generated from PV system with efficient MPP tracking

of the PV voltage portrayed in Fig. 8. The immediate force extricated from the PV exhibit utilizing the P&O calculation shows the proficiency of following force by the MPPT calculation compasses to its greatest cutoff for time stretch 1.5 to 6 s during its activity.

Figure 10 shows the yield force of the DC/DC help converter related to multi-PV cluster and BDC with battery. The force from the battery and the heap concerning load variety and the natural conditions can keep up with steady voltage as for changes happening it close by. The state-of-charge and the nominal voltage of the battery are shown respectively in Figs. 10 and 11.

The critical boundary of Battery energy stockpiling framework is state-of-charge (SOC). Its function (SOC) demonstrates the leftover power of a battery cell and is

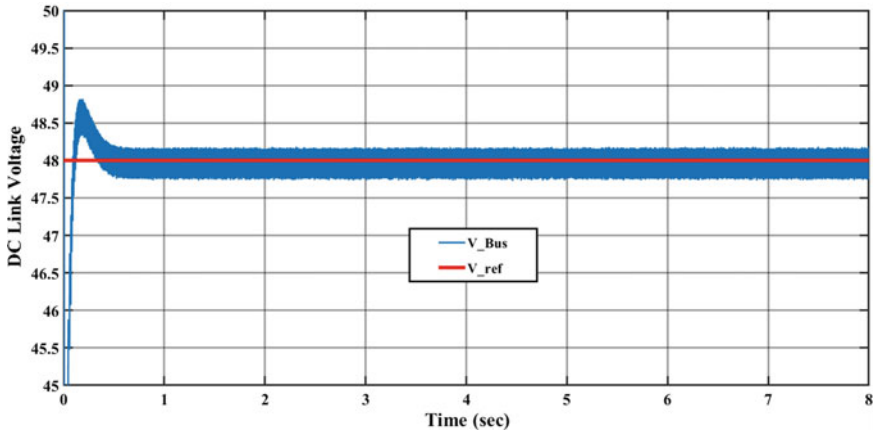


Fig. 10 DC link voltage by boost and battery during their operation

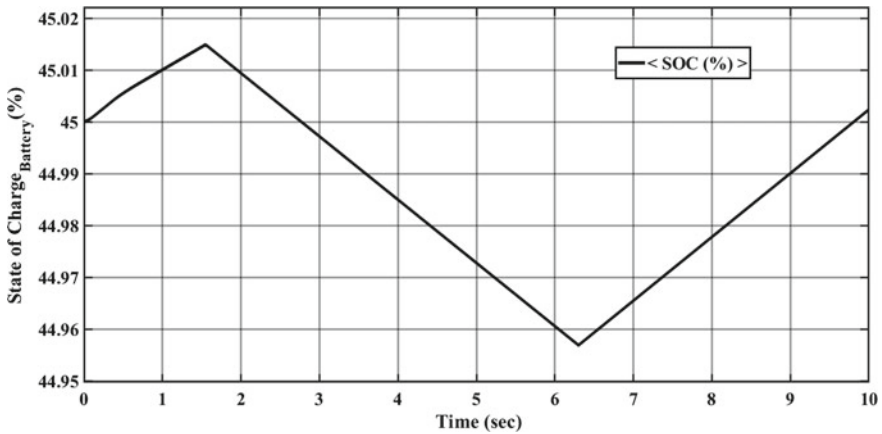


Fig. 11 The state of charge (SOC) of BESS

perhaps the main boundaries to pass judgment on battery states. An exact assessment of SOC can be utilized to deflect over-charging or over-releasing battery. Figure 12 shows the reproduction consequences of the SOC assessment by proposed framework at SOC half essentially cover the majority of the unique force from the PV frameworks.

In Fig. 13, it can be noticed that current drawn from and rush in the battery is able to track the reference battery current for an efficient flow of power from source and to the load while maintaining the constant DC link voltage during its operation for charging and discharging the battery.

When the system is able to maintain steady voltage for the BLDC motor for driving its load, solar irradiance is dynamic. Figures 13 and 14 shows the experimental

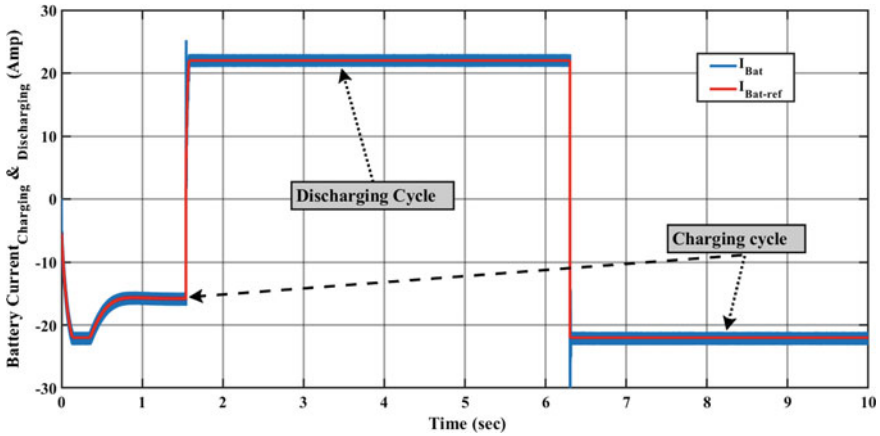


Fig. 12 Waveform showing charging and discharging cycles

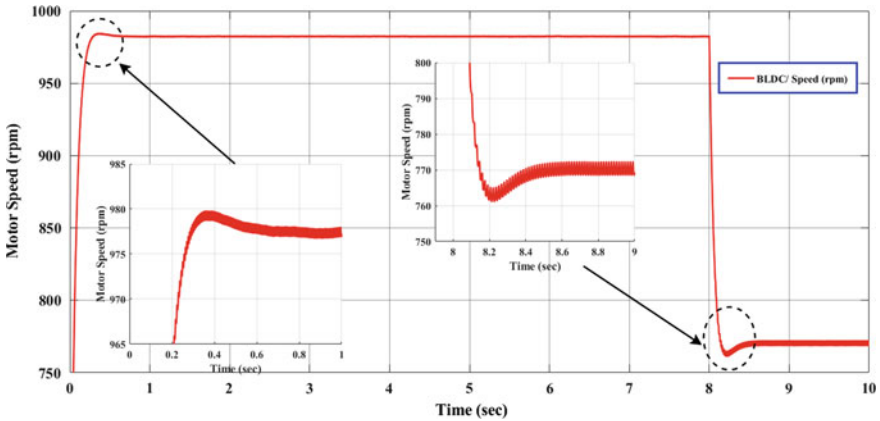


Fig. 13 Speed response of BLDC drive for dynamic load at 8 s

results of speed and torque response of the motor drive during the dynamic changes occurring in the system. It has been observed that the estimated and actual speeds differ of 40 rpm under free running condition of constant load of 2 Nm.

The above waveforms shown in Figs. 13 and 14 depicts the steady operation of BLDC drive for agricultural implements satisfies under various varying insolation conditions for proposed topology criteria.

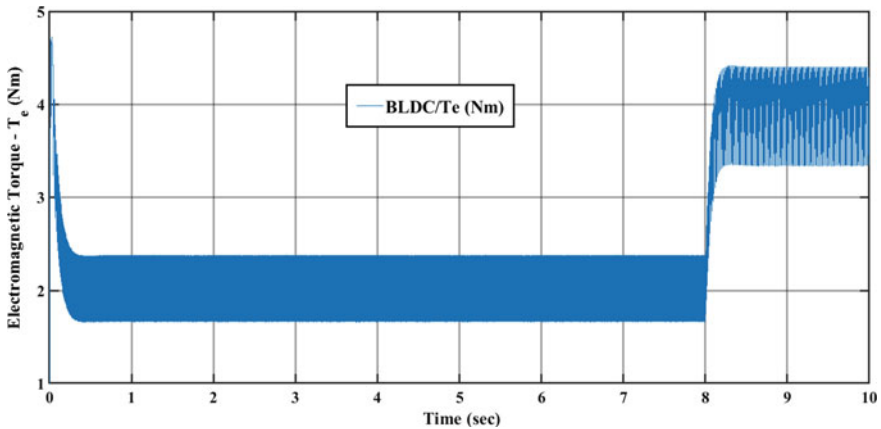


Fig. 14 Electro-magnetic torque for sudden change in load

## 6 Conclusion

The projected proposal in this paper indicates the integration of battery storage into PV systems which uses a dc-dc converters (i.e., boost and BDC), which are seen as of operating as charge controllers and MPPT features on a real-time basis. The key objective of this type of configuration is to enhance PV system production as total energy, improve PV system control to maintain dc-link voltage during power curtailment, and achieve high efficiency when compared to other conventional approaches that do use multiple power converters. Moreover, the suggested scheme provides a low cost for battery-integrated PV systems with motor drive to power agricultural loads. Through MATLAB/Simulink simulations of case studies with irradiance variation for a dynamic input, the detailed procedural aids of the stated configuration with respect to output produced by PV for smoothing the power and adapting the power generated while in operation were established. Finally, the steady-state operation as employed for Dc-Dc converter's and its performance, as well as the variation of various operation modes, were verified and validated, with the results provided. The framework is illustrated through stimulation in the MATLAB. The proposed configuration is capable of supplying uninterruptible power to dc motor drive loads, and ensures removal of surplus photovoltaic power into the BESS.

## References

1. Chander AH, Sahu LK, Ghosh S (2020) Stand-alone multiple input photovoltaic inverter for maximum power extraction and voltage regulation under mismatched atmospheric conditions. *IET Renew Power Gener* 14(9):1584–1595
2. Chander AH, Kumar L (2019) MIC for reliable and efficient harvesting of solar energy. *IET Power Electron* 12(2):267–275

3. Rallabandi V, Akeyo OM, Ionel DM (2016) Modeling of a multi-megawatt grid connected PV system with integrated batteries. In: 2016 IEEE International conference on renewable energy research and applications (ICRERA), Apr 2016, pp. 1146–1151
4. Villalva MG, Gazoli JR, Filho ER (2009) Comprehensive approach to modelling and simulation of photovoltaic arrays. *IEEE Trans Power Electron* 24(5):1198–1208
5. Nguyen XH, Phuong NM (2015) Mathematical modeling of photovoltaic cell / module/arrays with tags in MATLAB/Simulink. *Environmental Syst Res* 4:24
6. Kumar L, Jain S (2012) A novel dual input DC/DC converter topology. In: 2012 IEEE international conference on power electronics, drives and energy systems (PEDES), Dec 2012, pp 1–6
7. Luis MC, Rodríguez JR, Cecilia MC (2020) Modelling of PV systems as distributed energy resources for steady-state power flow studies. *Int. J. Electr. Power Energy Syst* 115. ISSN 0142-0615
8. Kumar J, Agarwal A, Singh N (2019) PV fed hybrid energy storage system supported DC microgrid. In: 2019 IEEE 1st International conference on energy, systems and information processing (ICESIP), pp 1–6
9. Hemant Kumar V, Pavan Kumar P, Patel RN (2019) Comparison of dynamic performance of solar PV fed BLDC motor drive with P&O and I.C. MPPT Algorithm. *Helix* 9(6):5889–5894
10. Koutroulis E, Blaabjerg F (2015) Overview of maximum power point tracking techniques for photovoltaic energy production systems. *Electr Power Components Syst* 43(12):1329–1351
11. Swarnkar R, Verma HK, Patel RN.: Comparative analysis of isolated and non-isolated bi-directional DC-DC converters for DC microgrid. In: 2019 3rd International conference on recent developments in control, automation and power engineering. RDCAPE 2019, pp 557–562
12. Singh SK, Kumar A, Sadhu PK (2019) A novel instantaneous mode switching Bi-directional DC-DC converter for dc grid voltage control. In: 2018 International conference on power energy, environment and intelligent control, PEEIC, pp 187–190. <https://doi.org/10.1109/PEEIC.2018.8665422>
13. Metin NA, Boyar A, Kabalci E (2019) Design and analysis of bi-directional DC-DC driver for electric vehicles. In: Proceedings of 2019 1st Global Power, Energy and Communication Conference (GPECOM), pp 227–232
14. Das R, Uddinchowdhury MA (2017) PI controlled Bi-directional DC-DC converter (BDDDC) and highly efficient boost converter for electric vehicles. In: 2016 3rd International Conference on Electrical Engineering and Information Communication Technology, iCEEiCT 2016
15. Sabatta D, Meyer J (2018) Super capacitor management using a Split-Pi symmetrical bi-directional DC-DC power converter with feed-forward gain control. In: 2018 International conference on domestic use of energy, DUE 2018, pp 1–5
16. Tulasi NP, Aithepalli LD (2019) Droop control of bi-directional DC-DC converter for improved voltage regulation and load sharing in DC microgrid. *Int J Intell Eng Syst* 12(3):228–243
17. Li X, Wang S (2019) A review on energy management, operation control and application methods for grid battery energy storage systems. *CSEE J Power Energy Syst* 1–15
18. Barchi G, Lollini R, Moser D (2016) Photovoltaic and battery energy storage systems in shopping malls: energy and cost analysis of an Italian case study. In: 2016 European PV solar energy conference and exhibition (EUPVSEC), June 2016
19. Vieira FM, Moura PS, de Almeida AT (2017) Energy storage system for self-consumption of photovoltaic energy in residential zero energy buildings. *Renew Energy* 103:308–320
20. Di Piazza MC, Luna M, Tona GL, Di Piazza A (2017) Energy management of hybrid PV-storage systems for improved grid integration. In: 2017 IEEE International conference on environment and electrical engineering and 2017 IEEE industrial and commercial power systems Europe (EEEIC/I&CPS Europe), 2017, pp 1–5
21. Ru Y, Kleissl J, Martinez S (2013) Storage size determination for grid connected photovoltaic systems. *IEEE Trans Sustain Energy* 4(1), 68–81



# Residential Feeder Energy Audit Analysis and Recommendation with Aid of Software



K. Anitha, Shailesh, L. Ramesh, and Murugananth Gopal Raj

## 1 Introduction

Energy demand is an increasing pressure for any government. For the developing country like India, the energy generated decides the economic growth of the country. It is found that India has increased its installed capacity from 271.722 GW in 2015 to 383.37 GW (31 May 2021). An energy audit can be defined as a process of inspecting a building to identify the areas of energy loss at different level in order to reduce the energy consumption and improve the overall efficiency.

The imbalance can be reduced by using the energy efficiently and by conserving the energy. During study, we found that in India nearly 24% of total generated electricity is consumed by domestic users which would increase to eight times higher in 2050. As a consequence, Bureau of Energy Efficiency (BEE) has implemented norms for appliance (star rating systems) for use of energy efficiently and conservation of energy. By which, we can save nearly 20–30% of total energy generated by energy conservation in residential houses by efficient way of usage lighting systems, air conditioners which use almost 15–25% of electricity used, washing machine which accounts about 25% of electricity bill and geyser/water heaters.

In order to reduce energy wastage with household users, several energy audits were undertaken at various home sites, and the results are provided in the next chapter.

---

K. Anitha (✉) · Shailesh

MGR IGEN ENSAV Club, Dr. M.G.R. Educational & Research Institute, Chennai, India

e-mail: [anniekumar25@gmail.com](mailto:anniekumar25@gmail.com)

L. Ramesh

Dr. M.G.R. Educational and Research Institute, Chennai, India

M. G. Raj

Ahalia School of Engineering and Technology, Palakkad, India

## 2 Literature Study on Residential Audit

A residential energy audit is one of the energy conservation methods in which the energy flow consumption of a house is analysed in order to provide the best strategies to increase home heating and air-conditioning efficiency. We would like to showcase a few residential audits in this chapter, which were conducted in a multitude of localities.

The audit has been conducted by Awanish and his team, at a residence in Maduravoyal, Chennai-600095, in 2015 [1]. And after all the data acquisition analysis, it is simulated in ETAP. There were two types of recommendations suggested : (1) with investment and (2) without investment, on implementation of the suitable SLD via ETAP. Through daily load analysis, a 50% reduction in load current was discovered. Priyanka and her team conducted the audit of a distribution feeder in Chennai in 2016 [2]. This audit proposed two sorts of recommendations for all 132 residences. The suggestions are tested with ETAP programme. Following their recommendations, they could see a 35% reduction in power losses.

The survey was conducted by R. Pieterse and his team in South Africa in 2004 [3] at smart buildings, where huge loads are consumed. This research proposes a wireless sensor network-based measuring system to monitor real-time energy usage of household appliances and show the results on a computer's graphical user interface. ACORD is an appliance coordination mechanism introduced by Kantarci et al. A minimum of 16.6% cost savings is recommended. An inspection was conducted by Luis pirez-lombard et al. on the buildings energy consumption information in the UK in 2007 [4]. The study or audit's final conclusion is that private initiative must be combined with government involvement to promote energy efficiency, new energy production technologies, reducing energy consumption and improving public knowledge of the rational use of energy. In 2013, Jukka Heinonen et al. conducted an energy audit of urban and rural households in Finland [5]. With the help of this audit, they proposed changing the way people consume power as well as implementing renewable energy sources. In order to lessen the environmental burdens associated with energy production, attention should be made to the fuels used.

Mathur et al. in 2018 [6] audited the U.A.E institutional building. (1) Schedule and adjust AHU connected to BMS system to reduce bulk energy consumption. (2) Manage the AHU's frequency to control the motor/fan speed without sacrificing the occupant's comfort. (3) Determine the temperature settings of various building zones to maximise energy efficiency. (4) Installation of tint windows and automatic revolving doors are recommended. Building audits can enhance energy efficiency and minimise energy consumption by improving HVAC and lighting systems. Pao and Chen [7] surveyed the relationship of developed countries between energy consumption, economic growth and carbon emission. The panel co-integration found that a long-term, balanced relationship exists between carbon emissions and economic growth, and these two variables are interdependent. During a time of economic growth and clean energy usage, the study hypothesises that feedback will occur.

To support energy research initiative towards United Nation Sustainable Development Goal 07, Dr. M.G.R. Educational and Research Institute has taken initiative in the year 2017 and inaugurated The Institution of Green Engineers-Energy Saving and Awareness club named “ IGEN GREEN9 MGR ENSAV CLUB”. The contributed research works under the pilot projects [8, 9] were published in Scopus publications, and the reports are published by the Research Forum IGEN-GREEN 9 (Energy Efficiency Research Group). This works in line with the energy conservation research on residential places. In the following chapter, we thoroughly examined the residential feeder data, to which 65 houses were connected. In order to reduce the tariff, three types of recommendations were proposed using ETAP simulation based on the category of houses.

### 3 Data Observation and Simulation

As a first step, we audited a home and provided the results in a single-line diagram made by ETAP simulation software, along with recommendations to minimise electricity usage and cut electricity bills.

### 4 Data Observation

Residential power utilisation surveys have been completed for 65 homes. The preliminary audit was carried out in accordance with procedure. The complete survey results are shown in a pie chart. Through the survey, we found that most of the electricity is wasted in the home, and that raising awareness about electricity is essential.

Figure 1 demonstrates that 1 ph current is used by 10.8% of persons in their homes, whereas 3 ph current is used by 89.2%. 47.7% of people are unsatisfied with the electricity bill they get, while 52.3% are happy with it.

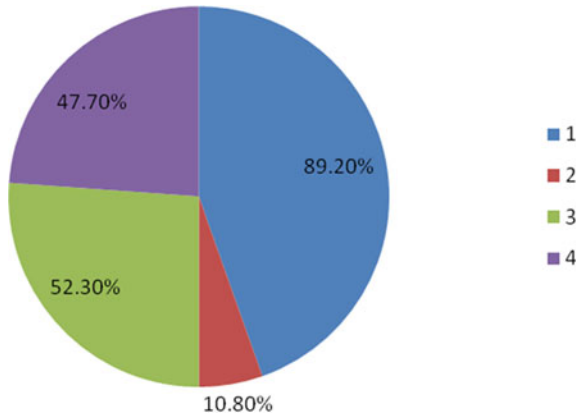
The graph in Fig. 2 reveals that just 10.9% of people have put their micro-renewable power generation to use, whereas 88.1% have. It also shows that 42.9% of people know about equipment with star ratings while 57.1% do not.

Figure 3 shows that about 13.8% people uses florescent bulb, whereas 86.2% uses CFL and 40% uses LED bulb.

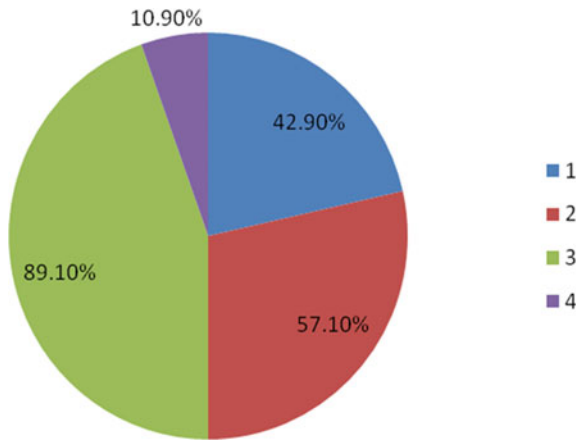
Figure 4 emphasises percentage of consumer who have awareness about the star-rated equipment.

#### Detailed Audit

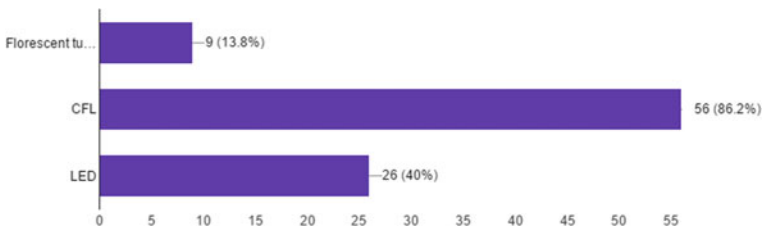
The second phase of work began with a feeder inspection and a preliminary audit of the house connected to that feeder. Underground power has been provided to the mini pillar box from a 500 KVA distribution transformer, and from there, three-phase and one-phase power have been provided to the consumers for two streets. One standby transformer has been installed at the intersections of both streets for backup purposes.



**Fig. 1** Ratios of different types of supply found in the home and the proportion of customers that are happy with EB

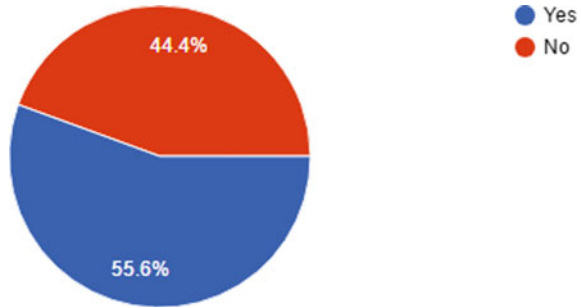


**Fig. 2** % of installed renewable power generation and % of star-rated machinery and apparatus



**Fig. 3** % of different lighting used by consumer

**Fig. 4** % consumers conscious of the star-rated equipment



A single-line diagram (SLM) of the feeder has been drawn as well as the necessary voltage and current details of R-Y-B-N phase (phase-to-phase and phase-to-neutral) of all the pillar boxes have noted (Fig. 5).

Finally, SLD was created with the support of electrical transient analysis programme (ETAP). ETAP simulation provided the various load flow analysis curves as well as the accuracy of power utilisation by the consumer in that feeder. Depending on the various outcomes based on the survey results, the house has been classified into three types based on the tariff and power consumption.

- (i) Tariffs starting at Rs. 4000 and going up from there.
- (ii) The price range is 2000–4000.
- (iii) Tariffs of Rs. 2000 or less.

We considered one house sample from each category, where significant waste was discovered and conducted an energy audit (Tables 1 and 2).

## 5 Recommendation

The residential feeder audit in Chennai, Kambhar Nagar, included feeder inspection as well as a preliminary audit of 65 houses connection to the inspected feeder. Two streets are served by a three-phase and a single-phase supply provided by an underground 500 KVA distribution transformer. Based on their energy consumption tariff, house were categorized into three types. A complete detailed house audit is conducted for one house from each categories. All of the houses' SLDs have been drawn and simulated in ETAP feeder. Individual and overall voltage and current graphs have been plotted and compared with and without audit.

The first house category (1-BHK): There is very little equipment in that house, but the losses are significant, which we identified and implemented in ETAP. In Fig. 6, we can see the variation of current comparison with and without recommendation.

Houses in the second category (2-BHK): There was a limited amount of equipment in that house, but it was accompanied by losses and old equipment, which we discovered and implemented in ETAP. Figure 7 shows voltage comparison with and without recommendation can be seen.

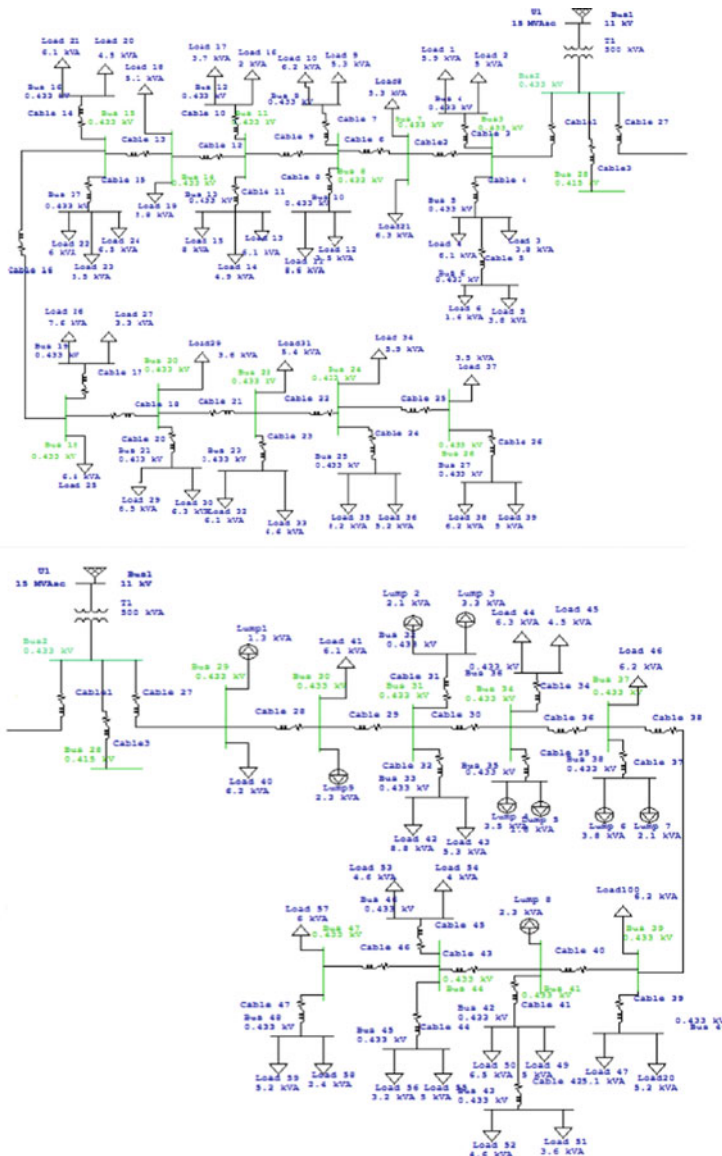


Fig. 5 Single-line diagram of the feeder

Houses in the second category (2-BHK): There was a limited amount of equipment in that house, but it was accompanied by losses and old equipment, which we discovered and implemented in ETAP. Figure 8 shows  $n$  voltage comparison with and without recommendation can be seen.

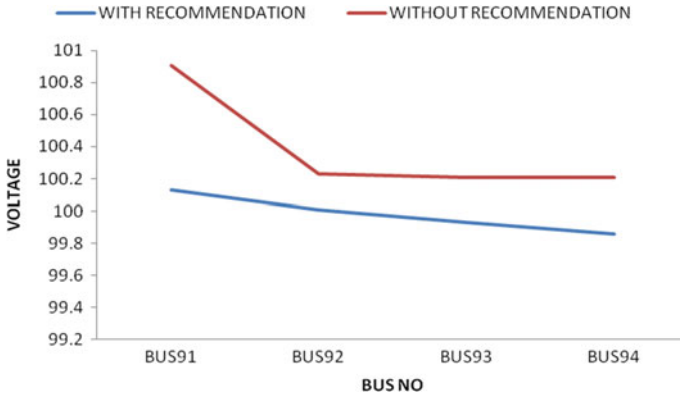


Fig. 6 Voltage in different bus compared with and without recommendation

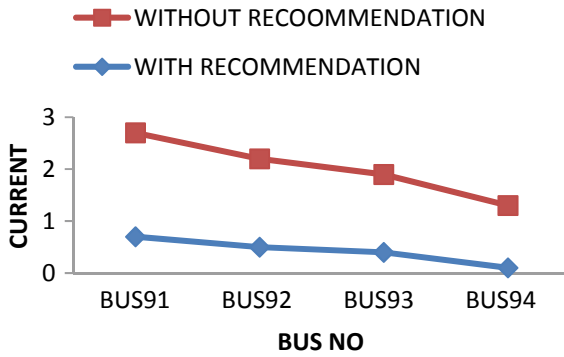


Fig. 7 Current in different bus compared with and without recommendation

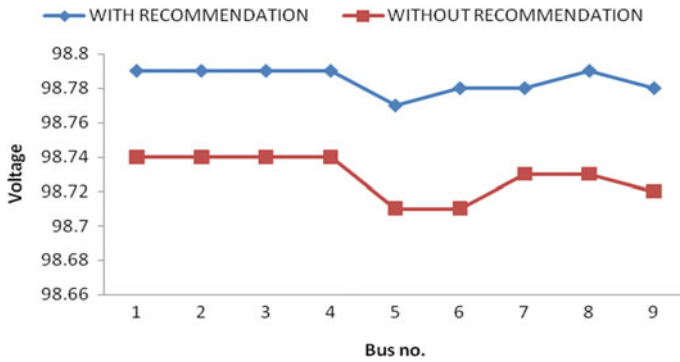


Fig. 8 Voltage in different bus compared with and without recommendation

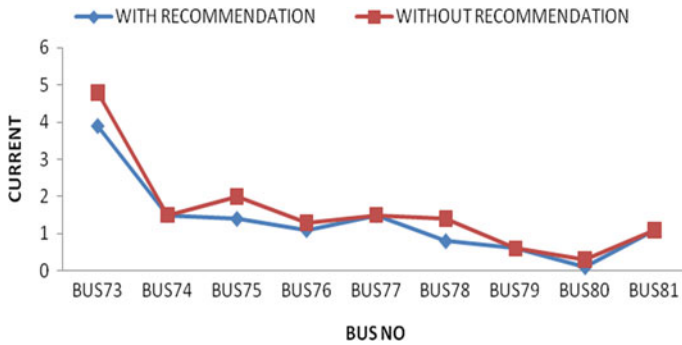


Fig. 9 Current in different bus compared with and without recommendation

Figure 9 compares the current of the feeder in different buses with and without recommendations. It has been determined that the fluctuation is in the feeder, which is caused by some consumers using the equipment incorrectly and equipment consuming more power than necessary. As a result, we proposed some general recommendations with and without investment, which were implemented in ETAP.

Figure 10 represents the voltage of the feeder in the various buses and compares it with and without recommendations. It has been determined that the fluctuation is in the feeder, which is caused by some consumers using the equipment incorrectly and equipment consuming more power than necessary. As a result, we proposed some general recommendations with and without investment, which were implemented in ETAP (Fig. 11).

When comparing before and after recommendations, the overall power savings are approximately 30% of the total power. The biggest savings are in the three-bedroom house, which had a lot of heavy and old equipment.

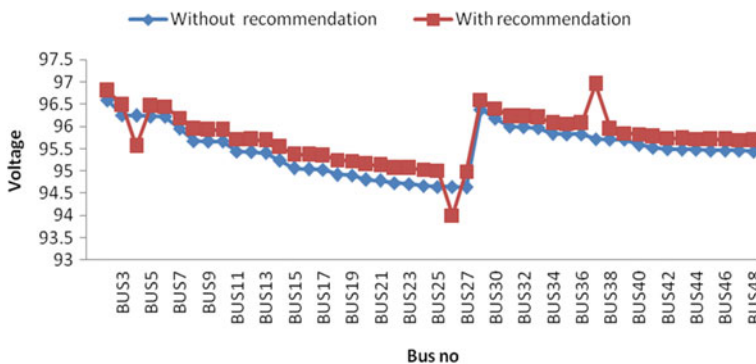


Fig. 10 Voltage in different bus compared with and without recommendation



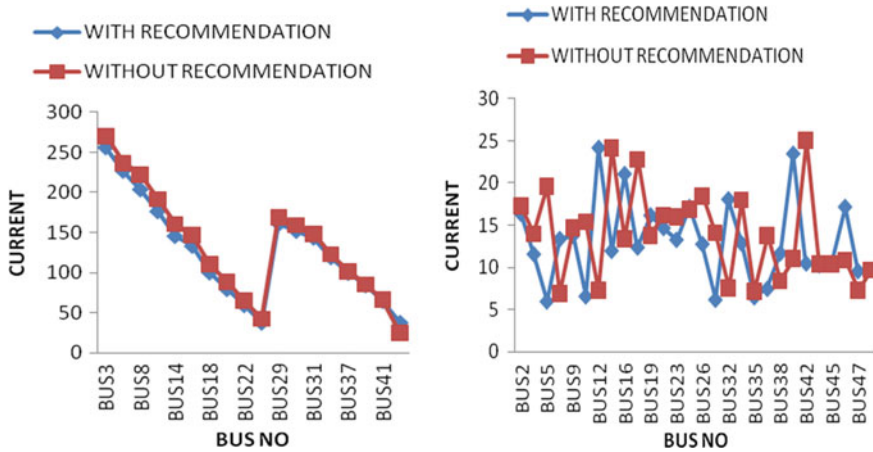


Fig. 11 Current conditions in various buses when compared with and without recommendations

## 6 Conclusion

Finally, we have come with different ideas to implement the renewable source of energy and also create awareness about the usages of equipment and saving the power. We also recommend aspects of using renewable resources, tree plantation around the building and make changes in the installation procedure for an effective, efficient, cleaner and greener environment. From the overall installed capacity, 24% power has been utilised by the residential, and the wastage of power is mostly done by the domestic users. After implementing the suggested recommendation, we observed that the total energy consumption was reduced upto 30%, which changes the environment more economically sustainable and freely available in the nature, that all wastage reduction and power management can be done through energy audit, which is very important to have better and balanced power with the sustainable range.

Table 1 Major equipment used in that 1-BHK house

Equipment	No. of. equipment	Watt	Running, time/h	Total units, used/month
Tube light	4	18	6	13
CFL	4	15	6	11.6
Fridge	1	200	3	144
Iron	1	1000	1	20
TV	1	80	8	19.6
Fan	2	70	4	17.36
LED	1	8	4	0.99
Fridge	1	200	24	148

**Table 2** Major equipment used in that 2-BHK house

Equipment	No. of equipment	Watt	Time	Total energy used
Tube light	8	18	20	89.28
Fridge	1	200	24	148
AC	1	1000	7	7
TV	1	80	5	19.6
Fan	4	70	20	173.6
CFL	4	17	4	8.432

## References

1. Kumar A, Ranjana S, Singh BK, Ramesh PKL (2015) Electrical energy audit in residential house. Elsevier.
2. Priyanka K, Sujana K, Pooja K, Neelakandan, Ramesh L (2016) Design of micro grid modeled distribution feeder through power audit recommendations. Published in National power system conference Dec 2016
3. Pieterse R, Kumar A (2004) Energy consumption audit system for smart building. South Africa anuj.kumar@ap.uc.za in 2004
4. Luis P, Ortiz J, Pout C (2007) A review on buildings energy consumption information. Publication Elsevier in Mar 2007
5. Heinonen J, Junnila S (2013) Residential energy consumption patterns and the overall housing energy requirements of urban and rural households in Finland. Publication Elsevier in Nov 2013
6. Mathur A, Muthukumaraswamy SA (2018) Investigation and analyses of energy-efficient technologies for HVAC and lighting systems via energy auditing processes. Published in 2018 International Conference on Smart Grid and Clean Technologies
7. Pao H-T, Chen C-C (2019) Decoupling strategies: CO<sub>2</sub> emissions, energy resources, and economic growth in the group of twenty. J Clean Prod
8. Janakiraman D, Hariharan J, Umamaheswaran N, Ramesh L, Velraj Kumar P (2021) My optimal residential house implementation through power audit. Published in 2021 2nd International conference for emerging technology (INCET)
9. Anitha K, Jyoti B, Ramesh L, Marsaline Beno M (2021) Comparative energy performance analysis at dyes and coating industry. Published in 2021 International conference on computing, communication and green engineering

# An Investigation into the Applications of Real-Time Simulator in Experimental Validation of PMSM-Based Electric Drive System



Anjaly Mohan, Meera Khalid, and A. C. Binojkumar

## 1 Introduction

Power electronics is one of the key technologies in any industry and is instrumental in the fast-growing renewable energy and electric vehicle automotive industries [1]. Performance, cost and reliability must be considerably improved to fully harness its potential. Simulation of testing systems by real-time digital simulation is not a very new idea. But in power electronics and drive industry, it is rare because of the difficulty in getting the small-time step required for simulation, since power electronic systems are switched dynamic systems requiring fast switching [1]. In academic research area, the experimental validation of the concepts is a challenging task due to the difficulty in setting up the hardware and the time and cost involved. The simulation software such as MATLAB and PSIM are used in research but they are not real-time simulators. The Hardware-In-Loop simulator developed by Typhoon is a promising solution to this problem. It is a real-time simulator in which the simulation time exactly matches with the real world-clock time. This enables the users to study and analyze the systems in real-time.

---

A. Mohan (✉) · M. Khalid · A. C. Binojkumar  
Electrical Engineering Department, Rajiv Gandhi Institute of Technology (Govt. Engineering College), Kottayam, India  
e-mail: [anjalymohan90@gmail.com](mailto:anjalymohan90@gmail.com)

APJ Abdul Kalam Technological University, Thiruvananthapuram, Kerala, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2022  
T. Sengodan et al. (eds.), *Advances in Electrical and Computer Technologies*,  
Lecture Notes in Electrical Engineering 881,  
[https://doi.org/10.1007/978-981-19-1111-8\\_84](https://doi.org/10.1007/978-981-19-1111-8_84)

1091

Many research works are carried out in various fields with the aid of real-time simulation tools and Typhoon HIL is a good platform with rich FPGA resources [2]. Various industrial applications of real-time simulation in the fields of avionics, motor drives, power systems and robotics are presented in [3]. Application of real-time simulator in modeling and testing of electric vehicles is detailed in [4–7]. PMSM drives are implemented and tested in HIL mode to validate respective findings in Field Oriented Controlled (FOC) drives, sensor less control, etc. in [8–11]. A wind turbine emulator based on squirrel cage induction motor is implemented and tested using HIL 402 in [12]. The problems like modeling accuracy, bandwidth and system stability associated with real-time simulation platforms are discussed in [13].

Hardly, there is any literature detailing the procedure to be followed while using such real-time platforms. Also, the HIL mode of operation of the real-time simulators is explained in almost all the past works whereas, the Rapid Control Prototyping (RCP) and Rapid Batch Simulation (RBS) modes are little explored. This paper aims at real-time analysis of a PMSM drive with space vector PWM control, by utilizing the RBS and RCP modes of operation. First part of this paper describes implementation of a PMSM drive system in RBS mode with a virtual motor. The next part validates experimentally the simulation results obtained in RBS mode using the Rapid Control Prototyping feature of the simulator.

The organization of the paper is as follows: In Sect. 2, a brief description about the real-time simulation and the three modes of operation is given. Section 3 details the features of Typhoon HIL. Section 4 deals with the implementation of the drive system in this platform and Sect. 5 deals with the simulation and experimental results. The concluding remarks are given in Sect. 6.

## 2 Real-Time Simulation

Real-time simulation refers to executing the mathematical model of a physical system in real-time, where the execution time is same as the wall-clock time. Configuring models to execute in real-time has many advantages and applications. The following are the different ways in which real-time simulation platform can be utilized in industrial, academic as well as research applications.

1. Rapid Batch Simulation.
2. Rapid Control Prototyping.
3. Hardware-in-Loop testing (HIL)

In massive batch run test, simulation is accelerated by rapid batch simulation method. In this technique, both the controller and the plant are modeled inside the simulator.

In rapid control prototyping scheme, a real-time simulator is used to quickly implement a controller which controls a real plant. This enables fast experimental testing in a plant with minimal implementation time. The controller is inside the simulator in RCP, which controls the actual plant.

The Hardware-in-Loop testing mode is just reverse of RCP mode. An actual controller is tested on a virtual real-time plant which is modeled inside the simulator. The controller is outside the simulator and the plant is inside the simulator. This is especially useful when the actual plant is large and complex like a ship or an aircraft and testing on the actual plant requires much time and cost.

### **3 Features of the Typhoon HIL Real-Time Simulator**

The real-time simulator HIL 402 used in this work has 20 ns PWM resolution which makes it highly suitable for wide band gap devices like SiC. It comprises mainly of two parts; one is a software part and the other one is a hardware part. The software part can either be run and viewed as an independent system or connected to the hardware part, to connect with the external plant.

#### ***3.1 Software Section***

Schematic editor in software section is used to draw the schematic diagram and the SCADA platform, which is a graphical environment helps to view the simulation results, using the built-in oscilloscope. Schematic Editor allows to create high-fidelity models of the power stage for real-time simulations. It has a library, from where the necessary components for creating the simulation model are selected. The waveform for analysis is viewed in the SCADA platform by connecting digital or analog probes.

#### ***3.2 The Hardware Module***

Hardware module is used to interface the simulator with a controller, as in the case of HIL mode of operation or to a real plant, as in the case of Rapid Control Prototyping. In this, 16 analog Input/Output pins (I/O) and 32 digital I/O pins are available for interfacing. I/O pins are configured using the SCADA panel.

### **4 Implementation of Drive System**

In this article implementation of a 1.07 kW PMSM drive in RBS and RCP modes of the HIL 402 real-time simulator is done and its performance analysis is carried out.

After compiling the model of the drive system created in the schematic editor, the control is passed to SCADA where necessary waveforms are observed. The

analog/digital input/output (I/O) pins of the Typhoon HIL simulator is configured appropriately, to connect to the external hardware.

In RBS operation while switching from schematic editor to SCADA, “Load the model to virtual device” mode is selected. In RCP operation, since hardware unit is involved, the control is automatically passed to SCADA.

#### 4.1 Rapid Batch Simulation

In rapid batch simulation, both the controller and the plant are modeled inside the simulator. The motor can either be modeled from the basic equations or the available library model can be used. In this example, model of the PMSM is developed using  $d$ - $q$  rotor reference frame theory and the equations used are as follows [14].

$$v_q = R_s i_q + \omega_r (L_d i_d + \lambda_f) + p(L_q i_q) \quad (1)$$

$$v_d = R_s i_d - \omega_r (L_q i_q) + p(L_d i_d + \lambda_f) \quad (2)$$

The developed electromagnetic torque in the motor is given by,

$$T_e = \frac{3}{2} \frac{P}{2} (\lambda_d i_q - \lambda_q i_d) \quad (3)$$

where  $v_d$  and  $v_q$ ,  $i_d$  and  $i_q$ ,  $L_d$  and  $L_q$ ,  $\lambda_d$  and  $\lambda_q$  are the  $d$ -axis and  $q$ -axis stator voltages, currents, inductances and flux, respectively.  $R_s$  is the stator resistance and  $\lambda_f$  is the rotor flux. The mechanical torque equation is,

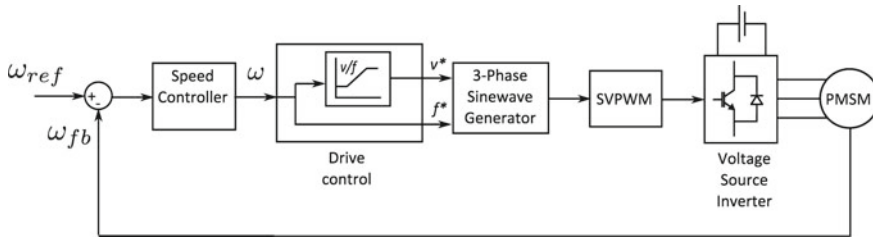
$$T_e = T_l + B\omega_m + J\{d\omega_m/dt\} \quad (4)$$

Solving for the rotor mechanical speed from Eq. (4).

$$\omega_m = \int \{(T_e - T_l - B\omega_m)/J\} dt \quad (5)$$

$$\omega_r = \omega_m \{P/2\} \quad (6)$$

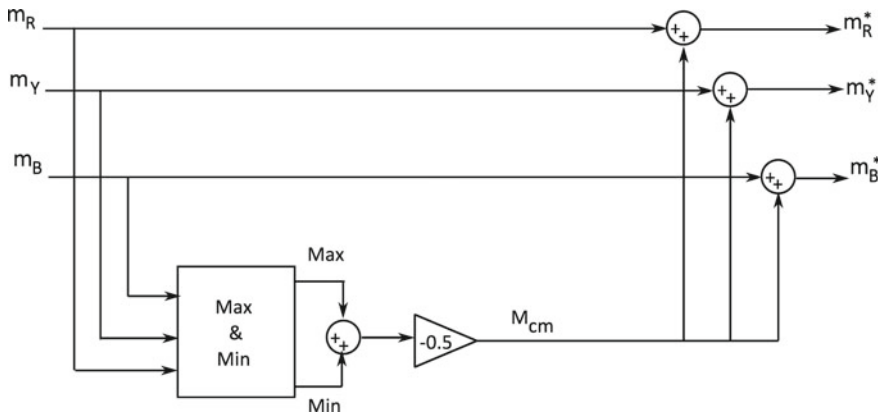
In the above equations,  $\omega_r$  is the rotor electrical speed and  $\omega_m$  is the rotor mechanical speed.  $T_e$  and  $T_l$  are the electromagnetic torque and load torque, respectively.  $B$  and  $J$  are the viscous damping coefficient and inertia of the motor. The basic block diagram of the control strategy used for the speed control of PMSM drive is presented in Fig. 1. The speed of the motor is feedback and is compared with the reference speed. The PI controller output represents the angular frequency, which is used to determine the frequency of the required stator voltage of the motor ( $f^*$ ). The drive



**Fig. 1** Block diagram for the control of PMSM drive

control block generates the voltage command using the frequency command as the input. The sine wave generator generates three 120° shifted sinusoidal modulating signals ( $m_R, m_Y, m_B$ ) using the commands from the drive control block. Modified modulating signals ( $m_R^*, m_Y^*, m_B^*$ ) required for carrier-based space vector PWM (SVPWM) are generated by the logic shown in Fig. 2.

Carrier-based space vector PWM generates equivalent switching patterns as conventional space vector PWM. Different PWM techniques used for inverter fed drives are detailed in [15–17]. In Fig. 2, the modified reference signals are generated by injecting a common mode signal ( $m_{cm}$ ) to the three-phase, sinusoidal modulating waves. These reference signals are then compared with a high frequency triangular carrier wave to obtain the pulses for each of the Insulated Gate Bipolar Transistor (IGBT) switches in the 3-phase inverter.



**Fig. 2** Modulating signal generation for carrier-based space vector PWM

### 4.2 Rapid Control Prototyping

In rapid control prototyping, the controller is implemented inside the simulator and the control signals are taken out with the help of the hardware module, through the I/O pins and is used to drive the actual inverter stack. The block diagram for Rapid Control Prototyping is shown in Fig. 3. PWM isolation module is used for providing isolation and amplification of the gate signals. In this experiment, a 3-phase IGBT-based Semikron inverter module is used to power the motor.

## 5 Simulation and Experimental Results

The simulation study is carried out in RBS mode and the same is experimentally validated in RCP mode. A 1.07 kW PMSM is used in both the cases. The motor used is a surface mounted type (having equal  $d$ -axis and  $q$ -axis stator inductance) and the parameters of the motor are listed in Table 1.

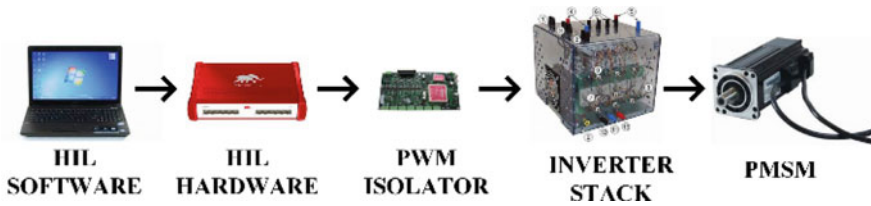


Fig. 3 Schematic diagram of the RCP mode of operation for a PMSM drive

Table 1 Parameters of motor

Symbols	Quantity	Value
Prated	Power	1.07 kW
$V$	Rated bus voltage	300 V
$N$	Rated speed	4000 rpm
$R_s$	Stator resistance	2.2 $\Omega$
$L_d, L_q$	$d$ -axis and $q$ -axis stator inductance	8.2 mH
$P$	Number of poles	4
$J$	Moment of inertia	$5.53 * 10^{-4}$ kg m <sup>2</sup>



### 5.1 Rapid Batch Simulation

The simulation diagram prepared in schematic editor of HIL for RBS mode is shown in Fig. 4. This study is carried out to analyze the performance of the drive in terms of its speed, torque, line voltages and stator current. SVPWM technique used here is an effective method to mitigate the harmonics in the output voltage of the inverter and to enhance the DC bus utilization. After compilation from schematic editor, a new SCADA panel is created to control the system. Capture/Scope in SCADA panel is an important widget used to view the waveforms at different parts of the model.

Continually running waveform is observed in scope, whereas it is captured using the capture mode. The entire control panel in SCADA is shown in Fig. 5. The motor is controlled by the commanded speed applied from the panel by means of slider. The load torque of the motor is varied by means of a torque slider provided in the panel. The controller gains  $K_p$  and  $K_i$  are designed and calculated using symmetric optimum principle. The values obtained are  $K_p = 0.9$  and  $K_i = 4.9$ .

Gages used for displaying the speed and torque shows exact tracking of the references given. The trace graph depicts the speed of the machine in rpm. All the waveforms depicted in this paper are plotted in MATLAB by exporting the data points to mat format. The 3-phase modulating signals shown in Fig. 6 corresponds to a fundamental frequency of 50 Hz. The shape of the modulating signals owes to addition of third harmonic components to the three-phase sinusoidal reference signals.

These signals are compared with triangular carrier of 5 kHz frequency (it is possible to develop switching signals up to 200 kHz) to generate the required switching signals. The switching pulses for the switches in the first leg of the inverter (S1 and S4) are given in Fig. 7 and these pulses are 180° phase shifted from each other. The pulses for top switches of the inverter (S1, S3 and S5) are given in Fig. 8. Here, 2 cases are considered for validation. One is the no load condition, where

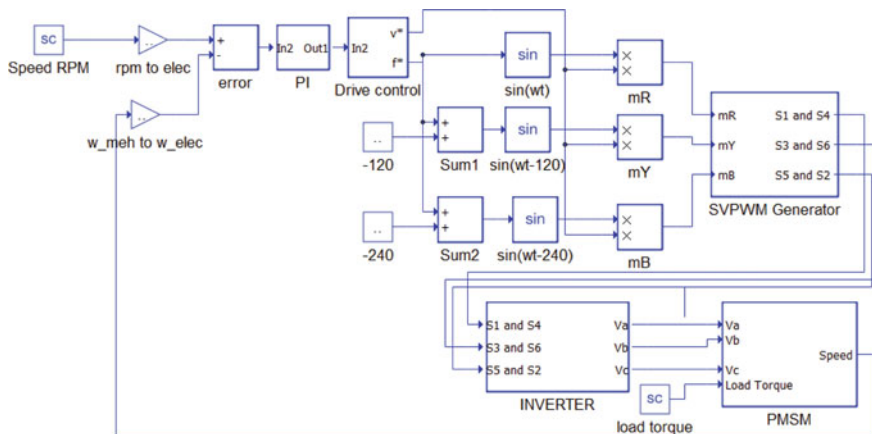


Fig. 4 Simulation diagram for rapid batch simulation

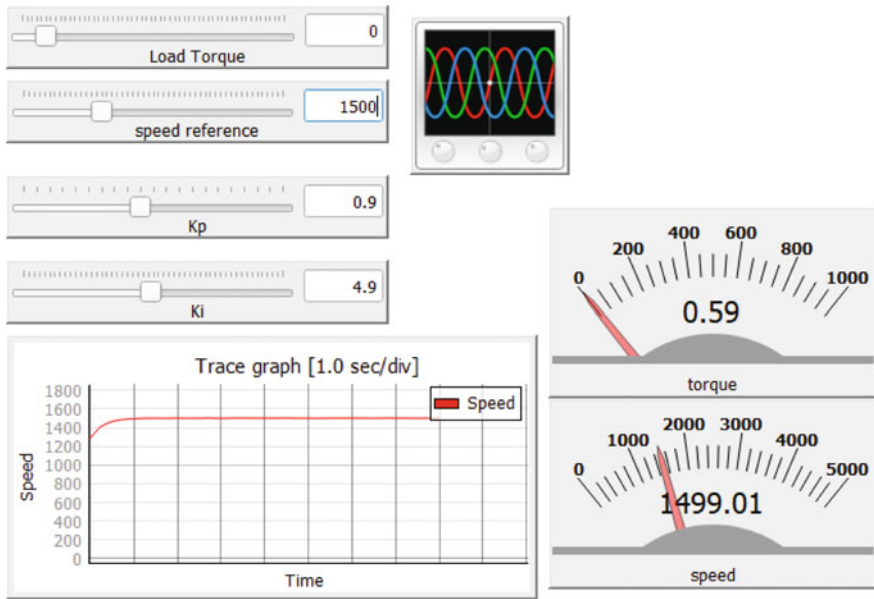


Fig. 5 SCADA panel for RBS mode of operation

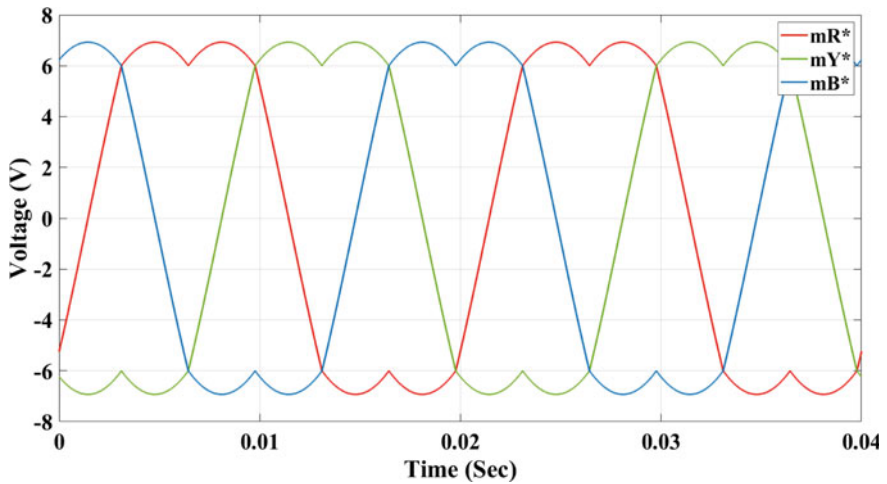
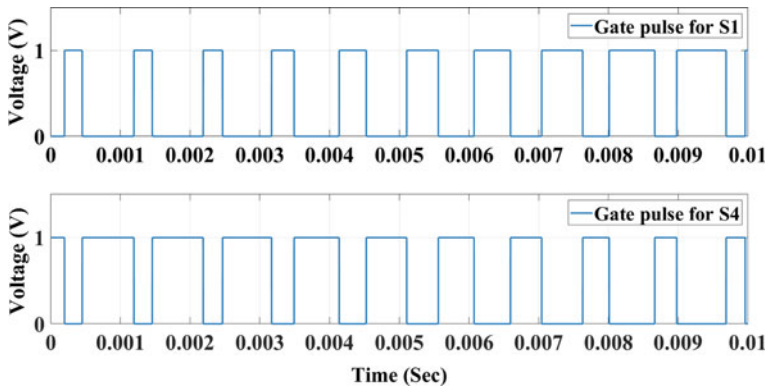
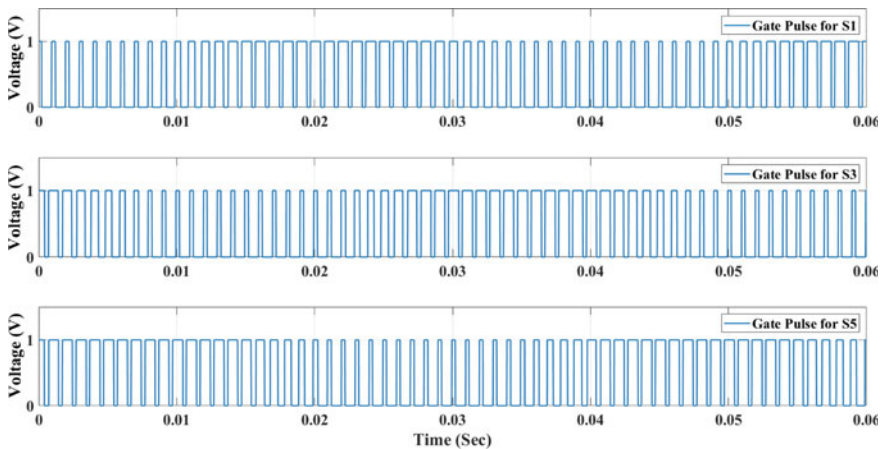


Fig. 6 Modulating signals for PWM generation obtained in simulation in RBS mode for a fundamental frequency of 50 Hz



**Fig. 7** Gate pulses to the switches in the first leg of the inverter obtained in simulation with RBS mode

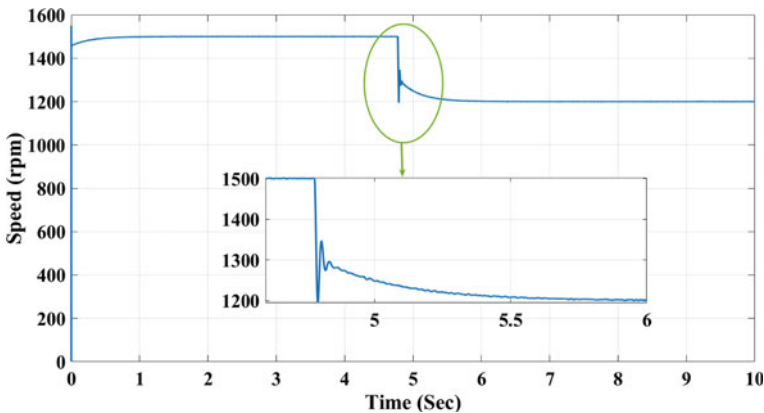


**Fig. 8** Gate pulses to the top switches of the inverter from simulation in RBS mode

speed reference is changed through a speed reference slider and second is the loaded condition, where load torque is applied using a torque slider at 1500 rpm.

The time variation of speed of the motor at no load condition is shown in the Fig. 9. Initially the motor is made to run at fundamental frequency of 50 Hz and after achieving steady state condition, the frequency is reduced to 40 Hz at 4.75 s. Correspondingly, the speed varies from 1500 to 1200 rpm. From the results, it is evident that the speed is following the given reference and is achieving the steady state quickly. Corresponding currents measured from the motor terminals are given in Fig. 10.

At 1500 rpm, the motor is drawing a sinusoidal stator current of 4.1 A (rms) and when the speed reference is reduced to 1200 rpm, the current reduces to 3.5 A (rms).

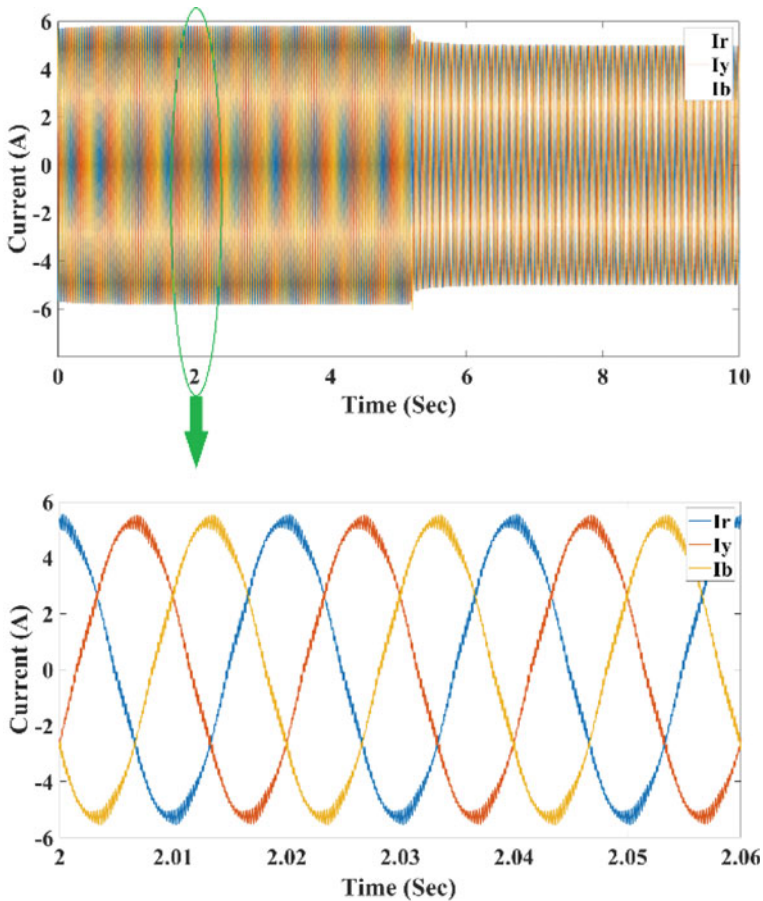


**Fig. 9** Simulated speed behavior for a step change from 1500 to 1200 rpm at 4.75 s under no load condition with RBS mode

The variation of electromagnetic torque developed, corresponding to a step change in load torque from no load to full load, at a speed of 1500 rpm is depicted in Fig. 11. The step change is given at 4.75 s. The non-sinusoidal components in the voltages at the inverter terminals induces ripples in torque. Two line voltages corresponding to 1500 rpm are depicted in Fig. 12 and it is clear that the fundamental line voltages are shifted by  $120^\circ$ .

## 5.2 Rapid Control Prototyping

The results obtained in simulation carried out in RBS mode is verified experimentally in RCP mode and the schematic diagram for the same is displayed in Fig. 13. The PWM strategy used here is the same carried-based Kim–Sul approach adopted in the simulation. The entire laboratory set-up for RCP mode of operation is shown in Fig. 14. A 1.07 kW PMSM is used as the target motor and is powered using Semikron make two level voltage source inverter module. Hardware module of HIL consists of the main module HIL 402, the Digital Signal Processor (DSP 100) and the TI controller (TMS320F28335). The control algorithm is simulated in the software part of HIL and the generated gate pulses are taken out through the hardware module. PWM isolation module used in this experiment functions as an isolator as well as an amplifier to boost the level of the pulses generated from the HIL real-time simulator. Figure 15 shows the modulating signals generated for the three legs of the inverter and are exactly matching with that of RBS mode. The injection of third harmonic component introduces flattening effect at the tip of the sinusoidal modulating wave which ensures more DC bus utilization. The given result corresponds to a fundamental frequency of 50 Hz. The experimentally generated gate pulses for the two switches on the first leg of the inverter with a switching frequency of 5 kHz is shown in Fig. 16.



**Fig. 10** Simulated motor stator current waveform for a step change in speed from 1500 to 1200 rpm at 4.75 s under no load condition with RBS mode

The pulses are of 15 V magnitude and are 180° out of phase with each other. A dead time is introduced between these pulses to avoid the shoot through problem in the inverter. Figure 17 depicts the PWM pulses given to the top switches of the two level inverter and are 120° in line with the simulation results. The pulses are 120° shifted to each other. The line voltage is measured using Tektronix P5122 high voltage probe and the waveforms obtained are as in Fig. 18. The voltages are maintaining a phase shift of 120°. The experimentally obtained stator current corresponding to a speed of 1500 rpm is measured using Tektronix TCP312A current probe and is shown in Fig. 19.

The current magnitude is seen to be 4.1 A with 1 A/V range selected in current probe. The stator current obtained is a sinusoidal function corresponds to a fundamental frequency of 50 Hz. Two line voltages  $V_{ry}$  and  $V_{yb}$  are shown in Fig. 18 in order to depict the exact shifting between them. The line voltage and stator current

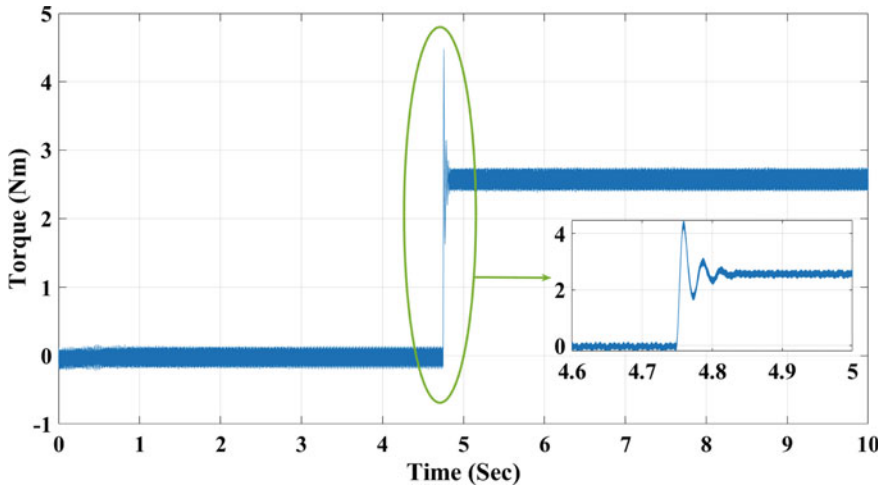


Fig. 11 Simulated electromagnetic torque developed for a speed of 1500 rpm with step change in load torque from no load to rated load at 4.75 s in RBS mode

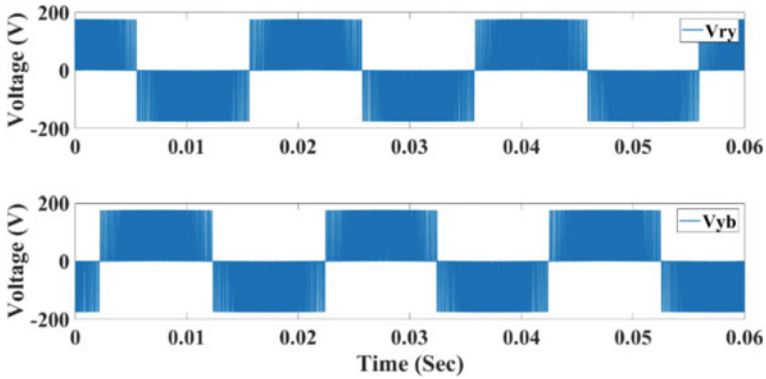


Fig. 12 Simulated stator line voltages  $V_{ry}$ ,  $V_{ry}V_{ry}$  and  $V_{yb}$  of the motor for a speed of 1500 rpm in RBS mode

waveforms obtained by experiment (RCP mode) are identical with the simulation results (RBS mode). The extensive simulation and experimental results validate the efficacy of Typhoon HIL real-time simulator in emulating any complex system even when the actual system is not available.

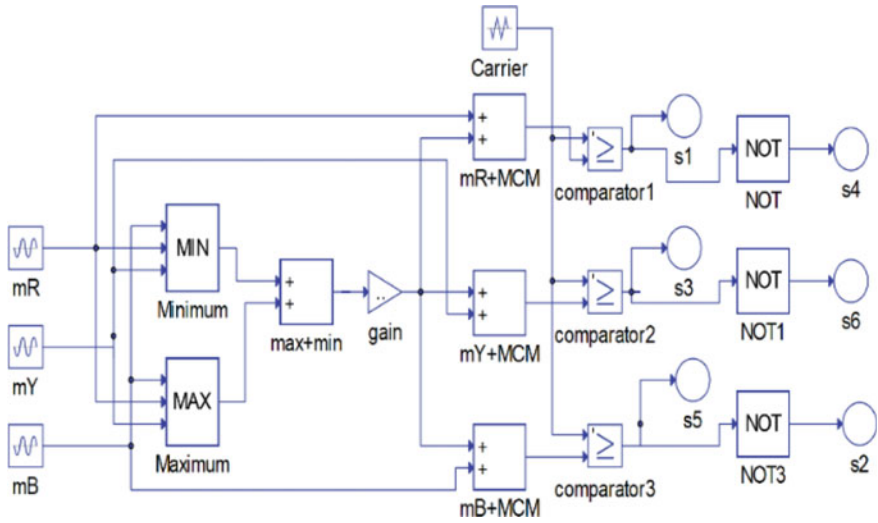


Fig. 13 Simulation schematic for the generation of control signals in RCP mode

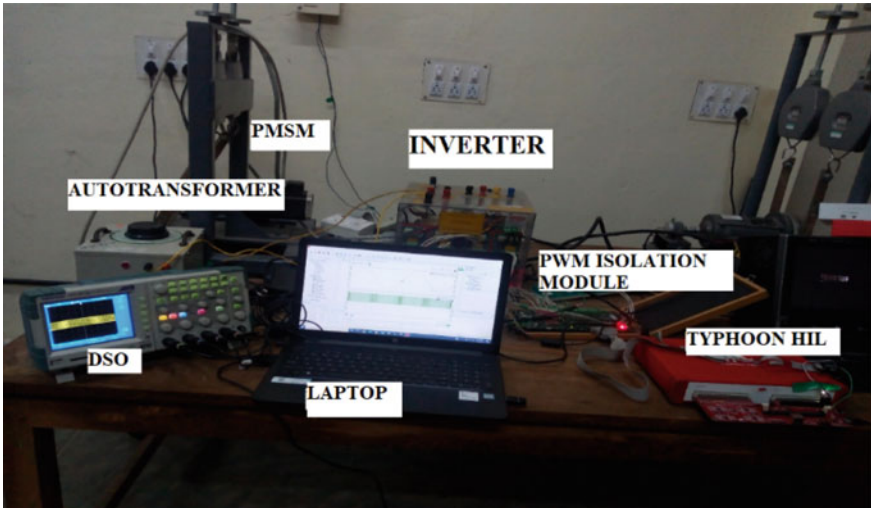


Fig. 14 Laboratory set-up for RCP mode of operation of PMSM drive with HIL 402

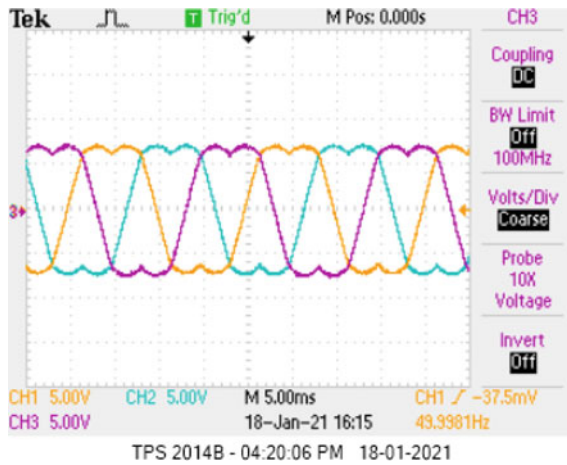


Fig. 15 Experimentally generated modulating signals in HIL with RCP

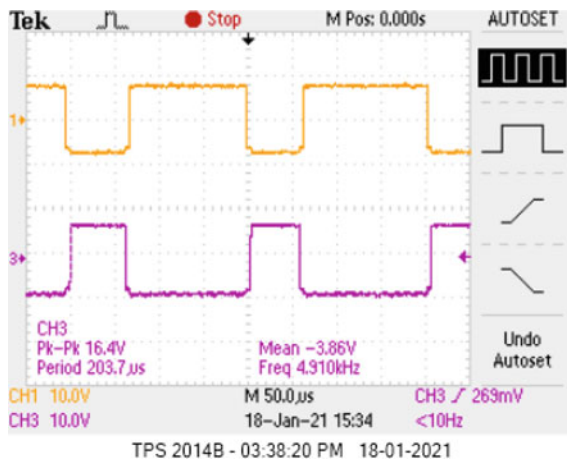


Fig. 16 Experimentally generated gate pulses for the first leg of the inverter in RCP mode of operation

## 6 Conclusion

Implementation and experimental validation of a PMSM-based Electric Drive system using Typhoon HIL 402 simulator is discussed in this paper. The motor is controlled using space vector PWM fed inverter. Two example models are created in Rapid Batch Simulation and Rapid Control Prototyping modes and implementation of the



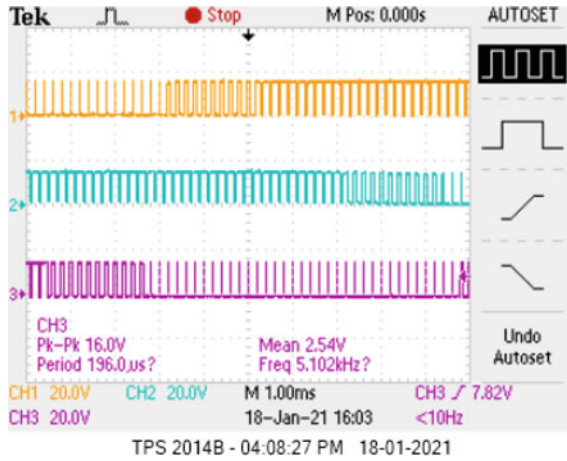


Fig. 17 Experimentally generated gate pulses for the top switches of the inverter in RCP mode

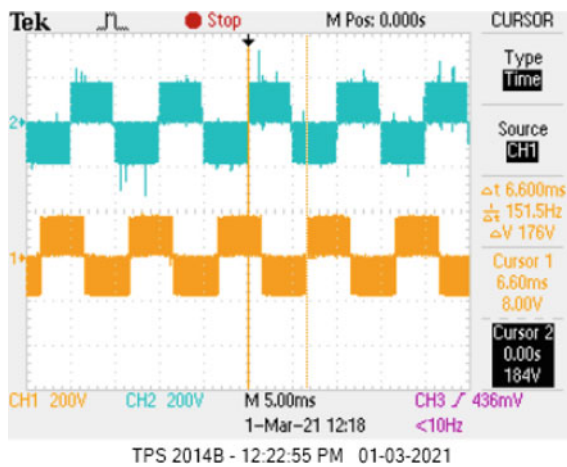
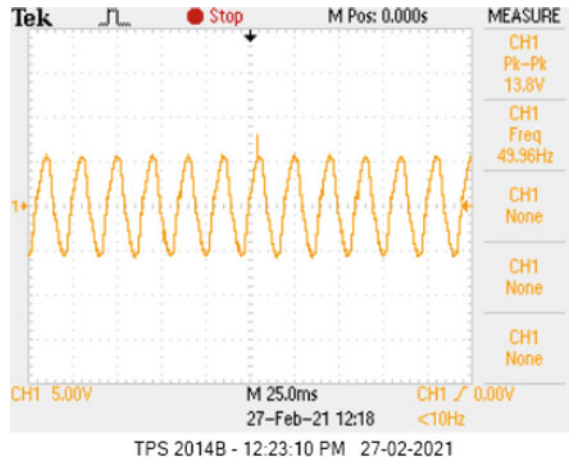


Fig. 18 Experimentally generated line voltages  $V_{ry}$   $V_{ry}$  and  $V_{yb}$  in RCP mode

same is explained. The experimental results are well matching with the theoretical prediction and the simulation results. Complex control schemes like Field Oriented Control, Direct Torque Control, Sensor less control, etc. can also be implemented in a similar manner. It is possible to model and test electric vehicles, microgrid and power system protection schemes using this platform.

**Fig. 19** Experimentally obtained R phase stator current in RCP mode



## References

1. Celanovic I, Haessig P, Carroll E, Katica V, Celanovic N (2009) Real-time digital simulation: enabling rapid development of power electronics. In: 15th International symposium on power electronics, Novisad, Serbia, pp 15–19
2. Jones JA (2014) State of charge based droop surface for optimal control of DC micro grids. Master's Thesis, Michigan Technological University
3. Dufour C, Andrade C, Bélanger J (2010) Real-time simulation technologies in education: a link to modern engineering methods and practices. In: 11th International conference on engineering and technology education (INTERTECH), Ilheus, Brazil, pp 114–118
4. Adžić EM, Adžić MS, Katic VA, Marcetić DP, Celanović NL (2013) Development of high-reliability EV and HEV IM propulsion drive with ultra-low latency HIL environment. *IEEE Trans Ind Inf* 9(3):630–639
5. Poon JJ, Kinsky MA, Pallo NA, Devadas S, Celanovic IL (2012) Hardware-in-the-loop testing for electric vehicle drive applications. In: IEEE applied power electronics conference and exposition—APEC, Orlando, USA, pp 2576–2582
6. Liu W, Luo G, Zhao N, Dou M (2009) Design and HIL simulation of proportional compression salient-pole permanent magnet synchronous motor for electrical traction vehicle. In: 5th IEEE vehicle power and propulsion conference, Dearborn, USA, pp 884–889
7. Abdelrahman AS, Algarny KS, Youssef MZ (2018) A novel platform for powertrain modeling of electric cars with experimental validation using real-time hardware in the loop (HIL): a case study of gm second generation Chevrolet volt. *IEEE Trans Power Electron* 33(11):9762–9771
8. Poon J, Chai E, Celanovic I, Genic A, Adzic E (2013) High-fidelity real-time hardware-in-the-loop emulation of PMSM inverter drives. In: IEEE Energy Conversion Congress and Exposition ECCE, pp 1754–1758
9. Dufour C, Lapointe V, Bélanger J, Abourida S (2008) Hardware in-the-loop closed-loop experiments with an FPGA-based permanent magnet synchronous motor drive system and a rapidly prototyped controller. In: IEEE International symposium on industrial electronics, Cambridge, UK, pp 2152–2158
10. Tang Z, Li X, Akin B, Jin H (2014) A high-fidelity integrated platform for rapid prototyping of AC motor drives. In: Industrial electronics conference, pp 3730–3735
11. Badini SS, Verma V (2019) A novel MRAS based speed sensorless vector controlled PMSM drive. In: Proceedings of 54th international universities power engineering conference, pp 1–6

12. Arafa OM, Abdallah ME, Abdel Aziz GA (2018) Realisation and HIL testing of wind turbine emulator based on DTC squirrel cage induction motor drive. *Int J Ind Electron Drives* 4(3):155–168
13. Li F, Wang Y, Wu F, Huang Y, Liu Y, Zhang X, Ma M (2020) Review of real-time simulation of power electronics. *J Mod Power Syst Clean Energy* 8(4):796–808
14. Krishnan R (2001) *Electric motor drives modeling, analysis, and control*. Oxford University Press, pp 525–529
15. Halasz S (2012) Continuous PWM strategies of multi-phase inverter-fed AC drives, *periodica polytechnica electrical engineering and computer*. *Science* 56(2):51–62
16. Khalid M, Mohan A, Binojkumar AC (2020) Performance analysis of vector controlled PMSM drive modulated with sinusoidal PWM and space vector PWM. In: *IEEE International power and renewable energy conference, Karunagappally, India*, pp 1–6
17. Chung DW, Sul JS, Sul SK (1998) Unified voltage modulation technique for real-time three-phase power conversion. *IEEE Trans Ind Appl* 34(2):374–380

# A Genetic Algorithm Approach for the Optimal Allocation of On-Site Generation in Radial Distribution Networks



Kasala Rajesh and J. Viswanatharao

## 1 Introduction

Electrical distribution system's power consumption is steadily increasing, resulting in increased load and voltage degradation. Restructuring idea has been already implemented in existing power networks; distributed generation (DG) has gained interest for its potential application. Distributed generation is when energy is created and delivered using small-scale technologies closer to its end users. Two main types of DG technologies used in distribution networks are renewable energy sources and fuel-based energy sources. Distribution network's operation is severely impacted by DG installation. Location of DG is essential because an inappropriate placement might result in power loss and voltage instability, and many studies have been conducted to decrease loss in distribution network [1].

The issue may appear to be a DG location issue, but it could also be a power quality (PQ) issue since it lowers voltage sag, which is definitely the most critical power quality issue. The magnitudes of voltage sags are strongly connected to the network's short circuit level. Because the failure level in distribution systems was relatively low, voltage sag is a key issue in distribution systems. DG connections raise the short circuit level, which tends to decrease voltage sag. As a result, changing the location of the DG connection may be an efficient strategy to decrease voltage sag.

The majority of sustainable power stations are expected to be connected as decentralized generators (DGs). They are defined in [2] as electric power producers or storage with capacities ranging from a few kilowatts to tens of megawatts that are not part of a major centralized power system and are situated closer to the consumers. As a

---

K. Rajesh (✉) · J. Viswanatharao  
Department of EEE, VNRVJIET, Hyderabad, Telangana 500090, India  
e-mail: [kasalarajesh566@gmail.com](mailto:kasalarajesh566@gmail.com)

J. Viswanatharao  
e-mail: [viswanatharao\\_j@vnrvjiet.in](mailto:viswanatharao_j@vnrvjiet.in)

result, they minimize the requirement for centralized energy production, transmission line losses, and voltage instability nearer to feeder terminals.

If DG is installed properly, voltage profile of system will be improved, and system losses will be reduced. Several variables, like as environmental, financial, and technical factors, should be considered before installing DG in distribution networks. [3] have established analytical methodologies for proper position of DG in distribution network for loss reduction.

Numerous optimization strategies are utilized in the literature [4]. Hung [5] presents an analytical approach for determining the best allocation of on-site generation. The majority of the papers use population-based meta heuristic algorithms as solution techniques. This includes particle swarm optimization, genetic algorithms, and evolutionary programming, in which many objective functions are aggregated to produce a single objective function and optimized.

The remaining sections of the paper are structured in the following order: Second section introduces fundamentals of genetic algorithms. Third section explains how to solve most favorable DG placement problem, and it covers load flow method. Section four gives an idea of an optimization problem and how to solve it with a genetic algorithm. Simulation results are analyzed, and conclusions are drawn in section five.

## 2 Problem Formulation

Major goal is to keep voltage limitations while minimizing overall active power losses ( $P_{\text{loss}}$ ). Recommended algorithm approach will be utilized to compute least distribution loss and find out appropriate DG size and best position [4].

Objective function is mathematically formulated as follows:

$$\text{Min. } f1 = \sum_{b=1}^{Nb} I_b^2 * R_b \quad (1)$$

where  $b$  represents number of branches and  $N_b$  and  $I_b$  are number of branches in total and branch currents, respectively.

$$f2(x) = \min \left( \frac{V_1 - V_j}{V_1} \right) \quad (2)$$

where  $j = 2, \dots, N_{\text{bus}}$  and  $V_1 = 1.05$ .

Following constraints apply to optimization problem:

(a) Following are the limits for DG power

$$P_{\min} \leq P_{DG} \leq P^{\max} \quad \forall_i = N_{DG}$$

(b) Limits for thermal are given as

$$I_{L_i} \leq I_{L_{\max_i}}$$

(c) At all nodes, voltage restrictions are as follows:

$$v^{\min} \leq v_i \leq v^{\max}$$

*Number and sizes of DGs:*

These are the constraints which fall under the DG and themselves. Generally, commercially available sizes of DG are discrete in nature [4]. By partial derivating Eq. (1), system total real power losses will be minimum with respect to injected power.

$$P_{DGn} = P_{DGn} - \frac{1}{\alpha_{nn}} \sum_m^n \sum_{m \neq 1} \alpha_{nm} P_{nm} - \beta_{nm} Q_m \tag{3}$$

$$Q_{DGn} = Q_{DGn} - \frac{1}{\alpha_{nn}} \sum_m^n \sum_{m \neq 1} (\alpha_{nm} Q_{nm} + \beta_{nm} P_m) \tag{4}$$

The sizes of DG can be determined by combining Eqs. (3) and (4). The optimal placement for DG is decided on the basis of lowest real power at the bus.

$$OPF = \frac{P_{DGn}}{\sqrt{P_{DGn}^2 + Q_{DGn}^2}} \tag{5}$$

*Optimal allocation*

As the method proposed in [4] is used for an optimal location, by this method first the various locations with respective optimal sizes are calculated for different DG, and their losses were calculated with optimal sizes for each case. Selection of optimal location of each DG is based on the case which gives minimum losses. The main advantage of this method is when dealing with the large distribution system, we can avoid the unnecessary computations which can save time.

Practically, complex distribution system consists of less number of sources, many buses, lines, and loads. As the power factor of each load will not be same, if each local DG supplies to each load, then each DG power factor will be equal to their respective load, where it results no current flow in lines. It shows total line power loss is zero.

So it is unnecessary to use transmission line in such locations, and it would be not recommended since capital investment cost of DG is very high; therefore, it reflects there should be a limited number of installed DG in the system.

### 3 Proposed Methodology

The backward/forward sweep way of calculating power flow is an iterative approach. The forward sweep is essentially a voltage drop computation with optional current or power flow updates. Nodal voltages are refreshed in a forward sweep, beginning with branches in the first layer and progressing to those in the final. The reverse sweep is essentially a current or power flow answer with voltage revisions. It begins with the branches in the final layer and progresses to the branches related to the root node. Forward and backward sweep power flow method is used in this research, and it is based on the usage of current injections that are similar [6]. Forward sweep and backward sweep steps of load flow method are as follows.

*Backward sweep:*

In this step, the load current of each node can be calculated with N nodes in the distribution system as equation (6):

$$\vec{I}_L(n) = \frac{P_L(n) - jQ_L(n)}{\vec{V}^*(n)} \tag{6}$$

[ $n = 1, 2, 3, 4 \dots N$ ], where  $Q_L(n)$  and  $P_L(n)$  denote reactive and active power demand at node  $n$ , respectively, and ( $\vec{x}$ ) indicates phasor values such as  $\vec{I}_L, \vec{V}^*$ , and the following formula is used to compute the current in each branch.

$$\vec{I}_L(nm) = \vec{I}_L(m) + \sum_n \vec{I}_L(n) \tag{7}$$

*Forward sweep:*

Following backward sweep, in this stage voltage is calculated at each node of a distribution network, as follows:

$$\vec{V}(m) = \vec{V}_n(m) + \vec{I}(nm)Z(nm) \tag{8}$$

where nodes  $m$  and  $n$  are receiving and transmitting end nodes for branch  $nm$ , respectively, and  $Z(nm)$  is branch's impedance. Equivalent current injections of node in distribution system define this power flow technology.

Complex load  $S_i$  of a node is described as follows:

$$S_i = P_i + jQ_i \tag{9}$$

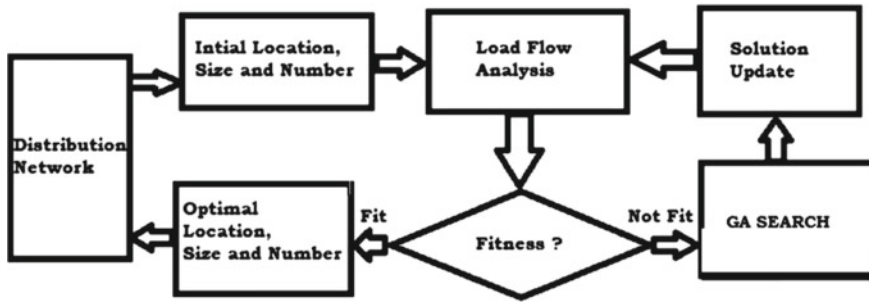


Fig. 1 Genetic algorithm approach for optimal placement of DG

Figure 1 depicts the model of the proposed method. In this work, an approach is developed for the location of several DGs and their size based on objective criteria, with the goal of improving performance of the system. A genetic algorithm method is used in the proposed structure. The suggested framework’s aims are to maximize voltage profile and reduce power loss [7]. The GA method is used for optimal placement of multi-DG with appropriate size based on the objective function.

### 4 Problem Modeling with Genetic Algorithm

GA is based on evolutionary concepts like characteristic selection and a generation of living organism. In this algorithm, the search method will be wisely randomized and can have the potential to verify multiple closely spaced answers in a single iteration. “It is based on Darwin’s Natural Selection and other evolutionary concepts” [7].

In general, there are three stages to search process in GA: Phase 1 is creation of a primary population; an evaluation of fitness function was done in phase 2, and phase 3 is generation of a new population. Fitness function is the only variable that GA optimizes. As a result, problem’s objective function and some of its constraints must be converted into fitness function. This function, which is closely related to objective function, measures quality of chromosomes [8–10].

A simplified version of entire MATLAB program is used for solving DG objective function, which includes load flow algorithm and genetic algorithm. Figure 2 shows the suggested optimization technique for choosing the best location and size of on-site generator in a radial distribution system using a genetic algorithm.

### 5 Implementation and Results

MATLAB platform has been used to implement suggested genetic algorithm technique for optimal DG placement. A 33-bus, 11 kV radial feeder is used to test



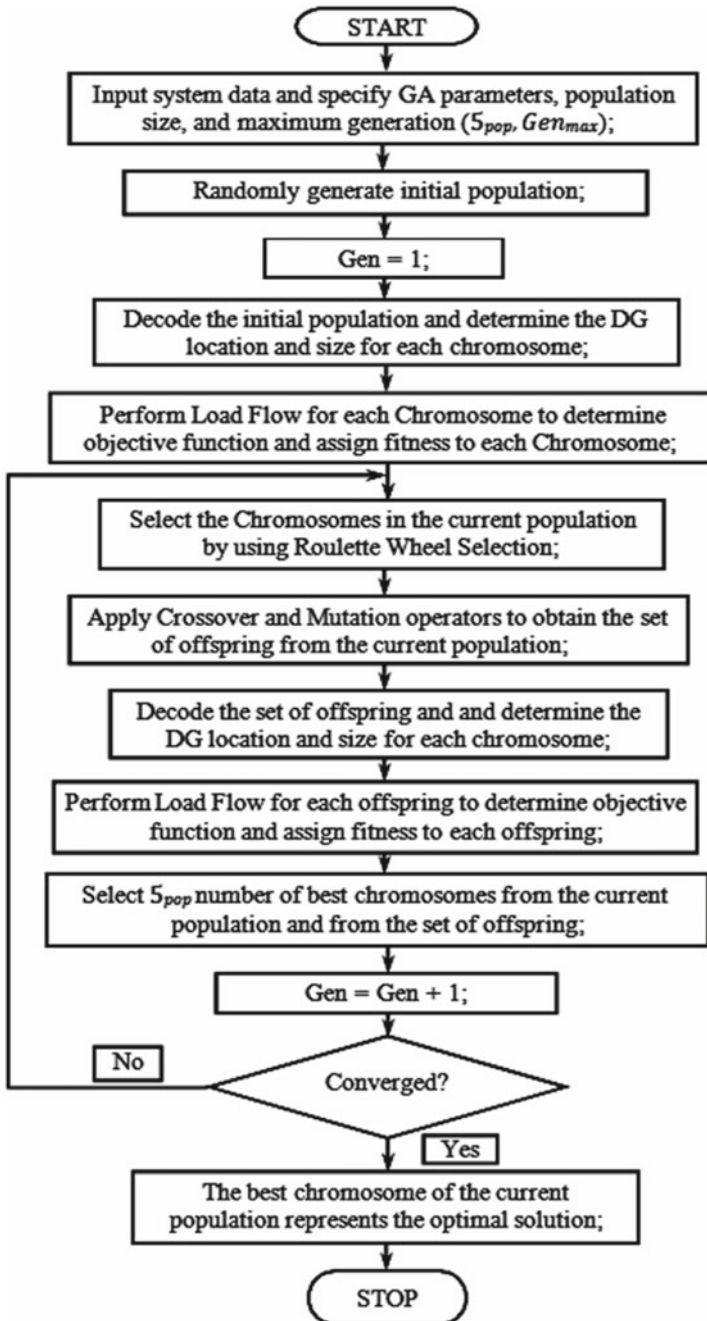


Fig. 2 Proposed genetic algorithm approach for optimal placement of DG

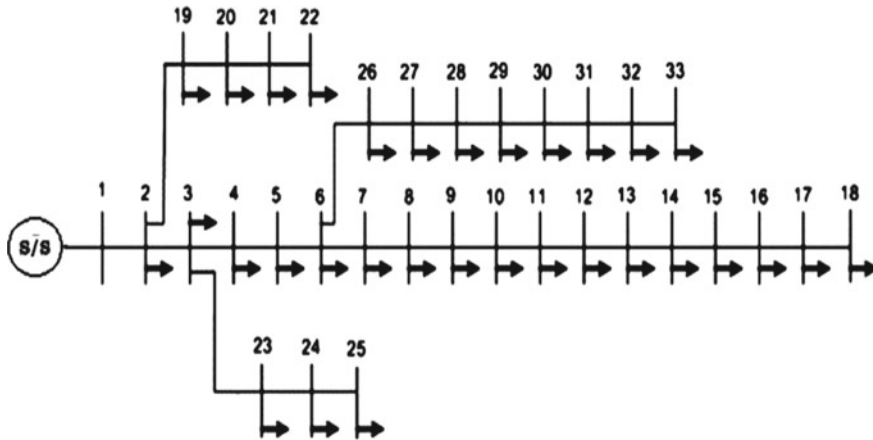


Fig. 3 Single line diagram of 33-bus radial distribution test network

suggested approach. Single line diagram of a 33-bus radial distribution test system is shown in Fig. 3. GA was repeated several times, each time with a different initial population. Because primary population which offers earliest genetic chromosome is generated at random, each experiment had a different set of results.

In general, method is entirely dependent on random processes.

### 5.1 Results of Computations Using Genetic Algorithm Approach

In this analysis, forward and backward sweep distribution load flow is used. To achieve required outcome for the 33-bus test system, a program/code was written in MATLAB R2016a. Figure 4 shows a graphical depiction of system voltage profile in basic scenario for a test system.

Figure 5 shows a voltage profile of 33-bus test systems with a disrupted load and without any DG, and the voltage profiles are violated due to the change in load. Active and reactive loads are modified as in [11]. When this case is implemented by using the genetic algorithm, then the voltage profile will be enhanced.

The total active power generation is 3690 kW, with a total reactive power generation of 2843 KVAR, and total active power loss is 206 kW, with a 137 KVAR reactive power loss. The minimum voltage without DG is 0.8823.

VSI is calculated for test network to determine appropriate node for DG placement, and bus numbers 6, 18, and 33 have lowest VSI value when compared to other buses, as given in Table 1. It is obvious that device should be installed at bus's weakest location to enhance potential benefit. In consideration of this, bus numbers 6, 18, and 33 have been chosen as best locations for DG for 33-bus test system.

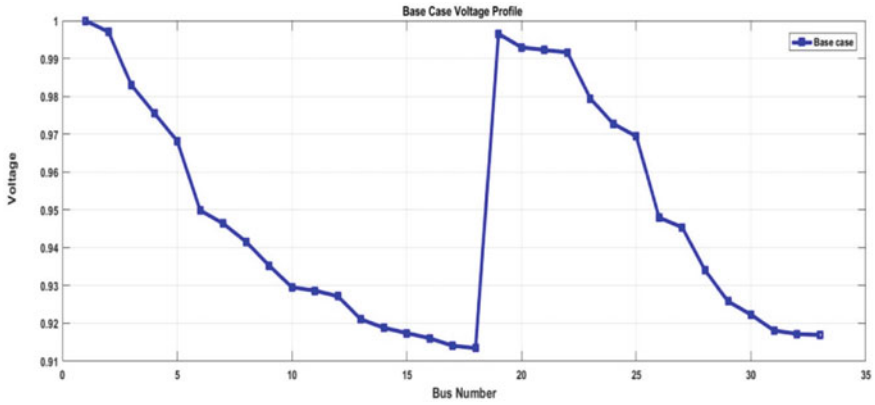


Fig. 4 Base case voltage profile of 33-bus radial distribution test system without DG

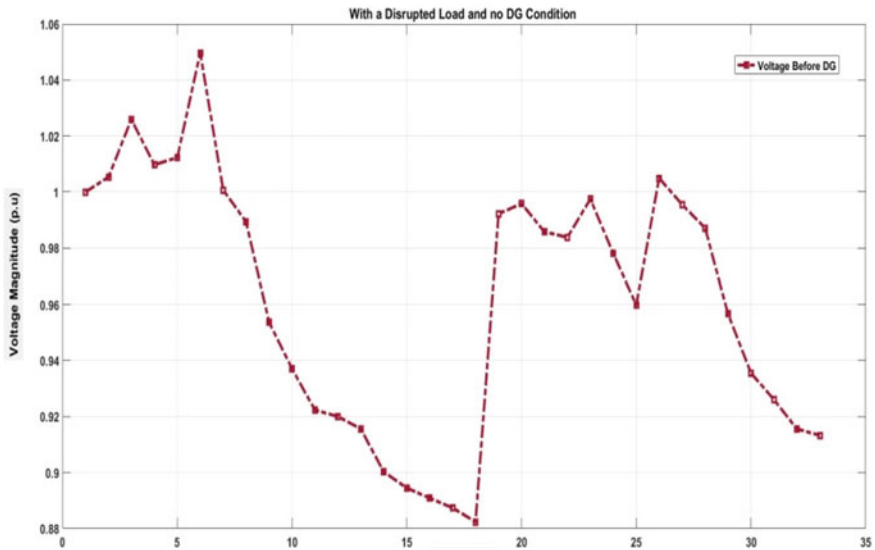


Fig. 5 Voltage profile of test system with a disrupted load and no DG condition

Table 1 Computational results with genetic algorithm approach

Bus location	Minimum stability index	DG size (Mw)	Total DG size (MW)	Loss before DG (kw)	Loss after DG (KW)	Power loss reduction in % (KW)
6	0.7146	1.018	2.1	336.166	175.87	47.6
18	0.8607	0.4895				
33	0.7882	0.662				

Figure 6 depicts a clear comparison of voltage profile in two different scenarios. Case 1 shows the system performance with a disrupted load with out DG condition and Case 2 shows the system performance was affected due to the disorganized loading condition of the system, the performance was reduced and the constraints were violated. When an on-site generator is placed in an appropriate location in the disturbed load condition of the test system, it reduces losses while also improving the system’s voltage profile.

Table 2 demonstrates outcomes of implementing DG for a 33-bus test system using proposed GA approach. System’s minimum voltage after DG is 0.9335; proposed

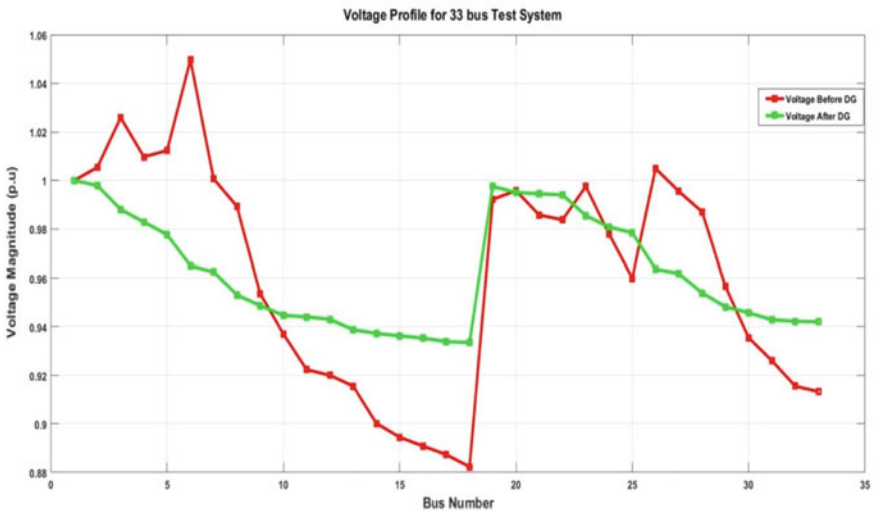


Fig. 6 Voltage profile for 33-bus test system with load disturbance

Table 2 Obtained results of 33-bus test system with and without DG

S. No	Bus No.	Voltage profile's (pu)		
		Base case	At load disturbance	After DG placement
1	1	1.0000	1.0000	1.0000
2	2	0.9975	1.0054	0.9979
3	3	0.9837	1.0259	0.9881
4	4	0.9755	1.0097	0.9829
5	5	0.9682	1.0123	0.9777
<b>6</b>	<b>6</b>	<b>0.9498</b>	<b>1.0497</b>	<b>0.9649</b>
7	7	0.9464	1.0007	0.9625
8	8	0.9416	0.9893	0.9531
9	9	0.9353	0.9537	0.9487

(continued)

**Table 2** (continued)

S. No	Bus No.	Voltage profile's (pu)		
		Base case	At load disturbance	After DG placement
10	10	0.9295	0.937	0.9447
11	11	0.9287	0.9223	0.9441
12	12	0.9272	0.9258	0.943
13	13	0.9211	0.9155	0.9388
14	14	0.9188	0.9002	0.9372
15	15	0.9174	0.8945	0.9362
16	16	0.9161	0.8909	0.9353
17	17	0.9141	0.8873	0.9339
<b>18</b>	<b>18</b>	<b>0.9135</b>	<b>0.8823</b>	<b>0.9335</b>
19	19	0.9965	0.9922	0.9975
20	20	0.9929	0.9958	0.995
21	21	0.9922	0.9858	0.9945
22	22	0.9916	0.9838	0.9941
23	23	0.9794	0.9975	0.9855
24	24	0.9728	0.978	0.9808
25	25	0.9694	0.9598	0.9785
26	26	0.9479	1.0049	0.9636
27	27	0.9454	0.9955	0.9618
28	28	0.934	0.9871	0.9539
29	29	0.9258	0.9567	0.9482
30	30	0.9223	0.9355	0.9458
31	31	0.9181	0.9261	0.9429
32	32	0.9172	0.9156	0.9422
<b>33</b>	<b>33</b>	<b>0.9169</b>	<b>0.9133</b>	<b>0.942</b>

Bold text in the table reflect the best placement of DG in the system with respect to minimum voltage.

GA approach has been found to reduce active power loss and improve voltage profile. Figure 4 depicts voltage profiles for a 33-bus test system.

Incorporation of DG leads to a significant voltage profile improvement in the system when compared to base case and disturbed load condition and obtained results are tabulated above.

## 6 Conclusion

DG installation in distribution networks is critical in terms of electricity generation. It improves voltage profile and helps distribution system to handle heavy loads while also optimizing voltage profile and reducing losses. Purpose of this research is to use a genetic algorithm to locate and size DGs in a radial distribution network, with total active power loss of system as target. Simulation results also demonstrated that good outcomes may be produced under disturbed loading situations, indicating that proposed approach is feasible. After evaluating data, it was fulfilled that appropriate on-site generator placement is favorable since it decreases system power loss and improves distribution system's voltage profile significantly.

## References

1. Keane A (2013) State-of-the-art techniques and challenges ahead for distributed generation planning and optimization. *IEEE Trans Power Syst* 28(2):1493–1502
2. Srinivas MS (2000) Distribution load flows: a brief review. In: 2000 IEEE Power Engineering Society Winter Meeting. Conference Proceedings (Cat. No.00CH37077), vol 2, pp 942–945. <https://doi.org/10.1109/PESW.2000.850058>
3. Biswas S, Goswami SK (2010) Genetic Algorithm based optimal placement of distributed generation reducing loss and improving voltage sag performance. *ACEEE Int J Electr Power Eng* 2
4. Ghiani CGE, Mocci S, Pilo F (2005) A multiobjective evolutionary Algorithm for the sizing and siting of distributed generation. *IEEE Trans Power Syst* 20:750–757
5. Hung DQ, Mithulananthan N, Bansal RC (2010) Analytical expressions for DG allocation in primary distribution networks. *IEEE Trans Energy Conversion* 25(3):814–820
6. Appa Rao A, Win Babu M (2013) Forward sweeping method for solving radial distribution networks. *IJAREEIE* 2(9)
7. Sattianadan D, Sudhakaran M, Dash SS, Vijayakumar K, Ravindran P (2013) Optimal placement of DG in distribution system using genetic algorithm. In: Panigrahi BK, Suganthan PN, Das S, Dash SS (eds) *Swarm, evolutionary, and memetic computing. SEMCCO 2013. Lecture notes in computer science*, vol 8298. Springer
8. Varikuti R, Damodar Reddy M (2009) Optimal placement of DG units using fuzzy and real coded genetic algorithm. *J Theor Appl Inf Technology* 145–151
9. Ni Ketut Aryani MA, Negara IMY, Soeprijanto A (2011) Optimal placement and sizing of distributed generation using quantum genetic algorithm for reducing losses and improving voltage profile. In: *TENCON 2011—2011 IEEE region 10 conference*, pp 108–112. <https://doi.org/10.1109/TENCON.2011.6129073>
10. De Souza BA, De Albuquerque JMC (2006) Optimal placement of distributed generators networks using evolutionary programming. In: 2006 IEEE/PES transmission & distribution conference and exposition, Latin America, pp 1–6. <https://doi.org/10.1109/TDCLA.2006.311571>
11. Baran ME, Wu FF (1989) Network reconfiguration in distribution systems for loss reduction and load balancing. *IEEE Trans Power Deliv* 4(2):1401–1407. <https://doi.org/10.1109/61.25627>

# Enhanced Thermophysical Properties and Productive Yield of Pyramid Solar Still Combined with Shallow Solar Pond by Incorporating ZnO/Al<sub>2</sub>O<sub>3</sub> Nanocomposites



B. Selvakumar, V. Shanmugapriya, K. Amudha, and P. Periasamy

## *Nomenclature*

$A$	Area, m <sup>2</sup>
$T$	Temperature, °C
$m_e$	Mass of the water, kg
$I$	Incident solar radiation, W/m <sup>2</sup>
$k$	Thermal conductivity, W/m °C
$\eta$	Efficiency (%)
$P$	Saturated Vapor Pressure, Pa
$h_v$	Latent Heat, kJ/kg
$\mu$	Viscosity Coefficient, kg/m s
$\rho$	Density of water vapor, kg/m <sup>3</sup>
$Q$	Heat transfer coefficient, W/m <sup>2</sup>

---

B. Selvakumar (✉)

Department of Physics, Sri Eshwar College of Engineering, Kinathukidavu, Coimbatore, Tamil Nadu, India

e-mail: [solarselva@gmail.com](mailto:solarselva@gmail.com)

V. Shanmugapriya

Department of Physics, International Research Centre, Kalasalingam Academy of Research and Education, Srivilliputtur, Tamil Nadu, India

K. Amudha

R.M.D Engineering College, Kavaraipeitai, Tamil Nadu, India

P. Periasamy

Nehru College of Engineering and Technology, T.M. Palayam, Coimbatore, Tamil Nadu, India

## *Greek and Suffix*

<i>a</i>	Air
<i>g</i>	Glass
<i>w</i>	Water

## **1 Introduction**

Provision of freshwater is fetching a notable issue around the world. Most abundant resources on earth are water, which covers about 3/4th of its surface. Only 3% (about 36 million km<sup>3</sup>) of it is the freshwater. Potable water is very scarce in arid and semi-arid region, the issue being resolved by desalination.

Human civilization mainly depends up on fresh water resources. Contamination of available fresh water resources by discharged sewage leads to further scarcity of water. Naturally and freely available sunlight is used for the desalination of saline (or) contaminated water. As a result of evaporation, yielded distillate water is without heavy metal and salt impurities. Naturally available sunlight can be used as an alternative for the desalination of water at placed having low rainfall, arid and semi-arid regions.

Hamdan et al. [1] showed that the productivity of freshwater mainly depends on the value of radiation and the period of exposure. Productive yield of greenhouse still for growing plants in the improved utilization of the available sunlight have been analyzed by Selvakumar et al. [2]. Tiwari [3, 4] investigated the basin type still which concludes that higher yield collector should be engaged only during daytime analysis. Selvakumar et al. [5] observed the better performance still coupled with charcoal absorber. Four-sloped basin solar still of area 0.986 m<sup>2</sup> have been analyzed by Jairaj et al. [6] and it is found that the observed efficiency is 33.99% while the predicted one is 35.29%. Long term distillate yield of typical solar still is studied by Mathioulakis et al. [7], Tripathi and Tiwari [8] and Singh and Tiwari [9, 10].

Shallow Solar Pond (SSP) is a thermal storage unit which is used for collecting and storing incident solar radiation in the form of heat. Noteworthy model and technological application of SSP to industrial application were studied by Casamajor [11] and Dickinson and Neifert [12]. Application of SSP in Military had been analyzed by Kudish and Wolf [13]. Results concluded that the water temperature depends on its level. SSP with various water level were analyzed by Garg [14] and the results inferred that better performance for the SSP with low water storage of around 4–15 cm. Heat extraction of SSP had been investigated by Varghese and Awari [15] and Kishore et al. [16]. Thermal performance of SSP with thermosyphon mode had been studied by Casamajor and Parsons [17], Bansal et al. [18] and Sabetta et al. [19]. Results concluded that thermosyphon mode is more efficient.

Chen et al. [20] studied role of ZnO/Al<sub>2</sub>O<sub>3</sub> in the decomposition of hazardous chemicals. Quynh et al. [21] demonstrated the COS conversion which reached



the maximum level of 81% using ZnO/Al<sub>2</sub>O<sub>3</sub> nanocomposites. Photocatalytic and antibacterial studies were reported by Yadav et al. [22] and results showed the optimal Photo-degradation of dye and improved antibacterial activity. Jensen et al. [23] reported ZnO–Al<sub>2</sub>O<sub>3</sub> nanoparticles and investigated their morphology.

In this report, for the first time, a shallow solar pond with heat extraction using a copper heat exchanger coupled with acrylic pyramid still to examine for high distillate yield with and without ZnO/Al<sub>2</sub>O<sub>3</sub> nanocomposites.

## 2 Materials and Methods

### 2.1 Synthesis of ZnO/Al<sub>2</sub>O<sub>3</sub> Nanocomposites

All the reagents used in the experiments were in analytic grade was purchased from SRL chemicals, India and used without further purification. The following chemicals were used in the preparation method: Aluminum acetate [C<sub>6</sub>H<sub>9</sub>AlO<sub>6</sub>], Zinc acetate [Zn(CH<sub>3</sub>CO<sub>2</sub>)<sub>2</sub>], ethylene glycol [C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>], Urea [CH<sub>4</sub>N<sub>2</sub>O] and Deionized water.

In a typical process, the ZnO/Al<sub>2</sub>O<sub>3</sub> nanocomposites were synthesized by using a hydrothermal method, which is a solution reaction-based approach. Nanomaterials which are not stable at preminent temperatures are synthesized by hydrothermal method. 0.02 M of Aluminum acetate [C<sub>6</sub>H<sub>9</sub>AlO<sub>6</sub>] and 0.02 M of Zinc acetate [Zn(CH<sub>3</sub>CO<sub>2</sub>)<sub>2</sub>] are dissolved in 110 ml of Deionized water and 10 ml of ethylene glycol dispersion agent is added along with it. 1 g of Urea via dropwise is added to the above-mixed suspension and kept stirring for 30 min. Mixed solution is heated for 8 h at 120 °C. Yielded sample is calcinated at 350 °C for 2 h to gain the ZnO/Al<sub>2</sub>O<sub>3</sub> nanocomposites.

### 2.2 Fabrication of Pyramid Solar Still (PSS)

Stainless steel material is used to design the PSS of area 0.50 m × 0.50 m. Height of water storage basin is around 0.15 m. ¼ inch pipe is provided at of 0.13 m from the base of the still for the inlet of water. Two pipes placed at the diagonally opposite side in the base of the still is used for removal of salt contamination and also for water circulation with SSP. Water collection segment of dimension 0.625 m × 0.0195 m × 0.015 m is used to collect the distillate. 3 mm thickness of the transparent acrylic sheet is used for top pyramid cover arrangement. PSS is placed in outer wooden box with sawdust and glass wool for preventing the heat loss.

### **2.3 Construction of Shallow Solar Pond (SSP)**

SSP of dimension 1.05 m × 1.05 m × 0.15 m is designed by using thermocouple and cement mortar. A base platform of a height of 0.1 m is laid with the help of cement mortar. Its sides are covered with the help of bricks and cement mortar. Thermocouples are placed at the bottom (0.15 m), middle (0.45 m) and surface (0.60 m) of the insulation layer to measure the temperature of the SSP. The gap between the insulation layer and the bricks wall is also filled by pieces of thermocoles. Water depth of 0.05 m is maintained and the surface is covered by a transparent polythene sheet. Transparent glass plate with wooden frames is tightened using bolts provided at the surface of the wall with cushions to reduce air leakage.

### **2.4 Experimental Arrangement**

Performance of the SSP is studied by filling the pond with a 5 cm thickness of water layer. The surface of the water is covered by using a transparent polythene sheet which acts as an evaporation suppresser. The total volume of the water filled in the storage basin is about 39 L. The hourly rise in temperature, incident solar insolation and room temperature is measured. PSS basin is filled with 10 L of saline water along with pre-calibrated thermocouples to observe the water and air temperature in the basin. Distillate yield is collected using the measuring jars placed at diagonally opposite ends.

In the combined performance study, the output of the still is the input of SSP and the output of SSP is the input of the still. Input and output are connected with CPVC pipes of diameter ¼ inch which is well insulated with glass wool of thickness 1¾ inch. This insulation is covered by using a PVC pipe of the inner diameter of 2 inches. Figure 1 shows the schematic representation of the combined performance of PSS with SSP.

### **2.5 Characterization Techniques**

Structural examination of the prepared sample was carried out by XRD, functional groups associated with as-prepared samples were studied by FTIR spectrum. Surface morphology of synthesized samples were investigated by SEM. Presence of elemental analysis was conducted by Energy Dispersive Spectroscopy Spectrum.



**Fig. 1** Combined performance of Pyramid Solar Still (PSS) with Shallow Solar Pond (SSP)

### 3 Thermophysical Properties

Toyoma et al. [24] proposed the thermophysical physical parameters of distillate yield using the condensation and evaporation temperatures as

$$k = 0.0244 + (0.7673 \times 10^{-4}) T_{av} \tag{1}$$

$$\mu = (1.718 \times 10^{-5}) + (4.620 \times 10^{-8}) T_{av} \tag{2}$$

$$\rho = \left( \frac{353.44}{273.35 + T_{av}} \right) \tag{3}$$

$$h_v = 2324.6 \left[ (1.0727 \times 10^3) - (1.0167) T_{av} + (1.4087 \times 10^{-4}) T_{av}^2 - (5.1462 \times 10^{-6}) T_{av}^3 \right] \tag{4}$$

SVP and PR of the still under two modes of analysis is calculated using the expression given by Brooker et al. [25]

$$P = 6893.03 \exp \left[ 54.63 - \left( \frac{12301.69}{T'} \right) - 5.17 \ln(T') \right] \tag{5}$$

$$PR = \left( \frac{m_e h_v}{I} \right) \tag{6}$$

## 4 Heat Transfer Modes in PSS

Thermal performance of PSS with SSP is enhanced by incorporating the heat transfer by various means using the expressions given by Dunkle [26] and Adhikari et al. [27].

### 4.1 Internal Heat Transfer Modes

$$Q_{ci} = 0.88 \left( (T_w - T_g) + \left( \frac{(P_w - P_g)(T_w + 273)}{268.9 \times 10^3 - P_w} \right)^{1/3} \right) (T_w - T_g) \quad (7)$$

$$Q_{ei} = 16.273 \times 10^{-3} h_{ci} \cdot R_1 (T_w - T_g) \quad (8)$$

$$Q_{ri} = \sigma \varepsilon \left[ (T_w + 273)^4 - (T_g + 273)^4 \right] \quad (9)$$

### 4.2 External Heat Transfer Modes

$$Q_{ce} = h_{ca} (T_g - T_a) \quad (10)$$

$$Q_{re} = \varepsilon_g \sigma \left[ (T_g + 273)^4 - (T_{sky} + 273)^4 \right] \quad (11)$$

$$Q_{be} = h_b (T_w - T_a) \quad (12)$$

## 5 Efficiency

Efficiency of PSS is calculated using latent heat and hourly distillate yield using the formula

$$\eta = \left( \frac{m_e h_v}{I A t} \right) \quad (13)$$

## 6 Results and Discussion

Structural analysis of as-synthesized  $\text{ZnAl}_2\text{O}_3$  nanocomposite analyzed by XRD is shown in Fig. 2. Characteristic peak of ZnO was observed (100), (002), (101), (102), (110), (103), (112) and (201), respectively, and all these peaks were indexed the JCPDS card No: 00-036-1451.

Similarly, for the  $\text{Al}_2\text{O}_3$  structure, some characteristic peaks were observed at (311), (400) and (440) and indexed by JCPDS card no: 00-010-0425 which showed cubic structure. Interestingly, the characteristic peaks of ZnO and  $\gamma\text{-Al}_2\text{O}_3$  were broadened due to the superposition of peaks at  $32.0^\circ$ , and  $36.4^\circ$  [28]. Our results were well supported by the previously reported work by Yadav et al. [29], He et al. [30] and Romcevic et al. [31].

Other than ZnO and  $\text{Al}_2\text{O}_3$ , none of the other elements was observed and it was confirmed by EDAX spectrum and elemental mapping analysis as shown in Fig. 3.

Fourier transform Infrared spectrum analysis of as-prepared  $\text{ZnAl}_2\text{O}_3$  nanocomposite was performed as shown in Fig. 4. Strong bond region  $450\text{--}950\text{ cm}^{-1}$  in the FTIR graph was ascribed to the inorganic network (Zn–O and Al–O) [32]. Small peak observed at  $682$  and  $778\text{ cm}^{-1}$  correlated to hexagonal phase of ZnO [33]. Peaks at  $1120$ ,  $1409$  and  $1566\text{ cm}^{-1}$  infer the C=C bond. Peak at  $1641\text{ cm}^{-1}$  correspond to the C=O bond attached to metallic ions. Strong peak at  $3446\text{ cm}^{-1}$  ascribed the presence of the H–O stretching vibrational group in the as-prepared nanocomposite [34–36].

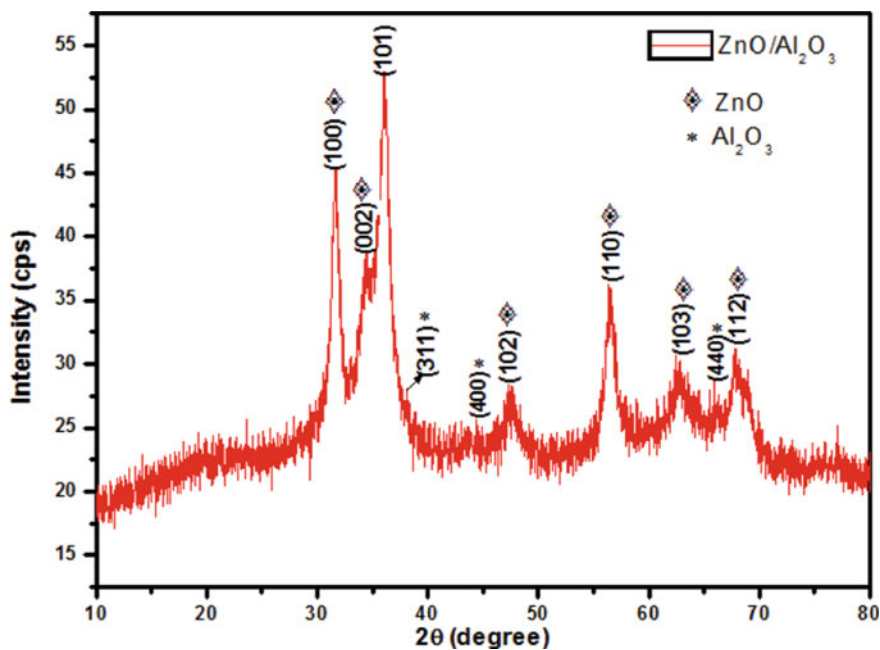
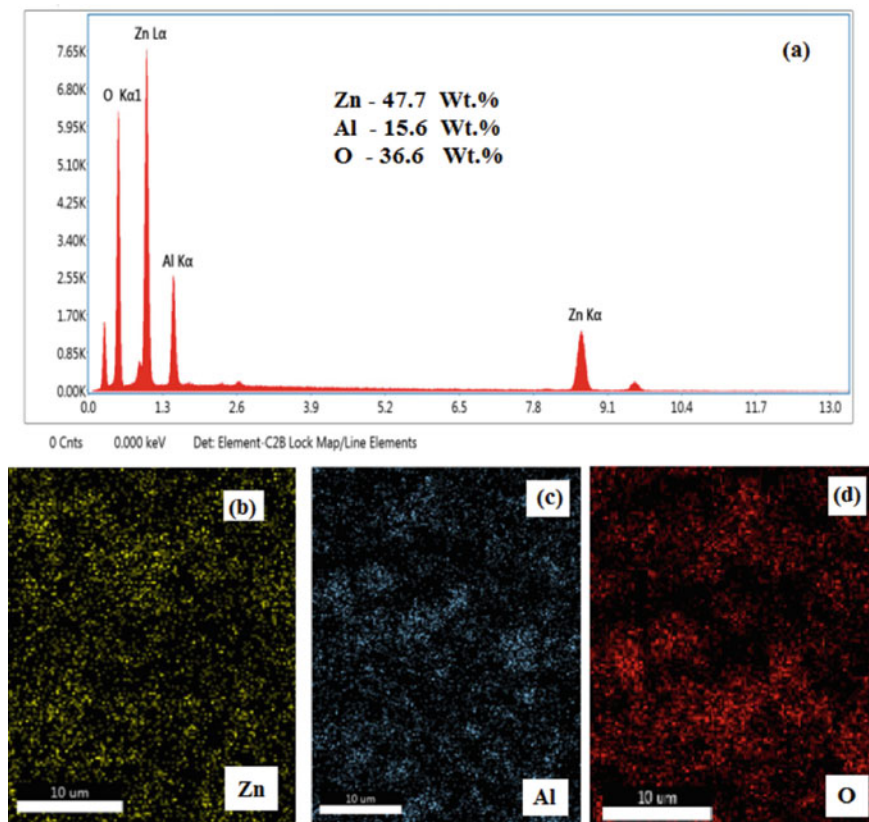


Fig. 2 XRD pattern of  $\text{ZnAl}_2\text{O}_3$  Nanocomposite



**Fig. 3** EDAX spectrum (a) and elemental mapping (b–d) of ZnAl<sub>2</sub>O<sub>3</sub> Nanocomposite

Surface morphology analyzed by SEM is shown in Fig. 5. High magnification images showed the flower shape with a small pedal. Elemental analysis of synthesized composites was examined using EDAX spectrum which inveterate the presence of Zn, Al and O. Atomic weight percentages were in the order of 47.7 wt.% (Zn), 15.6 wt.% (Al) and 36.6 Wt.% (O).

Thermal performance of PSS with SSP is examined without and with ZnAl<sub>2</sub>O<sub>3</sub> Nanocomposites.

Instantaneous efficiency (IE) observed is around 2.4–9.92% and 2.42–10.79% for the solar still performance study without and with ZnO/Al<sub>2</sub>O<sub>3</sub> nanocomposites. Performance Ratio (PR) obtained shows increasing trend with time. PR is found to be in the range of 2.11–10.60% and 2.38–13.08% for solar still study without and with ZnO/Al<sub>2</sub>O<sub>3</sub> nanocomposites. From the plot, it is inferred performance ratio sustains steady state as the basin water temperature rises to the maximum value. Plot of IE and PR is shown in Fig. 6.

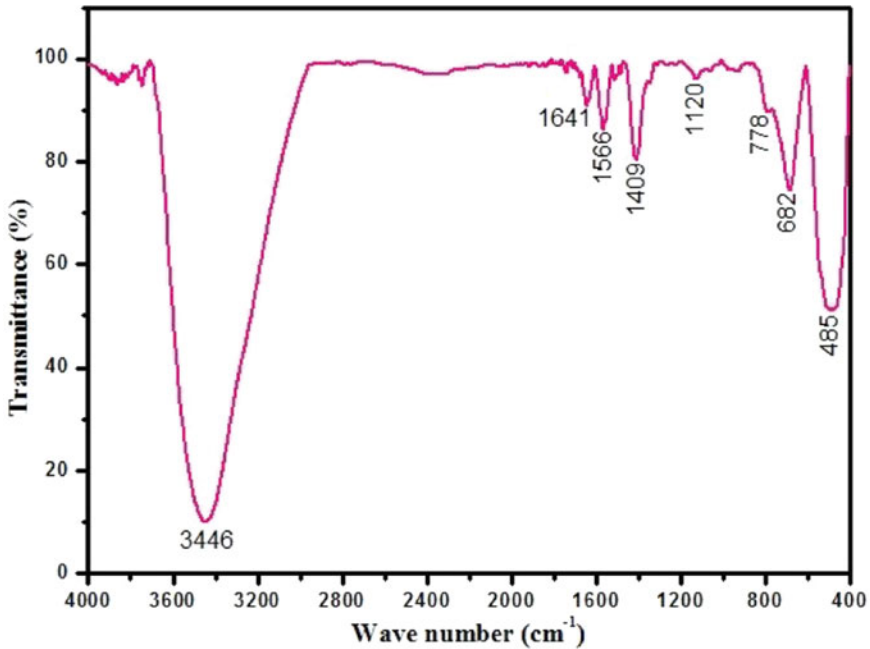


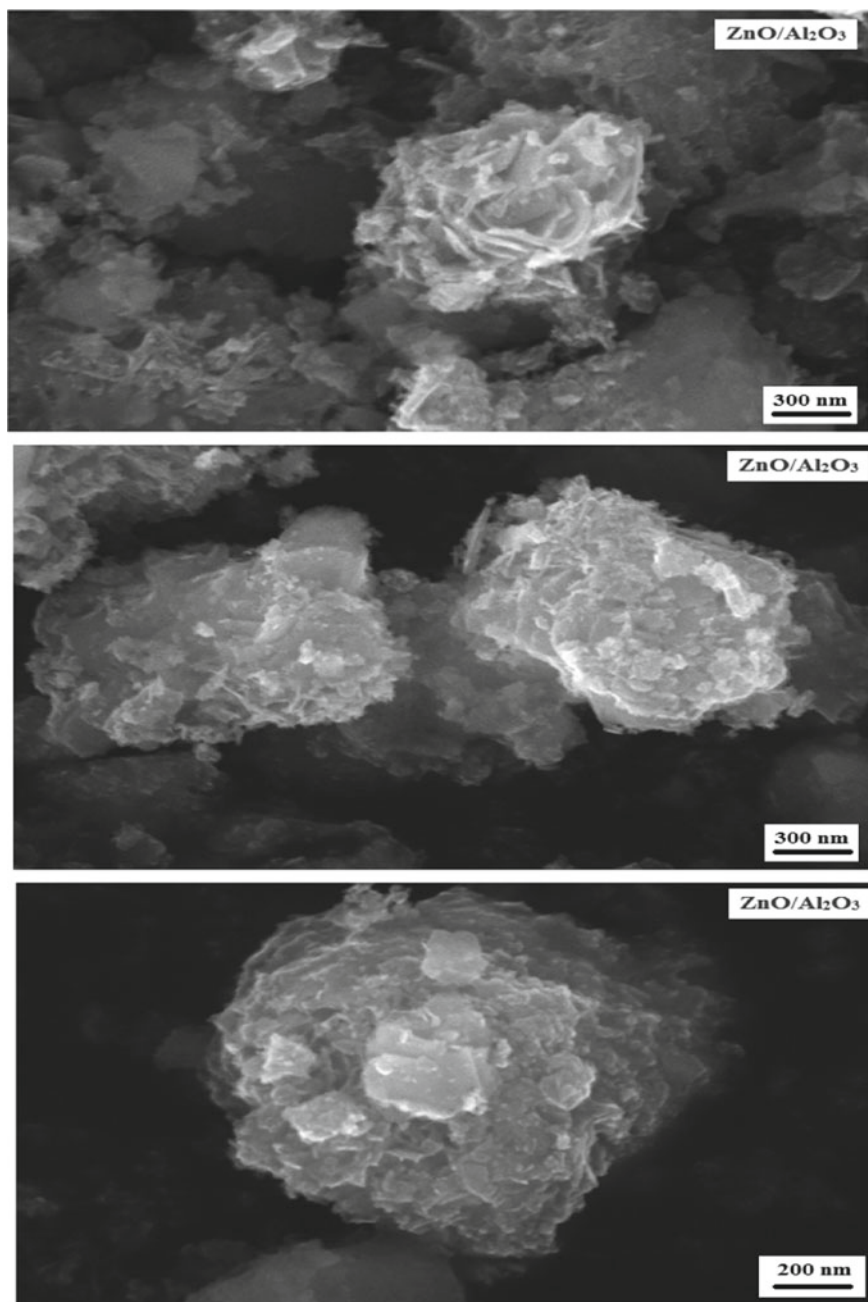
Fig. 4 FTIR Spectrum of ZnAl<sub>2</sub>O<sub>3</sub> Nanocomposite

Figure 7 shows the plot of solar radiation and water collection for PSS with SSP coupled without and with ZnO/Al<sub>2</sub>O<sub>3</sub> nanocomposites. Solar insolation observed is around 96.16–1050.7 W/m<sup>2</sup> and 46.61–1038.62 W/m<sup>2</sup> still analysis without and with ZnO/Al<sub>2</sub>O<sub>3</sub> nanocomposites.

Yielded distillate is found to be 0.02–0.048 kg/0.25 m<sup>2</sup> and 0.014–0.056 kg/0.25 m<sup>2</sup> for still without and with ZnO/Al<sub>2</sub>O<sub>3</sub> nanocomposites. Water collection tends to increase after attaining the steady state. Distillate yield is upheld till late hours as a result of the coupling effect of the shallow solar pond and ZnO/Al<sub>2</sub>O<sub>3</sub> nanocomposites in the combined performance study.

Observed values of SVP and LH under two modes of studies is shown in Fig. 8. SVP reaches a maximum value as the distillate yield reaches maximum and starts decline as the distillate is minimized. SVP predicted is found to be 4842.65 to 11,889.52 Pa and 5121.7 to 13,281.98 Pa for solar still study without and with ZnO/Al<sub>2</sub>O<sub>3</sub> nanocomposites. Due to the heat retention capacity of ZnO/Al<sub>2</sub>O<sub>3</sub> nanocomposites and shallow solar pond, SVP is sustained even at off daylight hours.

Latent heat detected is around 2,416,722.46–2,375,960.52 kg<sup>-1</sup> and 2,414,341.48–2,370,510.43 kg<sup>-1</sup> solar still study without and with ZnO/Al<sub>2</sub>O<sub>3</sub> nanocomposites. Results shows that latent heat reaches minimum value at higher temperature (Tables 1 and 2).



**Fig. 5** SEM Images of ZnAl<sub>2</sub>O<sub>3</sub> nanocomposite



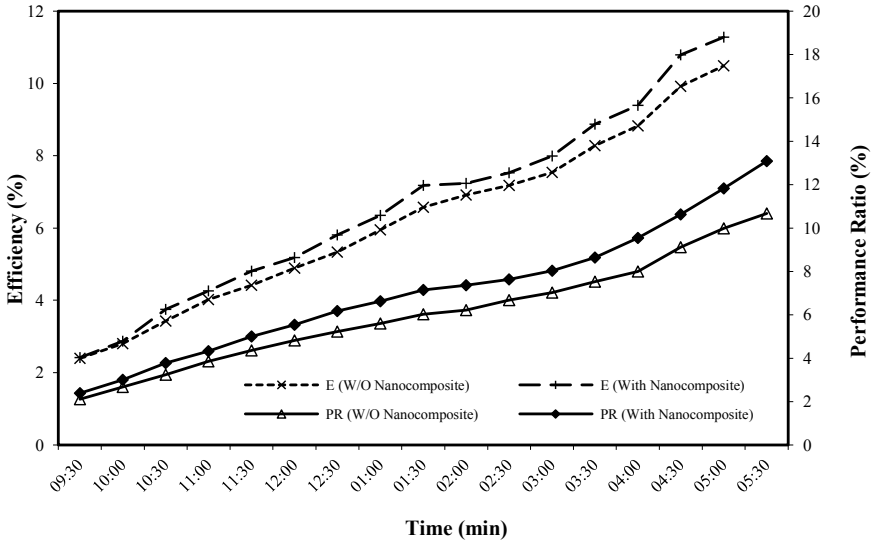


Fig. 6 Instantaneous efficiency and performance ratio analysis for two modes of study

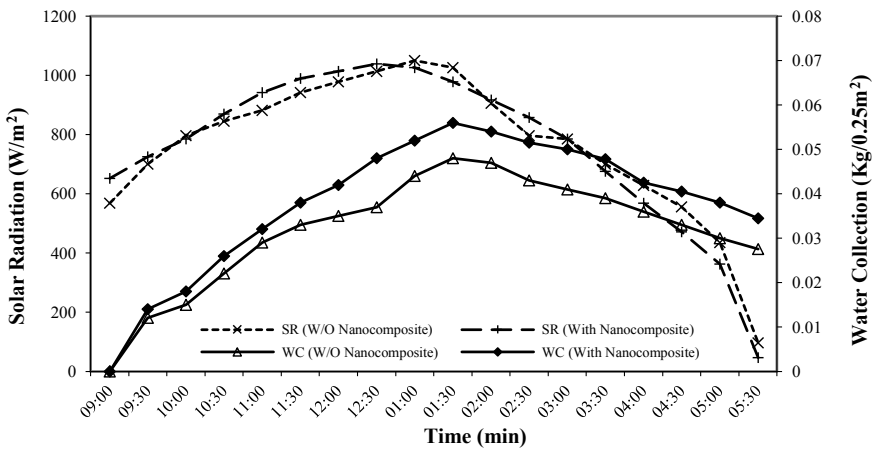


Fig. 7 Solar radiation and water collection with time for PSS with SSP under two modes of study

From the tabulation of thermophysical properties, it is determined that role of water temperature is vital. Water temperature is enhanced due to incorporating the solar radiation absorption effect of ZnO/Al<sub>2</sub>O<sub>3</sub> nanocomposites in solar still and SSP. It also concludes that the density of the distillate yield decreases as there is rise in water temperature and vice versa.

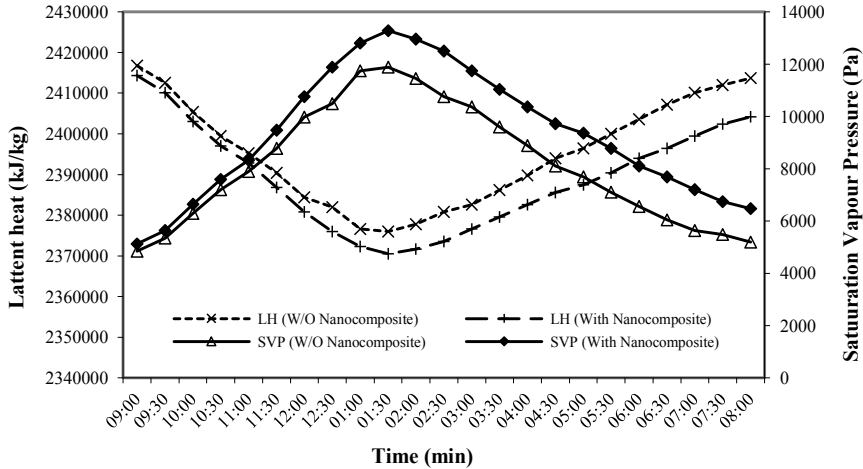


Fig. 8 Variation of SVP and LH with time for PSS with SSP under two modes of study

### 7 Conclusion

Results emphasize that distillate yield of PSS coupled with SSP is enhanced by the substitution of ZnO/Al<sub>2</sub>O<sub>3</sub> nanocomposites. Heat absorbing capacity of ZnO/Al<sub>2</sub>O<sub>3</sub> nanocomposites escalates in alleviating water temperature of PSS with SSP, which results in productive yield at later hours after the fall of solar insolation. It is found that the productivity yield is increased to 23% than that without the shallow solar pond.

**Table 1** Heat transfer coefficient

	$Q_{ci}$ (W/m <sup>2</sup> )	$Q_{ri}$ (W/m <sup>2</sup> )	$Q_{ei}$ (W/m <sup>2</sup> )	$Q_{ce}$ (W/m <sup>2</sup> )	$Q_{re}$ (W/m <sup>2</sup> )	$Q_{be}$ (W/m <sup>2</sup> )
PSS with SSP	17.22	51.52	127.47	37.11	78.93	4.65
ZnO/Al <sub>2</sub> O <sub>3</sub> Nanocomposites combined with Pyramid Solar Still and SSP	24.22	72.83	223.74	43.27	86.14	6.26

**Table 2** Thermophysical properties

	Pyramid solar still with SSP	ZnO/Al <sub>2</sub> O <sub>3</sub> nanocomposites combined with pyramid solar still and SSP
Thermal conductivity	$26.887 \times 10^{-3} \text{ W m}^{-2} \text{ }^{\circ}\text{C}^{-1}$ to $28.186 \times 10^{-3} \text{ W m}^{-2} \text{ }^{\circ}\text{C}^{-1}$	$26.968 \times 10^{-3} \text{ Wm}^{-2} \text{ }^{\circ}\text{C}^{-1}$ to $28.324 \times 10^{-3} \text{ Wm}^{-2} \text{ }^{\circ}\text{C}^{-1}$
Dynamic viscosity	$18.678 \times 10^{-6} \text{ Nsm}^{-2}$ to $19.458 \times 10^{-6} \text{ Nsm}^{-2}$	$18.738 \times 10^{-6} \text{ Nsm}^{-2}$ to $19.566 \times 10^{-6} \text{ Nsm}^{-2}$
Density	$11.552 \times 10^{-1} \text{ kgm}^{-3}$ to $10.948 \times 10^{-1} \text{ kgm}^{-3}$	$11.512 \times 10^{-1} \text{ kg m}^{-3}$ to $10.874 \times 10^{-1} \text{ kg m}^{-3}$

## References

1. Hamdan MA, Musa AM, Jubran BA (1999) Performance of solar still under Jordanian climate. *Energy Convers Mgmt* 40:495–503
2. Selvakuma B, Sanjaykumar, Jayaprakash R (2008) Performance analysis of single sloped green house solar still. *J Environ Res Dev* 2618–630
3. Tiwari GN (1985) Enhancement of daily yield in a double basin solar still. *Energy Convers Mgmt* 25:49–50
4. Raj SN, Tiwari GN (1983) Single basin solar still coupled with flat plate collector. *Energy Convers Mgmt* 23:145–149
5. Selvakumar B, Sanjaykumar, Jayaprakash R (2008) Performance analysis of “V” type solar still using charcoal absorber with boosting mirror. *Desalination* 229:217–30
6. Jairaj N, Selvakumar B, Jayaprakash R (2004) Design and thermal performance of four sloped glass surface basin type solar still. *J Arid Land Stud* 14S:179–182
7. Mathioulakis E, Voropoulos K, Belessiotis V (1999) Modelling and prediction of long term performance of solar stills. *Desalination* 122:85–93
8. Tripathi R, Tiwari GN (2004) Performance evaluation of a solar still by using the concept of solar fraction. *Desalination* 169:69–80
9. Singh HN, Tiwari GN (2004) Monthly performance of passive and active solar stills for different indian climatic conditions. *Desalination* 168:145–150
10. Sanjaykumar, Tiwari GN (1996) Estimation of convective mass transfer in solar distillation systems. *Sol Energy* 57:459–464
11. Casamajor AB (1999) Application of shallow solar ponds for industrial process heat: case histories. Technical Report; University of California: Report No. UCRL-82367
12. Dickinson WC, Neifert RDC (1975) Guide for calculation of collection efficiency for the shallow solar pond. Technical Report; University of California: Report No. UCRL-16446
13. Kudish AI, Wolf DA (1978) Compact shallow solar pond hot water heater. *Sol Energy* 21:317–322
14. Garg HP (1987) *Advances in solar energy technology*, vol I. D. Reidel Publishing Company, Dordrecht
15. Varghese JGK, Awari JK (2002) Distinct design of a batch type solar water heater: a detailed analysis. In ICRASECS, Maulana Azad Institute of Technology, India
16. Kishore VVN, Gandhi MR, Rao KS (1986) Experimental and analytical studies of shallow solar pond systems with continuous heat extraction. *Sol Energy* 36:245–256
17. Casamajor AB, Parsons RE (1975) Design guide for shallow solar ponds. Technical Report. University of California, Report No. UCRL-52385
18. Bansal PK, Hrishikesan DS, Garg HP (1984) Effect of heat exchanger on the performance of a shallow solar pond water heater. *Energy Convers Mgmt* 24:259–263
19. Sabetta F, Pacetti M, Princioi P (1985) An internal heat extraction system for solar pond. *Solar Energy* 34:297–302

20. Chen JC, Tang CT (2007) Preparation and application of granular ZnO/Al<sub>2</sub>O<sub>3</sub> catalyst for the removal of hazardous trichloroethylene. *J Hazard Mater* 142:88–96
21. Quynh NTN, Trang NTT, Thuy CT, Tai PA, Duc BM (2017) Synthesis of composites ZnO-Al<sub>2</sub>O<sub>3</sub>, ZnO-Al<sub>2</sub>O<sub>3</sub>-CuO and evaluation of its activity for COS conversion H<sub>2</sub>S sorption treatment. *Int J Chem Sci* 15:1–8
22. Yadav S, Mittal A, Sharma S, Kumari K, Chauhan NS, Kumar N (2020) Low temperature synthesized ZnO/Al<sub>2</sub>O<sub>3</sub> nano-composites for photocatalytic and antibacterial applications. *Semicond Sci Technol* 35:055008
23. Jensen JR, Johannessen T, Wedel S, Livbjerg H (2000) Preparation of ZnO–Al<sub>2</sub>O<sub>3</sub> particles in a premixed flame. *J Nanopart Res* 2:363–373
24. Tayoma S, Aragoki T, Murase K, Sumura T (1983) Simulation of a multi-effect solar distillers. *Desalination* 45:101–110
25. Brooker DB, Bakker-Arkma FW, Hall CW (1978) Drying cereal grain. AVI West Port, U.S.A
26. Dunkle RV (1961) Solar water distillation: roof type still and a multiple effect diffusion still. In: International development in heat transfer, A.S.M.E., Proc. International Heat Transfer, Part V, University of Colorado, p 895
27. Adhikari S, Kumar A, Sootha GD (1995) Simulation studies on a multi-stage staked solar still. *Solar Energy* 54:317–325
28. Ottone, C., Farías Rivera, V., Fontana, M., Bejtka, K., Onida, B., and Cauda, V.: Ultralong and Mesoporous ZnO and  $\gamma$ -Al<sub>2</sub>O<sub>3</sub> oriented nanowires obtained by template-assisted hydrothermal approach. *J. Mater. Sci. Technol.* 30 (2014) 1167–1173.
29. Yadav S, Mittal A, Sharma S, Kumari K, Chauhan NS, Kumar N (2002) Low temperature synthesized ZnO/Al<sub>2</sub>O<sub>3</sub> nano-composites for photocatalytic and antibacterial applications. *Semicond Sci Technol* 35:055008
30. He B, Fu X, Lian X, Jiang S, Xu P, Deng X, He C, Chen C (2019) Catalytic deacidification of vacuum gas oil by ZnO/Al<sub>2</sub>O<sub>3</sub> and its modification with Fe<sub>2</sub>O<sub>3</sub>. *Catalysts* 9:499
31. Romcevic N, Hadzic B, Romcevic M, Paunovic N, Sibera D, Narkiewicz U, Kuryliszyn-Kudelska I, Ristic-Djurovic JL, Dobrowolski WD (2020) Structural and optical properties of ZnO–Al<sub>2</sub>O<sub>3</sub> nanopowders prepared by chemical methods. *J Luminescence* 224:117273
32. Barroso MN, Gomez MF, Gamboa JA, Arrúa LA, Abello MC (2006) Preparation and characterization of CuZnAl catalysts by citrate gel process. *J Phys Chem Solids* 67:583–1589
33. Muhammad W, Ullah N, Haroon M, Abbasi BH (2019) Optical, morphological and biological analysis of zinc oxide nanoparticles (ZnO NPs) using Papaver somniferum. *L RSC Adv* 9:29541–29548
34. Dehghani A, Ranjbar M, Eliassi A (2018) Modification of Cu/Zn/Al<sub>2</sub>O<sub>3</sub> catalyst by activated carbon-based metal organic frameworks as precursor for hydrogen production. *J Inorg Organomet Polym Mater* 28:585–593
35. Mohammed HA, Rasool RA (2019) Structural Properties of zinc aluminate ZnAl<sub>2</sub>O<sub>4</sub> compound. *Rafidain J Sci* 28(2019):33–39
36. Sentosa JM, Karubaba JJ (2018) The combined process of pyrolysis and catalytic conversion from rice straw toward light olefin hydrocarbon with supported metal catalyst. In: E3S web of conferences, vol 67, p 02025

# Correction to: Design of High Speed Approximate Carry Select Adders Using RCPFA



T. Phaneendra, A. Anitha, and Rajasekhar Turaka

**Correction to:**  
**Chapter “Design of High Speed Approximate Carry Select Adders Using RCPFA” in: T. Sengodan et al. (eds.),**  
***Advances in Electrical and Computer Technologies, Lecture Notes in Electrical Engineering 881,***  
**[https://doi.org/10.1007/978-981-19-1111-8\\_58](https://doi.org/10.1007/978-981-19-1111-8_58)**

The original version of chapter “Design of High Speed Approximate Carry Select Adders Using RCPFA” was inadvertently published with wrong affiliation for the authors “A. Anitha” and “Rajasekhar Turaka”.

The affiliation for author “A. Anitha” has been changed from “ECE, Nalla Narasimha Reddy Education Society’s Group of Institutions, Hyderabad, Telangana, India” to “ECE, Velagapudi Ramakrishna Siddhartha Engineering College, Vijayawada, Andhra Pradesh, India”.

The affiliation for author “Rajasekhar Turaka” has been changed from “ECE, Velagapudi Ramakrishna Siddhartha Engineering College, Vijayawada, Andhra Pradesh, India” to “ECE, Nalla Narasimha Reddy Education Society’s Group of Institutions, Hyderabad, Telangana, India”.

The chapter and book have been updated with the changes.

---

The updated version of this chapter can be found at  
[https://doi.org/10.1007/978-981-19-1111-8\\_58](https://doi.org/10.1007/978-981-19-1111-8_58)