Lecture Notes in Networks and Systems 196

Santosh Kumar Das · Sourav Samanta ·
Nilanjan Dey · Bharat S. Patel ·
Aboul Ella Hassanien *Editors*

Architectural Wireless Networks Solutions and Security Issues

Springer

Lecture Notes in Networks and Systems

Volume 196

The series “Lecture Notes in Networks and Systems” publishes the latest developments in Networks and Systems—quickly, informally and with high quality. Original research reported in proceedings and post-proceedings represents the core of LNNS.

Volumes published in LNNS embrace all aspects and subfields of, as well as new challenges in, Networks and Systems.

The series contains proceedings and edited volumes in systems and networks, spanning the areas of Cyber-Physical Systems, Autonomous Systems, Sensor Networks, Control Systems, Energy Systems, Automotive Systems, Biological Systems, Vehicular Networking and Connected Vehicles, Aerospace Systems, Automation, Manufacturing, Smart Grids, Nonlinear Systems, Power Systems, Robotics, Social Systems, Economic Systems and other. Of particular value to both the contributors and the readership are the short publication timeframe and the world-wide distribution and exposure which enable both a wide and rapid dissemination of research output.

The series covers the theory, applications, and perspectives on the state of the art and future developments relevant to systems and networks, decision making, control, complex processes and related areas, as embedded in the fields of interdisciplinary and applied sciences, engineering, computer science, physics, economics, social, and life sciences, as well as the paradigms and methodologies behind them.

Indexed by SCOPUS, INSPEC, WTI Frankfurt eG, zbMATH, SCImago.

All books published in the series are submitted for consideration in Web of Science.

More information about this series at http://www.springer.com/series/15179

Santosh Kumar Das · Sourav Samanta · Nilanjan Dey · Bharat S. Patel · Aboul Ella Hassanien
Editors

Architectural Wireless Networks Solutions and Security Issues

Springer

Editors

Santosh Kumar Das
Department of Computer Science
and Engineering
Sarala Birla University
Birla Knowledge City
Ranchi, Jharkhand, India

Nilanjan Dey
Department of Computer Science
and Engineering
JIS University
Kolkata, India

Aboul Ella Hassanien
Department of Information Technology
Cairo University
Giza, Egypt

Sourav Samanta
Department of Computer Science
and Engineering
University Institute of Technology
Burdwan University
Burdwan, West Bengal, India

Bharat S. Patel
Yudiz Solutions Pvt. Ltd.
Ahmedabad, Gujarat, India

ISSN 2367-3370 ISSN 2367-3389 (electronic)
Lecture Notes in Networks and Systems
ISBN 978-981-16-0385-3 ISBN 978-981-16-0386-0 (eBook)
https://doi.org/10.1007/978-981-16-0386-0

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

Preface

In the last few decades, the application of wireless network increased rapidly along with its several variations based on diverse applications of the users and customers. Its main reason is flexibility and efficiency of the wireless network which is not available in the wired network. So, it brings a large number of jobs, applications, and opportunities for the students as well as customers. Although, the wireless network is an efficient and robustness platform for communication and data transmission, it has also some challenges and security issues in terms of several applications. Some of them are mentioned as limited hardware resources, unreliable communication, the dynamic topology of some wireless networks, vulnerability, unsecure environment, etc. Hence, it causes several types of attacks, data loss, replication, eavesdropping, overflow, etc., with respect to the architecture design of the wireless network. These issues cannot be controlled and managed directly, but it can model and reduce as an architectural solution. Therefore, to enhance the architecture model of the wireless network and enhance the security mechanism, some innovative as well as novel ideas are needed that reflected in this book.

Objective of the Book

This book contains some architectural solutions of wireless network and its variations. It deals with modeling, analysis, design, and enhancement of different architectural parts of the wireless network. The main aim of this book is to enhance the applications of the wireless network by reducing and controlling its architectural issues. This book is edited for wireless network's users, academicians, and researchers.

Organization of the Book

The book contains 17 chapters that are organized in four parts as follows. Before starting the parts, Chap. "Wireless Networks: Applications, Challenges and Security Issues" describes the overview of wireless network and its variation along with its several applications, challenges, and security issues. **Part One** contains four chapters that outline the modeling of some security issues with their solutions for enhancing the security part of the wireless network. **Part Two** contains four chapters that highlight some optimization models of the wireless network for enhancing the network lifetime. **Part Three** contains four chapters that outline the modeling of the aggregation system to control redundant information. **Part Four** contains four chapters that highlight some troubleshooting techniques that help to control and manage different issues of the network.

Part One: Modelling of Security Enhancements (Chaps. "An Authentication Model with High Security for Cloud Database"–"Linear Secret Sharing-Based Key Transfer Protocol for Group Communication in Wireless Sensor Communication")

This part outlines some security issues along with their solutions in the wireless network and its variations as wireless sensor network and cloud-based network. Short descriptions of these chapters are as follows.

Chapter "An Authentication Model with High Security for Cloud Database"

This chapter outlines an authenticate model that handles security and privacy problems of the cloud-based database. It helps to reduce malicious issues of the network and provides traceable services to the users. Finally, it helps to find an appropriate solution for the security issues at both administrator and customer levels in various directions.

Chapter "Design of Robust Smartcard-Based User Anonymous Authentication Protocol with AVISPA Simulation"

In this chapter, the author designed an extended user anonymous authenticated session key agreement protocol using a smartcard. The scalability of this scheme is measured in both formal and informal ways. Informal security analysis ensures that the proposed scheme resists to various kinds of fraudulent attacks. The proposed scheme does not only hold up security attacks, but also achieves some security features.

Chapter "Data Security in Cloud Computing Using Abe-Based Access Control"

This chapter discusses the dynamic access control model with the fusion of risk aware and hierarchical attribute set-based encryption. The combination of both methods provides a scalable and flexible services due to sub-domain hierarchy. It is also proved to be dynamic by permitting the user to access the data by risk evaluation using risk engine.

Chapter "Linear Secret Sharing-Based Key Transfer Protocol for Group Communication in Wireless Sensor Communication"

In this chapter, an intelligent protocol is proposed with the fusion of linear secret sharing and elliptic curve techniques. The combination of both techniques helps to overcome the drawback of traditional protocols. The proposed security protocol helps to reduce the overhead of the network and enhance the several security mechanisms against different conflicting attacks.

Part Two: Optimization Model for Network Lifetime (Chaps. "Fuzzy Rule-Based System for Route Selection in WSN Using Quadratic Programming–"Fuzzy Q-Learning-Based Controller for Cost- and Energy-Efficient Load Balancing in Cloud Data Center")

This part outlines some optimization models for enhancing the network lifetime of the wireless network or some variation of the wireless network by reducing uncertainty information and managing conflicting parameters of the networks. Short descriptions of these chapters are as follows.

Chapter "Fuzzy Rule-Based System for Route Selection in WSN Using Quadratic Programming"

In this chapter, a combination of intelligent technique as well as mathematical modeling is used where fuzzy logic as an intelligent technique and quadratic programming as mathematical modeling are used for solving the proposed goal. The combination of both provides a robustness technique that uses two basic parameters energy and distance for selecting the optimal route of the network.

Chapter "Wireless Sensor Network Routing Protocols Using Machine Learning"

This chapter consists of some machine learning algorithms to optimize the route of the wireless sensor network. This optimization helps the sensor nodes to learn the experience data to make appropriate routing decisions and respond to the changing environment using some learning techniques such as distributed regression, self-organizing map, and reinforcement learning.

Chapter "Distributed Traversal Based Fault Diagnosis for Wireless Sensor Network"

In this chapter, the author proposed a traversal-based diagnosis algorithm that seeks to diagnose both permanent as well as an intermittent fault in a sensor network. The proposed algorithm employs a special node called an anchor node to traverse the field. So, it is decided by a proposed traversal algorithm taking into consideration the length and breadth of the sensor field and the transmission range of the nodes. The anchor node stops at defined positions in the deployment field where it executes the fault diagnosis algorithm taking into consideration the normal sensor nodes which are in its range.

Chapter "Fuzzy Q-Learning-Based Controller for Cost and Energy Efficient Load Balancing in Cloud Data Center"

In this chapter, the author proposed a fuzzy Q-learning-based self-learning controller to optimize the load for a specific data center. The proposed method also helps to reduce uncertainty and solve the congestion issue efficiently through fuzzy linguistic behavior and membership function. In this proposal, the fuzzy output parameter is considered as reward value which is used to learn and update the state for each data center.

Part Three: Modelling of Aggregation Systems (Chaps. "Localization Techniques Using Machine Learning Algorithms"–"Analysis of Network Parameters for Network Lifetime in WSN: A Fuzzy Quadratic Programming Approach")

This part outlines some aggregation techniques that help to model several issues of the network and reduce redundancy of the wireless network efficiently. Short descriptions of these chapters are as follows.

Chapter "Localization Techniques Using Machine Learning Algorithms"

In this chapter, the author illustrates how the localization issue in wireless sensor networks can be solved using the three categorized machine learning algorithms such as supervised learning, unsupervised learning, and reinforcement learning algorithms. It also highlights that which machine learning algorithms conjointly evokes several sensible solutions for localization of nodes that maximize resource utilization and prolong the lifetime of the network.

Chapter "Vehicular Delay Tolerant Network-Based Communication Using Machine Learning Classifiers"

In this chapter, the authors highlight vehicular delay-tolerant network-based communication using machine learning classifiers. First the authors analyzed which machine learning classifier is the best solution for our problem. In this work, the authors used machine learning classifiers for filtering efficient vehicular nodes, so that packets can be delivered from source to destination.

Chapter "Applications of Big Data and Internet of Things in Power System"

This chapter highlights the use of big data and IoT for the power systems. IoT can be used in various areas of power system such as metering, transformer monitoring, prediction of demand, and planning for future consumption. The main objective of this chapter is to make a clear understanding of the use of big data and IoT in the power system and how it will improve customer service and social welfare.

Chapter "Analysis of Network Parameters for Network Lifetime in WSN: A Fuzzy Quadratic Programming Approach"

In this paper, a fuzzy quadratic programming is used to optimize network parameters efficiently. It is the fusion of fuzzy logic and quadratic programming. Fuzzy logic is a multi-values logic which is used to reduce uncertainty and estimate imprecise parameters efficiently. Quadratic programming is a nonlinear programming based on second order of mathematical polynomial for reducing the main objective. The combination of both helps to analyze conflicting network parameters and decide the optimal objective value along with constraints.

Part Four: Analyzing of Troubleshooting Techniques (Chaps. "IDS Detection Based on Optimization Based on WI-CS and GNN Algorithm in SCADA Network"–"Investigation of Memory, Nonlinearity and Chaos in Worldwide Monthly Mobile Data Traffic in Smartphones")

This part outlines different troubles in the wireless network in terms of intrusion, attack, and chaos and also provide their modeling methods. Short descriptions of these chapters are as follows.

Chapter "IDS Detection Based on Optimization Based on WI-CS and GNN Algorithm in SCADA Network"

In this chapter, it is identify and categorize the anomalies in a SCADA system through data optimization. At the initial stage, the collected real-time SCADA dataset is given as input. Then by using the aforementioned proposed machine learning algorithms, these data are clustered and optimized. Later to find the type of intrusion will remain as a further challenge, and for that, the authors proposed HNA-AA algorithm.

Chapter "Performance Analysis of MANET Under Grayhole Attack Using AODV Protocol"

In this chapter, the author analyzed the performance of the mobile ad-hoc network under grayhole attack as per AODV routing protocol using NS-2 simulation environment. Several attacks make the network pretty much risky to rely upon when scaling up on a large scale. Under the mobile ad-hoc network, all the transmissions between the mobile nodes occur wirelessly.

Chapter "Technique to Reduce PAPR Problem in Next-Generation Wireless Communication System"

In this chapter, a technique is design for reducing PAPR in next-generation wireless communication system. The main effect of strong PAPR is instability in the analog-to-digital converter and digital-to-analog converter, decreased its performance and raised costs. A PAPR reduction technique such as clipping and filtering greatly improves the efficiency compared to the initial GFDM signal PAPR.

Chapter "Investigation of Memory, Nonlinearity and Chaos in Worldwide Monthly Mobile Data Traffic in Smartphones"

In this chapter, the proposed chapter employs certain statistical signal processing techniques to realize the memory, self-similarity, self-organized criticality, nonlinearity, and chaos in the present time series of worldwide monthly mobile data traffic per smartphone. This study possibly indicates a persistent, self-similar, deterministic, nonlinear, and non-chaotic profile with no "soc" for the present time series.

Santosh Kumar Das
Department of Computer Science
and Engineering
Sarala Birla University
Birla Knowledge City
Ranchi, Jharkhand, India

Sourav Samanta
Department of Computer Science
and Engineering
University Institute of Technology
Burdwan University
Burdwan, West Bengal, India

Nilanjan Dey
Department of Computer Science
and Engineering
JIS University
Kolkata, India

Bharat S. Patel
Director and COO at Yudiz Solutions Pvt. Ltd.
India
Ahmedabad, India

Aboul Ella Hassanien
Founder and Head of the Egyptian Scientific
Research Group (SRGE)
Professor of Information Technology at
the Faculty of Computer and Artificial
Intelligence
Cairo University
Giza, Egypt

List of Reviewers

Abhishek Kumar, Swami Vivekananda Subharti University, Meerut
Amit Kumar Singh, Indian Institute of Technology (ISM), Dhanbad
Amitesh Kumar Pandit, Dr. Rammanohar Lohia Avadh University, Ayodhya, Uttar Pradesh
Arun Prasad Burnwal, GGSESTC, Bokaro, Jharkhand
Ashish Kumar Dass, National Institute of Science and Technology, Brahmapur, Odisha
Chandan Kumar Shiva, S. R. Engineering College, Ananthsagar, Hasanparthy, Warangal, Telangana
Harsh Nath Jha, Asansol Engineering College, Asansol, West Bengal
Jayraj Singh, Indian Institute of Technology (ISM), Dhanbad
Jeevan Kumar, R. V. S. College of Engineering and Technology, Jamshedpur
Kanhu Charan Gouda, Indian Institute of Technology, Roorkee
Mahendra Prasad, Indian Institute of Technology (ISM), Dhanbad
Manoj Kumar Mandal, Jharkhand Rai University, Ranchi
Mukul Majhi, Indian Institute of Technology (ISM), Dhanbad
Nabajyoti Mazumdar, Central Institute of Technology, Kokrajhar, Assam
Priyanka Jaiswal, Indian Institute of Technology (ISM), Dhanbad
Rakesh Ranjan Swain, National Institute of Science and Technology, Brahmapur, Odisha
Ruchika Padhi, National Institute of Science and Technology, Brahmapur, Odisha
Sagar Samal, National Institute of Science and Technology, Brahmapur, Odisha
Shalini Mahato, B.I.T, Mesra, Ranchi, Jharkhand
Siba Prasada Tripathy, National Institute of Science and Technology, Brahmapur, Odisha
Smita Rani Sahu, BPUT, Odisha
Sourav Samanta, University Institute of Technology, BU, Burdwan, West Bengal
Subhra Priyadarshini Biswal, National Institute of Science and Technology, Brahmapur, Odisha
Sunil Gautam, Institute of Advanced Research, Gandhinagar
Vishal Maheswari, RIT, Roorkee, Uttarakhand

Contents

Modelling of Security Enhancements

Optimization Model for Network Lifetime

Modelling of Aggregation Systems

Analysing of Troubleshooting Techniques

Editors and Contributors

About the Editors

Santosh Kumar Das received his Ph.D. degree in Computer Science and Engineering from Indian Institute of Technology (ISM), Dhanbad, India, in 2018 and completed his M.Tech. degree in Computer Science and Engineering from Maulana Abul Kalam Azad University of Technology (erstwhile WBUT), West Bengal, India, in 2013. He has about to three years teaching experience as Assistant Professor at School of Computer Science and Engineering, National Institute of Science and Technology (Autonomous), Institute Park, Pallur Hills, Berhampur, Odisha, India. He is currently working as Assistant Professor at Department of Computer Science and Engineering, Sarala Birla University, Birla Knowledge City, P.O.-Mahilong, Purulia Road, Ranchi, India. He has more than eight years teaching experience. He has authored/edited of five books with Springer in series as Lecture Notes in Networks and Systems, Tracts in Nature-Inspired Computing and Studies in Computational Intelligence. He has contributed more than 30 research papers. His research interests mainly focus on Ad-hoc & Sensor Network, Artificial Intelligence, Soft Computing, and Mathematical modelling. His h-index is 15 with more than 600 citations.

Sourav Samanta is currently working as Assistant Professor in the Department of Computer Science and Engineering at University Institute of Technology, The University of Burdwan, West Bengal, India. He has completed M.Tech. in Computer Science and Engineering from JIS College of Engineering, WBUT. He was Honorary Visiting Scientist at Global Biomedical Technologies Inc., CA, USA (2014–2015). He has published about 45 research papers in various reputed international journals and conference proceedings including five book chapters in books published by Elsevier and Springer, respectively. He is Co-editor of the book Design Frameworks for Wireless Networks published by Springer in Lecture Notes in Networks and Systems Series. He is a regular reviewer of *IEEE Access*, *IEEE Sensor Journals* and other various international journals. He serves as a Program/Technical Committee member for AISI2015, ICMCTI-2017, A2ICS-2017 and PerCAA-2019 International Conferences. He is a member of Computer Society of India, Institution of Engineers (India), Soft Computing Research Society and International Association of Engineers. His research area includes bio-inspired computing, quantum machine learning and information security. He has an interest in interdisciplinary research.

Nilanjan Dey is an Associate Professor, Department of Computer Science and Engineering, JIS University, Kolkata, India. He is a visiting fellow of the University of Reading, UK. He was an honorary Visiting Scientist at Global Biomedical Technologies Inc., CA, USA (2012–2015). He was awarded his Ph.D. from Jadavpur Univeristy in 2015. He has authored/edited more than 70 books with Elsevier, Wiley, CRC Press and Springer, and published more than 300 papers. He is the Editor-in-Chief of *International Journal of Ambient Computing and Intelligence*, *IGI Global*, Associated Editor of *IEEE Access* and *International Journal of Information Technology*, Springer. He is the Series Co-editor of Springer Tracts in Nature-Inspired Computing, Springer, Series Co-editor of Advances in Ubiquitous Sensing Applications for Healthcare, Elsevier, Series Editor of Computational Intelligence in Engineering Problem Solving and Intelligent Signal processing and data analysis, CRC. His main research interests include Medical Imaging, Machine learning, Computer Aided Diagnosis, Data Mining etc. He is the Indian Ambassador of International Federation for Information Processing—Young ICT Group and Senior member of IEEE.

Bharat S. Patel is Fellow at IEI, IETE and CSI. He is a member of CSI, International Red Cross, Association of British Scholars (a Division of British Council), British Business Group, GCCI, GESIA IT Association, and a founder member with Gujarat Innovation Society (GIS), ASSOCHAM and many more. He was Past Chairman of CSI, Ahmedabad, and Past Chairman of Gujarat State Centre of IEI. Currently, he is President with ABS (division of British Council), Chairman, Startups Mission, and Chairman, Startup and Innovation, ASSOCHAM for Western council, Chairman, Academia and Research Publications committee, GESIA, Vice Chairman, Gujarat Innovation Society, and a council member and Chairman, CPDB, at The Institution of Engineers (India). He is Director and COO at Yudiz Solutions Pvt. Ltd. India.

Aboul Ella Hassanein is Founder and Head of the Egyptian Scientific Research Group (SRGE) and Professor of Information Technology at the Faculty of Computer and Artificial Intelligence, Cairo University. Professor Hassanien has more than 1000 scientific research papers published in prestigious international journals and over 50 books covering such diverse topics as data mining, medical images, intelligent systems, social networks and smart environment. Prof. Hassanien won several awards including the Best Researcher of the Youth Award of Astronomy and Geophysics of the National Research Institute, Academy of Scientific Research (Egypt, 1990). He was also granted a scientific excellence award in humanities from the University of Kuwait for the 2004 Award and received the superiority of scientific in technology—University Award (Cairo University, 2013). Also he honored in Egypt as the best researcher in Cairo University in 2013. He was also received the Islamic Educational, Scientific and Cultural Organization (ISESCO) prize on Technology (2014) and received the state Award of excellence in engineering sciences 2015. He holds the Medal of Sciences and Arts from the first class from President of Egypt in 2017.

Contributors

Hifzan Ahmad Dr. A. P. J. Abdul Kalam Technical University (AKTU), Lucknow, India

Rifaqat Ali Department of Mathematics and Scientific Computing, National Institute of Technology, Hamirpur, Hamirpur, Himachal Pradesh, India

Rajanikanth Aluvalu Department of CSE, Vardhaman College of Engineering, Hyderabad, India

Subhra Priyadarshini Biswal School of Computer Science and Engineering, National Institute of Science and Technology (Autonomous), Berhampur, Odisha, India

Arun Prasad Burnwal Department of Mathematics, GGSESTC, Bokaro, Jharkhand, India

Preeti Chandrakar Department of Computer Science and Engineering, National Institute of Technology, Raipur, Raipur, India

Krishna Keerthi Chennam CSE Department, Muffakham Jah College of Engineering and Technology, Telangana State, Hyderabad, India

Chandrika Dadhirao SCOPE, VIT-AP University, Vellore Institute of Technology -AP University, Amaravathi, India

Santosh Kumar Das Department of Computer Science and Engineering, Sarala Birla University, Birla Knowledge City, Ranchi, Jharkhand, India

Sunil Gautam Department of Engineering and Physical Science, Institute of Advanced Research, Gandhinagar, India

Joydev Ghosh School of Computer Science and Robotics, National Research Tomsk Polytechnic University (TPU), Tomsk, Russia

Koushik Ghosh Department of Mathematics, University Institute of Technology, The University of Burdwan, Burdwan, West Bengal, India

Ramesh Chandra Goswami Department of Engineering and Physical Science, Institute of Advanced Research, Gandhinagar, India

Samiran Gupta Department of Computer Science and Engineering, Asansol Engineering College, Asansol, West Bengal, India

Priyanka Jaiswal Department of Computer Science and Engineering, IIT (ISM), Dhanbad, Jharkhand, India

P. V. Y. Jayasree GITAM University, Vizag, India

Harsh Nath Jha Department of Information Technology, Asansol Engineering College, Asansol, West Bengal, India

Hiren Joshi Department of Computer Science, Gujarat University, Ahmedabad, India

Manas Ranjan Kabat Department of Computer Science and Engineering, Veer Surendra Sai University of Technology, Burla, Sambalpur, India

Pabitra Mohan Khilar National Institute of Technology Rourkela, Rourkela, India

Abhishek Kumar Department of Electronics and Communication Engineering, Swami Vivekananda Subharti University, Meerut, India

B. Praveen Kumar Department of EEE, Bharat Institute of Engineering and Technology, Hyderabad, India

Deepak Kumar National Institute of Technology Rourkela, Rourkela, India

B. K. Mahatha Amity School of Engineering and Technology, Amity University Jharkhand, Ranchi, India

Vishal Maheswari Department of Computer Science, RIT, Roorkee, Uttarakhand, India

Manoj Kumar Mandal Department of Mathematics, Jharkhand Rai University, Ranchi, India

Divya Mishra Department of Computer Science Engineering, Swami Vivekananda Subharti University, Meerut, India

Vishwas Mishra Swami Vivekananda Subharti University, Meerut, India

Hari Om Department of Computer Science and Engineering, Indian Institute of Technology (ISM), Dhanbad, India

Rajendra Pamula IIT (ISM) Dhanbad, Jharkhand, India

Dinesh Kumar Sah Indian Institute of Technology (ISM), Dhanbad, India

Satya Prakash Sahoo Department of Computer Science and Engineering, Veer Surendra Sai University of Technology, Burla, Sambalpur, India

Priyanka Saini Swami Vivekananda Subharti University, Meerut, India

Swetadri Samadder Department of Mathematics, Fakir Chand College, Diamond Harbour, India

RaviSankar Sangam SCOPE, VIT-AP University, Vellore Institute of Technology -AP University, Amaravathi, India

K. Sangeetha Department of CSE, Sri Satya Sai University of Technology and Medical Sciences, Sehore, Madhya Pradesh, India

N. Satheesh Department of CSE, St. Martin's Engineering College, Hyderabad, India

Nikhil Saxena University of Cincinnati, Cincinnati, OH, USA

Biswa Ranjan Senapati National Institute of Technology Rourkela, Rourkela, India

Aditya Sharma Institute of Nanoengineering and Microsystems, National Tsing Hua University, Hsinchu, Taiwan R.O.C.

S. Shitharth CSE Department, Vardhaman College of Engineering, Telangana State, Hyderabad, India

Chaya Shivalingagowda Kalsekar Engineering College, New Panvel Mumbai and GITAM University, Vizag, India

Amit Kumar Singh IIT (ISM) Dhanbad, Jharkhand, India

Rakesh Ranjan Swain Department of CSE, ITER, Siksha O Anusandhan (Deemed to be University), Bhubaneswar, India

Sachin Tripathi Department of Computer Science and Engineering, IIT (ISM), Dhanbad, Jharkhand, India

Shobhit Tyagi Swami Vivekananda Subharti University, Meerut, India

V. Uma Maheswari Department of CSE, Vardhaman College of Engineering, Hyderabad, India

Wireless Networks: Applications, Challenges, and Security Issues

Santosh Kumar Das, Vishal Maheswari, and Aditya Sharma

Abstract Nowadays, wireless technology is an essential part of communication. Most of the organizations benefitted by adopting wireless technology solutions may lead to higher productivity. Today, globally, several customers are using this technology for resolving various business issues and create advantages over competitors. This technology helps to achieve high customer satisfaction with lesser complexity. It also assists various types of exciting applications such as sensor networks, Bluetooth, mobile communication systems, and Internet of Things (IoT). Wireless technology makes the use of radio waves to transfer data without cables or wiring. In this proposed paper, several applications of wireless networks and its variations are illustrated along with their challenges and security issues. It provides a guideline about upcoming inventions in the area of wireless technology.

Keywords Wireless ad-hoc network · Wireless sensor network · Security issues · Challenges · Internet of Things · Attacks

1 Introduction

In the last few decades, the applications of wireless networks and their variations have increased rapidly due to the widespread use in the developing wireless techniques [1–3]. Wireless in its simple form can be expressed as the automation process in which transfer of data and information takes place without using any wired media. One might be thinking how can data be transferred without using wires and if so, then

S. K. Das (✉)
Department of Computer Science and Engineering, Sarala Birla University, P.O.-Mahilong Purulia Road, Birla Knowledge City, Ranchi, Jharkhand, India

V. Maheswari
Department of Computer Science, RIT, Roorkee, Uttarakhand 247667, India

A. Sharma
Institute of Nanoengineering and Microsystems, National Tsing Hua University, No. 101, Sec. 2, Guang Fu Road, Hsinchu 30013, Taiwan R.O.C.

S. K. Das et al. (eds.), *Architectural Wireless Networks Solutions and Security Issues*, Lecture Notes in Networks and Systems 196,
https://doi.org/10.1007/978-981-16-0386-0_1

what is the medium? Air is the only medium for the transfer of data through wireless mode which in return uses electromagnetic waves for the transmission of signal from the transmitter to the receiver [4]. One might be able to understand that for short-ranged communications, one can use the wireless technology very smoothly but what about the long-range communication? Therein comes the concept of receiving and transmitting data through the waves, i.e., radio waves, which in it provide some energy for the transmission to occur over longer distances. Herein, it is cleared out the use of wireless technology which is applicable and widely used for both short as well as long-distance communication. Figure 1 shows types of wireless network. Wireless network are categorized as three major types which are: Wireless ad-hoc network (WANET) [5, 6], wireless sensor network (WSN) [7, 8], and other wireless network. WANET is a collection of dynamic nodes that are deployed at a particular location for any operation. It has several variations or types such as mobile ad-hoc network (MANET) [9], vehicular ad-hoc network (VANET) [10, 11], and hybrid ad-hoc network (HANET) [12]. MANET is a collection of mobile nodes that are simply movable based on the requirement of the users or customers. VANET is a collection of different vehicles that are connected dynamically to provide the services to the driver as well as the passenger for an automated system. HANET is a combination of static as well as dynamic nodes. The combination of both helps the user in both static and dynamic purpose of the services. WSN is a collection of wireless sensor nodes. The purpose of these sensor nodes is to sense environmental information and send it to the base station (BS). BS analyzes this information for future processing and forwards it to the sink node. WSN is also used in HANET with the fusion of VANET and smart ad-hoc network to make use the services of Internet of Things (IoT) [13]. In HANET, several physical objects are connected with digital technology to make an efficient and appropriate communication services in HANET.

IoT is nothing more than a collection of wide range of software, systems, and users via the Internet technology; having a built-in ability of transferring data over a network without having a human interaction [14, 15]. Talking about first generation of IoT, SCADA [16] is an acronym for "supervisory control and data acquisition". SCADA provides a bundle full of different types of software-based application program to perform a particular task which can be accessed from remote location. It includes both hardware as well as software components. The use of hardware

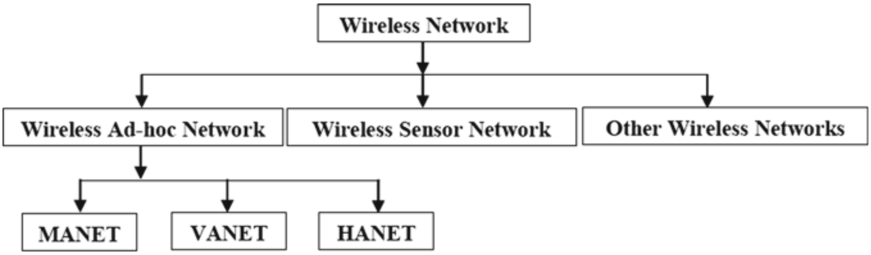


Fig. 1 Types of wireless network

component is to gather the data and then feed it into the computer, wherein the next step is carried out by respective software according to the situation. A SCADA system is used to gather information, like from where the smoke is coming from a building, then it transfers the information gathered back to the central site, warning the home station that the smoke has occurred, carrying out the necessary analysis and further controlling the scenario, gathering some more information for determining whether the smoke is caused by fire, and displaying the gathered information in a proper logical and organized manner. Other areas where SCADA system can be used include municipal water supply, in a small building and many more. The next section is illustrated with developmental strategy of IoT and other variations of the wireless network.

The roadmap of the paper as follows. Unit 2 describes some applications of wireless network and its variations. Unit 3 illustrates some constraints of ad-hoc and sensor network. Unit 4 describes some security and major design issues. And Unit 5 concludes the paper.

2 Applications of Wireless Network and Its Variations

IoT is one of the novel variations of the wireless network. First of all, an ecosystem of IoT is developed. This ecosystem is different from the typical ecosystem containing some planets and stars. This ecosystem contains in it a huge number of hardware and software devices that are connected to a Web-enabled network source which encapsulates a number of embedded processors, sensors, hardware, and software based on the task they are going to be helping with. The gathered data is shared by the IoT devices by bridging of the data to the gateway of IoT or sending the data to the cloud based systems where it can be easily analyzed and the output can be made to be accessed and performed as required. All the objects which are having an in-built sensor are connected to an IoT based platform, which gathers the data from various hardware devices and share the unmatched information with the software to meet the required data analytics.

The ecosystem of IoT platform can itself decide which information is to be taken into consideration and which can be safely ignored without any loss or manipulation of data. The information collected is based on a preprogrammed software which include some patterns and recommendations used to find out some possible problems or issues before they take place. For an example, a person is the owner of a supermarket store, and he/she must be aware of the products which are most popular. Sensors can be placed in the supermarket to detect the most popular areas, and where customers wait around or stay for a longer span of time [17–19]. The faster selling products can be identified by checking the daily sales data, in case the most selling product must not go out with not on stock board; automatically align sales data with supply, so that popular items don't go out of stock.

The information gathered by the smart connected devices can guide one in making smart decisions on the products to have a higher stock, which would be completely based on the people's shopping bucket list and would help in saving the man-power required to take up the stock and check out over people's activity. It is obvious that the data gathered by the devices will bring more efficiency and accuracy as compared to the traditional means; and likewise, more efficiency leads to doing work in a smarter and more controlled manner and resulting in work. By the help of smart objects and systems, one can automate certain tasks, particularly when those tasks are bulky, repetitive, mundane, time-taking, or dangerous. Let us have a look over some examples to make the scenario clearer and more accurate. In one's daily monotonous life routine, everyone has to work for having a meal and having a pending or delayed work can make one lose one's job. In this competitive era where technology is faster than human, many times one faces a scenario where a person woke up on time, but it's raining outside or his/her car engine is not working, someone has flattened his/her car tier, he/she had to get off in traffic and many others. In all such cases, there is a fixed prepared reason to be used by human for his/her delay. Here comes the role of IoT where the delay can be easily postponed and one need not have to blame one's luck over it. Let us summarize some of the benefits of IoT taking the above scenario into consideration.

(a) Save time and money
(b) Ease of service
(c) Enhance working experience
(d) Increase productivity
(e) Low investment high returns
(f) Taking smart business decisions
(g) Easy to monitor the business.

IoT helps companies and individuals to take smart decisions, adopt smart technologies, and allow them to work more productively and efficiently. The major concern of developer is how they are going to secure the use of such an enormous amount of data, where all the devices are connected to Internet. For the use of IoT based devices, the only thing which needs to be taken into consideration is the security and privacy issues. The IoT based devices needs to be always connected to a network, the hacker has to simply gain an access to any single device and manipulate all the data, and for a solution to it, you can provide security patch on a regular time interval. But how many manufacturers are there who will update it to the latest firewall? Apart from, WANET, MANET, VANET, HANET, and WSN, several wireless networks are used based on customer requirements such as cellular network, mesh network, delay tolerant network, and software defined network. The stated variations of the wireless network have become a major and important part of our life and real-life applications. The combination of all variations gives a lot of efficient and reliable benefits to the users and customer in terms of mobility and remote areas. It is low cost, low time consuming, more efficient, and intelligent compared to wired network. It is also simple for use and license free and also deployable. Wireless network is a location-depended service that is a replacement of wired network and helps to the users and customers in emergency situation, business, offices, traveling salesman, etc. with combinations of some devices such as Wi-Fi, GPS, and cordless

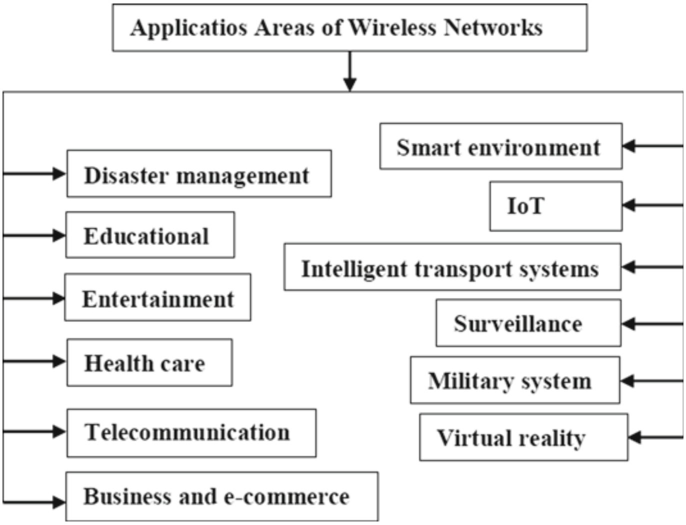


Fig. 2 Applications areas of wireless network

telephones. There are several applications of wireless networks and its variations which is shown in Fig. 2. Some of them are artificial intelligence, enforcement and control systems, environmental monitoring, intelligent transport systems, IoT, military applications, person locator services, smart environment, telecommunications system, traffic avoidance, virtual reality visual surveillance sensor networks, etc.

3 Constraints of Ad-Hoc and Sensor Networks

The fusion of WANET and WSN is known as ad-hoc sensor network. Although, both have some similar features, they also both have some differences like a number of nodes in WSN are more as compared to WANET. The nodes in WSN, known as sensor nodes, are more prone to failure and energy drain. Although there are several applications and usage in terms of wireless network-based infrastructure and infrastructure-less, static, and dynamic topologies, combination of both have some limitations that differ from classic network such as limited energy supply, limited computing power, limited bandwidth of the wireless links connecting data, routing challenges, data aggregation, coverage and scalability, and data reporting methods and protocols. Summarized limitations are described as follows.

(a) **Limited hardware resources**: Due to several issues of WSN such as limited storage, computational system, limited energy, long distance from receiver, it is limited by the hardware resources.

(b) **Unreliable communication**: Due to limited bandwidth, dramatically dependency, temporary, and variable channel, the communication is always unreliable.

(c) **Dynamic topology**: In terrestrial sensor network, nodes are deployed densely, and in underwater sensor network due to flow of water, sensor nodes are mobile.

(d) **Vulnerability and insecure environment**: There are several applications of sensor nodes such as monitoring, sensing, target tracking, and detecting hostile object and region. So, nodes become susceptible to attacks and threats.

The several networks and their variations have been described in the above-mentioned section. In each variation, security is one of the most crucial parts in every sector of real-life application. There are several limitations and constraints described in the above section, in which limited energy is the crucial part. Due to the above constraints, several threats are detected in the network as shown in Fig. 3. Basic types of these threats are as follows [20, 21].

(a) **Passive attacks**: This attack is done by the malicious nodes without interrupting the main operation by receiving information about network and data transmission, e.g., message distortion, unnecessary message reply, leakage or trap secret information, interfering, and eavesdropping.

(b) **Active attacks**: This attack is done by some external or internal nodes. It can destroy or delete the important data and information, and sometimes it tries to modify, inject, or drop data packet.

  (i) **Compromise attacks**: In this attack, attacker may compromise the node for modifying or reading the secret data or information.

  (ii) **Routing attack**: This attack consists of unreliable data transferred to the destination node. It is also known as rushing attack. Examples are packet dropping, packet replication, routing table poisoning, and overflow.

  (iii) **DoS attack**: It this attack, attacker tries to prevent the resources from accessing the data. It is more difficult to detect and handle. Sometimes, it handles with encryption method of the cryptography.

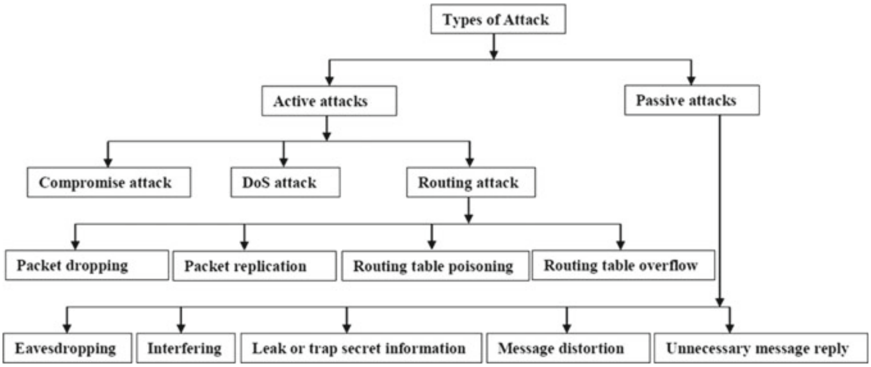


Fig. 3 Attacks and its type

4 Security and Major Design Issues

Ad-hoc and sensor network and its variations have different capabilities in terms of topology and network parameters. In the above sections, several limitations and constraints discussed that motivate for designing an efficient model that care about the following paradigms.

(a) **Modeling of security enhancements:** The nodes of ad-hoc and sensor networks are dynamic and autonomous. They act as routers and help in sending and transmitting the data packets. It greatly relies on the environment of the modern technology. It also has several limitations like limited energy supply, limited computing power, limited bandwidth of the wireless links connecting data, routing challenges, data aggregation, coverage and scalability, data reporting methods and protocols, unreliable communication, vulnerability, and unsecure environment. These stated limitations cause two types of attacks: passive and active attacks. Examples of passive attacks are message distortion, unnecessary message reply, leakage or trap secret information, interfering, and eavesdropping. Examples of active attacks are modify, inject or drop data packet, modify or read secret data information, packet dropping, packet replication, routing table poisoning and overflow, etc. So, network needs an intelligent and efficient security modeling with the help of any artificial intelligence, soft computing, and machine learning techniques. Sometimes cryptographic technique is also mixed with any of the stated techniques to make the network more secure in terms of privacy in both systems like network-based data as well as cloud-based data.

(b) **Optimization Model for Network Lifetime Enhancements**: The nature of the ad-hoc and sensor network is dynamic and autonomous. Each node behaves as router and acts as an intelligent agent that plays the role of data transferring agent between source and destination nodes. Due to this intelligence characteristic, several types of interferences occur. So, there is need of some optimization techniques to model the network and enhance the lifetime of the wireless network. Network lifetime is the time duration between when the network is started and when half of the nodes are exhausted. The optimization technique is used to find an optimal as well as feasible solutions. The optimal solution is the best solution among all of the solution, and feasible solution is the solution nearby optimal solution. In ad-hoc and sensor network, optimal solution indicates the solution when all network metrics are outperformed in terms of traditional worst metrics. It helps to increase and decrease the network metrics based on network lifetime such as packet delivery ration, throughput, goodput, and residual energy are increases and end-to-end delay, packet loss, jitter, overhead, are decrease. The combination of both changes helps in overall network performance.

(c) **Modeling of Aggregation Systems**: Ad-hoc and sensor network is a collection of large number of small nodes. The purpose of the wireless sensor nodes is to sense the main requirement phenomena from the environment and send it to the required places. The purpose of the network is that it should be useful in several applications such as military, security maintenance, disaster management, and habitat monitoring. In each application, a node plays an important role, and each ad-hoc node or sensor node consists of limited energy capacity or battery which is not sufficient during any operation. Both the networks have high density due to several variations of sensor nodes or ad-hoc nodes. Same data packets are sensed by multiple nodes and raising the redundancy or duplicate data packets. Data aggregation is used to control this issue efficiently and in an intelligent way. This data aggregation technique is rapidly used in ad-hoc and sensor network and their several variations. It helps to enhance the network lifetime as well as network metrics efficiently.

(d) **Analysis of Troubleshooting Techniques**: The above-mentioned sections and paragraphs contain several applications and uses of ad-hoc and sensor network. In each application, there are several types of randomness and uncertainties. It raises multiple interferences between one node and another node, source node to destination node or among multiple neighbor nodes. These interferences and uncertainties are the main cause of imprecise information and network troubles. These results in of several network security issues and cause different attacks that are mentioned in the above section. Hence, there is a need of some intelligent technique using artificial intelligence, soft computing, machine learning, or any other intelligent technique. Sometimes a single technique is efficient for handling any trouble. Sometimes there is need of some fusion between multiple techniques. The combination of multiple techniques provides more robustness for handling uncertainty of the network and estimate imprecise information efficiently.

The stated inherent paradigms required some necessary precautions shown in Fig. 4 which help to overcome some major design issues such as coverage that indicates communications between two or multiple nodes in term of data acquisition.

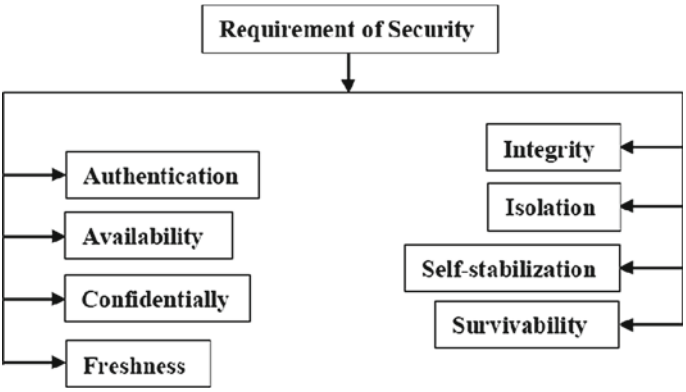


Fig. 4 Requirement of security

Coverage has multiple types like target based and areas based. Target-based coverage indicates based on destination node cover or sink node cover, and area-based coverage indicates cover communication range with source node, sink or destination node, BS, and multiple neighbor nodes. Network lifetime is the key point of the life cycle of any wireless network or ad-hoc and sensor network. It also indicates the duration when half of the nodes have exhausted their energy. Network traffic and network connectivity are two basic points of enhancement of the network lifetime. Sometimes connectivity is slow, sometimes it is fast or moderate, it deals with the fuzzy value that handles multiple logic between actually true or actually false values. Network traffic indicates gathering of nodes for communication or data transferring services. Although, network traffic and network connectivity both are basic key points of the network lifetime, combination of both has some minor metrics that also affect the variation of network lifetime. These metrics named as packet delivery ration, packet loss, average end-to-end delay, throughput, goodput, jitter, etc. These metrics are affected by control data packets, ratio between data sent and received, and different inherent elements required during data transfer. The nature of some of the metrics is same or opposite based on the network behavior.

5 Conclusions

The proposed paper analyzed the details of wireless network and its variations with their applications, security challenges, and issues. The paper briefly describes the working principles of wireless network along with its variations such as WANET, MANET, VANET, WSN, and IoT and how these variations help the user and customer in context with their real-life applications and requirements. The limitations along with the constraints give guidelines to the readers and researchers for enhancing the field of wireless network and their inherent elements. It also guides modeling and optimization models for security enhancement, network lifetime enhancement, data aggregation, and troubleshooting techniques.

References

1. Lin Z, Wang P (2019) A review of data sets of short-range wireless networks. Comput Commun 147:138–158
2. Wang W (2020) Deployment and optimization of wireless network node deployment and optimization in smart cities. Comput Commun 155:117–124
3. Chen Q (2020) Wireless network signal monitoring based on LAN packet capture and protocol analysis on grid programming. Comput Commun,15:45–52
4. Chen X, Gao L, Chen J, Lu S, Zhou H, Wang T, Wang A, Zhang Z, Guo S, Mu X, Wang, ZL (2020). A chaotic pendulum triboelectric-electromagnetic hybridized nanogenerator for wave energy scavenging and self-powered wireless sensing system. Nano Energy 69:104440
5. Vasudeva A, Sood M (2018) Survey on sybil attack defense mechanisms in wireless ad hoc networks. J Netw Comput Appl 120:78–118

6. Zhong X, Chen F, Guan Q, Ji F, Yu H (2020) On the distribution of nodal distances in random wireless ad hoc network with mobile node. Ad Hoc Netw 97:102026
7. Pang C, Xu G, Zhang Y (2020) A new energy efficient management approach for wireless sensor networks in target tracking. Defence Technol. https://doi.org/10.1016/j.dt.2020.05.022
8. De D, Mukherjee A, Das SK, Dey N (2020) Nature inspired computing for wireless sensor networks
9. Goyal P, Parmar V, Rishi R (2011) Manet: vulnerabilities, challenges, attacks, application. IJCEM Int J Comput Eng Manage 11(2011):32–37
10. Hartenstein H, Laberteaux K (eds) (2009) VANET: vehicular applications and inter-networking technologies, vol 1. Wiley
11. Karnadi FK, Mo ZH, Lan KC (2007) Rapid generation of realistic mobility models for VANET. In 2007 IEEE wireless communications and networking conference. IEEE, pp 2506–2511
12. Qiu T, Chen N, Li K, Qiao D, Fu Z (2017) Heterogeneous ad hoc networks: architectures, advances and challenges. Ad Hoc Netw 55:143–152
13. Ai Y, Peng M, Zhang K (2018) Edge computing technologies for Internet of Things: a primer. Digit Commun Netw 4(2):77–86
14. Hao R, Yang H, Zhou Z (2019) Driving behavior evaluation model base on big data from internet of vehicles. Int J Ambient Comput Intell (IJACI) 10(4):78–95
15. Shinde GR, Olesen H (2018) Beacon-based cluster framework for internet of people, things, and services (IoPTS). Int J Ambient Comput Intell (IJACI) 9(4):15–33
16. Igure VM, Laughter SA, Williams RD (2006) Security issues in SCADA networks. Comput Secur 25(7):498–506
17. Yang W, Wang X, Song X, Yang Y, Patnaik S (2018) Design of intelligent transportation system supported by new generation wireless communication technology. In: Intelligent systems: concepts, methodologies, tools, and applications. IGI Global, pp 715–732
18. Chandrakar P (2019) A secure remote user authentication protocol for healthcare monitoring using wireless medical sensor networks. Int J Ambient Comput Intell (IJACI) 10(1):96–116
19. Wu J, Huang Z, Guan Y, Cai C, Wang Q, Xiao Z, Zheng Z, Zhang H, Zhang X (2011) An intelligent environmental monitoring system based on autonomous mobile robot. In 2011 IEEE international conference on robotics and biomimetics. IEEE, pp 138–143
20. Poornima IGA, Paramasivan B (2020) Anomaly detection in wireless sensor network using machine learning algorithm. Comput Commun 151:331–337
21. Yang G, Dai L, Si G, Wang S, Wang S (2019) Challenges and security issues in underwater wireless sensor networks. Procedia Comput Sci 147:210–216

Modelling of Security Enhancements

An Authentication Model with High Security for Cloud Database

Krishna Keerthi Chennam, Rajanikanth Aluvalu, and S. Shitharth

Abstract The cloud computing standards are gaining an increased research interest due to various benefits they offer. Though there are so many influences with cloud computing, security and privacy problems are various issues handling with the extensive adaption by the model. Malicious problem of service provider is one more issue which cannot be traceable by data proprietors. Hence, finding the appropriate solutions to these security issues at both administrator level and customer level is very attractive in various directions. Cryptographically enforced access control for securing electronic pathological records (CEASE) is formulated by extending the proposed ciphertext-based attribute-based encryption (CP-ABE) with advanced encryption standard (AES) through limited-shuffle techniques. The main objective of CEASE is to provide data confidentiality, and access control limited-shuffle protects the data from inference attacks and protects the data confidentiality for hot data. In the next step, this research works design a multistage encrypt-or model by differentiating the users as public and personal. Two separate algorithms such as Vigenere encryption algorithm and two-fish encryption are applied in personal and public domain, respectively. Further, where, hierarchical agglomerative clustering (HAC) algorithm is also processed for clustering of users in the public domain by which the overhead decreases effectively. As a final system, this work develops an integrated framework by combining the CP-ABE with AES, multistage encryptor and limited-shuffle. As it is combined with individual methods, this method achieves an efficient performance in the provision of security and data confidentiality.

Keywords Cloud computing · Data security · Access control models · Encryption · Clustering algorithm · Limited-shuffle

K. K. Chennam
CSE Department, Muffakham Jah College of Engineering and Technology, Telangana State, Hyderabad, India

R. Aluvalu · S. Shitharth (✉)
CSE Department, Vardhaman College of Engineering, Telangana State, Hyderabad, India

S. K. Das et al. (eds.), *Architectural Wireless Networks Solutions and Security Issues*, Lecture Notes in Networks and Systems 196,
https://doi.org/10.1007/978-981-16-0386-0_2

1 Introduction

Cloud computing standards are gaining an increased research interest by the different influences. The major benefit involves time savings, with reduced cost and efficient utilization of computing resources. Though there are so many influences with cloud computing, security and privacy problems are the important problems holding back the extensive adaption of this automation. The general characteristic of cloud computing technology requires the clients to store their data on third-party cloud service providers, which can also be termed as outsourcing of data. The security and privacy are generally maintained by the CSP where the data proprietors do not have complete control on the data security, malicious nature of service provider and third-party users is one more issue which cannot be traceable by data proprietors. Hence, finding the appropriate solutions to these security issues at both administrator level and client level is very important in various directions.

Earlier research is based on standard encryption algorithms like AES, data encryption standard (DES), etc. However, the advancement in the technology makes these approaches ineffective because of the lack of control on authorization and authentication. In contrast, the attribute-based encryption (ABE) was the new research which has the desire to give the maximum to handle by the data proprietors who can give the data and also provide an efficient management for the cloud service provider. However, the ABE-based approaches provide security at the cost of execution. Therefore, the challenge of achieving the dual goals of privacy preserving with effective cloud data sharing remains unresolved.

In summary, the major significant addition in the section is to influence by the benefits of the ABE application to carry out the real-time answers to security and privacy problems experienced in the cloud computing environments.

Section 1 discusses the introduction about cloud computing, data security and access control schema. Section 2 discusses the CP-ABE with AES, Sect. 3 discusses the CEASE, and Sects. 4 and 5 talk about the partial shuffling with two-stage encryption and integration model. Last section discusses the results and conclusion.

2 Problems in Data Security

Security, privacy and trust issues are existing and given importance since the evolution of Internet, and they are widely spoken these days because of cloud computing. Cloud's dynamic nature demands higher security levels. Users or organizations subscribed to cloud for running their business processes are strikes to acquiring the next level of endanger because of expanded applications. A cloud user while saving the data on the cloud, which wants to make sure if the data is correctly stored and can be retrieved later. The service provider must ensure the secure infrastructure to protect the data and applications of its clients and the users. Various security strategies proposed earlier have become ineffective due to advancements in technology.

This is not the usual CSP and the user for both imaginations. What is required is a mechanism that assures data consistency to the cloud user and protects that the user is not some malicious hacker. Hence, the necessity for developing trust-based security model is the need of the hour.

3 Objectives of Data Security

Cloud computing applications have to ensure security of the data stored in the cloud. Existing approaches are suffering from various drawbacks and require improvement. In particular, the proposed scheme has the following objectives:

(a) Dual optimization: Data confidentiality and processing time are the two main constraints which are not achieved simultaneously. More processing time (encryption time + decryption time) is required to achieve an efficient data security for data stored in the cloud. On the other hand, the less computational operations to encrypt the data will reduce the data confidentiality and result in an information loss. To meet these two constraints simultaneously, this research work focuses on developing an effective cloud computing technique based on the ABE [1] and multistage encrypt-or. By adding some more standard techniques (AES and limited-shuffle) with these approaches, this work tries to achieve the data confidentiality and less processing time.

(b) Increase data confidentiality: To achieve increased hot data confidentiality and preservation of privacy, this work proposes a CEASE. In this approach, an advanced encryption standard accomplishes an encryption algorithm to reduce the effect of curious/malicious administrator.

(c) Resilient to inference attacks: To make the system more secure from inference attacks and from malicious authority attacks, this work proposes a single-level block index method along with limited-shuffle, by which the system acquires data accomplished models off the record without reducing the querying process.

(d) Reduced computational overhead: To reduce the unnecessary computational overhead in the large-scale cloud storages, this work accomplishes a clustering mechanism, called HAC supports based with the place of utilizers.

3.1 CP-ABE with AES

This section proposed access control within database strategy, CP-ABE combined with standard AES algorithm. Here CP-ABE [2] achieves the authenticated accessing of only legal users and AES ensures the data security. Before uploading data to the cloud, it is encrypted through AES algorithm by which the data user will be relaxed about the data security. Further in CP-ABE, proprietor accommodates attributes set, when the user wants the data accomplishment which needs the attributes set and requires the secret key for decrypting the data, where encrypt-or accommodates the key with the strategy of the access control plan of action. Though administrators are curious about the data, due to the non-availability of key, the data cannot be accessed by that malicious authority [3]. Hence, this method protects the security from malicious authority more effectively [4].

The proposed strategy gives cloud document for the security space with respect to performance metrics like key generation time, encryption and decryption time. The key generation time is computed with various secret keys with the identified set attributes. To produce non-public key in CP-ABE with AES is not exactly the same by CP-ABE with bilinear mapping. It is observed that, for every attributes set, the obtained key generation time is less when contrast to the conventional CP-ABE with bilinear. The encryption time and decryption time are computed with various no. of policy of leaf nodes, which is limited in CP-ABE with AES contrast with CP-ABE with bilinear mapping. CP-ABE with AES gives protection and security for data records for the information of the cipher and store in cloud. The CP-ABE with AES gives limited key generation, encryption time when contrasted by CP-ABE with bilinear mapping.

4 CEASE

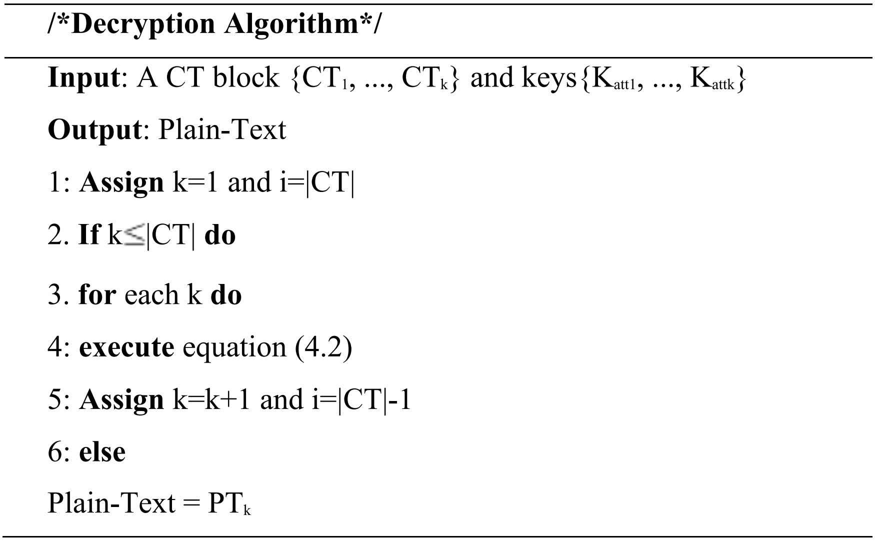
The CEASE is outlined in this section; main objective of CEASE is to provide data confidentiality and access control of outsourced CS information over the security threats. The proposed CEASE framework comprises three constituents to protect the cloud data security:

(a) Accomplishment of AES on sensitive patient health records.
(b) Secure information retrieval through a data accomplishes and direct technique and query encryption.
(c) Data confidentiality for hot data through limited-shuffle to protect the data from inference attacks.

Initially, the holder of data modifies the loyal proxy server by extending AES on the health data before transferring it to CSP. Ordinal, the proxy server is the important attribute set administration recognizes the individuals applying the set of attributes and overdrives access control plan of action on electronic information inward cloud. The encrypted queries retrieve the encrypted data from the cloud and to decrypt the data using attributes in the proxy server before delivering information to the final consumer. Nevertheless, retrieving encrypted information of ciphertext assures high confidentiality of every patient record in the cloud, and there is a possibility of inference attacks. Thirdly, the CEASE techniques apply the limited-shuffle within a single block of the data that contains the sensitive health records and protects the data confidentiality aside from swift retrieval. Thus, the recommended CEASE algorithm protects malicious authority of cloud unable to take or change (hot) information one of two is treasure delicate health files or encrypted query execution along with the faster querying process [5]. The performance metrics such as querying cost, storage overhead and hot data confidentiality are examined on the recommended method. The decryption algorithm decrypts the data and sends the plain text to the client when the set of attributes are matched according to the CP-ABE with AES.

The performance evaluation of recommended CEASE is carried out on the JAVA platform on a personal health records. The performance metrics such as querying cost, storage overhead and hot data confidentiality are measured for varying data sizes. From the simulation results, it is proved that the recommended approach shows slight increase in the querying cost but reduced storage overhead and finally an improvement in the hot data confidentiality contrast with existing approach.

Algorithm 1 Decryption algorithm



```
/*Decryption Algorithm*/
Input: A CT block {$CT_1$, ..., $CT_k$} and keys{$K_{att1}$, ..., $K_{attk}$}
Output: Plain-Text
Assign k=1 and i=|CT|
If k$\leq$|CT| do
    for each k do
        execute equation (4.2)
        Assign k=k+1 and i=|CT|-1
else
    Plain-Text = $PT_k$
```

5 Multistage Encrypt-Or for Securing Data Records

This section outlines the recommended multistage encrypt-or strategy of protecting the personal health records (PHR) of third-party database storage. The main objective of this approach is to provide security for the PHR in the cloud with less computational overhead. The framework differentiated by the multiple regions partitioned by public and personal domains is discussed according to the client's data to give access permissions.

To ensure security in the non-public domain, this approach uses Vigenere encryption algorithm, and for the public domain, it uses the two-fish-based encryption algorithm. For every user in the non-public domain (PSD), the clients, relatives or nearby people are connected in a chain fashion, and they are able to get PHR in glimpse of getting opportunities designated from the sick person. Here each client achieves the Vigenere encryption-based system to manage the decoding by receiving awards of customers in his/her PSD. In public domain (PUD), two-fish encryption is used by the attendant of diverse AAs, each one directing a disjoint subset of characters [6, 7]. To regulate approaches by the PUD wards and let on to reflect role-oriented fine-grained approaches for their PHR documents, while they do not require the sanctioned users at the time of encryption. The PUDs contain the maximum number of wards. By coming through the difficulty, here this approach groups the ward's duty in the PUD with HAC algorithm. Wards of PUDs get back attribute designated encryption keys supported with the ward functions. The observational maps about encryption, decryption time, clustering accuracy and storage requirements are evaluated using various data set sizes. The observational effect shows that the recommended method has more clustering quality, less encryption and decryption time.

6 Raising the Security with Fine-Grained Access Control Plan of Action Using Two-Stage Encryption with Limited-Shuffle in the Cloud

This section integrates the CP-ABE with AES and two-stage encrypt-or with limited-shuffle [8]. The primary goal of CP-ABE with AES is to recognize the malicious clients and data proprietors who can access data from the cloud. Next, the multistage encrypt-or helps in reducing the extra computational overhead [9]. The electronic records are protected from inference attacks by applying limited-shuffle as shown in Fig. 1. The data proprietors are maintaining the keys distribution authority, certificate verification and attribute authority and send the data to the proxy server. The proxy server applied two-stage encryption techniques based on the domains mentioned above, while doing encryption the key pairs are received from elliptic curve. The encrypted data is stored in the cloud database. The proxy server encrypts data before storing in cloud and plain queries also encrypted by proxy server before retrieving data from cloud, where there is no possibility of plain data to the malicious admin in cloud and in network or in proxy server.

Two-fish algorithm uses with different and random key length of variable size of 128 bits, 192 bits and 256 bits. Two-fish is a symmetric algorithm with quick encryption great with AES due to its speed, adaptability and protection outline. For every query, the database needs to be searched line by line in the table, where the questioning time is expanded as the information size is expanded straightly. To address this issue, a record is made by information by examining the file is decent as opposed to examining the entire information base. The entire information base records are organized consecutively with the Customer ID. Before storing new data into database by examine place by identification and a short time later by embed new segment with the objective that masterminding demand should be kept up. Single-level information square relies upon activity key, and the information is kept in

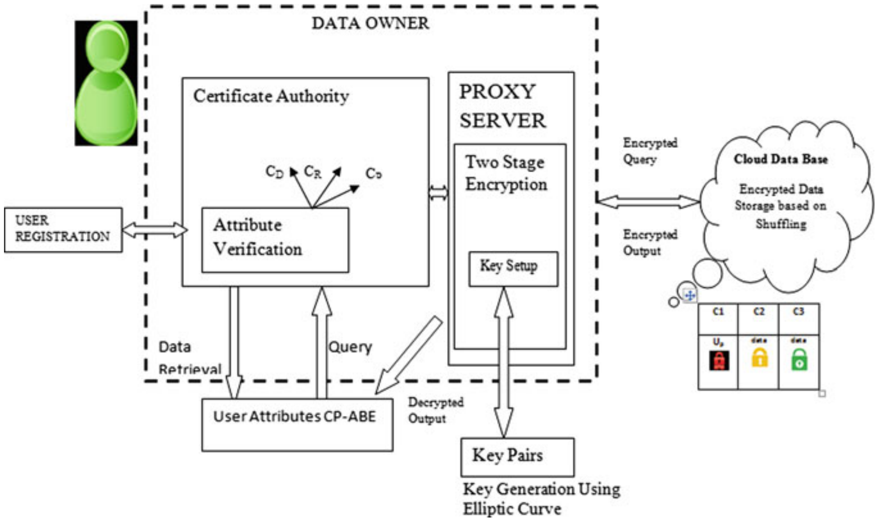


Fig. 1 Proposed methodology

squares. All information kept in the database is named as transparent record. Right when the record is recovered from the information base, the information is changed to dark. The flooded list rearranging is not required where transparent stamped records are not recovered, and it is highly unlikely of spillage with transparent records. Dark checked information is revamped high for entire single list information base after each rearranging is finished. By rearranging the dark records is a constrained mix strategy with the different information squares which outfit information mystery and brisk questioning with the ordering [10, 11].

Protecting the pathological information initially by the access control plan of action is used based on the user attributes which is CP-ABE and the information is encrypted [12] by AES techniques by separating the security domain into multiple areas one is non-public domain and another is public domain where cardinal-independent encryption schemes are used for different domain, one is Vigenere encryption used for non-public domain, and two-fish algorithm is used for PUD, respectively, as shown in Fig. 2. The chance of information spillage of third-party database provider of regular avenue example of records, to beat that the restricted mix, is utilized with single square stockpiling and high security is given with the method [13]. This strategy includes less key generation time, encryption time and unscrambling period much as appeared differently in relation to spare CP-ABE plans, eventually, centered on the distinctive encryption calculation to make sure about pathological information.

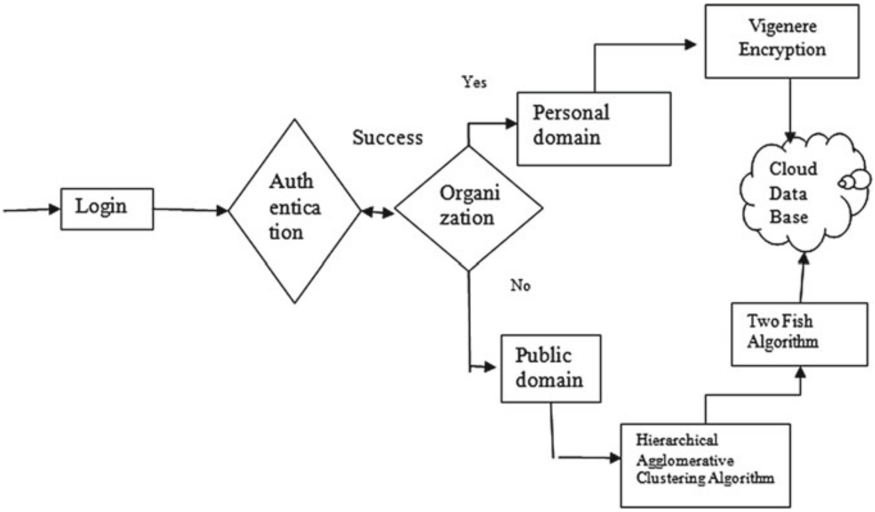


Fig. 2 Multistage encrypt-or

7 Result Analysis

This section outlines results of the recommended schemes on the personal health records (PHR). The entire recommended methodology is accomplished over the PHR data set, and its performance is evaluated through the performance metrics such as encryption time, decryption time, time taken to generate non-public key and hot data confidentiality [14].

The overall research work is implemented in four phases to meet the defined objectives.

1. Dual optimization through hybridizing CP-ABE and AES.
2. CEASE—Improving data confidentiality and developing resistance to inference attacks through hybridizing CP-ABE, AES and implementing limited-shuffle.
3. Multistage encryption—Reduction in computational overhead by using multi-stage encryption on hybridized CP-ABE with AES.
4. Integrating the multistage encryption model with limited-shuffle to further reduce the computational overhead [15].

The information is scrambled before re-appropriating onto the cloud with symmetric encryption using AES. This mechanism will restrict the unauthorized users from accessing the data, and the administrator cannot decrypt the data as they are not given access to keys. By utilizing this recommended model, the information is made sure about AES encryption and CP-ABE containment strategy. CP-ABE with bilinear mapping is in contrast to CP-ABE with AES on different parameters. The key age time is decreased utilizing the recommended system. It is seen that CP-ABE with bilinear mapping is procuring tremendous time to generate key than the CP-ABE with AES. The plain information is scrambled before re-appropriating the information in cloud to shield the information from the pernicious manager. The encryption time is diminished in CP-ABE with AES in contrast to CP-ABE with bilinear mapping and KP-ABE. The customer needs to unscramble the information, and the decoding times for CP-ABE with AES are lessened in contrast to the CP-ABE with bilinear mapping and KP-ABE. Ciphertext varies less and has more safety measures in both recommended and existing techniques. Furthermore, the recommended CEASE calculation makes sure that the vigorous admin of third-party database cannot recover any (hot) information from the delicate records.

The CEASE scheme enforces the recommended method performance. This method is resolved by various levels in the access control plan of action, encrypted database to store in third-party database and limited-data shuffling. The performance of CEASE scheme is in contrast to the encryption scheme integrated with an access control (EIAC). Querying cost is defined as the time taken to fetch the query result against encrypted database, data encryption and decryption time. The querying cost is slightly increased with the database size, but the storage overhead is less and hot data confidentiality is in more contrast to the existing methods.

Thirdly multistage encrypt-or model is tested on the personal health records. As the number of users on the public domain may be high, securing the data access is a complex issue. Hence, two-stage encryption model is developed. For a user located in the personal domain, this approach adopts Vigenere encryption algorithm, and for a user located in the public domain this approach adopts two-fish encryption algorithm. For each personal domain, the data proprietor is connected in a chain fashion through his/her generations and dear one, which may retrieve personal records in view of access given by the data proprietor. Here every data proprietor uses Vigenere encryption algorithm, maintains the decrypting key and requires sanctions of his/her wards in his/her personal. The key generation is completely carried through the elliptic curve method. The generation of key pairs is only allowed after the authentication of the user.

In the public domain, the users are clustered through HAC algorithm. Based on the roles and responsibilities of the users, they are clustered into some groups in a hierarchical fashion. Finally, the performance is measured through the performance metrics such as encryption time, decryption time, storage requirement and clustering accuracy for varying data sizes.

Multistage with two-fish and Vigenere encryption is in contrast to the existing blowfish algorithm. Further the multistage method is evaluated through clustering accuracy. Here the clustering accuracy is measured as the number of users grouped into public and personal domains. Since the clustering also plays an important role in the security provision, the performance of recommended approach is measured by varying the data size, and for every instant the clustering accuracy is measured and formulated. Two-fish and Vigenere encryption and decryption time are in contrast to the existing blowfish algorithm, and results are tabulated [16, 17].

Finally, an integrated approach using CP-ABE with AES and multistage encrypt-or exhibited high performance through limited-shuffle. The final model is constructed by merging the CP-ABE with AES, multistage encrypt-or and limited-shuffle. The main objective of CP-ABE with AES is to perceive the malicious clients or administrator and not giving access for unauthorized users to analyze the data in the cloud. This has demonstrated that the recommended strategy encryption time and unscrambling time are not exactly the other CP-ABE plans. To encode the information before outsourcing by isolating the PUD and PSD, encryption is diminished, comparably unscrambling time likewise decreased in this recommended strategy.

In key generation time, secret keys with various numbers of set of attributes which match the equivalent sets based on the leafy nodes, and the keys are generated as shown in Fig. 3. Figures 4 and 5 show the encryption and decryption time for various database sizes, respectively, with the matched attributes sets. Figure 6 shows the

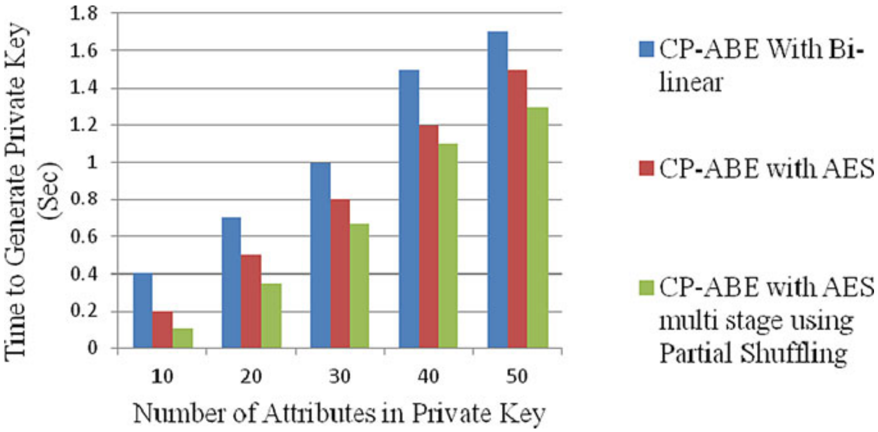


Fig. 3 Key generation time (MS)

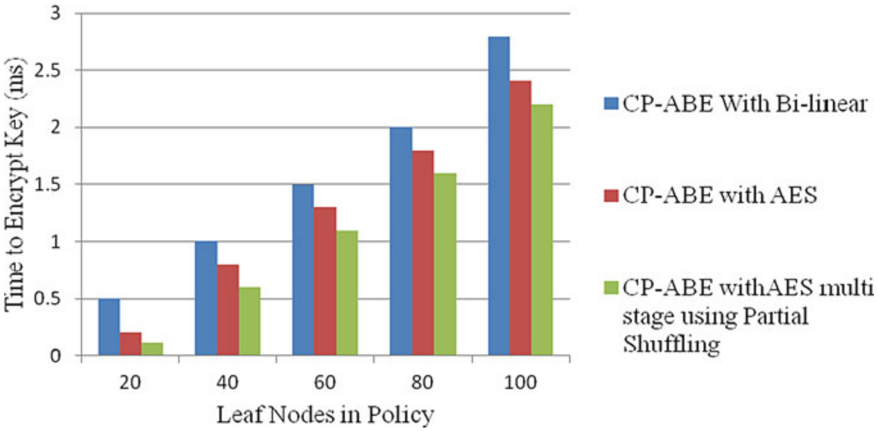


Fig. 4 Encryption time (MS)

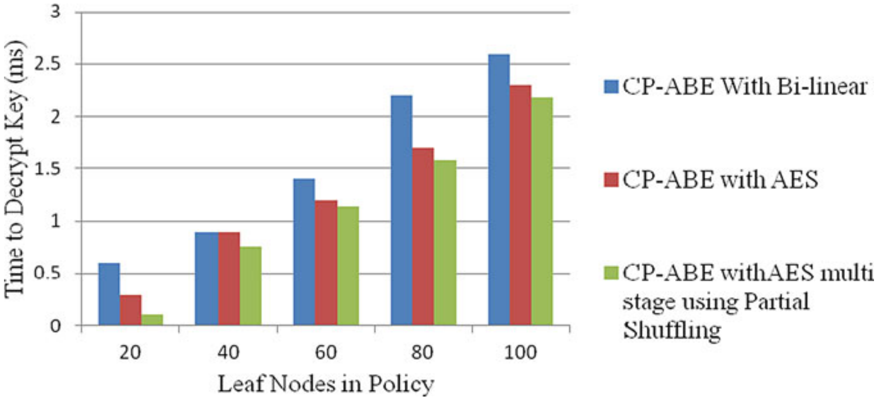


Fig. 5 Decryption time (MS)

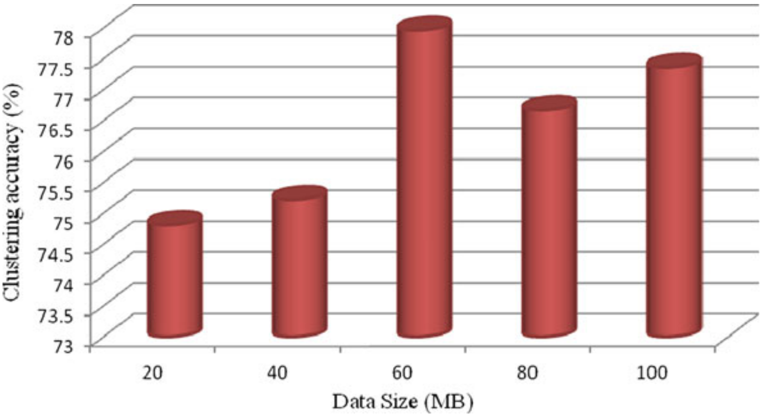


Fig. 6 Data size versus clustering accuracy

clustering accuracy for different database sizes with HAC algorithm. The multi-stage encrypt-or helps in reducing the additional computational overhead which is acquired by separating all users into clusters. This methodology endeavors to shield the electronic records from inference attacks through the accomplishment of limited-shuffle [18, 19]. The key generation and information retrieval time are limited in the developed model in contrast to various CP-ABE and KP-ABE techniques.

8 Conclusions

This considered the implementation of CP-ABE with AES and two-stage encrypt-or exhibiting high performance through limited-shuffle. In this work, initially CP-ABE with AES is developed and proved reduced key generation time. Secondly, CEASE is developed to use query encryption method to retrieve results from database. Thirdly, multistage encryption model using two-fish and Vigenere is developed.

Here the users are divided into non-public and public domains, and HAC clustering is used to further divide the users into groups. This work is evaluated by performance metric and clustering accuracy [20]. Performance of HAC approach is also measured by varying the data size for every instance. Finally, the integrated model is developed using CP-ABE with AES and multistage encryption model through limited-shuffle. Final work will endeavor to shield the electronic records from inference attacks through the accomplishment of limited-shuffle. This model has majorly addressed the below challenges, namely dual optimization to meet the equality between the data confidentiality, and processing time is achieved with AES and limited-shuffle-based ABE. Increased hot data confidentiality and privacy preservation are achieved by reducing the effect of curious/malicious authority using an encryption algorithm CEASE. Resilient to inference attacks from malicious authority is attained by information access arrangement familiarity by not changing the querying process using single-level block index method along with limited-shuffle. Reduced computational overhead to reduce the unnecessary computational overhead is achieved by HAC algorithm mechanism based on the roles of users. Future work is to reduce the cost of recommended model by increasing the security in data.

Acknowledgements Thank you for your cooperation and the contribution of co-authors and Springer for publishing the manuscript.

References

1. Lewko A, Waters B (2004) Decentralizing attribute-based encryption, EUROCRYPT, pp 223–238.
2. Waters B (2011) ciphertext-policy attribute-based encryption: an expressive, efficient, and provably secure realization, public key cryptography–PKC. Springer, Berlin, pp 53–70
3. Boneh D, Boyen X, Goh EJ (2005) Hierarchical identity based encryption with constant size cipher text. Advances in cryptology—EUROCRYPT vol 3493, pp 440–456
4. Bettencourt J, Sahai A, Waters B (2003) Ciphertext-policy attribute based encryption. In: IEEE Symposium on security and privacy (SP), pp 321–334
5. Bobba R, Khurana H, Prabhakaran M (2009) Attribute-sets: A practically motivated enhancement to attribute-based encryption. In: European symposium on research in computer security, pp 587–604
6. Chase M (2007) Multi-authority attribute based encryption. Spring Theory Cryptograph 4392:515–534
7. Ferretti L, M Colajanni, M.Marchetti (2013) Access control enforcement on query-aware encrypted cloud databases, IEEE 5th international conference on cloud computing technology and science (CloudCom), vol 2, pp 219–219
8. Goyal V, Pandey O, Sahai A, Waters B (2006) Attribute-based encryption for fine-grained access control of encrypted data. In: Proceedings of the 13th ACM conference on computer and communications security, ACM, pp 89–98
9. Muller S (2008) Distributed attribute based encryption. J Inf Secur Cryptol 4:20–36

10. Mell P, Grance T (2011) The NIST definition of cloud computing. National Institute of Standards and Technology Special Publication. 53:1–7
11. Sahai A, Waters B (2005) Fuzzy identity-based encryption. In: Advances in cryptology, EUROCRYPT-2005, Springer, pp 557–573
12. Muller S, Katzenbeisser S, Eckert C (2009) On multi-authority ciphertext-policy attribute-based encryption. Bull Korean Math 46(4):803–819
13. Chennam KK, Muddana L (2018) An efficient two stage encryption for securing personal health records in cloud computing. Int J Serv Oper Inf 9(4):277–296
14. di Vimercati SD, Foresti S, Paraboschi S, Pelosi G, Samarati P (2014) Protecting access confidentiality with data distribution and swapping. In: Proceedings of IEEE 4th international conference on big data and cloud computing (BDCLOUD), pp 167–174
15. Yang K, Zhang J, Zhang W, Qiao D (2011) A light-weight solution to preservation of access pattern privacy in un-loyal clouds. In: Proceedings of the European conferences on research in computer security (ESORICS), pp 528–547
16. Wan Z, Liu J, Deng RH (2012) HASBE: A Hierarchical Attribute-Based Solution for Flexible and Scalable Access Control in Cloud Computing. IEEE Trans Inf Forensics Secur 7(2):743–754
17. Hao R, Yang H, Zhou Z (2019) Driving behaviour evaluation model base on big data from internet of vehicles. Int J Ambient Comput Intell 10:78–95. https://doi.org/10.4018/IJACI.2019100105
18. Chandrakar P (2019) A secure remote user authentication protocol for healthcare monitoring using wireless medical sensor networks. Int J Ambient Comput Intell 10:96–116. https://doi.org/10.4018/IJACI.2019010106
19. Das SK, Samanta S, Dey N, Kumar, R (2020) Design frameworks for wireless networks, lecture noted in network and systems, Springer.
20. De D, Mukherjee A, Kumar Das S, Dey N (2020) Nature Inspired computing for wireless sensor networks. Springer Tracts in Nature-Inspired Computing, Springer.

Design of Robust Smartcard-Based User Anonymous Authentication Protocol with AVISPA Simulation

Rifaqat Ali and Preeti Chandrakar

Abstract Recently, Byun presented a privacy maintaining smartcard-based authentication protocol with provable security. We analyze and identify that his scheme is suffering from online password guessing threat, replay threat, and privileged insider threat. It is also not providing user-anonymity and password change phase. To eliminate these above-mentioned security issues, we have designed an extended user anonymous authenticated session key agreement protocol using smartcard. The scalability of our scheme is measured in both formal and informal ways. The formal validation of our scheme has done using Burrows-Abadi-Needham (*BAN*) logic. Also, simulation is done by automated validation of Internet security protocols and applications (*AVISPA*) tool. Informal security analysis ensures that our scheme resists to various kinds of fraudulent attacks. The proposed scheme does not only hold up aforementioned security attacks, but also achieves some security features like user-anonymity and easy-to-use password change phase. Our protocol is comparatively more efficient than other schemes in the terms of costs and estimated time.

Keywords Authentication · *AVISPA* · *BAN* logic · Security attacks

1 Introduction

Authentication is the procedure of verifying the legitimacy of involved entities in any communication protocol. Authentication is based on the evidence presented by the claimed identity [1–4]. This evidence is known as an authentication factor.

R. Ali (✉)
Department of Mathematics and Scientific Computing, National Institute of Technology, Hamirpur, Hamirpur 177005, Himachal Pradesh, India
e-mail: rifaqatali27@gmail.com

P. Chandrakar
Department of Computer Science and Engineering, National Institute of Technology, Raipur, Raipur 492010, India
e-mail: preet29.chandrakar@gmail.com

S. K. Das et al. (eds.), *Architectural Wireless Networks Solutions and Security Issues*, Lecture Notes in Networks and Systems 196,
https://doi.org/10.1007/978-981-16-0386-0_3

There are multiple types of factors used to authenticate users nowadays. Schemes that use the knowledge factor such as passwords or pins are known as one-factor schemes. Two-factors schemes utilize something owned by a user or possession factors like some token or smartcard [5, 6]. Whereas three-factor authentication schemes take advantage of a user's biometric data such as fingerprints, voice, iris, or voice pattern [7]. The beneficial of biometric key is that it cannot be forgotten, difficult to reproduce and distribute. Biometric key is very hard to guess and not easy to break than password. So, biometric along with password-based authentication schemes have main role in information security's field.

In very beginning, an authentication scheme is projected by Lamport [8] in which server can validate user on basis of identity and password via unreliable channel. So, many researchers have been designed safe and reliable two factor based authentication scheme using idea of Lamport's article. Since then, Li and Hwang [9] developed a well-organized authentication protocol based on biometric and smartcard. They asserted that computation cost is comparatively less than other schemes and very confident to defend all types of wicked security threats. In 2010, Li et al. [10] recognized that Li and Hwang's scheme [9] is not capable to withstand man-in-the-middle and impersonation attacks. To cure these security problems, Li et al. developed an improved authentication scheme which is more brawny and applicable for real life applications. But, Das [11] had done cryptanalysis of Li et al.'s scheme and find out not correct authentication and also unable to modify new password properly. To overcome these vulnerabilities, Das projected an improved biometric-based authentication protocol.

Turkanovic et al. [12] and Karuppiah et al. [13] offered an authentication and key agreement schemes using a smartcard for remote login. The use of smartcards provides two-factor authentication which requires a legal smartcard along with a password for a successful login. In year 2015, Kalra and Sood [14] put forward an ECC-based authentication and key agreement scheme for IoT and cloud server using encrypted cookie. Later in 2016 Farash et al. [15] identified that [12] experience stolen smart card attack and man-in-the-middle attack. Furthermore, this scheme lacks security features like untraceability and forward/backward secrecy; therefore, they came up with a more secure scheme to enhance Turkanovic's scheme. Kaul et al. [16] addressed that the scheme in [17] is weak as many of the security parameters like the user's password and the secret key of the server can be easily obtained by an adversary, and also, it does not offer perfect forward secrecy. Hence, they suggested an enhanced version of the scheme [17].

Kumari et al. [18] build up a symmetric key, password and smartcard based authentication scheme and declared an improved protocol provides anonymity while hold up all acknowledged attacks. But, Chaudhry et al. [19] identified that the scheme [18] is still suffering from user-anonymity problem and smartcard stolen attack. Since then, they build up an updated scheme to surmount the problems of Kumari et al.'s protocol. They proved that the scheme is able to defend all types of known attacks and also provides confidentiality and anonymity. Later, Radhakrishan et al. [20] showed that [21] lacks the user anonymity feature and local password verification. Besides, it is prone to replay attack and offline password guessing attack. In the same year, Wu et al. [22] presented an authentication protocol for e-healthcare application. Kumari et al. [23] identified some shortcomings in Kalra and Sood's scheme [14] and offered an enhanced authentication framework for IoT and cloud servers. Karuppiah et al. [24] pointed out that [16] suffers from several limitations like user anonymity, perfect forward secrecy, and it is susceptible to offline password attack. Also Karuppiah et al. [24] found some security flaws in [25].

In 2015, Byun [26] proposed privacy maintaining smartcard-based authentication protocol. He asserted that his scheme provides a session-key and mutual-authentication. Despite that, the protocol guarantees the privacy of user. However, we have scrutinized Byun's scheme and pointed out that his protocol is not capable to hold up online password guessing pitfall, replay threat, privileged insider attack, and also not providing user-anonymity. Moreover, password change or update is not available in Byun's scheme. However, updating or changing password is broadly suggested for making high secure applications. So, keeping in mind these before-mentioned security weaknesses, we have introduced an extended user anonymous authenticated session-key agreement protocol using smartcard.

1.1 Our Contributions

(i) In this manuscript, we have analyzed Byun's protocol [26] and discussed its security pitfalls such as on-line password guessing attack, replay attack, privileged insider attack and does not provide user anonymity. It is also unable to provide password change or update phase as quick demand of the user.

(ii) To eliminate these security pitfalls, we have presented an extended user anonymous authenticated session-key agreement protocol using smartcard.

(iii) We have done formal security analysis using *BAN* logic. Moreover, simulation verification is done by *AVISPA* software.

(iv) Authors have proved resilience of possible security attacks of presented protocol in informal security analysis.

(v) Performance evaluation shows that the presented scheme withstands several kinds security attacks and also provides better complexities in terms of overheads and time than others [23, 26–30].

1.2 Layout of This Manuscript

Sections 2 and 3, show revisiting and important security threats in Byun's scheme. In Sect. 4, proposed scheme is explained. In Sect. 5, security analysis is done. In Sect. 6, informal security analysis presented. In Sect. 7, performance evaluation is given. At the last, conclusion is presented in Sect. 8.

2 Revisiting Byun's Scheme

Byun's scheme [26] is made of three phases like registration, login, and authentication. All phases are summarized with complete informations sequentially.

2.1 Registration Phase

In this part, suppose that there are n users. Server S selects secret keys $s_i \in \mathbb{Z}_q^\star$ for user $U_i (1 \leq i \leq n)$. All keys are free to each other and used for encryption of an identifier I_i. S keeps n records such as $\psi = (I_i, s_i, s)$ for $1 \leq i \leq n$ in a database.

Step 1: U_i put forwards I_i and password pw_i to S via reliable channel.

Step 2: Upon getting (I_i, pw_i) from U_i, S calculates two parameters $\lambda = (\alpha, \beta)$ such as $\alpha = g^{s_i} + \mathcal{H}(\text{pw}_i)$ and $\beta = g^s + \mathcal{H}(\text{pw}_i)$.

Step 3: S stores (I_i, λ, H) into memory of smartcard SC_i and produces SC_i to U_i. Here, H is set of hash functions such as $H = (\mathcal{F}(.), \mathcal{G}(.), \mathcal{H}(.), \mathcal{J}(.))$ and $\mathcal{H}$: $\{0, 1\}^* \rightarrow \mathbb{G}$. Although, all $\mathcal{F}$, $\mathcal{G}$ and $\mathcal{J}$ are cryptographic hash functions from $\{0, 1\}^* \rightarrow \{0, 1\}^l$.

2.2 Login and Authentication Phase

First of all, user U_i inserts smartcard SC_i into device then U_i keys own identity I_i and password pw_i. Some steps are executed between SC_i and S.

Step 1: SC_i computes $\alpha^{'} = \alpha - \mathcal{H}(\text{pw}_i)$ and $\beta^{'} = \beta - \mathcal{H}(\text{pw}_i)$. If inputing values are correct, then $\alpha^{'} = g^{s_i}$ and $\beta^{'} = g^s$. For finding an encryption key K, SC_i uses $\beta^{'}$ and also creates a random number a_1, a_2 from $\mathbb{Z}_q^*$ and then computes g^{a_1}, g^{a_2} and $K = \mathcal{J}(S, g^{a_1}, \beta^{'}, (\beta^{'})^{a_1})$. SC_i also determines $(\alpha^{'})^{a_2} = g^{a_2 s_i}$ and then creates a set of messages $M = [A, B, C] = [g^{a_2 s_i}, I_i, (g^{a_1}, g^{a_2})]$. After that SC_i encrypts M with K such that $E = E_{\text{K}}(g^{a_2 s_i}, I_i, g^{a_1}, g^{a_2})$ and sends it to S together with $C = (C_1, C_2) = (g^{a_1}, g^{a_2})$.

Step 2: After getting E and C, S calculates $K = \mathcal{J}(S, C_1, g^s, (C_1)^s)$ and decrypt $E_{\text{K}}(M)$ and then S gets $[A, B, C] = [g^{a_2 s_i}, I_i, (g^{a_1}, g^{a_2})]$. With the help of I_i, S computes secret key s_i from the list ψ and matches the condition $(C_2)^{s_i} = A$. If this true, then S also matches the obtained C is equal to decrypted C. If both are true, then S succeeds to authenticate U_i. Otherwise, session is terminated. Afterwards, S chooses a random value $b \in \mathbb{Z}_q^*$ and calculates the values of F, C' and K_s such that $F = \mathcal{F}(I_i \parallel S \parallel C_2 \parallel C^{'} \parallel K_s \parallel g^{s_i})$, $C^{'} = g^b$, $K_s = (C_2)^b = g^{a_2 b}$ and transmits $(F, C_{'})$ to U_i.

Step 3: After receiving $(F, C^{'})$ from S, U_i verifies $\mathcal{F}(I_i \parallel S \parallel C_2 \parallel C^{'} \parallel K_u \parallel \alpha^{'})$ = F, where $K_u = g^{a_2 b}$. If this is true then S is authenticated. Otherwise, session rejected. Finally, U_i and S agreed for mutual authentication and procreate session key $SK = \mathcal{G}(I_i \parallel S \parallel g^{a_2} \parallel g^{b} \parallel g^{a_2 b} \parallel g^{s_i})$.

3 Important Security Threats Found in Byun's Scheme

In this segment, we have scrutinized Byun's protocol [26] and identified several vulnerabilities which are described in subsequent subsections.

3.1 *On-Line Password Guessing Attack*

We have discovered that Byun's scheme is unable to resist online password guessing attack because if an attacker $\mathcal{A}$ take out all hidden parameters $(I_i, \lambda = (\alpha, \beta), H)$ from SC$_i$ by power analysis and also eavesdrops all messages such as (E, C) and $(F, C^{'})$ transmitting via public channel. By using these records, $\mathcal{A}$ can perform some steps as follows.

Step 1: An attacker $\mathcal{A}$ guesses password pw_i^* then determines $\alpha^* = \alpha - \mathcal{H}(\mathrm{pw}_i^*)$ and $\beta^* = \beta - \mathcal{H}(\mathrm{pw}_i^*)$.

Step 2: $\mathcal{A}$ chooses a_1^* and a_2^* as a random numbers from $\mathbb{Z}_q^*$ and enumerates encryption key $K^* = \mathcal{J}(S, g^{a_1^*}, \beta^*, (\beta^*)^{a_1^*})$ then produces $M^* = [A^*, B^*, C^*] = [g^{a_2^* s_i}, I_i, (g^{a_1^*}, g^{a_2^*})]$.

Step 3: $\mathcal{A}$ encrypts M^* by K^* such that $E^* = E_{\mathrm{K}^*}(g^{a_2^* s_i}, I_i, g^{a_1^*}, g^{a_2^*})$ then transmits E^* and $C^* = (C_1^*, C_2^*) = (g^{a_1^*}, g^{a_2^*})$ to the S.

Step 4: After acquiring the message from $\mathcal{A}$, S determines $K^* = \mathcal{J}(S, C_1^*, g^s, (C_1^*)^s)$.

Step 5: S decrypt $D_{\mathrm{K}^*}(E^*)$ and get $[A^*, B^*, C^*] = [g^{a_2^* s_i}, I_i, (g^{a_1^*}, g^{a_2^*})]$. Now, by using I_i which is already stored in server's database, S computes s_i from the list ψ and matches the condition $(C_2^*)^{s_i} = A^*$. If this is true, then S also matches the received C^* is equal to decrypted C^*. If both conditions are true then S authenticates $\mathcal{A}$ and accept login request. It means that the guessed password is correct. Otherwise, $\mathcal{A}$ repeats from Steps 1–5 until or unless get success.

3.2 *User-Anonymity*

User-anonymity certified the confidentiality of user (such as identity) from an attacker $\mathcal{A}$. Additionally, anonymity builds more strong of an authentication scheme from $\mathcal{A}$

because key role method for providing anonymity is to hidden real information like identity, password, etc., during communication. But, in Byun's scheme, user's identity is directly kept in smartcard which shows the privacy of user to $\mathcal{A}$. So, by above reasons, Byun's scheme does not equip user anonymity.

3.3 Replay Attack

Let us presume that an attacker $\mathcal{A}$ has eavesdropped a past login message (E, C), where $E = E_{\mathrm{K}}(g^{a_2 s_i}, I_i, g^{a_1}, g^{a_2})$ and $C = (C_1, C_2) = (g^{a_1}, g^{a_2})$. She/he tries to resend eavesdropped message (E, C) to S. After getting this message (E, C), S computes $K = \mathcal{J}(S, C_1, g^s, (C_1)^s)$ with the help of secret value s and decrypts $E_{\mathrm{K}}(M)$ and then get $[A, B, C] = [g^{a_2 s_i}, I_i, (g^{a_1}, g^{a_2})]$. With the help of I_i, S computes secret key s_i from the list ψ and matches the condition $(C_2)^{s_i} = A$. If this condition is true then S also matches the received C is equal to decrypted C. If both condition are true, S authenticates to $\mathcal{A}$. It means that $\mathcal{A}$ is successful to login S by replaying previous eavesdropped login message. Note that, this attack is happened due to use of random values only. Therefore, S is unable to verify originality of the received login message. So, from this justification, we can say that Byun's scheme cannot withstand replay attack.

3.4 Privileged Insider Attack

In registration phase of Byun's scheme, U_i put forwards own I_i and pw_i directly to S because S is considered as truthful, but there may be possibility that S can be as an insider attacker $\mathcal{A}$ and uses pw_i of U_i for other applications. Let U_i uses same pw_i to accessing several other applications then S can pretend as legitimate U_i in this scenario. So from this justification, we can say that Byun's scheme is not capable to hold up privileged insider attack.

3.5 Lack of Password Change Phase

Password change or update method is mandatory in authentication system because it is extensively recommended security issue for providing security from an attacker $\mathcal{A}$. But, Byun's scheme does not provide password change or update method whenever or wherever required by user. This is main security problem in Byun's scheme. **Note that:** Fixed password is absolutely more susceptible than change or update password.

4 Proposed Scheme

There are four phases in proposed protocol like registration; login; authentication; and password update or change. The overall idea of the proposed protocol is delineated in Fig. 1 and meaning of notations which are employed in proposed protocol are shown in Table 1.

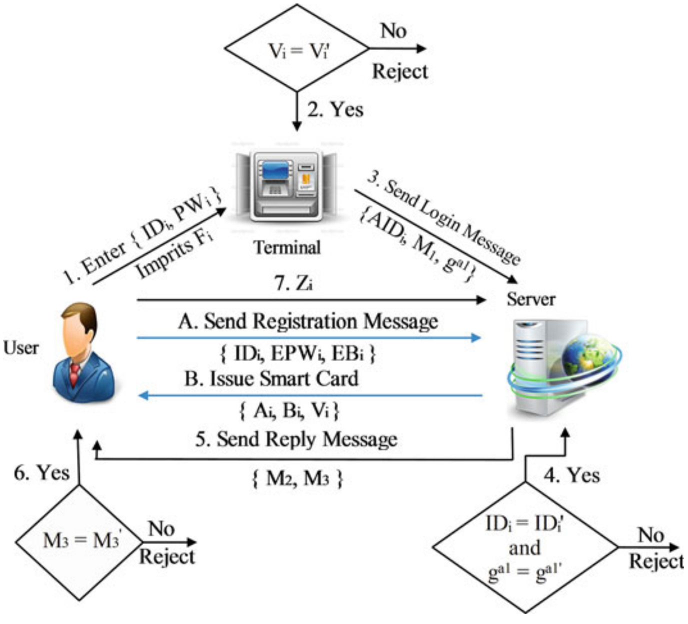

$V_i = V_i'$
No
Reject
2. Yes
Terminal
1. Enter { ID_i, PW_i } Imprits F_i
3. Send Login Message {AID_i, M_1, g^{a1}}
7. Z_i
Server
User
A. Send Registration Message
{ ID_i, EPW_i, EB_i }
B. Issue Smart Card
{ A_i, B_i, V_i }
5. Send Reply Message
{ M_2, M_3 }
6. Yes
$M_3 = M_3'$
No
Reject
4. Yes
$ID_i = ID_i'$ and $g^{a1} = g^{a1'}$
No
Reject

Fig. 1 Overview diagram of proposed scheme

Table 1 Notations list

Notation	Description
U_i, S	User and server
$\mathcal{A}$	An attacker
ID_i, PW_i, F_i	Identity, password and biometric of U_i
r	Random number
a_1	U_i's nonce
b_1	S's nonce
s	Server's secret key
$h(.)$, $H(.)$	Hash and bio-hash functions
$\oplus$ and $\|$	XOR and concatenation operations
$E(.)$	Symmetric key encryption

4.1 Registration Phase

Step 1: In this phase, the user U_i opts own identity ID_i, password PW_i and imprints his/her personal biometric F_i. Now, U_i computes $\mathrm{EPW}_i = h(\mathrm{PW}_i \parallel \mathrm{ID}_i \parallel r)$, $\mathrm{EB}_i = H(F_i \parallel r)$ and sends $\{\mathrm{ID}_i, \mathrm{EPW}_i, \mathrm{EB}_i\}$ to the server S via secure channel.

Step 2: After enlisting the parameters from U_i, S evaluates $A_i = g^s + h(\mathrm{EPW}_i \parallel \mathrm{ID}_i)$, $B_i = h(s_i) + h(\mathrm{EPW}_i \parallel \mathrm{EB}_i)$ and $V_i = h(\mathrm{ID}_i \parallel \mathrm{EPW}_i \parallel \mathrm{EB}_i)$.

Step 3: S creates a database and stores

User identity	Secret key
ID_{i1}	$E_s(s_{i1})$
ID_{i2}	$E_s(s_{i2})$
–	–
–	–
–	–
ID_{in}	$E_s(s_{in})$

Step 4: At the end, S issues smartcard storing information $\{A_i, B_i, V_i, H(.), h(.)\}$ and put forwards to U_i by using reliable channel.

Step 5: After obtaining smartcard from S, U_i computes $r_{\text{new}} = r \oplus h(\mathrm{ID}_i \parallel \mathrm{PW}_i \parallel H(F_i))$ and stores it into smartcard.

Step 6: Finally, smartcard holds information $\{A_i, B_i, V_i, r_{\text{new}}, h(.), H(.)\}$.

4.2 Login Phase

Step 1: In this part, U_i wants to login S, the following operations execute after inputting ID_i, PW_i and imprints F_i and then computes $r' = r_{\text{new}} \oplus h(\mathrm{ID}_i \parallel \mathrm{PW}_i \parallel H(F_i))$, $\mathrm{EPW}'_i = h(\mathrm{PW}_i \parallel \mathrm{ID}_i \parallel r')$, $\mathrm{EB}'_i = H(F_i \parallel r')$, $V'_i = h(\mathrm{ID}_i \parallel \mathrm{EPW}'_i \parallel \mathrm{EB}'_i)$.

Step 2: Now, smartcard reader compares $V'_i = V_i$. If this matching holds then U_i is legitimate. Otherwise, abolished the session.

Step 3: The smartcard reader computes $g^s = A_i - h(\mathrm{EPW}_i \parallel \mathrm{ID}_i)$ and $h(s_i) = B_i - h(\mathrm{EPW}_i \parallel \mathrm{EB}_i)$. Subsequently, it creates a nonce a_1 and computes $A\mathrm{ID}_i = \mathrm{ID}_i \oplus h((g^s)^{a_1}) = \mathrm{ID}_i \oplus h(g^{sa_1})$, $K = h(h(s_i) \parallel \mathrm{ID}_i)$, and $M_1 = E_K[\mathrm{EPW}_i, \mathrm{ID}_i, g^{a_1}]$.

Step 4: At last, U_i sends $\{\mathrm{AID}_i, M_1, g^{a_1}\}$ to S by public channel.

4.3 Authentication Phase

Step 1: Upon enlisting login message from U_i, S computes $(g^{a_1})^s = g^{a_1 s}$ and $\mathrm{ID}_i = \mathrm{AID}_i \oplus h(g^{a_1 s})$.

Step 2: Now, S checks whether or not ID_i exists in database. If not then request of login message is rejected by S. Otherwise, retrieves corresponding s_i from database.

Step 3: S computes $K = h(h(s_i) \parallel \mathrm{ID}_i)$ and $M_1 = D_K[\mathrm{EPW}_i, \mathrm{ID}_i, g^{a_1}]$. Then, S compares decrypted g^{a_1} with received g^{a_1} and decrypted ID_i with computed ID_i. If it holds then S believes the authenticity of U_i. Otherwise, expired the session.

Step 4: S calculates $M_2 = E_{h(\mathrm{EPW}_i \parallel h(s_i))}[g^{b_1}]$, $M_3 = h[\mathrm{EPW}_i \parallel \mathrm{ID}_i \parallel (g^{a_1})^{b_1} \parallel g^s]$ and sends $\{M_2, M_3\}$ to U_i.

Step 5: After enlisting this message $\{M_2, M_3\}$ from S, U_i calculates $D_{h(\mathrm{EPW}_i \parallel h(s_i))}[M_2] = g^{b_1}$ and $M_3' = h[\mathrm{EPW}_i \parallel \mathrm{ID}_i \parallel (g^{b_1})^{a_1} \parallel g^s]$. Now, U_i compares $M_3' = M_3$, if this comparison is true, then mutual authentication is exist. Otherwise, quits the session.

Step 6: U_i computes $\mathrm{SK} = [(g^{b_1})^{a_1} \parallel \mathrm{EPW}_i \parallel h(s_i) \parallel g^s]$ and $Z_i = h(\mathrm{SK} \parallel \mathrm{ID}_i)$. Thereafter, U_i sends $\{Z_i\}$ to S.

Step 7: After enlisting the message $\{Z_i\}$ from U_i, S determines $\mathrm{SK}' = [(g^{a_1})^{b_1} \parallel \mathrm{EPW}_i \parallel h(s_i) \parallel g^s]$ and $Z_i' = h(SK' \parallel \mathrm{ID}_i)$.

Step 8: Now, S compares $Z_i' = Z_i$, if this comparison holds, then SK is verified by S.

4.4 Password Update or Change Phase

If suppose user's password divulges or steals or leaks because of many reasons by third party, then this phase is used. In this circumstances, it is necessary to update password promptly. So, our protocol facilitates to update or change password user-friendly without intervene of S. Some steps are described below to perform this phase.

Step 1: When U_i enters own ID_i, PW_i and personal biometric F_i, then card-reader computes $r' = r_{\mathrm{new}} \oplus h(\mathrm{ID}_i \parallel \mathrm{PW}_i \parallel H(F_i))$, $\mathrm{EPW}_i' = h(\mathrm{PW}_i \parallel \mathrm{ID}_i \parallel r')$, $\mathrm{EB}_i' = H(F_i \parallel r')$ and $V_i' = h[\mathrm{ID}_i \parallel \mathrm{EPW}_i' \parallel \mathrm{EB}_i']$.

Step 2: Now, smartcard reader checks whether or not $V_i' = V_i$, if this is true then performs next step. Otherwise, quits slot.

Step 3: U_i enters new password $\mathrm{PW}_i^{\mathrm{new}}$ and smartcard reader computes $r_{\mathrm{new}}' = r_{\mathrm{new}} \oplus h(\mathrm{ID}_i \parallel \mathrm{PW}_i) \parallel H(F_i)) \oplus h(\mathrm{ID}_i \parallel \mathrm{PW}_i^{\mathrm{new}} \parallel H(F_i))$, $\mathrm{EPW}_i' = h(\mathrm{PW}_i^{\mathrm{new}} \parallel \mathrm{ID}_i \parallel r')$, $\mathrm{EB}_i' = h(F_i \parallel r')$, $V_i' = h(\mathrm{ID}_i \parallel \mathrm{EPW}_i' \parallel \mathrm{EB}_i')$, $A_i' = A_i - h(\mathrm{EPW}_i \parallel \mathrm{ID}_i) \oplus h(\mathrm{EPW}_i' \parallel \mathrm{ID}_i)$ and $B_i' = B_i - h(\mathrm{EPW}_i \parallel \mathrm{EB}_i) \oplus h(\mathrm{EPW}_i' \parallel \mathrm{EB}_i')$.

Step 4: Finally, smartcard replaces $\{A_i', B_i', V_i', r_{\mathrm{new}}'\}$ in place of $\{A_i, B_i, V_i, r_{\mathrm{new}}\}$.

5 Security Analysis

In this part, first we did security analysis using *BAN* logic. Then, simulation verification is done by *AVISPA* tool.

5.1 Validation of Proposed Scheme Using BAN logic

In this part, the *BAN* (Burrows-Abadi-Needham) logic [31] is used to validate the proposed scheme. *BAN* logic is well recognized conventional. It is used to examine the security of protocol on the basis of mutual-authentication and session-key agreement.

Step 1. Primarily, let authentication goals of our scheme and according to *BAN* logic, these goals demonstrate the legitimacy of our scheme and then proves the protocol is more secure from several sort of known and unknown threats.

Goal 1. $U_i| \equiv (U_i \overset{\mathrm{SK}}{\longleftrightarrow} S)$
Goal 2. $U_i| \equiv S| \equiv (U_i \overset{\mathrm{SK}}{\longleftrightarrow} S)$
Goal 3. $S| \equiv (S \overset{\mathrm{SK}}{\longleftrightarrow} U_i)$
Goal 4. $S| \equiv U_i| \equiv (S \overset{\mathrm{SK}}{\longleftrightarrow} U_i)$

Step 2. Idealized forms of the presented scheme is as follows.

Message (1): $\{\mathrm{AID}_i, g^{a_1}, M_1 :< g^{a_1} >_{h(h(s_i)\|\mathrm{ID}_i)}\}$
Message (2): $\{M_3, M_2 :< g^{b_1} >_{h(\mathrm{EPW}_i\|h(s_i))}\}$

Step 3. Further, let some suppositions of the presented protocol are as follows.

$A_1 : U_i| \equiv \#\{g^{a_1}, g^{b_2}\}$
$A_2 : S| \equiv \#\{g^{a_1}, g^{b_2}\}$
$A_3 : S| \equiv S \xleftrightarrow{h(h(s_i)\|\mathrm{ID}_i)} U_i$
$A_4 : U_i| \equiv U_i \xleftrightarrow{h(\mathrm{EPW}_i\|h(s_i))} S$
$A_5 : S| \equiv U_i \Rightarrow \{g^{a_1}\}$
$A_6 : U_i| \equiv S \Rightarrow \{g^{b_1}\}$

Step 4. With the help of some above steps and predefined rules of *BAN* logic [31], the main proof is demonstrated as follows.

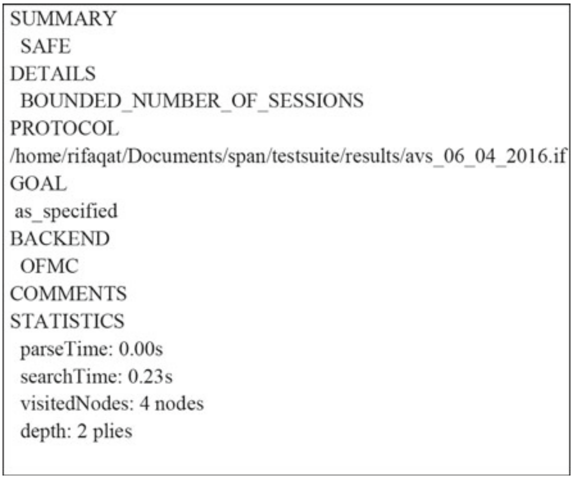
From **Message (1)**, we can see that
$S_1 : S \lhd \{\mathrm{AID}_i, g^{a_1}, M_1 :< g^{a_1} >_{h(h(s_i)\|\mathrm{ID}_i)}\}$
From A_3, S_1 and using ***Message meaning rule***, we can write
$S_2 : S| \equiv U_i| \sim \{g^{a_1}\}$

According to A_2 and by using ***Freshness conjuncatenation rule***, we get
$S_3 : S| \equiv \#\{g^{a_1}\}$.
With the help of S_2, S_3 and predefined ***Nonce verification rule***, we could obtain
$S_4 : S| \equiv U_i| \equiv \{g^{a_1}\}$
We have seen that from A_5, S_4 and ***Jurisdiction rule***
$S_5 : S| \equiv \{g^{a_1}\}$
We can get from S_4, A_2 and ***Session key rule***
$S_6 : S| \equiv (S \overset{\text{SK}}{\longleftrightarrow} U_i)$ **Goal 3**
From A_2, S_6 and predefined ***Nonce verification rule***, we obtain
$S_7 : S| \equiv U_i| \equiv (S \overset{\text{SK}}{\longleftrightarrow} U_i)$ **Goal 4**

Further, we can write from **Message (2)**
$S_8 : U_i \triangleleft \{M_3, M_2 :< g^{b_1} >_{h(\text{EPW}_i \| h(s_i))}\}$
We have seen that from A_4, S_8 and predefined ***Message meaning rule***
$S_9 : U_i| \equiv S| \sim \{g^{b_1}\}$
According to A_2, S_9 and predefined ***Nonce verification rule***, we get
$S_{10} : U_i| \equiv S| \equiv \{g^{b_1}\}$
From ***Jurisdiction rule***, A_6 and S_{10}, we can write
$S_{11} : U_i \equiv \{g^{b_1}\}$
By using A_1, S_{10} and ***Session key rule***, we see
$S_{12} : U_i| \equiv (U_i \overset{\text{SK}}{\longleftrightarrow} S)$ **Goal 1**
From A_1, S_{12} and predefined ***Nonce verification rule***, we get
$S_{13} : U_i| \equiv S| \equiv (U_i \overset{\text{SK}}{\longleftrightarrow} S)$ **Goal 2**

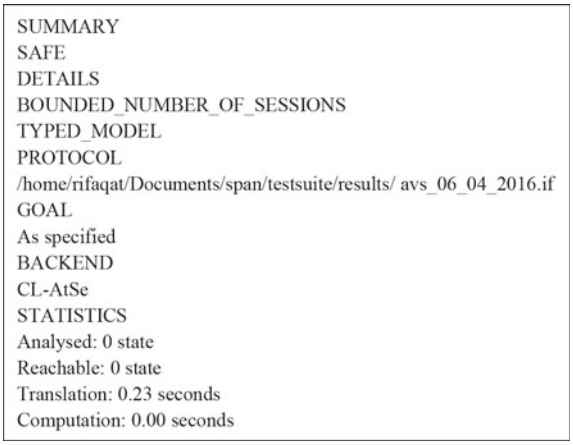
5.2 Simulation Verification by Using AVISPA

In this part, we have done the simulation verification of proposed scheme by *AVISPA* tool [32] which indicates that our scheme is safe from replay and man-in-the-middle attacks. Firstly, we have implemented our protocol using *HLPSL* (high-level protocol specification language) in three roles such as user U_i; server S_j; and session, goals, and environment. Then, *HLPSL* code is executed with the help of *AVISPA* tool. The simulation outputs of our protocol in the context of *OFMC* (on-the-fly Model Checker) and *CL-AtSe* (constraints logic-based attack searcher) back-ends are delineated in Figs. 2 and 3. The outputs are clearly showing that our scheme is safe.



```
SUMMARY
  SAFE
DETAILS
  BOUNDED_NUMBER_OF_SESSIONS
PROTOCOL
/home/rifaqat/Documents/span/testsuite/results/avs_06_04_2016.if
GOAL
 as_specified
BACKEND
  OFMC
COMMENTS
STATISTICS
  parseTime: 0.00s
  searchTime: 0.23s
  visitedNodes: 4 nodes
  depth: 2 plies
```

Fig. 2 Simulation result in terms of *OFMC*



```
SUMMARY
SAFE
DETAILS
BOUNDED_NUMBER_OF_SESSIONS
TYPED_MODEL
PROTOCOL
/home/rifaqat/Documents/span/testsuite/results/ avs_06_04_2016.if
GOAL
As specified
BACKEND
CL-AtSe
STATISTICS
Analysed: 0 state
Reachable: 0 state
Translation: 0.23 seconds
Computation: 0.00 seconds
```

Fig. 3 Simulation result in terms of *CL-AtSe*

6 Informal Security Analysis

Here, we showed the presented protocol is able to resist various sorts of attacks.

6.1 Identity and Password Guessing Attack

Let us assume that U_i uses a low entropy identity ID_i and password PW_i which is easily guessable or breakable in polynomial time. However, in our protocol, an attacker $\mathcal{A}$ is unable to guess ID_i and PW_i of U_i with the help of smart card parameters $\{A_i, B_i, V_i, r_{\text{new}}\}$ and communicated messages $\{\text{AID}_i, M_1, g^{a_1}, M_2, M_3, Z_i\}$ between

U_i and S. The brief explanation for resisting identity and password guessing attack as follows.

(i) If suppose $\mathcal{A}$ verifies guessed ID_i and PW_i from $A_i = g^s + h(\mathrm{EPW}_i \parallel \mathrm{ID}_i)$, where $\mathrm{EPW}_i = h(\mathrm{PW}_i \parallel \mathrm{ID}_i \parallel r)$. Now, $\mathcal{A}$ has to speculate four unknown parameters $\{\mathrm{ID}_i, \mathrm{PW}_i, r, s\}$ at one time, which is not feasible in polynomial time. Moreover, ID_i, PW_i and r all are protected by hash function which does not exist inverse.

(ii) If $\mathcal{A}$ tries to achieve ID_i and PW_i from $B_i = h(s_i) + h(\mathrm{EPW}_i \parallel \mathrm{EB}_i)$, where $\mathrm{EPW}_i = h(\mathrm{PW}_i \parallel \mathrm{ID}_i \parallel r)$, $\mathrm{EB}_i = H(F_i \parallel r)$. It is clear that $\mathcal{A}$ cannot acquire secret key s_i due to non-invertible hash function which is stored in server's database only. Furthermore, $\mathcal{A}$ cannot estimate four unknown values $\{\mathrm{ID}_i$, PW_i, F_i, $r\}$ at same time which is not possible in polynomial time.

(iii) Similarly, $\mathcal{A}$ cannot speculate ID_i and PW_i from this equation $V_i = h(\mathrm{ID}_i \parallel \mathrm{EPW}_i \parallel \mathrm{EB}_i)$, where $\mathrm{EPW}_i = h(\mathrm{PW}_i \parallel \mathrm{ID}_i \parallel r)$, $\mathrm{EB}_i = H(F_i \parallel r)$. Here, $\mathcal{A}$ has to enumerate four unknown parameters, i.e., $\{\mathrm{ID}_i, \mathrm{PW}_i, F_i, r\}$ at same time which not feasible in polynomial time. All these unknown parameters also protected by hash function.

(iv) From $r_{\mathrm{new}} = r \oplus h(\mathrm{ID}_i \parallel \mathrm{PW}_i \parallel H(F_i))$, $\mathcal{A}$ cannot guess ID_i, PW_i, F_i and r at one time. In addition, biometric feature F_i protected by bio-hash function and also secured by hash function.

(v) From $A\mathrm{ID}_i = \mathrm{ID}_i \oplus h((g^s)^{a_1}) = \mathrm{ID}_i \oplus h(g^{sa_1})$, $\mathcal{A}$ cannot predict three unknown values i.e. ID_i, server's secret key s and nonce a_1 at same time.

(vi) Similarly, $\mathcal{A}$ cannot enumerate ID_i and PW_i from remaining communicated messages, i.e., $\{$ M_1, g^{a_1}, M_2, M_3, $Z_i\}$ cause of same reason which we have explained in above steps.

6.2 User Impersonation Attack

We have supposed that $\mathcal{A}$ trapped login message $\{A\mathrm{ID}_i, M_1, g^{a_1}\}$ and then do some modification in login message. he/she tries to imitate as a genuine user U_i. But in our scheme, $\mathcal{A}$ cannot impersonate as a authorized user because of subsequent reasons.

(i) $\mathcal{A}$ attempts to calculate $\mathrm{AID}_i = \mathrm{ID}_i \oplus h((g^s)^{a_1}) = \mathrm{ID}_i \oplus h(g^{sa_1})$. It is very obvious that AID_i depends on server's secret key s and a nonce a_1. So, $\mathcal{A}$ cannot enumerate AID_i without the knowing these three unaware parameters $\{\mathrm{ID}_i$, s, $a_1\}$.

(ii) For enumerating equation $M_1 = E_K[\mathrm{EPW}_i, \mathrm{ID}_i, g^{a_1}]$, where $K = h(h(s_i) \parallel \mathrm{ID}_i)$ and $\mathrm{EPW}_i = h(\mathrm{PW}_i \parallel \mathrm{ID}_i \parallel r)$. Here, $\mathcal{A}$ has to know K, EPW_i but these values rely on ID_i, PW_i, r and a_1 which is not possible to guess these values at one time in polynomial time.

6.3 Server Impersonation Attack

If $\mathcal{A}$ wants to imitate as a server S, then he/she attempts to find out the values of communicated messages $\{M_2, M_3\}$ and $\{ Z_i\}$. After that S does some modification and tries to act as legal server. However, our protocol is capable to resist server impersonation attack due to various reasons as given below.

(i) For calculating $M_2 = E_{h(\text{EPW}_i \| h(s_i))}[g^{b_1}]$, where $\text{EPW}_i = h(\text{PW}_i \parallel \text{ID}_i \parallel r)$. Now, $\mathcal{A}$ has to know ID_i, PW_i, nonce b_1 and secret key s_i. Since all values are unknown for $\mathcal{A}$ which is not possible to guess at one time, so $\mathcal{A}$ cannot speculate M_2.

(ii) To compute this $M_3 = h[\text{EPW}_i \parallel \text{ID}_i \parallel (g^{a_1})^{b_1} \parallel g^s]$, $\mathcal{A}$ has to know values of parameters $\{\text{PW}_i, \text{ID}_i, a_1, b_1, s\}$. But, without knowledge of these parameters, $\mathcal{A}$ cannot calculate M_3.

(iii) Now for evaluating equation $Z_i = h(SK \parallel \text{ID}_i)$, where $SK = [(g^{b_1})^{a_1} \parallel \text{EPW}_i \parallel h(s_i) \parallel g^s]$. This equation depends on s, s_i, ID_i,PW_i, ID_i, a_1 and b_1 which is infeasible to guess in polynomial time. Therefore, $\mathcal{A}$ cannot computes Z_i.

6.4 Privileged Insider Attack

This is very serious attack in authentication system. Several schemes is broken by insider attack. In most cases, U_i utilized one password to approach different applications for his/her amenity. If malicious administrator knows user's PW_i, then he/she tries to approach another account of U_i. But, in our scheme, U_i put forwards only ID_i for registration but not PW_i directly to S. So, S is not aware about U_i's password. Therefore, our scheme is not easily breakable by insider attack.

6.5 Replay Attack

We presume that $\mathcal{A}$ snooped login message $\{\text{AID}_i, M_1, g^{a_1}\}$, communicated messages and attempts to impersonate as a valid U_i by sending this snooped login message after some while. But, authors have used nonce as a common countermeasure to prevent this attack. There are several reasons for withstanding replay attack as follows given below.

(i) In our protocol, login message $\{\text{AID}_i, M_1, g^{a_1}\}$ incorporates nonce a_1. So, by using property of nonce, login message is unique and valid for one session.

(ii) Now, communicating message $\{M_2, M_3\}$ also includes nonce a_1, b_1 and according to property of nonce communicated message is unrepeated and valid only for one session.

(iii) The reply message $\{Z_i\}$ also encompasses nonce a_1, b_1. But, by the characteristic of nonce, the reply message is also unique and authentic for one session only.

6.6 User Un-Traceability Attack

We have assumed that $\mathcal{A}$ eavesdropped two login messages $\{\text{AID}_i, M_1, g^{a_1}\}$ and $\{\text{AID}'_i, M'_1, g^{a'_1}\}$ and try to discover legal U_i. If any value from both eavesdropped messages are same, then $\mathcal{A}$ will know that both messages sent by same U_i. But, in our scheme, this situation is not possible by some reasons. First, AID_i is computed with nonce a_1 as well as server's secret key s. So, it is unique in each session and valid only for it. Furthermore, value of M_1 is also computed using nonce a_1. Thus, it is uncommon in each session and valid only for that session.

6.7 Smartcard Stolen Attack

This attack is very influential in password-based authentication scheme. If suppose an attacker $\mathcal{A}$ theft smartcard of user and take out parameters $\{A_i, B_i, V_i, r_{\text{new}}, h(.), H(.)\}$ from smartcard. Then, $\mathcal{A}$ attempts to guess either password PW_i of U_i or tries to enumerate valid login message for impersonating as a genuine user. Moreover, $\mathcal{A}$ may also produce new smartcard using own PW_i and F_i and tries to access different applications of user. But, in our protocol, there are following reasons for resisting smartcard stolen attack.

(i) If suppose $\mathcal{A}$ verifies guessed identity and password from equation $A_i = g^s + h(\text{EPW}_i \parallel \text{ID}_i)$, where $\text{EPW}_i = h(\text{PW}_i \parallel \text{ID}_i \parallel r)$. Now, $\mathcal{A}$ has to speculate four unknown parameters $\{\text{ID}_i, \text{PW}_i, r, s\}$ at one time, which is not feasible in polynomial time. Moreover, ID_i, PW_i and r all are protected by hash function which does not exist inverse.

(ii) If $\mathcal{A}$ tries to achieve ID_i and PW_i from $B_i = h(s_i) + h(\text{EPW}_i \parallel \text{EB}_i)$, where $\text{EPW}_i = h(\text{PW}_i \parallel \text{ID}_i \parallel r)$, $\text{EB}_i = H(F_i \parallel r)$. It is clear that $\mathcal{A}$ is not able to acquire secret key s_i due to non-invertible hash function which is stored in server's database only. Furthermore, $\mathcal{A}$ cannot estimate four unknown values $\{\text{ID}_i, \text{PW}_i, F_i, r\}$ at same time which is not possible in polynomial time.

(iii) Likewise, $\mathcal{A}$ is unable to speculate identity and password from $V_i = h(\text{ID}_i \parallel \text{EPW}_i \parallel \text{EB}_i)$, where $\text{EPW}_i = h(\text{PW}_i \parallel \text{ID}_i \parallel r)$, $\text{EB}_i = H(F_i \parallel r)$. Here, $\mathcal{A}$ has to enumerate four unknown parameters, i.e., $\{\text{ID}_i, \text{PW}_i, F_i, r\}$ at same time which not feasible in polynomial time. All these unknown parameters also protected by hash function.

(iv) In addition, from $r_{\text{new}} = r \oplus h(\text{ID}_i \parallel \text{PW}_i \parallel H(F_i))$, $\mathcal{A}$ cannot guess ID_i, PW_i, F_i and r at one time and then biometric feature F_i protected by bio-hash function and also secured by non-invertible hash function.

Table 2 Security features comparison

Schemes	Ref. [27]	Ref. [28]	Ref. [29]	Ref. [33]	Ref. [34]	Ref. [30]	Ref. [35]	Ref. [23]	Ref. [26]	Our
S1	N	N	N	N	N	N	Y	Y	N	Y
S2	N	Y	N	N	Y	N	Y	N	Y	Y
S3	N	Y	Y	N	N	N	Y	Y	N	Y
S4	Y	Y	N	Y	Y	Y	Y	N	N	Y
S5	N	N	N	N	Y	N	Y	Y	Y	Y
S6	N	–	Y	Y	N	Y	Y	N	N	Y
S7	N	Y	N	N	Y	N	Y	Y	N	Y
S8	–	Y	Y	Y	Y	N	Y	Y	–	Y
S9	N	N	Y	Y	Y	Y	N	Y	N	Y
S10	N	Y	Y	Y	Y	Y	Y	Y	Y	Y
S11	N	N	Y	Y	Y	N	Y	Y	N	Y
S12	–	Y	N	Y	Y	N	Y	N	–	Y

Note $Y \Rightarrow$ yes, $N \Rightarrow$ no, *S1* ⇒ resisting password guessing attack, *S2* ⇒ resisting user and server impersonation attack, *S3* ⇒ resisting privileged insider attack, *S4* ⇒ resisting replay attack, *S5* ⇒ resisting user un-traceability attack, *S6* ⇒ provide forward secrecy, *S7* ⇒ provide session-key verification, *S8* ⇒ flaws in password change phase, *S9* ⇒ provide correct authentication, *S10* ⇒ resist to smartcard stolen attack, *S11* ⇒ preserve user-anonymity, *S12* ⇒ resist to known session-key temporary information attack

6.8 Stolen Verifier Attack

In our scheme, server S maintains a database. Suppose if database of S is hacked or its confidential informations are achieved by an attacker $\mathcal{A}$ cause of some other means. Instead of that, $\mathcal{A}$ cannot obtain value of user's secret key s_i because it is encrypted by own secret key s of S. Thus, from this justification, proposed protocol resists to stolen verifier attack.

6.9 User-Anonymity

User-anonymity means that the confidentiality or privacy or secrecy of U_i (like ID_i and PW_i) is not disclosed from an attacker $\mathcal{A}$. Furthermore, user-anonymity also makes more strong of any authentication protocol. The main merit of user anonymity is to hidden all real information like identity and password during communication via public channel. But, in our scheme, user's privacy is indirectly stored in smartcard and also not sending directly to server S via unreliable channel. So, by these reasons, we can state that proposed scheme facilitates to user anonymity.

Table 3 Performance comparison

⇓ Schemes	SCSC(in Bits)	CC(in Bits)	CCRP	CCLAP	TCC	ET(in Sec)
Ref. [27]	5280	6624	$1T_{\mathrm{H}} + 2T_{\mathrm{E}}$	$4T_{\mathrm{H}} + 6T_{\mathrm{E}}$	$5T_{\mathrm{H}} + 8T_{\mathrm{E}}$	4.1785
Ref. [28]	3392	3008	$2T_{\mathrm{H}} + 1T_{\mathrm{E}}$	$6T_{\mathrm{H}} + 6T_{\mathrm{E}}$	$8T_{\mathrm{H}} + 7T_{\mathrm{E}}$	3.6580
Ref. [29]	2688	2688	$4T_{\mathrm{H}} + 1T_{\mathrm{E}}$	$12T_{\mathrm{H}} + 10T_{\mathrm{E}}$	$16T_{\mathrm{H}} + 11T_{\mathrm{E}}$	5.8220
Ref. [33]	2368	1120	$1T_{\mathrm{H}} + 1T_{\mathrm{E}}$	$7T_{\mathrm{H}} + 5T_{\mathrm{E}}$	$8T_{\mathrm{H}} + 6T_{\mathrm{E}}$	3.1360
Ref. [34]	2688	1520	$3T_{\mathrm{H}}$	$12T_{\mathrm{H}} + 2T_{\mathrm{E}}$	$15T_{\mathrm{H}} + 2T_{\mathrm{E}}$	1.0515
Ref. [30]	4256	3008	$9T_{\mathrm{H}}$+$2T_{\mathrm{E}}$	$9T_{\mathrm{H}}$+$8T_{\mathrm{E}}$	$11T_{\mathrm{H}}$+$10T_{\mathrm{E}}$	5.2255
Ref. [35]	4736	3712	$3T_{\mathrm{H}}$	$8T_{\mathrm{H}} + 6T_{\mathrm{E}}$	$11T_{\mathrm{H}} + 6T_{\mathrm{E}}$	3.1375
Ref. [23]	480	1760	$7T_{\mathrm{H}} + 2T_{ECM}$	$7T_{\mathrm{H}} + 8T_{ECM}$	$14T_{\mathrm{H}} + 10T_{ECM}$	0.63775
Ref. [26]	2368	3744	$2T_{\mathrm{H}} + 2T_{\mathrm{E}}$	$7T_{\mathrm{H}} + 11T_{\mathrm{E}}$	$9T_{\mathrm{H}} + 13T_{\mathrm{E}} + 2T_S$	6.8079
Our	1824	2528	$6T_{\mathrm{H}} + 1T_{\mathrm{E}}$	$16T_{\mathrm{H}} + 5T_{\mathrm{E}} + 4T_S$	$22T_{\mathrm{H}} + 6T_{\mathrm{E}} + 4T_S$	3.2126

Note SCSC ⇒ smart card storage cost, *CC* ⇒ communication cost, *CCRP* ⇒ computation cost of registration phase, *CCLAP* ⇒ computation cost of login and authentication phase, *TCC* ⇒ total computation cost, *ET* ⇒ estimated time

7 Performance Evaluation

In this section, we have presented performance evaluation of our protocol among other protocols [23, 26–30, 33–35] in the context of costs like smart card storage, computation, and communication and then estimated time also. We compare securities and functionalities of our protocol along with other protocols. This evaluation shows influence of our protocol among others which shown in Tables 2 and 3. For real-life applications, our proposed scheme is capable to withstand several sort of known attacks.

Now, for computation cost calculation, we considered some symbols such as T_{H}, T_{E}, T_{S}, T_{ECM} for hash function, modular exponentiation, symmetric key encryption or decryption, and elliptic curve point multiplication operation, respectively. Therefore, our scheme needs this computation cost $6T_{\mathrm{H}} + 1T_{\mathrm{E}}$, $16T_{\mathrm{H}} + 5T_{\mathrm{E}} + 4T_{\mathrm{S}}$ for registration phase, login and authentication phase which is low as compared to other schemes [26–30, 35].

Subsequently, for communication cost evaluation, we presumed length of identity ID_i, password PW_i, nonce, elliptic curve point, and hash function $h(.)$ all are 160 bits [36]. But, symmetric key encryption/decryption takes 512 bits. In Table 3, our protocol achieves comparatively low communication cost than other protocols [26–30, 35]. It is observed that proposed scheme is secure against several types of known and unknown attacks. We demonstrated smartcard storage and communication costs of our work and others which is comparatively less than other works [26–30, 33–35].

Furthermore, for evaluating estimated time, we assume some operations like hash function, symmetric key encryption or decryption, point multiplication of elliptic curve and modular exponential takes 0.0005 s, 0.0087 s, 0.063075 s, and 0.522 s, respectively [37]. We can observe that estimated time of our protocol is comparatively less than other protocols [26–30].

The comparison of relevant security protocols [23, 26–30, 33–35] with the proposed protocol has been displayed in Table 2. The schemes are compared to several essential security features. From the table, it is evident that the proposed scheme can withstand several attacks. We can see that our suggested scheme incorporates all the essential security features. Hence, it is assured that our proposed scheme is robust against various criminal attacks.

8 Conclusion

The proposed paper observes Byun's scheme and recognized the number of vulnerabilities like online password guessing threat, privilege insider attack, replay attack, and user anonymity. Additionally, it also does not have password change or update phase, which is very highlight problem in Byun's scheme. So, the authors proposed an extended user anonymous authenticated session-key agreement scheme using smartcard. The presented protocol is verified on the basis of formal and informal analyses, which ensures for applying on real-life applications. It removes all the vulnerabilities of Byun's scheme. Moreover, the presented scheme also facilitates password change phase, which is very user-friendly as quick demand of user.

References

1. Roy S, Karjee J, Rawat U, Dey N et al (2016) Symmetric key encryption technique: a cellular automata based approach in wireless sensor networks. Procedia Comput Sci 78:408–414
2. Dey N, Ashour AS, Shi F, Fong SJ, Sherratt RS (2017) Developing residential wireless sensor networks for ecg healthcare monitoring. IEEE Trans Consum Electron 63(4):442–449
3. Chandrakar P, Sinha S, Ali R (2019) Cloud-based authenticated protocol for healthcare monitoring system. J Ambient Intell Hum Comput: 1–17
4. Ali R, Chandrakar P, Kumar A (2020) On the security weaknesses in password-based anonymous authentication scheme for e-health care. In: Design frameworks for wireless networks. Springer, pp 23–40
5. Chandrakar P (2019) A secure remote user authentication protocol for healthcare monitoring using wireless medical sensor networks. Int J Ambient Comput Intell (IJACI) 10(1):96–116
6. Chandrakar P, Om H (2018) An efficient two-factor remote user authentication and session key agreement scheme using rabin cryptosystem. Arab J Sci Eng 43(2):661–673
7. Ali R, Pal AK (2018) An efficient three factor-based authentication scheme in multiserver environment using ECC. Int J Commun Syst 31(4):e3484
8. Lamport L (1981) Password authentication with insecure communication. Commun ACM 24(11):770–772

9. Li C-T, Hwang M-S (2010) An efficient biometrics-based remote user authentication scheme using smart cards. J Netw Comput Appl 33(1):1–5
10. Li X, Niu J-W, Ma J, Wang W-D, Liu C-L (2011) Cryptanalysis and improvement of a biometrics-based remote user authentication scheme using smart cards. J Netw Comput Appl 34(1):73–79
11. Das AK (2011) Cryptanalysis and further improvement of a biometric-based remote user authentication scheme using smart cards. arXiv preprint arXiv:1103.3159
12. Turkanović M, Brumen B, Hölbl M (2014) A novel user authentication and key agreement scheme for heterogeneous ad hoc wireless sensor networks, based on the internet of things notion. Ad Hoc Netw 20:96–112
13. Karuppiah M, Saravanan R (2014) A secure remote user mutual authentication scheme using smart cards. J Inf Secur Appl 19(4–5):282–294
14. Kalra S, Sood SK (2015) Secure authentication scheme for iot and cloud servers. Pervasive Mob Comput 24:210–223
15. Farash MS, Turkanović M, Kumari S, Hölbl M (2016) An efficient user authentication and key agreement scheme for heterogeneous wireless sensor network tailored for the internet of things environment. Ad Hoc Netw 36:152–176
16. Kaul SD, Awasthi AK (2016) Security enhancement of an improved remote user authentication scheme with key agreement. Wirel Pers Commun 89(2):621–637
17. Kumari S, Khan MK, Li X (2014) An improved remote user authentication scheme with key agreement. Comput Electr Eng 40(6):1997–2012
18. Kumari S, Gupta MK, Khan MK, Li X (2014) An improved timestamp-based password authentication scheme: comments, cryptanalysis, and improvement. Secur Commun Netw 7(11):1921–1932
19. Chaudhry SA, Farash MS, Naqvi H, Kumari S, Khan MK (2015) An enhanced privacy preserving remote user authentication scheme with provable security. Secur Commun Netw 8(18):3782–3795
20. Radhakrishnan N, Karuppiah M, Pandi V, Bhuiyan MZA (2017) Security on a lightweight authentication scheme with user untraceability. International conference on security, privacy and anonymity in computation, communication and storage. Springer, pp 489–496
21. Yeh K-H (2015) A lightweight authentication scheme with user untraceability. Front Inf Technol Electron Eng 16(4):259–271
22. Wu F, Xu L, Kumari S, Li X, Das AK, Shen J (2018) A lightweight and anonymous rfid tag authentication protocol with cloud assistance for e-healthcare applications. J Ambient Intell Hum Comput 9(4):919–930
23. Kumari S, Karuppiah M, Das AK, Li X, Wu F, Kumar N (2018) A secure authentication scheme based on elliptic curve cryptography for iot and cloud servers. J Supercomput 74(12):6428–6453
24. Karuppiah M, Das AK, Li X, Kumari S, Wu F, Chaudhry SA, Niranchana R (2019) Secure remote user mutual authentication scheme with key agreement for cloud environment. Mob Netw Appl 24(3):1046–1062
25. Qi M, Chen J (2017) An efficient two-party authentication key exchange protocol for mobile environment. Int J Commun Syst 30(16):e3341
26. Byun JW (2015) Privacy preserving smartcard-based authentication system with provable security. Secur Commun Netw 8(17):3028–3044
27. Awasthi AK, Srivastava K, Mittal R (2011) An improved timestamp-based remote user authentication scheme. Comput Electr Eng 37(6):869–874
28. Islam SH (2016) Design and analysis of an improved smartcard-based remote user password authentication scheme. Int J Commun Syst 29(11):1708–1719
29. Khan MK, Kumari S (2013) An authentication scheme for secure access to healthcare services. J Med Syst 37(4):9954
30. Li X, Niu J, Khan MK, Liao J (2013) An enhanced smart card based remote user password authentication scheme. J Netw Comput Appl 36(5):1365–1371

31. Burrows M, Abadi M, Needham RM (1989) A logic of authentication. Proc R Soc Lond A 426(1871):233–271
32. Chandrakar P, Om H (2017) Cryptanalysis and improvement of a biometric-based remote user authentication protocol usable in a multiserver environment. Trans Emerg Telecommun Technol 28(12):e3200
33. Chen B-L, Kuo W-C, Wuu L-C (2014) Robust smart-card-based remote user password authentication scheme. Int J Commun Syst 27(2):377–389
34. Bin Muhaya FT (2015) Cryptanalysis and security enhancement of zhu's authentication scheme for telecare medicine information system. Secur Commun Netw 8(2):149–158
35. Chaturvedi A, Mishra D, Mukhopadhyay S (2013) Improved biometric-based three-factor remote user authentication scheme with key agreement using smart card. In: International conference on information systems security. Springer, pp 63–77
36. Islam SH, Khan MK (2014) Cryptanalysis and improvement of authentication and key agreement protocols for telecare medicine information systems. J Med Syst 38(10):135
37. Jiang Q, Ma J, Li G, Yang L (2014) An efficient ticket based authentication protocol with unlinkability for wireless access networks. Wirel Pers Commun 77(2):1489–1506

Data Security in Cloud Computing Using Abe-Based Access Control

Rajanikanth Aluvalu, V. Uma Maheswari, Krishna Keerthi Chennam, and S. Shitharth

Abstract Business organizations and individual users are using cloud storage for storing their data and files. Cloud storage is managed by cloud service provider (CSP) being third party person to the data owners. Cloud storage consists of user's confidential data. After storing data in cloud, the owner of data cannot have control over data, where owner cannot trust the CSP because possibility of a malicious administrator. Based on this, different schemes are proposed. Security is a major concern for cloud stored data, and CSP has to provide trust to the data owner on security of the cloud stored data. In general, security to data and applications is provided through authentication and authorization. Security through authentication is provided by distributing user name and password to data users. However, the organizational user is not allowed to access all the organizational data. Authorization for accessing the data is provided by using access control models. Regular models are not enough to use the CSP based on the models uses dynamic method and proposed different models using attribute-based encryption (ABE). Earlier access control models cannot be used because of multiple disadvantages. This chapter will discuss dynamic access control model named as RA-HASBE. This model is proved to be scalable and flexible, due to sub-domain hierarchy. It is also proved to be dynamic by permitting user to access the data by risk evaluation using risk engine.

Keywords Cloud computing · Data security · Access control models · Encryption · Risk analysis

R. Aluvalu · V. Uma Maheswari · S. Shitharth (✉)
Department of CSE, Vardhaman College of Engineering, Hyderabad, India

K. K. Chennam
Department of CSE, Muffakham Jah College of Engineering and Technology, Hyderabad, India

S. K. Das et al. (eds.), *Architectural Wireless Networks Solutions and Security Issues*, Lecture Notes in Networks and Systems 196,
https://doi.org/10.1007/978-981-16-0386-0_4

1 Introduction

Cloud computing has become a widely accepted computing model in the next years. The success of cloud computing depends on the security provided to the data stored in the cloud. Cloud computing being widely distributed delivers services over the Internet [1]. CSP is main to control the cloud environment and has to ensure trust and security of the data stored by the data owner on the cloud. Confidentiality of stored data can be protected by providing access control models as an authorization mechanism. An access control model helps in restricting unauthorized access to sensitive data by users [2]. Allowing user to access application is called authentication and permitting user to access data and files is called authorization. Various traditional access control models such as discretionary access control (DAC), mandatory access control (MAC), and role-based access control (RBAC) are defined and used widely.. These models are not sufficient for providing security to data in cloud computing environment [3]. Later attribute-based encryption schemes are proposed for providing security to outsourced data. Sahai and Waters (2005) proposed an ABE model. In ABE, data is encrypted and decrypted using user attributes. User's secret key and the cipher text are dependent upon attributes. To decrypt user's data, ciphertext attributes should match with attributes of the user key. The major disadvantage with ABE scheme is that data owner needs to use the public key of every authorized user to encrypt data [4]. ABE demands data provider and client to be online for exchanging keys. Various ABE-based access control schemes to overcome the above-discussed issues are proposed as follows:

(a) KP-ABE: It was proposed by Goyal et al. [3]. KP-ABE is for attributes of users and secret keys which is linked with the model. Only the user who is linked with the model, which has same set of attributes, can decrypt the data.

(b) CP-ABE: It was proposed by Sahai [5]. CP-ABE is linked with the model, secret key and attributes set from the user. Only the person satisfies the attributes set by the user who have similar access structure can decrypt the data

(c) Hierarchical ID-based encryption (HIBE): It was proposed by c.gentry [6]. In HIBE, private key generation (PKG) arranges the load and authentication of identity to minor range of PKG. Practically, the HIBE has maximum resistance collusion and high secured by any eves dropping. Most disadvantage of this HIBE cannot control the identity length based on hierarchy.

(d) Risk-aware access control (RAAC): It was proposed by khalid [7]. A risk-aware access control model uses a proper risk-estimation technique suitable to a particular context and appropriate mechanism to utilize risk for access decision making [8].

(e) Hierarchical attribute set-based encryption (HASBE): It was proposed by Zhiguo [9]. The combined structure of HIBE and CP-ABE is HASBE. The hierarchical order only the users are showed, if not it is similar to CP-ABE. The root (top level authority) will be on the top followed by domain masters. Domain masters consist of user sets [10]. Access control model is highly scalable because of its hierarchical structure. Like CP-ABE and KP-ABE, it also stores data in encrypted format on the untrusted server. However, HASBE still suffered from various drawbacks like handling compound attributes, lack of flexibility in the authorization, lack of efficient key management mechanism. HASBE is extended for supporting sub-domain level hierarchy. The extended model supports the distribution of keys with secure way to access the files that are stored in the cloud based on user roles. In the extended model, it is not required for the data owner to be always online. Key distribution will be handled by trusted authority (TA) (always online) in the more secured way. Data owner will share keys and specific role-based policy with trusted authority. TA will distribute keys to data consumers on request if they satisfy the data owner's predefined policy. HASBE schemes use attributes-based encryption access control with user-level domain hierarchy. We can enhance this scheme by creating sub-domains within the user-level domain hierarchy.

2 Major Issues

There is demand for risk-aware flexible and dynamic access control models with encryption to handle complex and on demand work of organizations. The users due access restrictions are unable to access the data stored in cloud. Existing access control models are static in nature and do not take dynamic decisions to cater the user access requests. The above issues are addressed by allowing data owner to encrypt the data and define access policies [8]. A mechanism is developed using user attributes to evaluate the risk of allowing access to user who had failed to satisfy the access policy defined by the data owner. This makes the system more dynamic. Data owner is allowed to apply the fine-grained access control models while sending data to cloud, where the data is secured by the fine-grained access control models [4]. The model is named as risk-aware hierarchical attribute set-based encryption access control model (RA-HASBE). It is a combination of HASBE and RAAC access control models.

The model is named as risk-aware hierarchical attribute set-based encryption access control model (RA-HASBE). It is a combination of HASBE and RAAC access control models.

3 Major Objectives of RA-HASBE

To support data owners in preventing CSP and other third party users from accessing the contents of their confidential data files stored on remote servers, every such stored file owns a privacy policy.

The minimum requirements of RA-HASBE access control model are:

1. Only authorized users should be able to access the file.
2. One must be allowed to access restricted file in an emergency, provided the risk factor is below the threshold.

In particular, the model has the following objectives:

(a) **Fine-grained access control**: Allows users to access different files and holds a different authorization for every single user.
(b) **User revocation**: Allows data owner to revoke user's access privileges from further access, whenever required.
(c) **Flexible policy specification**: Allows generating complex data access policies by combining simple policies easily.
(d) **Scalability**: Supports managing a large number of users, to store without any problem, key management, and computation.
(e) **Dynamic**: Allowing access to restricted files dynamically based on their risk calculation.

4 Design of RA-HASBE Access Control Model

To address the Issues discussed, it is proposed to develop risk-aware access control model using attribute-based encryption.

To define proposed model, mathematical design model is defined initially.

(a) Organization Employees list Emp = {e1, e2, e3, e4 …}
(b) 'Emp' is the complete set of employees working in the organization. List of attributes used to describe the complete set of registered organizational employees

    1. Attrib = {a1, a2, a3, a4……}

(c) List of attributes requested by other employees
(d) Oattrib = {o1, o2, o3……}
(e) List of employees who had requested attributes of other employees

    1. OAEmp = {oe1, oe2, oe3……}

(f) Identify attribute key of newly joined employee
(g) NEmp = {NE1, NE2, NE3….}

(h) Where ‘NE’ is the main set of attribute key of user’s NE1, NE2, NE3…

(i) Set of transactions T = {T1, T2, T3, T4……}

  i. Each process is set of T1 = {emp1, emp2, emp3….}
  where
  ii. {emp1 = wants to upload data file}
  iii. {emp2 = wants to make entry in data file stored on cloud}
  iv. {emp3 = requesting for new attributes}
  v. {emp4 = requesting for information of employee transferred}

(j) Transaction failures TF = {f1, f2, f3…}

  a. f1 = = {failed to upload data, due to Internet connection failure}

(k) Transaction success TS = {s1, s2, s3…..}

  a. S1 = {data uploaded successfully with good Internet connectivity}
  b. S2 = {s|“s“ if data is added to data file/database}
  c. S2 = {data retrieved successfully from cloud/server}

(i) Initial conditions = I

  i. Good Internet connectivity to user
  ii. Good Internet connectivity to admin

Consider cloud as universal set denoted by ‘U.’
U = {Emp, Attrib,Us, Rs}
Emp = employee set
Attrib = attribute set
Us = Userset
Rs = registered set

Initial State

Us = {r, ur}

r = list of registered users

ur = list of unregistered users.

Next State

Request for more attributes

M = request for a list of new attributes

N = contains all the required attributes

R = provide the list of attribute requested

$S1 = M \cap N$.

Hierarchy

$U = \{u1, u2, u3, u4\}$

Where

U is cloud

u1 is director
u2 is assistant director

u3 is the list of heads

u4 is the list of employees.

Flexibility

$U = \{N1, N2, N3\}$

Where

N1 = Earlier company work location of the employee

N2 = Employee transferred

N3 = New company work location of transferred employee

Where,

D2 = Employee data is available for access only to the new work location

D2 = (N1-N2) U N3.

Scalability

$B = \{u1, u2, u3, u4\}$

$B' = \{u1, u2\}$

$B'' = \{u3, u4\}$

B' = Online user list

B'' = Offline user list

$S3 = B'\ U\ B''$

Final State

Transactions as T

T = {Set of transactions}

T = {T1, T2, T3, T4……}

Where

T1 = {N1, N2, N3}

Where,

{N1 = Received new attributes after request}

{N2 = Received new employee information when employee get transfer.}

{N3 = Privilege to access of data/file}

Enhanced HASBE

U = {u1, u2, u3, u4}

u1 = Data owner

u2 = Data consumer

u3 = Domain authority

u4 = Trusted authority

t1 = u1 encrypted data 'D,' prepared access policy 'P' using attributes 'at' and uploaded data on the cloud.

t2 = u2 want to access 'D' and requested u3 for granting access. Then,

u3 = online AND u2 'at' satisfy u1 'at,' grant 'access' otherwise 'deny' OR.

u3 = offline, pass on the request to u4. u4 will perform verification and grant 'access.'

RA-HASBE

U = {u1, u2, u3, u4}

RE = Risk engine

u1 = Data owner

u2 = Data consumer

u3 = Domain authority

u4 = Trusted authority

t1 = u1 encrypted data 'D,' prepared access policy 'P' using attributes 'at' and uploaded data on the cloud

t2 = u2 want to access 'D' and requested u3 for granting access. Then,

u3 = online AND u2 'at' satisfy u1 'at,' grant 'access' otherwise 'deny'

OR

u3 = offline, pass on the request to u4. u4 will perform verification and grant ‘access.’

If u2 = failed to satisfy access policy ‘P’ Then request ‘RE’ for access.

RE = Key_at is ‘TRUE’ AND AT=satisfy policy>AT=Not satisfy policy, then grant access, otherwise deny.

5 System Design

Data owner will define access control policies, which enables cloud users to encrypt the data, fix the tolerance threshold, store on the cloud, and access the same. Access control schemes will ensure security and improves the performance of tasks on cloud data and processes with fine-grained access control defined using attributes by the cloud owner (Fig. 1).

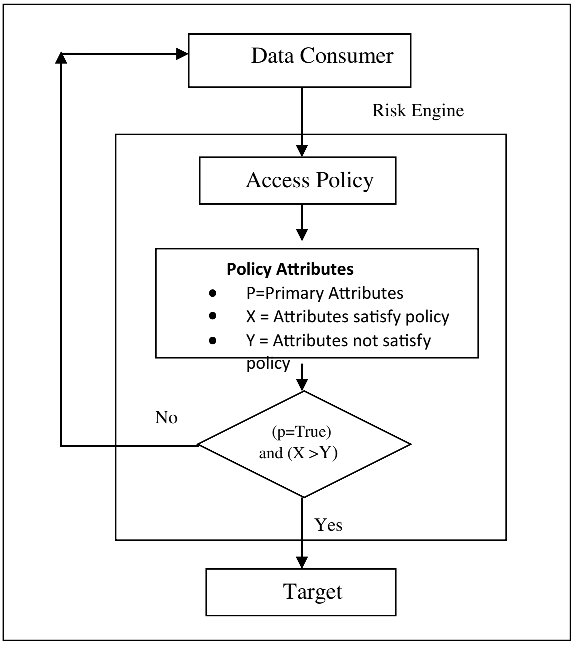


Fig. 1 Risk assessment process

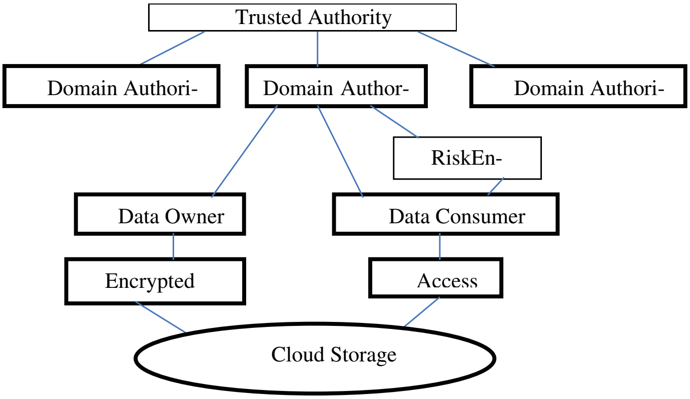


Fig. 2 RA-HASBE architecture

The shown Fig. 2 represents the HASBE scheme associated with risk engine [9, 11]. This makes the scheme dynamic. Risk engine using the values of attributes defined in access policies will evaluate the risk and compare the risk with tolerance factor defined to grant or deny the access to the data and files. The user of data asks for the risk engine to receive the permissions by the basic access control model [7]. By the dynamic nature of the proposed model, the threshold analysis is the important point for the model.

Figure 1 represents the mechanics for the risk in risk engine and changes to be made by representing the security of the models. Based on this model, the user identifies each attribute to define the key attributes.

(a) *Risk threshold calculation:*

The validation based on the risk engine with the user data before the registration and identifies the threshold of the risk by accepting and rejecting the permissions based on the given procedure.

- *Request granted:*

1. Prime attrib = True AND
2. Number of attribs values satisfying with values of access policies is larger than the no. of attrib values not satisfying the access policy values.

- Request denied:

1. Prime attrib = False

OR

1. Prime attribute = True AND

Prime attrib = False

2. Number of attribs values satisfying with values of access policies is smaller than the no. of attrib values not satisfying the access policy values.

By giving permission for an individual file with the help of the above method, the threshold cutoff shows the permission or rejects the access. With this Fig. 1, the personal authorization and giving security by this risk engine can be integrated. Accessing methods for calculating risk, finding the tolerance of risk levels and sharing the information by control by the risk-aware system [11].

6 Algorithm Design

Algorithm-1: Enhanced-Hierarchical Attribute Set based Encryption Access control model.

```
Input: Data Consumer Attributes
Output: Data Access:  Grant/deny
 Procedure:
 Start
  1.  Input user Attribute values.
  2.  Input Access policy.
  3.   Input Request Access grant to file.
  4.  Function Enhanced –HASBE
          a.   for file access grant
          b.    Condition  Access Policy verification
          c.  If  user attributes satisfy access policy  then
          d.  Grant  file access to User
          e.  Else Deny file access grant.
          f.  End.
  5.  Exit.
```

Algorithm-2 : Risk Aware-Hierarchical Attribute Set based Encryption Access control model.

```
Input: Data Consumer Attributes
Output: Data Access:  Grant/deny
 Procedure:
 Start
  1. Input user Attribute values.
  2. Input Access policy and Primary Key attribute.
  3.  Input Request Access grant to file.
  4. Function RA –HASBE
        a.  for file access grant Request domain Authority
        b.   Condition  Access Policy verification
        c. If  user attributes satisfy access policy  then
        d. Grant  file access to User
        e. Else Deny file access grant.
              i. Request Risk Engine for Access grant.
             ii. Function RISK ENGINE
                    1. For file access grant
                    2. Risk Assessment  if ( Primary Key
                       Attribute =TRUE)  AND (No of
                       Attributes satisfy Access policy > No of
                       Attributes do not satisfy Access policy)
                       then
                    3. Grant Access to the Request User AND
                       Acknowledge to file owner.
                    4. Else Deny file Access grant.
                    5. End.
        f. End
  5. Exit.
```

7 Implementation of RA-HASBE

Step-1: Implementing Enhanced-HASBE

HASBE access control model is improved by maintaining the sub-domains under domain, e.g., if the domain is section, then below the section of domain, the particular sub-sections like software, hardware, testing, supply chain and management and production, and planning, based on the given data, where the user connected with sub-section and the main section, this will help us to improve the different problems in finding data related and queries based on the sub-section and display of data only from the sub-section related and permissions to show with the particular sections[12, 13]. Sub-domain creation will help in decreasing the time taken for fetching the data, query processing, and overall operational time [9].

Step 2: Implementing Dynamic Attribute-based Access Control Model. (DAAC)
Access permissions are granted through risk calculation based on user attribute values. Risk engine is developed to serve the purpose. Risk engine will assess the cost of risk in granting the user to access the file. Risk engine after evaluation, if the risk factor is below threshold, will grant permission otherwise will deny the request to access. For this, risk engine is developed. Risk engine consists of risk assessment functionality [7].

Step-3: Integrating Enhanced HASBE with RAAC: RA-HASBE
The model is further enhanced by integrating enhanced HASBE with risk engine developed for DA-RAAC. Risk engine will assess the cost of risk in granting the user to access the file. Risk engine after evaluation, if the risk factor is below threshold, will grant permission otherwise will deny the request to access [14].

Whenever data consumer fails to satisfy access policy, he can send request to risk engine, requesting access to the data. This flexibility and dynamic behavior will help the organizations to complete transactions in emergency without waiting for someone. Here, risk of granting access is accounted for making access control decisions [15, 16]. As mentioned earlier, this dynamic approach is particularly useful to allow some risky access in an emergency situation. Majorly proposed RA-HASBE will perform three core functionalities, namely

(a) Measuring the risk of granting access,
(b) Define risk tolerance level and compare with risk of granting access.
(c) Finally restrict or grant access based on the result of risk comparison with tolerance level.

8 Result Analysis

Computation complexity of RA-HASBE is calculated theoretically considering all the operations such as system setup, file access, encryption time, and decryption time and later analyzed the performance with other models.

With enhanced HASBE scheme *scalability, flexibility, user revocation, and flexible policy specification* objectives are met [9]. With sub-domain grant, the system has become scalable to handle increased number of users. The model allows user to move from one sub-domain to the other easily during employee transfers and promotions, and this proves the flexibility of the model. As when user changes the domain, the access privileges will be revoked by the system, the user has to request for new access keys, and this proves the objective user revocation [10, 17]. As explained during design, phase system allows user to create complex access policies from simple access policies[18–20].

Enhanced HASBE scheme is further improved by adding risk engine and named it as RA-HASBE. The change in access policy with RA- HASBE scheme in comparison with traditional HASBE scheme is evaluated [21, 22]. RA-HASBE grants permission to access files by the user when the risk factor is below the threshold, else the access grant is denied. Rest all remains same as enhanced-HASBE. RA-HASBE is developed by integrating enhanced HASBE with DA-RAAC [16]. It is already proved that enhanced HASBE is *scalable, flexible, supports fine-grained access control* [23], *and user revocation* [24, 25]. By integrating DA-RAAC, the model has become highly *dynamic* in nature [11].

The security of RA-HASBE proved by analyzing its functionality with enhanced -HASBE.

i. RA-HASBE scheme, when the consumer fails to satisfy access policy and the risk value is below threshold, had allowed the consumer to access the data, whereas access is denied by enhanced HASBE. The dynamic objective of RA-HASBE scheme is proved.
ii. RA-HASBE denies access to the data consumer, when the consumer fails to satisfy access policy and the risk value is above threshold, even enhanced HASBE denies access.

The performance of RA-HASBE with enhanced HASBE is evaluated. It is proved from the results that RA-HASBE is dynamic in nature.

9 Conclusions

This chapter considered the implementation of the proposed RA-HASBE access control model which is highly scalable, flexible, and dynamic and supports user revocation and fine grain access control in terms of allowing access to the users. In this work initially, HASBE access control model is enhanced by adding sub-domain Hierarchy and achieved scalability, flexibility, fine grain access control, and user revocation objectives. Later by integrating RAAC with enhanced HASBE model, the scheme is transformed into dynamic model by granting access through risk factor calculation and named the model as RA-HASBE. This model is secured from the privilege escalation because of the role hierarchy management with least privilege grants. RA-HASBE supports secure data sharing on remote cloud servers with authorized users. The hybrid risk-aware attribute-based user privilege control supports unique users and maintains a hierarchical control structure of users [26–28]. Implemented model is highly scalable and flexible in terms of user access management.

The present model is best useful for an organization with hierarchical roles and caters to emergency works. The current implementation and evaluation are based on the already collected users attributes and access monitoring results. In the future, the model can be monitored on user behavior in a real environment in order to evaluate their trust level and can be deployed in a real-time system.

References

1. Hao R, Yang H, Zhou Z (2019) Driving behavior evaluation model base on big data from internet of vehicles. Int J Ambient Comput Intell (IJACI) 10(4):78–95
2. Das SK, Sachin T (2018) Intelligent energy-aware efficient routing for MANET. Wirel Netw 24(4):1139–1159
3. Kandukuri R, Paturi VR, Rakshit A (2009) Cloud security issues. In: Proceedings of the 2009 IEEE international conference on services computing, pp 517–520
4. Neena Antony A. Melvin AR (2012) A survey on encryption schemes in the clouds for access control. Int J Comput Sci Manage Res 1(5):1135–1139
5. Wan Z, Liu J, Deng RH (2012) HASBE: a hierarchical attribute-based solution for flexible and scalable access control in cloud computing. IEEE Trans Inf Foren Secur 7(2):743–754
6. Kandala S, Sandhu R, Bhamidipati V (2011) An attribute based framework for risk-adaptive access control models, availability, reliability and security (ARES). In: 6th international conference, pp 236–241
7. Aluvalu R, Lakshmi M (2016) A dynamic attribute-based risk aware access control model (DA-RAAC) for cloud computing. In: 2016 IEEE international conference on computational intelligence and computing research (ICCIC). IEEE
8. Goyal V, Pandey O, Sahai A, Waters B (2006) Attribute-based encryption for fine-grained access control of encrypted data. In: Proceedings of the 13th ACM conference on computer and communications security, pp 89–98
9. Aluvalu, R, Lakshmi M (2015) Access control model with enhanced flexibility and scalability for cloud. In: 2015 international conference on green computing and Internet of Things (ICGCIoT). IEEE
10. Wang G, Liu Q, Wu J (2010) Hierarchical attribute-based encryption for fine-grained access control in cloud storage services. In: Proceedings of the 17th ACM conference on computer and communications security, pp 735–737
11. Aluvalu R, Lakshmi M (2016) Risk aware hierarchical attribute set-based encryption (RA-HASBE) access control model. Annals. Comput Sci Ser 14(2)
12. Aluvalu R, Muddana L (2015) A survey on access control models in cloud computing. In: Emerging ICT for bridging the future-proceedings of the 49th annual convention of the computer society of India (CSI) vol 1, no 1, pp 653–664
13. Bobba R, Khurana H, Prabhakaran M (2009) Attribute-sets: a practically motivated enhancement to attribute-based encryption. In: European symposium on research in computer security, pp 587–604
14. Gentry C, Silverberg A (2002) Hierarchical ID-based cryptography. In: Proceedings of Asiacrypt, vol 2501. LNCS, pp 548–566
15. John B, Sahai A, Waters B (2007) Ciphertext-policy attribute-based encryption. In: 2007 IEEE symposium on security and privacy (SP'07), pp 321–334
16. Aluvalu R, Kamliya V, Muddana L (2016) HASBE access control model with secure key distribution and efficient domain hierarchy for cloud computing. Int J Electr Comput Eng 6(2):770
17. Sarwar A, Khan MN (2013) A review of trust aspects in cloud computing security. Int J Cloud Comput Serv Sci (IJ-CLOSER) 2(2):116–122. . ISSN: 2089–3337
18. Zissis D, Lekkas D (2012) Addressing cloud computing security issues. Future Gener Comput Syst 28(3):583–592
19. Buyya R, Yeo CS, Venugopal S, Broberg J, Brandic I (2008) Cloud computing and emerging IT platforms: vision, hype, and reality for delivering computing as the 5th utility. Future Gener Comput Syst 25(6):599–616
20. Devi D, Arun PS (2014) A design for secure data sharing in cloud. Int J Eng Res Gener Sci 2(5):72–77
21. Liang C, Jason C (2011) Risk-aware role-based access control. In: 7th international workshop, STM 2011, Copenhagen, Denmark, pp 140–156

22. Thirumaleshwari Devi B, Shitharth S (2020) An Appraisal over intrusion detection systems in cloud computing security attacks. In: 2nd International Conference on Innovative Mechanisms for Industry applications, ICIMIA -2020, ConferenceProceedings, pp 122
23. Shitharth S, Sangeetha PK (2019) Integrated probability relevancy classification (IPRC) for IDS in SCADA. In: Design Framework for wireless network, Lecture notes in network and systems, vol 82, issue 1, Springer, pp 41–64
24. Shitharth S, Shaik M, Sangeetha S, Mining of intrusion attack in SCADA network using clustering and genetically seeded flora based optimal classification algorithm. Infor Sec IET 14(1):1–11
25. Shitharth DP, Winston D (2017) An enhanced optimization algorithm for intrusion detection in SCADA network. J Comput Sec, Elsevier 70:16–26
26. Shitharth DP, Winston, (2016) A new probabilistic relevancy classification (PRC) based intrusion detection system (IDS) for SCADA network. J Elect Eng 16(3):278–288
27. Kumar, Parmar Vipul J, Aluvalu RK (2015) Key policy attribute-based encryption (KP-ABE): a review. Int J Innov Emerg Res Eng 2:49–52
28. Aluvalu R, Chennam KK, Uma Maheswari V, Jabbar MA (2021) A novel and secure approach for quantum key distribution in a cloud computing environment. In: Balas VE, Semwal VB, Khandare A, Patil M (eds) Intelligent Computing and Networking. Lecture Notes in Networks and Systems, vol 146. Springer, Singapore

Linear Secret Sharing-Based Key Transfer Protocol for Group Communication in Wireless Sensor Communication

Priyanka Jaiswal and Sachin Tripathi

Abstract Wireless sensor network (WSN) is a collection of autonomous nodes which are used to sense environmental information for a particular operation or goal. Each node of the sensor network is directly or indirectly connected with base station (BS). The purpose of the BS is to collect required information from the sensor nodes and process its future purpose. Each node of the network contains low capacity of battery. This battery does not fully complete any operation due to its energy capacity, and in the middle of the operation, communication is fails. This event occurs frequently due to sensor nodes energy capacity. It degrades the performance of the network as well as network metrics and raises several types of interference and noise. It causes several types of attacks and hacking. So, to prevent this, in this paper, an intelligent protocol is proposed with the fusion of linear secret sharing (LSS) and elliptic curve techniques. The combination of both techniques helps to overcome the drawback of traditional protocols. Finally, this security protocol helps to reduce the overhead of WSN and enhances several security mechanisms against different conflicting attacks.

Keywords Wireless sensor network · Linear secret sharing · Vandermonde matrix · Group communication · Authentication

1 Introduction

Day by day the applications of wireless network increase rapidly for its variable natures and usages. There are several works proposed in this area like Yang et al. [1] designed an intelligent system for transportation system in wireless network. This is based on an existing transportation system based on processed structured system. Finally, it helps to enhance the network capabilities and services of the network. It also helps the user functions and usages in the network and network metrics properly to maintain the network. Loganathan and Subbiah [2] designed an energy-based

P. Jaiswal (✉) · S. Tripathi
Department of Computer Science and Engineering, IIT (ISM), Dhanbad, Jharkhand, India

S. K. Das et al. (eds.), *Architectural Wireless Networks Solutions and Security Issues*, Lecture Notes in Networks and Systems 196,
https://doi.org/10.1007/978-981-16-0386-0_5

communication system for device-to-device communication in the network. It is based on multi-criteria decision-making system where multiple criteria are involved in integrating the network metrics efficiently. Finally, it helps to enhance the network lifetime and helps in communication system. Jat et al. [3] designed an intelligent technique for QoS in WLAN. This proposal is based on video delivery system. This is based on multimedia application for video data processing and analyzing. The data is analyzed here based on real-time data generated by the Internet. It also helps in video data transmission, storage, evaluating, and broadcasting. In WSN, data is gathered from multiple homogeneous or heterogeneous sources, because real-life data is connected with different IoT, IoV, or cloud environment. So, it is difficult to keep the natures of the data in the same structure. Information retrieval [4] is a very important part in modern research areas which indicates to collect information that is stored in unstructured form based on multiple local languages and process it in particular patterns after observing. Hao et al. [5] designed an evaluation system for big data analysis. This data is based on IoV where it means Internet of the vehicle. This proposal is based on K-means algorithm that is used here as a clustering. In this work, different behavior of the driving is involved for controlling vehicle. Finally, it helps in reducing fuel consumption and helps in transportation globally.

The abovementioned literature is based on the wireless network. But there are so many variations of wireless network based on its design frameworks that described in [6]. This work contains several frameworks such as optimization, security and privacy, localization and network lifetime enhancement. This book provides the basic framework ideas of the users and new researchers. The proposed work is based on wireless sensor network (WSN). In [7], the authors proposed nature-inspired-based methods in WSN. This book contains several popular nature-inspired algorithms that help to the readers and researchers both in designing as well as innovating any algorithm. Das and Tripathi [8] proposed a method for software-defined network which is based on ad hoc manner. The main key element of this work is nonlinear formulation method which is used to optimize the network by using objective function and their constraints. Finally, it helps to manage conflicting strategies of the network efficiently.

The nodes of ad hoc and sensor networks are dynamic and autonomous. It acts as router and helps in sending and transmitting the data packets. It greatly relied on the environment of the modern technology. It has several limitations also like limited computing power, limited bandwidth and unreliable communication, limited energy supply, etc. These stated limitations cause two types of attacks like passive and active attacks. So, the network needed an authenticated security mechanism for reducing several attacks and interferences like in [9] designed an authentication system for the users with wireless network. This is basically based on healthcare systems and used for medical purpose. This proposal is used for sensing patient body information and sending to the doctor for treatment and diagnosis purpose. It also helps in user authentication, privacy and data security purpose, so that efficient result comes from the diagnosis system.

To avoid abovementioned limitations and uncertainties, most efficient security goal is needed, which is inherent in some paradigms such as “key freshness”, “key confidentiality", and "key authentication". First inherent element indicates freshness of the communication system, so that session key is not used previously. It uses previously, and then it creates ambiguity and uncertainty in the group communication. Second, inherent element indicates authorized access of the session key that is only accessed by the group members. The group members only encrypt and decrypt the information of the group communication. So, that no other person, it can access unauthorized way. Third element indicates that authorized member only can access the session key, although it initiates by one initiator who is a part of the group member. The proposed method is the combination of the stated three elements. It is the fusion of linear secret sharing (LSS) and elliptic curve techniques. The combination of both techniques helps to overcome the drawback of traditional protocol.

As said by Boyd [10], the key establishment protocols are categorized into key transport protocols (also called as key distribution or group key transfer protocol) [11–13] and key exchange protocols (also called a group key agreement protocol) [14–16]. The first group key transfer protocol has been introduced by Ingemarsson et al. [17] which extended the two parties DH to a group Diffie–Hellman. After that, some group communication protocol extended the DH protocol, which inherited DH properties such as Steer et al. [18], Steiner et al. [19], Burmester, and Desmedet [20].

A traditional group communication protocol can be partitioned into two groups: One is the centralized-based group communication protocol, and other is the distributed group communication protocol. In a centralized approach, a central server/ entity is used to construct the group key and transfer it to other group users [21, 22]. The distributed key management approach [17–19] involves dynamically selecting group participants, which play the role of distributed server. The distributed key management approach can be categorized into DH distribution [14, 20, 23–25] and non-DH distribution [25, 26], and the non-DH key agreement approach generally comes with fault tolerance features. The term fault tolerance means the system even works whenever the power failure occurs. Steiner et al. [19] proposed a distributed DH-based group key exchange protocol, which shows the natural extension of two party DH protocols [27]. After that an authentication service is incorporated with the scheme to provide more security. Bohli [23] designs a protocol to provide robustness in the scheme. Then one of the protocols has been established by Katz and Yang in 2007 which shows scalability with constant round feature.

Tzeng [15] designed a group key protocol with the hardness capability of discrete logarithm (DL) assumption which shows a non-DH approach. The Tzeng [15] protocol has been modified by Cheng and Laih [26] using the concept of bilinear pairing. A non-interactive protocol has been developed by Huang et al. [27] which incorporates the features of DL assumption to provide better efficiency. Roy [28] presented a symmetric-based key agreement. Zhao et al. [29] designed a group protocol in 2010 which uses the concept of RSA cryptography for improving the efficiency of the protocol. Dey [30] shows a wireless sensor network-based protocol. Since the session key is computed by the contribution of all group participants, so the time required calculating the group key will be greater, probably in case of larger group size.

1.1 Motivation and Contribution

To avoid encryption and decryption operations, one of the popular techniques is used called as secret sharing [29], the mechanism secret sharing is important because it involves various advantages to share a secret. Blackly [31] and Shamir [32] independently introduced secret sharing (SS) as a solution to the key management problem. They represent a mechanism that splits a secret into multiple shares and secret may be recovered by only authorized set of participants. Secret t generally consists of three phases: sharing, distribution, and reconstruction. In sharing phase, a leader (one of the member) divides secret into multiple parts called shares, in distribution phase dealer send shares to group participants, and in reconstruction a qualified set of parties combines their shares to reconstruct the secret.

SS scheme can be implemented by one of the two ways [32, 33]. The first method says that the, at initialization, a trusted offline server does the work. The second one assumes a trusted third party called as KGC works as an online server in all the process [34]. The offline type of secret sharing is called as key pre-distribution secret sharing scheme. In this type of scheme, the offline server computed and distributed secret information to the entire members. One of the major problems of key pre-distribution scheme is that each member stores large information and then later is used, which causes a burden on the system. In the second one, an online server to be active is to choose a group session key and broadcast it to all group members [34, 35].

Laih et al. [34] use first SS scheme called as (t, n) SS in 1989. After that some papers follow the Laih et al. [34] approach to share the secret like [35–37]. Harn et al. [38] design a key distribution protocol in which the property of confidentiality and authentication have been achieved by security analysis phase and the concept of Lagrange interpolation polynomial is used to calculate t degree interpolation. However, the new improvement is given by [39, 40] and pointed out that Harn et al. [38] do not protect from malicious user. The protocol [39, 40] show the problems associated with the group and given the synthesis of group communication.

The remaining paper is divided as follows. Unit 2 highlights basic preliminaries information of the paper. Unit 3 illustrates design frameworks of the proposed method. Unit 4 describes the proposed method briefly. Unit 5 describes a security analysis of the model. Finally, Unit 6 concludes the paper.

2 Background and Preliminaries

The mathematics background and fundamental preliminaries are given below.

2.1 Fundamental of Elliptic Curve Group (ECG)

Suppose E (F_p) denote an elliptic curve E over finite field (F_p).

The elliptic curve equation is given as.

$Y^2 \bmod p = (x^3 + ax + b) \bmod p$, where F_p is a prime finite field and $a, b \in F_p$ with the discriminant $\Delta = (4a^3 + 27b^2) \bmod p \neq 0$.

The elliptic curve equation over a group can be represented as below: $G = \{f(x, y) : x, y \in F_p \text{ and } (x, y) \in E(F_p)\} \cup \{o\}$ an additional point O called infinity point.

2.1.1 Point Multiplication

Let the given scalar is j, then the scalar point multiplication over ECG is given by the equation:

$$jA = jA = A + A + \cdots + A(j\text{Times})$$

The deep study about elliptic curve is found in [41, 42].

2.1.2 Discrete Logarithm Problem (DLP) on ECG

If the generator P of the group (G) and an element $D \in {Z_p}^*$ is given, then it will be difficult to discover an integer $a \in {Z_p}^*$ such that $D = a.P$.

2.2 Linear Secret Sharing Schemes

The secret sharing (SS) scheme states that an information (secret) s can be separated into n small information called as shares and when the shares are splitted among n participants through a trusted shareholder in the way that only valid (authorized) set of users can recover the information but invalid (unauthorized) set of members (participants) cannot recover the actual information. The SS scheme is called as perfect SS scheme, if only authorized set of users obtain the shares and all the authorized set of users are called as access structure Γ. The unauthorized set of participants is called as prohibited structure.

Let $\rho = \{1,..., n\}$ is the set of participants (shareholders), by the use of Shannon's functional equation [33]. The secret sharing scheme holds the following conditions, with respect to access structure Γ.

Let S denotes a secret information (domain) and P_i denotes a share information of participant i, where $1 \leq i \leq n$. If a trusted shareholder (dealer) D desires to share a piece of information (secret) $s \in S$ to the set of members $\rho = \{1,..., n\}$, the dealer provides a share $p_i \in P_i$ to each participant. The pieces of information (shares) are transferred securely for which every member of the group cannot identify the shares given to the remaining group members. A subset of group members may find their shares to recover the secret s later. The following conditions are mandatory if a scheme is a *SS scheme* [43].

1. Correctness: Any subset of group members wants to find secret s can calculate s. For r any $A \in \Gamma$, it holds $H(S|A) = 0$.
2. Security: Any subset of group members does not want to find secret s cannot reproduce s, while they pool all of their information pieces together. Formally, for any $A \notin \Gamma$, it holds $0 < H(S|A) \leq H(S)$.

If $H(S|A) = H(S)$, members in A keep their (pieces of information) shares together and get no information on S. We can say that this is the case where the scheme is perfect.

If $|S| = |P_i\ |$ holds for $1 \leq i \leq n$, we say that it is *linear*, if its secret domain $S = \kappa$ is a finite field, $P_i - S$ are linear spaces over κ, and the recovery operations are linear [24].

The monotone span programs (MSP) have been developed by Karchmer and Wigderson [37]. LSS-S based on Vandermonde matrix (VM) has been developed by Hsu et al. [44].

If there are $(n + 1)$ shareholders $P = P_0, P_1, \ldots, P_n$ and a mutually trusted dealer (D), the following two algorithms (share generation and secret reconstruction) are used to design LSS-S based on VM.

(1) **Share Generation Algorithm (SGA)**: In SGA, the initiator or dealer (D) first takes a VM V_{n+1} and a random vector $\mathbf{r} = (r_0, r_1, r_2, \ldots, r_n) \in \overline{V}$ and let $\mathbf{r}$ be public, in this case all the computations are performed in the finite field κ.The dealer computes:

$$\begin{pmatrix} 1 & x_1 & x_1^2 & \ldots & x_1^n \\ 1 & x_2 & x_2^2 & \ldots & x_2^n \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 1 & x_{n+1} & x_{n+1}^2 & \ldots & x_{n+1}^n \end{pmatrix} \begin{pmatrix} r_0 \\ r_1 \\ \ldots \\ r_n \end{pmatrix} = \begin{pmatrix} s_0 \\ s_1 \\ \ldots \\ s_n \end{pmatrix}$$

Then, the algorithm outputs a list of $(n + 1)$ shares$(x_1, \ldots, x_{n+1})$ and distributes each share x_i to corresponding shareholder P_i secretly.

(2) **Secret reconstruction algorithm (SRA)**: The SRA takes all $(n + 1)$ shares $(x_0, x_1, \ldots, x_n)$ and the public vector $\mathbf{r}$ as inputs and outputs the secret $S = s_0, s_1, \ldots, s_n$ by computing each inner product $(\mathbf{v}\ (x_i), \mathbf{r}) = s_i$
Since computational assumptions are not used to fulfill the security requirements, so LSS-S-based VM is theoretically secure.

3 The Proposed Protocol

The proposed scheme can be designed using two phases as given below:

(1) Secret establishment phase
(2) Session key transfer phase.

A group can have n number of members, suppose that the n members of the group $U_i = \{1,...,n\}$ wants to design a group protocol for secure communication. The group members, including leader/ initiator, should design a (public, private) a pair of key (*puk*, *prk*) where *puk* = *prk.Q*, and *Q* represents an elliptic curve generating point. Here, it should be noted that a trusted authority certifies the pair of key (*puk*, *prk*) by providing corresponding certificate to each group member. In the proposed protocol, there is *n* number of group member, and one of the group member selected as a group leader, the group leader has the power to choose a group key and to start the communications.

3.1 Secret Establishment Phase:

(1) The leader/ initiator chooses a random integer $r_n \in Z_p^*$ and sends the following message to every group participant for the announcement of the group key generation protocol.
$(r_n, puk_n, 1, \ldots, n)$.

(2) The participants $(1, \ldots, n)$ for $i = 1, \ldots, n-1$, choose the random challenge (number) $r_i \in Z_p^*$ after getting the messages from announcement, and the participants also computed the following information.

- $R_i = r_i.puk_i$
- $\overline{R_i} = r_i.puk_n$
- $s_i = \overline{R_i}.prk_i$
- $Auth_i = h(s_i||r_n)$

and after computing these values, the group member sends the reply messages to the initiator as: $\{R_i, Auth_i\}$.

(3) The initiator computed $s_i^{'} = R_i.prk_n$, after getting the informative message from participants. The leader checks the value of $Auth_i$ using s_i' and r_n where $Auth_i \overset{?}{=} h(s_i^{'}||r_n)$. The initiator checks whether value of *Auth*i is valid or not, and if found it valid, then only the initiator trusts that the secret $s_{i} = r_i.prk_n.prk_i.Q$ is shared by the legitimate user; otherwise, he declares that the user *i* is not the actual user, he is/are fake and after that revives the system.

3.2 The Session Key Transfer Phase

The session key transfer phase uses the following steps to generate the group key. In this phase, the group users perform the following steps of operations.

We assume a linear space over κ, where κ is a finite field and $\overline{V} = \kappa^n$, V is a n is a dimensional vector with the characteristic char $(\kappa) = p$ in this phase. Assume $\{e_1,\ldots, e_n\}$ of $\overline{V}$ with $\overrightarrow{e_j} = (0, \ldots 0, \overset{j}{1}, 0, \ldots 0) \in \kappa^n$ for $j = (1,\ldots, n)$ the transition $\mathrm{v} : \kappa \rightarrow \overline{V}$ given by $\mathbf{v}(x) = \sum_{j=1}^{n} x^{j-1} e_j = \left(1, x, x^2, \ldots, x^{n-1}\right)$ is defined. The initiator (leader of the group) is denoted by n, and the other group participants are i $(i\epsilon\{1,\ldots,n - 1\})$ in the proposed scheme. For creating a secure session among the group members, the leader and the other participants of the group execute some steps as described below. For sharing the secret, firstly, the leader of the group selected a group key and transferred this key to all participants through secure way. The steps of operation for key transfer are as below.

(1) The leader partitioned each secret s_i into two pieces x_i and y_i, where $(x_i||y_i) = s_i$ for $i = 1,\ldots, (n - 1)$ and chooses a group key called as session key $K_G\epsilon\ Z_p^*$. After that, he calculated $(n - 1)$ additional public values, $U_i = (K_{\mathrm{G}} - K_i)$ for $i = 1,\ldots, n - 1$, and the value $Auth = h(K_{\mathrm{G}}, 1, \ldots, n, r_1, \ldots r_n, U_1, \ldots, U_{n-1})$, where h is a non-invertible function. The initiator broadcast $\{Auth,\ U_i\}$ for $i = 1, \ldots, n - 1$, to the members.

(2) The group member who knows the public values U_i can compute the inner product $(y_i v(x_i),\ \vec{r_i}) = K_i$ and regenerate the session key $K_G = (U_i + K_i)$ where the vector $\vec{r_i} = (r_i, \ldots, r_n) \in \kappa^n$. Then $(i\epsilon\{1,\ldots,n - 1\})$ needs to calculate $h(K_{\mathrm{G}}, 1, \ldots, n, r_1, \ldots r_n, U_1, \ldots,\ U_{n-1})$, and verify the value of *Auth,* i.e., whether the calculated value of *Auth* is same as the value of previous value of *Auth* or not. If he found that the values are same, then the user authenticates the group key that the key is sent from the initiator.

The group (session) key K_{G} is established after the successful implementation of the above designing steps. All the members of the group use this group key K_{G} for the establishment of common session between them.

Remark 1 The proposed scheme uses the concept of ECDLP assumption. In this protocol, the initiator/ leader (dealer) distributes a secret with the group member (i) *where there exist n numbers of group member*. The join/leave operation of the group members does not depend on the updating of the shared secret. In this protocol, sharing secret information to all the members depends upon broadcasting the informative message to every member. The remaining group participant except initiator computed the inner product using random challenge, and from that inner product $(y_i v(x_i),\ \vec{r_i}) = K_i$ and public values, the members are able to compute the group key. The group key can be recovered by using the function of group participant's random challenge and the secret shared among group members by the initiator. The used linear secret sharing is ideal and perfect sharing because it has no need to compute the inner product of two vectors to construct session key.

4 Security Discussion

The security discussion (analysis) of the scheme is verified here. It ensures that the proposed scheme secures against various securities attacks and provides the security goals as abovementioned. It also resists insider and outsider attacks.

The security attacks in the cryptography can be categorized into insider attacks and outsider attacks. Outsider (adversaries) attacks on the confidentiality features of the system. The outside attacker cracks the confidentiality to know the secret information of the system, whereas the outside attacker is unauthorized to know the common session key of the system.

The second adversary inside attackers can be the active member of the group who belongs to the particular session of the group. The insider knows the group session key for a particular session, for which he is authorized. The inside attacker tries to find the individual member secret. In the proposed protocol, suppose that the group has n participants, $\{1,\ldots, n\}$, and each participant has a shared secret like s_1, …, s_n. The given security analysis demonstrates that the proposed scheme resists insider and outsider attacker and accomplishes the following security attributes:

(1) ***key freshness***
(2) ***key confidentiality***
(3) ***key authentication.***

Key freshness As the name key freshness suggests that the fresh session key should be used for each and every session. The key freshness feature assures that the fresh (new) session key is required for each service request, and it is done by invoking a service request by the leader/ initiator to randomly choose a session key. The proposed protocol ensures the key freshness because group key is choosing by the initiator randomly for each session, and the session key is a function of randomly chosen number chosen by every participant and the longtime secret shared among the respective participants and the leader.

Key confidentiality The term confidentiality means that the private information kept secret from the attackers. In this protocol, the confidentiality feature is achieved by using the security assumptions ECDLP and the ideal and perfect LSS-S. In the given LSS-S, the initiator/ leader chooses a session key K_G at random and takes $n - 1$ values, $U_i = (K_\mathrm{G} - K_i)$ for $i = 1, \ldots, n - 1$, which are publicly known. Only the certified participants of the group have knowledge about $(y_i v(x_i), \vec{r_i}) = K_i$, inner products, and so the legal group participants are only capable of constructing the secret key as $K_\mathrm{G} = U_i + K_i$. However, on the other hand, the unauthorized member or adversaries is not capable of constructing the group key because they have only $n - 1$ values of U_i and he cannot get any extra information on K_i and $\sum_{1\le i\le n-1} K_i$. Thus, the adversary cannot recover the session key. We can say that the proposed LSS-S is perfect LSS-S, and information theoretically secures and provides confidentiality.

Key authentication The term authentication of the session key ensures the identity of the participants that only authorized member can only be participated on the group key communication, and in the proposed protocol this feature is ensured by the function of $Auth_i$ and $Auth$. In the second phase, $Auth_i$ shows a non-invertible hash function of secret s_i shared between the user and the leader of the group say initiator. The initiator randomly chosen a number as input, and the secret s_i is confidential so, the adversary cannot compromise this value.

Outsider Attack *Suppose that the adversary wants to impersonate as a legal group member to participate in the communication,* and *then the adversary can neither find the group key nor share it with other participants.*

Proof If any adversary can grasp the communication information between leader and the participants of the group, the adversary cannot find the shared information between the initiator (leader) and the participating group member and so he cannot obtain shared information s_i because it is impossible to know the private key prk_i of any user i. Moreover, the group key K_{G} can only be recovered by a legal member since the legal member can only have the right shared secret s_i and also the group key is constructed from the ideal secret sharing LSS-S. Thus, the adversary cannot impersonate as a legitimate participant of the group to find session key. The adversary also cannot impersonate as the server because it is impossible for the attacker to find the server's private key and so he cannot impersonate as server.

Theorem 5 (Insider attack)

Suppose the group communication protocol executes several times, then long-term secret $s_i$of each participant shared by the initiator to the corresponding participants would be unknown to others 1.

Proof The initiator of the group randomly chosen a session key and takes $n - 1$ public values represented as (U_i). The legitimate group participants who have corresponding secret s_i shared by initiator and the public values can only construct the inner product $(y_i \mathrm{v}(x_i), \vec{r_i}) = K_i$ and so only legitimate group member can only regenerate the group key K_{G} by the use of public values U_i and the corresponding inner product K_i.

The group members (insiders), who, know the group key K_{G} and the public values Ui, then from these values the insider can obtain K_i, whereas the insider cannot find the shared secret s_i of other group member from the product $(y_i \mathrm{v}(x_i), \vec{r_i}) = K_i$ and also the shared secret s_i of each group member depends on the long-term private key (prk_i, prk_n) and the random challenge (r_i, r_n) and so it is impossible to find the one time shared secret $(x_i || y_i) = s_i$.

5 Performance Comparison

The performance comparison of the proposed scheme shows that the designed scheme enhanced the performance in terms of computational comparisons to other related schemes. This section is divided into two parts: The first part is performance analysis based on computational comparison given in Sect. 5.1, and the second part performance of proposed protocol based on functionality comparison is given in Sect. 5.2.

5.1 Performance Analysis Based on Computational Comparison

Most of the traditional secret sharing protocols use KGC to pre-shared secret among group members, and those schemes totally depend on trust of KGC that KGC should never be compromised, which is a drawback of the traditional scheme. The other performance advantage of the proposed scheme is that it does not pre-share secret between KGC and group members, while it established the secret using ECDLP assumption. ECDLP is used in place of modular exponentiation to reduce the computational overheads. The present scheme decreases the system computations and communicational cost, and the protocol is more suitable and practical for groupware applications.

Second, the use of linear secret sharing in place of threshold secret sharing scheme (TSS) is used which is more computation-efficient than TSS. In LSS, instead of calculating interpolation polynomial as in TSS, we computed the inner product of the vectors in the field.

Parameters used are as follows:

t: no of group.

Tm: modular multiplication.

Tx: modular exponentiation.

Th: hash operation1.

Ti: modular inverse operation.

Te: elliptic curve multiplication.

Suppose that there are t number of group exist in the proposed protocol and the given shared secret is s_1, s_2, …, $s(t-1)$. The group key KG belongs to k and the vector r belongs to k^t. The time taken by the initiator in secret establishment phase in the proposed protocol is $(t-1)Te + 1Th,$ and the time taken by the each group participant in the same phase is $3(t-1)Te + 1Th,$ whereas the time taken by the Hsu protocol in secret establishment phase by initiator is $2(t-1)Tm + 1Tx + 1Th,$ and by each group participants is $2(t)Tm + 1Tx + 1\ h$, which is greater than the proposed scheme.

In first (key transfer) phase, the time taken by initiator in distribution of session key in the proposed scheme is $2t(t-1)Tm + Th,$ and the time taken by the initiator in recovery of the group key is by each member which is $2(t)Tm + Th.$ However, the Harn's schemes take more time in distribution and the recovery of the group key, which shows that the proposed protocol performs better from the other related protocol (Table 1).

Table 1 Performance comparison

Protocols/computational comparison	Harn and Lin [38]	Hsu [45]	Proposed
CCSE	–	$2(t-1)Tm + 1Tx + 1Th$	$(t-1)Te + 1Th$
CCSE	–	$2(t)Tm + 1Tx + 1\,h$	$3(t-1)Te + 1Th$
CCKD	$t^2(t+1)(Tm + Ti) + Th$	$2t(t-1)Tm + Th$	$2t(t-1)Tm + Th$
CCKR	$t(t+1)(Tm + Ti) + Th$	$2(t)Tm + Th$	$2(t)Tm + Th$

CCST computation cost in secret establishment phase by initiator, *CCST* computation cost in secret establishment phase by group participants, *CCKD* computation cost in key distribution, *CCKR* computation cost in key recovery, *CCKR* computation cost in key recovery

Table 2 Functionality comparisons of the proposed protocol with the other related protocol

Protocols/functionality	Harn and Lin [38]	Hsu [45]	Proposed
Without registration with a trusted server	N	Y	Y
Without an online KGC	N	Y	Y
Group key generated by users	N	Y	Y
Excludes malicious participants	Y	Y	Y
No additional time required	Y	Y	Y
Mutual authentication	Y	Y	Y
Session key agreement	Y	Y	Y
Ecc is used to reduce extra overheads	Y	N	Y

5.2 *The Protocol Performance on Functionality Comparison*

Table 2 shows the comparison of the major functionalities and security attributes used from other related protocols. The result of the functionality comparison shows that the proposed protocol fulfills all the functionality mentioned in Table 2 and which is discussed in security analysis section while the other compared protocol mere fulfills the functionality attributes. The functionality used in this chapter is discussed below.

5.2.1 Without Registration with a Trusted Server

Pre-sharing secret is not required in the proposed protocol at initialization of the protocol. The traditional group key transfer protocol priory shares the secret between group users and trusted third party through safe channel. However, the proposed initiator does some calculation on the secret and share the secret through public channels.

5.2.2 No Need of Online KGC

The session key for the group is generated and distributed with the help of group members, and so there is no need of KGC and the overhead of KGC is reduced from the overall overhead of the protocol. However, other traditional key transfer protocol requires online KGC for key selection and transport.

5.2.3 Session Key Generated by Group Participants

One of the major features of the protocol is that all group members commonly created the group key by each other contributions.

5.2.4 Excludes Malicious Participants

The unauthorized users who want to participate in the group communication cannot involve in the communication. The proposed protocol excludes the malicious participants, and it excludes insider and outsider adversary from the system.

5.2.5 No Need of Additional Synchronization Time

The traditional session key transport protocol requires extra time for the synchronization of the start-up of the protocol. However, there is no need of additional time synchronization required in the proposed protocol.

5.2.6 Mutual Authentication

The mutual authentication feature is provided between initiator and other group participants, and the feature is hold by ECDLP hardness problem.

5.2.7 Session Key Agreement

The term session key agreement is one of the common functionality of this scheme, this functionality helps to design a secure session key for all the participants, and the participants of the group can share/ exchange top secret information among each other.

6 Conclusion and Future Scope

The proposed scheme uses the concept of LSS-S and elliptic curve cryptography to design the protocol for group-oriented applications in wireless networks. The used LSS-S is ideal and theoretically secures as discussed earlier in the security analysis section. The security feature, like authentication and confidentiality, is ensured by the LSS-S and hardness of ECDLP assumptions. The proposed scheme shows a new result in the field of group-oriented applications since the protocol uses the linear secret sharing scheme which is information theoretically secure as well as the protocol uses the concept of elliptic curve cryptography which reduces the overall overheads of the system and provides confidentiality and information security. In this scheme, one of the group members called initiator selected a common group key and distributed it to the other group member which also reduces the computation cost of the KGC. The proposed scheme achieves essential security features, and the security feature of the protocol is discussed in the security analysis section. The performance comparison of the protocol shows that the proposed scheme is more efficient than other referred protocol. The overall performance and the security discussion present that the protocol is more suitable for other groupware applications.

References

1. Yang W, Wang X, Song X, Yang Y, Patnaik S (2018) Design of intelligent transportation system supported by new generation wireless communication technology. In: Intelligent systems: concepts, methodologies, tools, and applications. IGI Global, pp 715–732
2. Jayakumar L, Subbiah J (2020) Energy aware dynamic mode decision for cellular D2D communications by using integrated multi-criteria decision making model. Int J Ambient Comput Intell 11(3). (7 February 2020, IGI Global)
3. Jat DS, Bishnoi LC, Nambahu S (2018) An intelligent wireless QoS technology for big data video delivery in WLAN. Int J Ambient Comput Intell (IJACI) 9(4):1–14
4. Rasheed I, Banka H (2018) Query expansion in information retrieval for Urdu language. In: 2018 4th international conference on information retrieval and knowledge management (CAMP). IEEE, pp 1–6
5. Hao R, Yang H, Zhou Z (2019) Driving behavior evaluation model base on big data from internet of vehicles. Int J Ambient Comput Intell (IJACI) 10(4):78–95
6. Das SK, Samanta S, Dey N, Kumar R (2020) Design frameworks for wireless networks. Springer
7. De D, Mukherjee A, Das SK, Dey N (2020) Nature inspired computing for wireless sensor networks.
8. Das SK, Tripathi S (2020) A nonlinear strategy management approach in software-defined Ad hoc network. In: Design frameworks for wireless networks. Springer, Singapore, pp 321–346
9. Chandrakar P (2019) A secure remote user authentication protocol for healthcare monitoring using wireless medical sensor networks. Int J Ambient Comput Intell (IJACI) 10(1):96–116
10. Boyd C (1997) On key agreement and conference key agreement. In: Australasian conference on information security and privacy. Springer, Berlin, pp 294–302
11. Sáez G (2003) Generation of key predistribution schemes using secret sharing schemes. Discrete Appl Math 128(1):239–249

12. Jaiswal P, Tripathi S (2017) An authenticated group key transfer protocol using elliptic curve cryptography. Peer-To-Peer Netw Appl 10(4):857–864
13. Sun Y, Wen Q, Sun H, Li W, Jin Z, Zhang H (2012) An authenticated group key transfer protocol based on secret sharing. Procedia Eng 29:403–408
14. Katz J, Yung M (2003) Scalable protocols for authenticated group key exchange. In: Annual international cryptology conference. Springer, Berlin, pp 110–125
15. Tzeng WG (2002) A secure fault-tolerant conference-key agreement protocol. IEEE Trans Comput 51(4):373–379
16. Diffie W, Hellman M (1976) New directions in cryptography. IEEE Trans Inf Theory 22(6):644–654
17. Ingemarsson I, Tang D, Wong C (1982) A conference key distribution system. IEEE Trans Inf Theory 28(5):714–720
18. Steer DG, Strawczynski L, Diffie W, Wiener M (1988) A secure audio teleconference system. In: Conference on the theory and application of cryptography. Springer, New York, pp 520–528
19. Steiner M, Tsudik G, Waidner M (1996) Diffie-Hellman key distribution extended to group communication. In: Proceedings of the 3rd ACM conference on computer and communications security, pp 31–37
20. Burmester M, Desmedt Y (1994) A secure and efficient conference key distribution system. In: Workshop on the theory and application of cryptographic techniques. Springer, Berlin, pp 275–286
21. Fiat A, Naor M (1993) Broadcast encryption. In: Annual international cryptology conference. Springer, Berlin, pp 480–491
22. Canetti R, Garay J, Itkis G, Micciancio D, Naor M, Pinkas B (1999) Multicast security: a taxonomy and some efficient constructions. In: IEEE INFOCOM'99. Conference on computer communications. Proceedings. 18th annual joint conference of the IEEE computer and communications societies. The Future is Now (Cat. No. 99CH36320) (vol 2, pp 708–716). IEEE
23. Bohli, J. M. (2006, May). A framework for robust group key agreement. In: International conference on computational science and its applications. Springer, Berlin, pp 355–364
24. Hsu CF, Zeng B, Cheng Q, Cui G (2012) A novel group key transfer protocol. IACR Cryptology ePrint Archive 2012:43
25. Klein B (1995) Conference key distribution protocols in distributed systems. Proc Codes Ciphers Crypt Coding IV 1995:225–242
26. Cheng JC, Laih CS (2009) Conference key agreement protocol with non-interactive fault-tolerance over broadcast network. Int J Inf Secur 8(1):37–48
27. Huang KH, Chung YF, Lee HH, Lai F, Chen TS (2009) A conference key agreement protocol with fault-tolerant capability. Comput Stand Interf 31(2):401–405
28. Roy S, Karjee J, Rawat US, Dey N (2016) Symmetric key encryption technique: a cellular automata based approach in wireless sensor networks. Procedia Comput Sci 78:408–414
29. Zhao J, Gu D, Li Y (2010) An efficient fault-tolerant group key agreement protocol. Comput Commun 33(7):890–895
30. Dey N, Ashour AS, Shi F, Fong SJ, Sherratt RS (2017) Developing residential wireless sensor networks for ECG healthcare monitoring. IEEE Trans Consum Electron 63(4):442–449
31. Blakley GR (1979) Safeguarding cryptographic keys. In: The AFIPS national computer conference, NCC'79
32. Shamir A (1979) How to share a secret. Commun ACM 22(11):612–613
33. Blundo C, De Santis A, Vaccaro U (1994) Randomness in distribution protocols. In: International colloquium on automata, languages, and programming. Springer, Berlin, pp 568–579
34. Laih CS, Lee JY, Harn L (1989) A new threshold scheme and its application in designing the conference key distribution cryptosystem. Inf Process Lett 32(3):95–99
35. IEEE (2004) IEEE 802.11i-2004: Amendment 6: medium access control (MAC) security enhancements

36. Berkovits S (1991) How to broadcast a secret. In: Workshop on the theory and application of of cryptographic techniques. Springer, Berlin, pp 535–541
37. Li CH, Pieprzyk J (1999) Conference key agreement from secret sharing. In: Australasian conference on information security and privacy. Springer, Berlin, pp 64–76
38. Harn L, Lin C (2010) Authenticated group key transfer protocol based on secret sharing. IEEE Trans Comput 59(6):842–846
39. Nam J, Kim M, Paik J, Won D (2012) Security weaknesses in Harn-Lin and Dutta-Barua protocols for group key establishment. KSII Trans Internet Inf Syst 6(2)
40. Liu Y, Cheng C, Cao J, Jiang T (2012) An improved authenticated group key transfer protocol based on secret sharing. IEEE Trans Comput 62(11):2335–2336
41. Hankerson D, Menezes AJ, Vanstone S (2006) Guide to elliptic curve cryptography. Springer
42. Stinson DR (2005) Cryptography: theory and practice. CRC Press
43. Farràs O, Martí-Farré J, Padró C (2012) Ideal multipartite secret sharing schemes. J Cryptol 25(3):434–463
44. Hsu C, Zeng B, Cui G, Chen L (2014) A new secure authenticated group key transfer protocol. Wirel Pers Commun 74(2):457–467
45. Hsu C, Zeng B, Zhang M (2014) A novel group key transfer for big data security. Appl Math Comput 249:436–443

Optimization Model for Network Lifetime

Fuzzy Rule-Based System for Route Selection in WSN Using Quadratic Programming

Manoj Kumar Mandal, Arun Prasad Burnwal, B. K. Mahatha, Abhishek Kumar, Santosh Kumar Das, and Joydev Ghosh

Abstract Wireless sensor network (WSN) is a part of wireless network which has flexible and dynamic nature in context of real-life applications. It has several usages in terms of user requirements. It consists of several nodes having limited energy capacity. Energy capacity of the nodes does not completely fulfil the requirement of the services. During transaction or transmission, data is dropped and fails to reach the destination node or base station (BS). This BS also suffers several types of difficulties for sending or receiving data packets. So, there is need of some techniques or modelling that help to protect this issue. Apart from energy, distance is also one important parameter for transmitting data successfully. Although energy is the crucial parameter, but, combination of both energy and distance plays an important role for managing efficient route of the network. The proposed method is the combination of intelligent technique as well as mathematical modelling that uses fuzzy logic as an intelligent technique and quadratic programming as a mathematical modelling for solving the proposed goal. The combination of both provides a robustness technique that uses two basic parameters, energy and distance, for selecting optimal route of

M. K. Mandal (✉)
Department of Mathematics, Jharkhand Rai University, Ranchi 835222, India

A. P. Burnwal
Department of Mathematics, GGSESTC, Bokaro, Jharkhand 827013, India

B. K. Mahatha
Amity School of Engineering and Technology, Amity University Jharkhand, Ranchi 834001, India

A. Kumar
Department of Electronics and Communication Engineering, Swami Vivekananda Subharti University, Meerut 250005, India

S. K. Das
Department of Computer Science and Engineering, Sarala Birla University, P.O.-Mahilong Purulia Road, Birla Knowledge City, Ranchi, Jharkhand, India

J. Ghosh
School of Computer Science and Robotics, National Research Tomsk Polytechnic University (TPU), Tomsk, Russia

S. K. Das et al. (eds.), *Architectural Wireless Networks Solutions and Security Issues*, Lecture Notes in Networks and Systems 196,
https://doi.org/10.1007/978-981-16-0386-0_6

the WSN. The proposed method is validated in LINGO optimization software for formulating and validating the model efficiently.

Keywords Wireless sensor network · Quadratic programming · Fuzzy logic · Rule-based system · Routing

1 Introduction

Wireless sensor network (WSN) is a part of wireless network which is also known as subset of a wireless network. In [1], there are several frameworks discussed in terms of optimization, localization, troubleshooting, and security analysis. This design frameworks influence several variations of the wireless network based on applications. WSN is one of the variations that influences rapidly for solving different issues and problem. In [2–4], several works are proposed for WSN based on nature-inspired techniques or algorithms. Figure 1 shows an illustration of WSN communication between multiple users that consist of several types of sensor nodes, base station

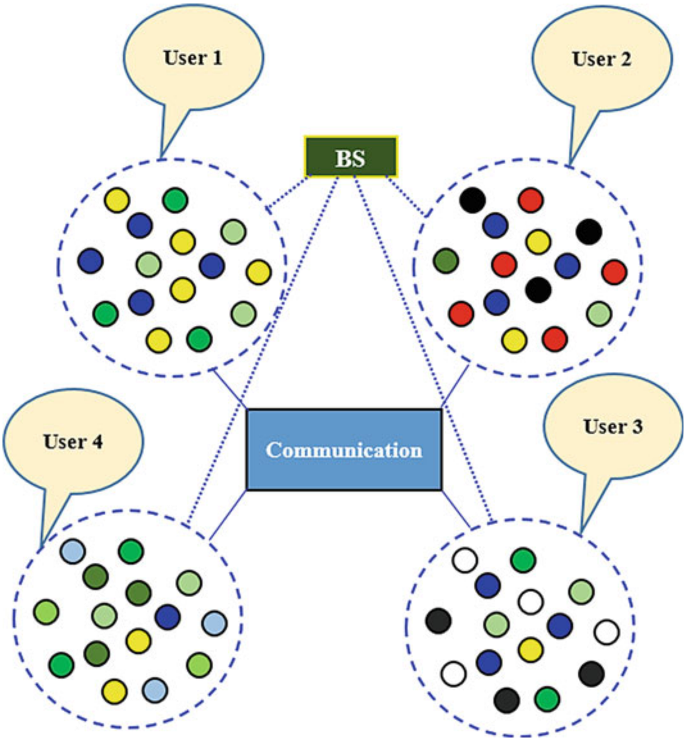


Fig. 1 Wireless sensor network communication between multiple users

(BS), and users. The purpose of these entities is to establish an efficient communication. Wireless sensor nodes are used to sense environmental information and send it to the BS for future references. BS previously stores the user queries in its database, receives sensed information from the sensor nodes, analyses it based on received user queries, and provides responses to the users. In this diagram, there are four users present to establish communication, but in real life, the number of users is more based on number of sensors used in the operational region. In this figure, variation of sensor nodes is more, because battery capacity of each sensor node is different as “Low”, “Medium”, “Poor”, “Bad”, “High”, “Very High”, “Sufficient”, etc. These terms are known as linguistic variables in which assumptions are changed based on user or administrator.

The above-mentioned figure is one example, but there are several real-life applications of WSN such as emergency situation, disaster management, business, offices, entertainment, school and colleges. In each application, there are several types of randomness and uncertainties. It raises multiple interferences between one node to another node and source node to destination node or amongst multiple neighbour nodes. These interferences and uncertainties are main causes of imprecise information and network troubles. Hence, there is a need of robustness technique that helps to reduce the above issues by controlling network parameters and select optimal path for communication between multiple users. The proposed technique is a combination of intelligent technique as well as mathematical modelling that uses fuzzy logic as an intelligent technique and quadratic programming as a mathematical modelling for solving the proposed goal. The combination of both provides a robustness technique that uses two basic parameters as energy and distance for selecting optimal route of the WSN.

The roadmap of the paper is as follows. Unit 2 described some information about existing works. Unit 3 illustrated the basic preliminaries information related to the proposed method. Unit 4 describes the details of the proposed method. Unit 5 describes the simulation analysis. And Unit 6 concludes the paper.

2 Literature Review

In several years, various works are proposed in the context of WSN along with its variations. Some works are discussed in this section as follows. Movassagh and Aghdasi [5] designed a game theory-based scheduling algorithm for WSN. In this method, some nodes are active, and some other nodes are in sleep for optimizing the network lifetime and reducing the redundancy in coverage system. Finally, it helps in enhancing the network lifetime and network metrics of the WSN using the strategy management technique of WSN. Chen et al. [6] proposed a method for WSN based on game theory technique. This game theory technique is based on evolutionary system that is used to control and manage selfish nodes of the network. In WSN, the number of nodes is more for handling any operation. So, behaviour of the selfish nodes fluctuated frequently. The proposed method helps to manage packet forwarding system of the node by optimizing strategy of the network and increase fitness of the WSN. Sun et al. [7] proposed a technique for spectrum sharing in WSN where the nature of the network is heterogeneous. This is based on game theory optimization technique. In this network, several types of nodes are available where each node tries to enhance its own profit which degrades the performance of the network. The game theory optimization technique helps to establish a strategy where no one deviates the rule of the game and optimizes the network metrics efficiently. Yang et al. [8] designed an intelligent system for transportation system in wireless network. This is based on an existing transportation system based on process-structured system. Finally, it helps to enhance network capabilities and services of the network. It also helps to the user function and usages in the network and network metrics properly to maintain the network. Loganathan and Subbiah [9] designed an energy-based communication system for device-to-device communication in the network. It is based on multi-criteria decision-making system where multiple criteria are involved for integrating the network metrics efficiently. Finally, it helps to enhance the network lifetime and helps in communication system. Shen et al. [10] designed a predictable-based routing method for ad hoc network. In this work, topology is organized by the helps to static and dynamic topology distribution with the help of not completely predictable method. Here, incomplete predictable technique is initiated by anti-pheromone system. Finally, it achieves energy efficiency and node utilization both for enhancing the network lifetime. Chatterjee and Das [11] designed an ACO-based routing technique for MANET. The main aim of this routing method for enhancing the QoS is by increasing ratio of packet delivery ratio and decreasing network delay by using ant. This method uses DSR routing as a base routing protocol. The basic route packets like RREQ and RREP are used here—"request ant" and "reply ant" packet for managing the network. Finally, it determined the level of pheromone for each route to decide optimal route of the network. Fatemidokht and Rafsanjani [12] designed an anomaly detection method for VANET based on clustering approach. In this paper, VANET contains some malicious nodes that act as several vehicles in the area of transportation. The nodes in this work disconnect and organize frequently in terms of changes of topology. The clustering method in this work is used for decision of gateway selection, proper neighbour selection, and also cluster head selection. Finally, it helps in packet delivery ratio and reduction of end-to-end delay. Jat et al. [13] designed an intelligent technique for QoS in WLAN. This proposal is based on video delivery system. This is based on multimedia application for video data processing and analysing. The data is analysed here based on real-time data generated by the Internet. It also helps in video data transmission, storage, evaluating, and broadcasting. In WSN, data is gathered from multiple homogeneous or heterogeneous sources because real-life data is connected with different IoT, IoV, or cloud environment. So, it is difficult to keep the natures of the data in same structure. Information retrieval [14] is very important part in modern research areas which indicates collect information that is stored in unstructured form based on multiple local languages and processes it in particular pattern after observing. Hao et al. [15] designed an evaluation system for big data analysis. This data is based on IoV where it means Internet of vehicle. This proposal is based on *K*-means algorithm that is used here as a clustering. In this work, different behaviours of the driving are involved for controlling vehicle. Finally, it helps in reducing fuel consumption and helps in transportation globally. Das and Tripathi [16] proposed a method for software-defined network which is based on ad hoc manner. The main key element of this work is nonlinear formulation method which is used to optimize the network by using objective function and their constraints. Finally, it helps to manage conflicting strategies of the network efficiently. Singh et al. [17] designed an optimized-based localization system for WSN. This work is based on communication between anchor and target nodes. It helps to optimize several issues such as localization, organization, security, scheduling of task, routing, lifetime of the network, and computation of data. These fusions are handled and optimized with the help of PSO and H best PSO, where H indicates Hilbert trajectory technique for the optimization. Kotary and Nanda [18] proposed a distributed-based optimization for WSN. It is based on diffusion system of the WSN which indicates *K*-means clustering algorithm. This diffusion technique is mixed with PSO optimization technique and in identifying optimal clustering based on intra-distance system of the sensor node. The proposed method is easily helped as a robustness technique and employed as detection of outlier of the network. Mohammed et al. [19] designed an optimization technique for WANET fuzzy logic. In this system, fuzzy logic is basically used to reduce the uncertainty of the network. This work is based on clustering system where clustering is designed with the help of constraints of the fuzzy logic. The several network parameters are used as a design of fuzzy constraints such as hop count, speed, movement of the nodes, position of the nodes, and residual energy of the nodes. Finally, with the help of the stated network parameters, fitness function is designed that helps to evaluate optimal route of the network and increase productivity and throughput [20].

3 Preliminaries

In this section, basic preliminaries are described that help to understand the proposed method efficiently in terms of intelligent technique as well as mathematically. Short descriptions are as follows.

3.1 Linear Programming

Linear programming is used to solve linear relationship amongst objective function and constraints based on the problem and issue. In this model, objective function and constraints both are linear in nature. It easily helps to optimize different parameters in terms of finding optimal solution.

3.2 Quadratic Programming

Quadratic programming is a part of nonlinear programming which relates with objective function and constraints nonlinearly. This is based on second-order polynomial technique. In this mathematical modelling, the objective function is always nonlinear in nature but constraints are linear or nonlinear based on the situation.

3.3 Fuzzy Logic

Fuzzy logic is a multi-valued logic which is based on the relation between partial truth and partial false depending on degree of truth value. Degree of truth value is evaluated based on relation of universe of discourse and degree of membership function. Fuzzy logic deals with linguistic variables for reducing uncertainty of information and estimates imprecise parameters of the system.

3.4 Rule-Based System

Rule-based system is a fusion of fuzzy logic and knowledge-based system. It is basically used to solve uncertainty of the system using If–Then statement. It has two basic components such as antecedent and consequent. Antecedent handles "If" clause, and consequent handles "Then" clause. It is rapidly used in several applications such as engineering and science for reducing the uncertainly of the information.

4 Proposed Method

In this section, the main proposed method is illustrated with the help of two basic network parameters such as energy and distance. The purpose of these parameters is to design an optimal strategy for evaluating optimal route that enhances the network lifetime efficiently. The nature of the energy parameter is conflicting with the parameter distance because if energy is increased, then network lifetime is also increased, and if distance is increased, then network lifetime is decrease. So, energy is equal to inverse of distance parameter. In this paper, energy is considered as 500 unit, and distance is considered as 1600 unit. The membership functions of both parameters are shown in Tables 1 and 2. Network lifetime of the WSN is evaluated with these two parameters and optimization models of the proposed method shown in Eqs. (1)–(4).

Table 1 Membership functions of energy

Linguistic variable	Notation	Notation of range	Value
Low	E_{L}	$[E_{\mathrm{L}-}, E_{\mathrm{L}+}]$	(0–150)
Medium	E_{M}	$[E_{\mathrm{M}-}, E_{\mathrm{M}+}]$	(100–250)
High	E_{H}	$[E_{\mathrm{H}-}, E_{\mathrm{H}+}]$	(180–375)
Very High	E_{VH}	$[E_{\mathrm{VH}-}, E_{\mathrm{VH}+}]$	(240–500)

Table 2 Membership functions of distance

Linguistic variable	Notation	Notation of range	Value
Low	D_{L}	$[D_{\mathrm{L}-}, D_{\mathrm{L}+}]$	(0–400)
Medium	D_{M}	$[D_{\text{M-}}, D_{\mathrm{M}+}]$	(350–800)
High	D_{H}	$[D_{\text{H-}}, D_{\mathrm{H}+}]$	(500–1200)
Very high	D_{VH}	$[D_{\mathrm{VH}-}, D_{\mathrm{VH}+}]$	(950–1600)

$$\begin{array}{ll}
\text{Maximize:} & \mathrm{Obj}_1 = (x_1)^2 + (x_2)^2; \\
\text{Subject to constraints:} & e_1x_1 + d_1x_2 \geq 500; \\
 & e_2x_1 + d_2x_2 \geq 500; \\
 & x_1 >= 0;\ x_2 \geq 0; \\
 & e_1 \geq 0;\ e_1 \leq 150; \\
 & e_2 \geq 0;\ e_2 \leq 150; \\
 & d_1 \geq 950;\ d_1 \leq 1600; \\
 & d_2 \geq 950;\ d_2 \leq 1600;
\end{array} \tag{1}$$

$$\begin{array}{ll}
\text{Maximize:} & \mathrm{Obj}_2 = (x_1)^2 + (x_2)^2; \\
\text{Subject to constraints:} & e_1x_1 + d_1x_2 \geq 1500; \\
 & e_2x_1 + d_2x_2 \geq 1500; \\
 & x_1 >= 0;\ x_2 \geq 0; \\
 & e_1 \geq 100;\ e_1 \leq 250; \\
 & e_2 \geq 100;\ e_2 \leq 250; \\
 & d_1 \geq 500;\ d_1 \leq 1200; \\
 & d_2 \geq 500;\ d_2 \leq 1200;
\end{array} \tag{2}$$

$$\begin{array}{ll}
\text{Maximize:} & \mathrm{Obj}_3 = (x_1)^2 + (x_2)^2; \\
\text{Subject to constraints:} & e_1x_1 + d_1y_1 \geq 2500; \\
 & e_2x_1 + d_2y_1 \geq 2500; \\
 & x_1 >= 0;\ x_2 \geq 0; \\
 & e_1 \geq 180;\ e_1 \leq 375; \\
 & e_2 \geq 180;\ e_2 \leq 375; \\
 & d_1 \geq 350;\ d_1 \leq 800; \\
 & d_2 \geq 350;\ d_2 \leq 800;
\end{array} \tag{3}$$

$$\begin{array}{ll} \text{Maximize:} & \mathrm{Obj}_4 = (x_1)^2 + (x_2)^2; \\ \text{Subject to constraints:} & e_1x_1 + d_1x_2 \geq 3500; \\ & e_2x_1 + d_2x_2 \geq 3500; \\ & x_1 >= 0;\ x_2 \geq 0; \\ & e_1 \geq 240;\ e_1 \leq 500; \\ & e_2 \geq 240;\ e_2 \leq 500; \\ & d_1 \geq 0;\ d_1 \leq 400; \\ & d_2 \geq 0;\ d_2 \leq 400; \end{array} \tag{4}$$

In above-mentioned Eqs. (1)–(4), the optimization models are illustrated for network lifetime using energy and distance parameters. In these models, decision variables x_1 and x_2 are used for energy and inversely distance parameters based on two constraints where energy is used "Low" to "Very High" and distance is used "Very High" to "Low" due to contradictory nature of the distance parameter. In these models, a number of sensor nodes are used such as 500, 1500, 2500, and 3500 with difference of 1000 nodes. Each model is evaluated in ten iterations based on same linguistic variation relation of energy and distance parameters. In the proposed model, energy and distance are considered as input parameters and output parameters considered as network lifetime. The different values of each iteration is shown in Tables 3, 4, 5, and 6, and fuzzy relation between input and output parameters is shown in Table 7 as rule-based system.

In Tables 3, 4, 5, and 6, network lifetimes are illustrated based on different rounds and iterations. It shows that when a number of nodes are increased then network lifetime also increases based on rounding of sensor nodes. It indicates, during data transmission, network topology is changed, sometime route is bad, sometime route is good, or sometime route is moderate that varies based on the theorem of the fuzzy logic. Table 7 shows rule-based system including antecedents and consequents that are attached with two input parameters, i.e. energy and distance and output parameter, i.e. network lifetime (NL). In this system, distance parameter is mapped reverse way

Table 3 Values of ten iterations for first model based on "Low" energy and "Very High" distance

S. No.	Low energy	Very High distance	x_1	x_2	Obj$_1$
1	**30,150**	**950,1600**	**0.1660206E−01**	**0.5257915**	**0.2767323**
2	50,140	1000,1500	0.2492657E−01	0.4987537	0.2493766
3	80,120	1050,1550	0.3607258E−01	0.4734421	0.2254486
4	90,110	1150,1450	0.3382011E−01	0.4321358	0.1878852
5	70,100	1250.1350	0.8842397E−01	0.3987496	0.1594998
6	60,130	1200,1390	0.2075928E−01	0.4156287	0.1731782
7	10,90	1110,1490	0.4062012E−01	0.45041398	0.2028891
8	40,80	1150,1300	0.1510622E−01	0.4342572	0.1888075
9	0,100	1100,1500	0.5304115E−04	0.4545455	0.2066116
10	**20,110**	**960,1580**	**0.1082386E−01**	**0.5206078**	**0.2711497**

Table 4 Values of ten iterations for second model based on “Medium” energy and “High” distance

S. No.	Medium energy	High distance	x_1	x_2	Obj_2
1	**120,240**	**520,1100**	**0.632002**	**2.738769**	**7.900261**
2	140,230	540,1150	0.6747605	2.602840	7.230077
3	150,200	640,1000	0.5207142	2.221708	5.207128
4	130,190	600,1050	0.5173460	2.387908	5.969753
5	**100,250**	**510,1200**	**0.5553337**	**2.832288**	**8.330248**
6	150,220	700,900	0.4390244	2.048781	4.390244
7	160,210	750,950	0.4080676	1.912946	3.825880
8	110,180	650,1030	0.3796273	2.243448	5.177174
9	125,170	550,930	0.5893653	2.593326	7.072692
10	115,230	540,1170	0.5658695	2.657269	7.381284

Table 5 Values of ten iterations for third model based on “High” energy and “Medium” distance

S. No.	High energy	Medium distance	x_1	x_2	Obj_3
1	**190,370**	**370,750**	**2.745668**	**5.346819**	**36.12717**
2	200,350	390,700	2.602854	5.075459	32.56514
3	210,320	410,710	2.474132	4.830322	29.45335
4	220,300	400,740	2.639204	4.798438	29.99040
5	**180,375**	**350,800**	**2.905053**	**5.648830**	**40.34861**
6	230,295	380,720	2.914363	4.814991	31.67765
7	240,300	390,620	2.861168	4.649537	29.80448
8	260,350	380,680	3.066017	4.481146	29.48113
9	215,290	360,780	3.057014	5.118728	35.54671
10	195,360	395,720	2.5122871	5.088871	32.20819

Table 6 Values of ten iterations for fourth model based on “Very High” energy and “Low” distance

S. No.	Very High energy	Low distance	x_1	x_2	Obj_4
1	280,350	120,300	10.56028	4.526017	132.0043
2	**250,450**	**100,350**	**12.06896**	**4.827602**	**168.9655**
3	260,480	110,380	11.41781	4.830625	153.7014
4	290,490	50,200	11.72056	2.020736	141.4550
5	300,500	30,250	11.55113	1.155396	134.7635
6	240,300	150,350	10.48691	6.554280	152.9338
7	**260,360**	**90,260**	**12.02112**	**4.161206**	**161.8230**
8	280,460	130,370	10.28331	4.774409	128.5414
9	290,430	20,210	12.01184	0.8283709	144.9704
10	315,410	40,320	10.93481	1.388375	121.4976

Table 7 Rule-based system of the proposed method

Rule No.	Description
Rule 1	If Energy is E_{L} and Distance is D_{VH} then Network Lifetime is NL_1
Rule 2	If Energy is E_{L} and Distance is D_{H} then Network Lifetime is NL_2
Rule 3	If Energy is E_{L} and Distance is D_{M} then Network Lifetime is NL_3
Rule 4	If Energy is E_{L} and Distance is D_{L} then Network Lifetime is NL_4
Rule 5	If Energy is E_{M} and Distance is D_{VH} then Network Lifetime is NL_5
Rule 6	If Energy is E_{M} and Distance is D_{H} then Network Lifetime is NL_6
Rule 7	If Energy is E_{M} and Distance is D_{M} then Network Lifetime is NL_7
Rule 8	If Energy is E_{M} and Distance is D_{L} then Network Lifetime is NL_8
Rule 9	If Energy is E_{H} and Distance is D_{VH} then Network Lifetime is NL_9
Rule 10	If Energy is E_{H} and Distance is D_{H} then Network Lifetime is NL_{10}
Rule 11	If Energy is E_{H} and Distance is D_{M} then Network Lifetime is NL_{11}
Rule 12	If Energy is E_{H} and Distance is D_{L} then Network Lifetime is NL_{12}
Rule 13	If Energy is E_{VH} and Distance is D_{VH} then Network Lifetime is NL_{13}
Rule 14	If Energy is E_{VH} and Distance is D_{H} then Network Lifetime is NL_{14}
Rule 15	If Energy is E_{VH} and Distance is D_{M} then Network Lifetime is NL_{15}
Rule 16	If Energy is E_{VH} and Distance is D_{L} then Network Lifetime is NL_{16}

with energy parameter. Hence, linguistic behaviour of the NL is shown in Table 8 and feasible, and optimal values of each rounds are shown in Tables 9, 10, 11, and 12.

In Table 8, several linguistic variables are shown based on chronological order of their increment behaviour of degree of membership or degree of truth value. It shows rule number 16 is the highest priority, i.e. NL_{16} having “Very High” energy and “Low” distance associates linguistic variable is “Too Excellent”. In Tables 9, 10, 11, and 12, feasible and optimal values are shown based on round 1–4 based on

Table 8 Linguistic variable of the output parameter network lifetime

Network Lifetime	Linguistic variable	Network Lifetime	Linguistic variable
NL_1	Very poor	NL9	Marginally Good
NL_2	Poor	NL_{10}	Good
NL_3	Bad	NL_{11}	Perfect
NL_4	Slightly bad	NL_{12}	Very good
NL_5	Average	NL_{13}	Highly good
NL_6	Medium	NL_{14}	Outstanding
NL_7	Satisfactory	NL_{15}	Excellent
NL_8	Fair	NL_{16}	Too excellent

Table 9 Dataset for round 1 under 500 sensor nodes for rule 1

Obj$_1$	Residual energy (x_1)	Distance (x_2)	Low residual energy	Very high distance
0.2767323	0.1660206E-01	0.5257915	$e_1 = 30$	$d_1 = 950$
			$e_2 = 150$	$d_2 = 1600$
0.2711497	0.1082386E-01	0.5206078	$e_1 = 20$	$d_1 = 960$
			$e_2 = 110$	$d_2 = 1580$

Table 10 Dataset for round 2 under 1500 sensor nodes for rule 6

Obj2	Residual energy(x_1)	Distance (x_2)	Medium residual energy	High distance
8.330248	0.5553337	2.832288	$e_1 = 100$	$d_1 = 510$
			$e_2 = 250$	$d_2 = 1200$
7.900281	0.632002	2.738769	$e_1 = 120$	$d_1 = 520$
			$e_2 = 240$	$d_2 = 1100$

Table 11 Dataset for round 3 under 2500 sensor nodes for rule 11

Obj3	Residual energy (x_1)	Distance (x_2)	High residual energy	Medium distance
40.34861	2.905053	5.648830	$e_1 = 180$	$d_1 = 350$
			$e_2 = 375$	$d_2 = 800$
36.12717	2.745668	5.346819	$e_1 = 190$	$d_1 = 370$
			$e_2 = 370$	$d_2 = 750$

Table 12 Dataset for round 4 under 3500 sensor nodes for rule 16

Obj4	Residual energy (x_1)	Distance (x_2)	Very high energy	Low distance
168.9655	12.06896	4.827602	$e_1 = 250$	$d_1 = 100$
			$e_2 = 450$	$d_2 = 350$
161.8230	12.02112	4.161206	$e_1 = 260$	$d_1 = 90$
			$e_2 = 360$	$d_2 = 260$

sensor nodes 500, 1500, 2500, and 3500 using rule numbers 1, 6, 11, and 16. Here, highest value indicates optimal decision, and lowest value indicates feasible value. Hence, optimal route is based on energy having “Very High” and distance having “Low”.

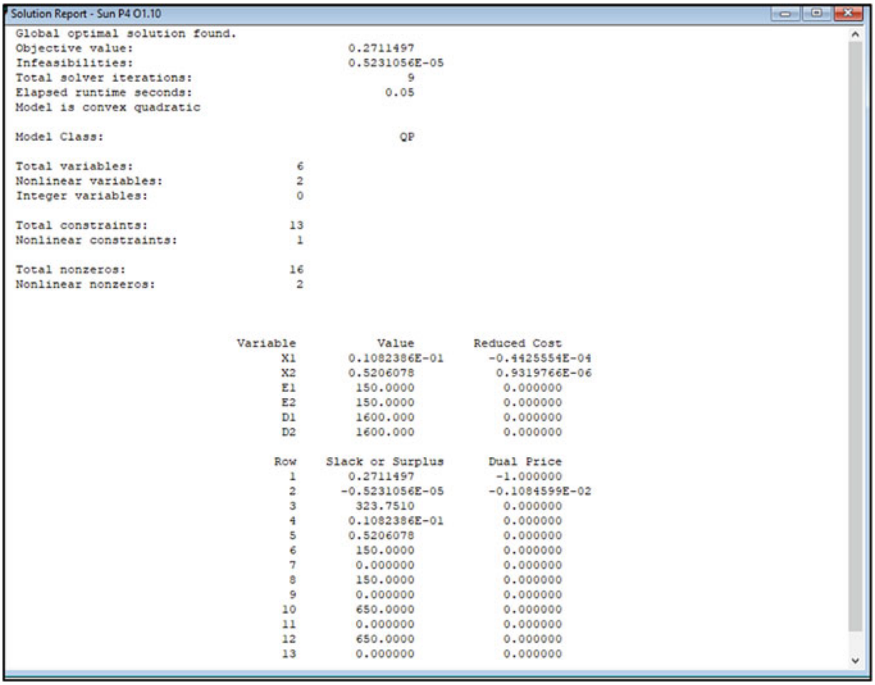
5 Simulation and Analysis

In this section, details of simulation and analysis are illustrated. The proposed method is simulated in LINGO optimization software based on fusion of linear and nonlinear formulations. The proposed method is simulated and verified in four optimization models based on four rounds where each round is repeated ten times based on different data set of same linguistic variable. The parameters details are shown in Table 13.

Figures 2, 3, 4, and 5 show feasible solutions of the optimization model based on sensor nodes 500 in round 1, 1500 in round 2, 2500 in round 3, and 3500 in round 4. In each round, data is tested 10 times based on same rule-based system. Data is shown in Tables 3, 4, 5, and 6, and their feasible and optimal data set is shown in Tables 9, 10, 11, and 12 based on rule-based system mentioned in Table 7 with linguistic behaviour mentioned in Table 8. First feasible value of network lifetime is 0.2711497 based on "Low" energy and "Very High" distance. Second feasible value of network lifetime is 7.900281 based on "Medium" energy and "High" distance. Third feasible value of network lifetime is 36.12717 based on "High" energy and "Medium" distance. Fourth feasible value of network lifetime is 161.8230 based on "Very High" energy and "Low" distance. Based on all feasible values, it is observed that when a number of nodes are increased as 500, 1500, 2500, and 3500, then network lifetime is also increased.

Table 13 Simulation parameter details

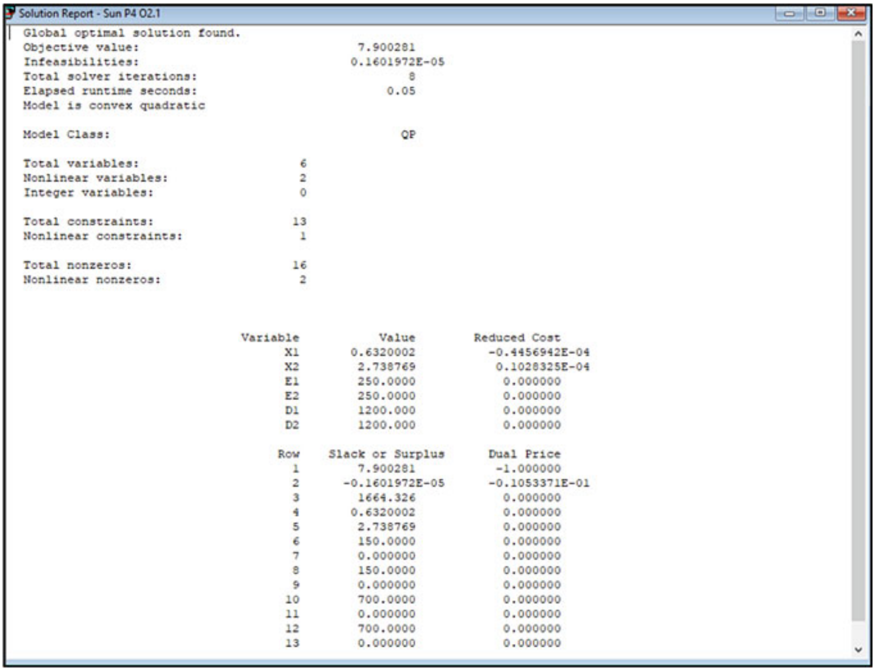
Parameter	Description
Windows	Windows 10 pro
MS Office	2013
Optimization software	LINGO
Energy	500 unit
Distance	1600 unit
Optimization	Maximization
Input parameters	Two (energy, distance)
Output parameter	Network lifetime
Total rules	16
Rounds	4
Iterations	4 × 10
Objective functions	4
Constraints	4 × 2
Nature of objectives	Nonlinear
Nature of constraints	Linear
Number of nodes	500–3500
Linguistic variables of energy	4
Linguistic variables of distance	4



```
Global optimal solution found.
Objective value:                              0.2711497
Infeasibilities:                              0.5231056E-05
Total solver iterations:                              9
Elapsed runtime seconds:                           0.05
Model is convex quadratic

Model Class:                                         QP

Total variables:                      6
Nonlinear variables:                  2
Integer variables:                    0

Total constraints:                   13
Nonlinear constraints:                1

Total nonzeros:                      16
Nonlinear nonzeros:                   2



                              Variable           Value        Reduced Cost
                                    X1    0.1082386E-01     -0.4425554E-04
                                    X2    0.5206078          0.9319766E-06
                                    E1     150.0000           0.000000
                                    E2     150.0000           0.000000
                                    D1     1600.000           0.000000
                                    D2     1600.000           0.000000

                                   Row    Slack or Surplus     Dual Price
                                     1    0.2711497          -1.000000
                                     2   -0.5231056E-05      -0.1084599E-02
                                     3     323.7510           0.000000
                                     4    0.1082386E-01       0.000000
                                     5    0.5206078           0.000000
                                     6     150.0000           0.000000
                                     7     0.000000           0.000000
                                     8     150.0000           0.000000
                                     9     0.000000           0.000000
                                    10     650.0000           0.000000
                                    11     0.000000           0.000000
                                    12     650.0000           0.000000
                                    13     0.000000           0.000000
```

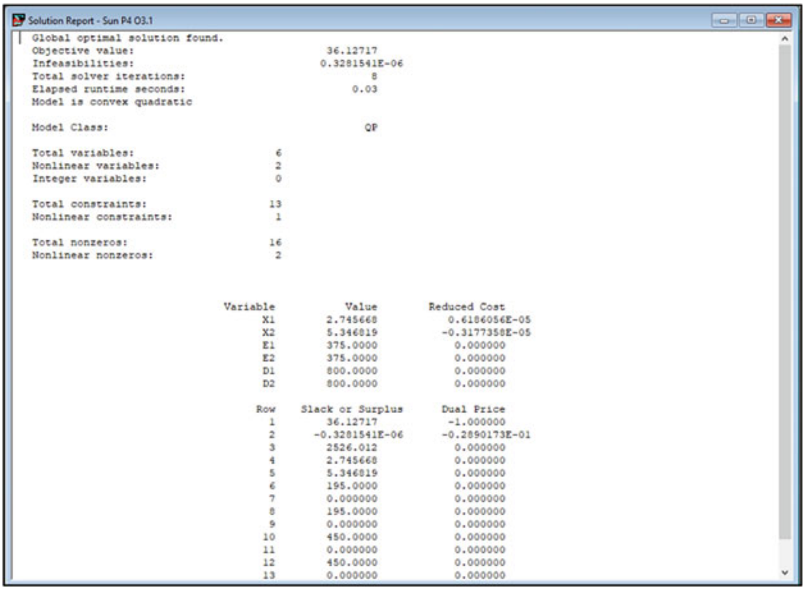
Fig. 2 Feasible solution for network lifetime in round 1 under 500 nodes

Figures 6, 7, 8, and 9 show optimal solutions of the optimization model based on sensor nodes 500 in round 1, 1500 in round 2, 2500 in round 3, and 3500 in round 4. In each round, data is tested 10 times based same rule-based system. Data is shown in Tables 3, 4, 5, and 6, and their feasible and optimal data set is shown in Tables 9, 10, 11, and 12 based on rule-based system mentioned in Table 7 with linguistic behaviour mentioned in Table 8. First optimal value of network lifetime is 0.2767323 based on “Low” energy and “Very High” distance. Second optimal value of network lifetime is 8.330248 based on “Medium” energy and “High” distance. Third optimal value of network lifetime is 40.34861 based on “High” energy and “Medium” distance. Fourth optimal value of network lifetime is 168.9655 based on “Very High” energy and “Low” distance. Based on all feasible values, it is observed that when a number of nodes are increased as 500, 1500, 2500, and 3500, then network lifetime is also increased.



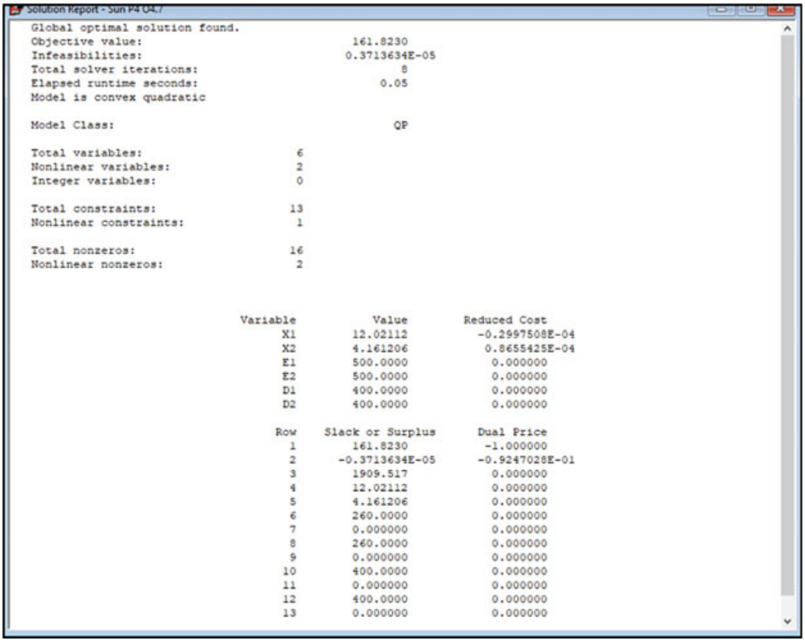
```
Global optimal solution found.
Objective value:                              7.900281
Infeasibilities:                             0.1601972E-05
Total solver iterations:                             8
Elapsed runtime seconds:                          0.05
Model is convex quadratic

Model Class:                                        QP

Total variables:                     6
Nonlinear variables:                 2
Integer variables:                   0

Total constraints:                  13
Nonlinear constraints:               1

Total nonzeros:                     16
Nonlinear nonzeros:                  2


                       Variable           Value        Reduced Cost
                             X1       0.6320002       -0.4456942E-04
                             X2        2.738769        0.1028325E-04
                             E1        250.0000        0.000000
                             E2        250.0000        0.000000
                             D1        1200.000        0.000000
                             D2        1200.000        0.000000

                            Row    Slack or Surplus      Dual Price
                              1        7.900281        -1.000000
                              2     -0.1601972E-05     -0.1053371E-01
                              3        1664.326         0.000000
                              4       0.6320002         0.000000
                              5        2.738769         0.000000
                              6        150.0000         0.000000
                              7        0.000000         0.000000
                              8        150.0000         0.000000
                              9        0.000000         0.000000
                             10        700.0000         0.000000
                             11        0.000000         0.000000
                             12        700.0000         0.000000
                             13        0.000000         0.000000
```

Fig. 3 Feasible solution for network lifetime in round 2 under 1500 nodes



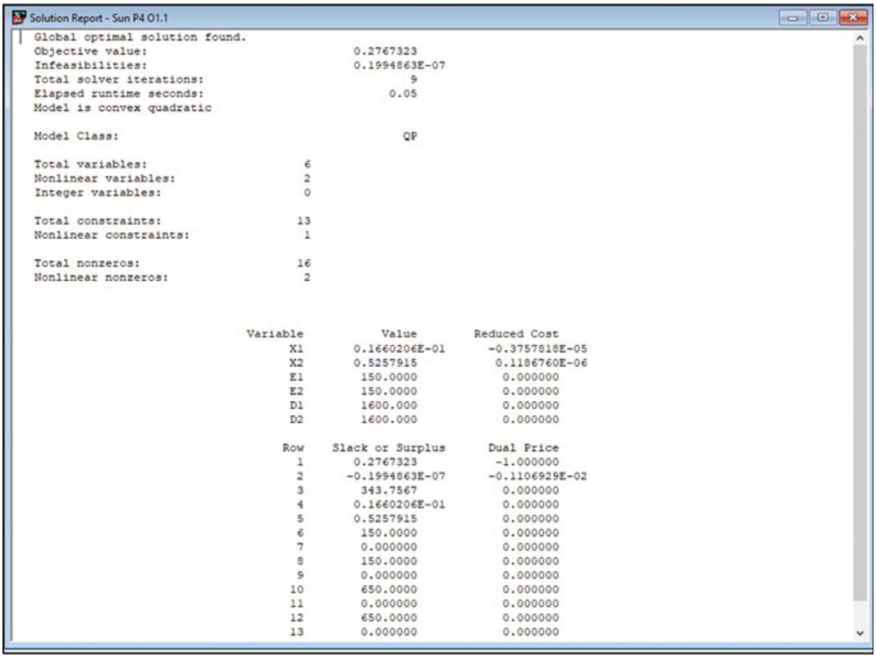
```
Global optimal solution found.
Objective value:                              36.12717
Infeasibilities:                             0.3281541E-06
Total solver iterations:                             8
Elapsed runtime seconds:                          0.03
Model is convex quadratic

Model Class:                                        QP

Total variables:                     6
Nonlinear variables:                 2
Integer variables:                   0

Total constraints:                  13
Nonlinear constraints:               1

Total nonzeros:                     16
Nonlinear nonzeros:                  2


                       Variable           Value        Reduced Cost
                             X1        2.745668        0.6186056E-05
                             X2        5.346819       -0.3177358E-05
                             E1        375.0000        0.000000
                             E2        375.0000        0.000000
                             D1        800.0000        0.000000
                             D2        800.0000        0.000000

                            Row    Slack or Surplus      Dual Price
                              1        36.12717        -1.000000
                              2     -0.3281541E-06     -0.2890173E-01
                              3        2526.012         0.000000
                              4        2.745668         0.000000
                              5        5.346819         0.000000
                              6        195.0000         0.000000
                              7        0.000000         0.000000
                              8        195.0000         0.000000
                              9        0.000000         0.000000
                             10        450.0000         0.000000
                             11        0.000000         0.000000
                             12        450.0000         0.000000
                             13        0.000000         0.000000
```

Fig. 4 Feasible solution for network lifetime in round 3 under 2500 nodes



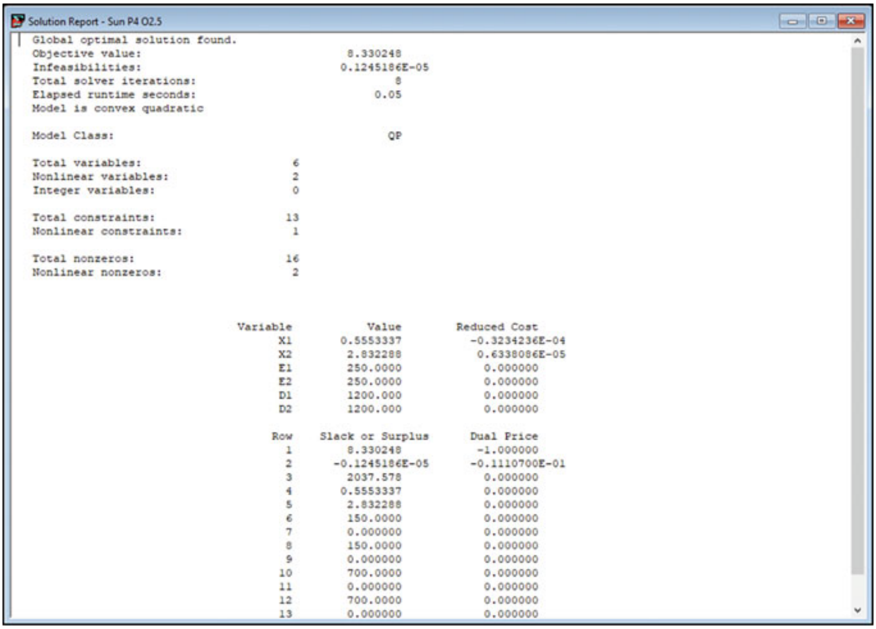
```
Global optimal solution found.
Objective value:                              161.8230
Infeasibilities:                             0.3713634E-05
Total solver iterations:                             8
Elapsed runtime seconds:                          0.05
Model is convex quadratic

Model Class:                                        QP

Total variables:                     6
Nonlinear variables:                 2
Integer variables:                   0

Total constraints:                  13
Nonlinear constraints:               1

Total nonzeros:                     16
Nonlinear nonzeros:                  2


                              Variable           Value        Reduced Cost
                                    X1        12.02112       -0.2997508E-04
                                    X2        4.161206        0.8655425E-04
                                    E1        500.0000         0.000000
                                    E2        500.0000         0.000000
                                    D1        400.0000         0.000000
                                    D2        400.0000         0.000000

                                   Row    Slack or Surplus      Dual Price
                                     1        161.8230        -1.000000
                                     2     -0.3713634E-05     -0.9247028E-01
                                     3        1909.517         0.000000
                                     4        12.02112         0.000000
                                     5        4.161206         0.000000
                                     6        260.0000         0.000000
                                     7        0.000000         0.000000
                                     8        260.0000         0.000000
                                     9        0.000000         0.000000
                                    10        400.0000         0.000000
                                    11        0.000000         0.000000
                                    12        400.0000         0.000000
                                    13        0.000000         0.000000
```

Fig. 5 Feasible solution for network lifetime in round 4 under 3500 nodes



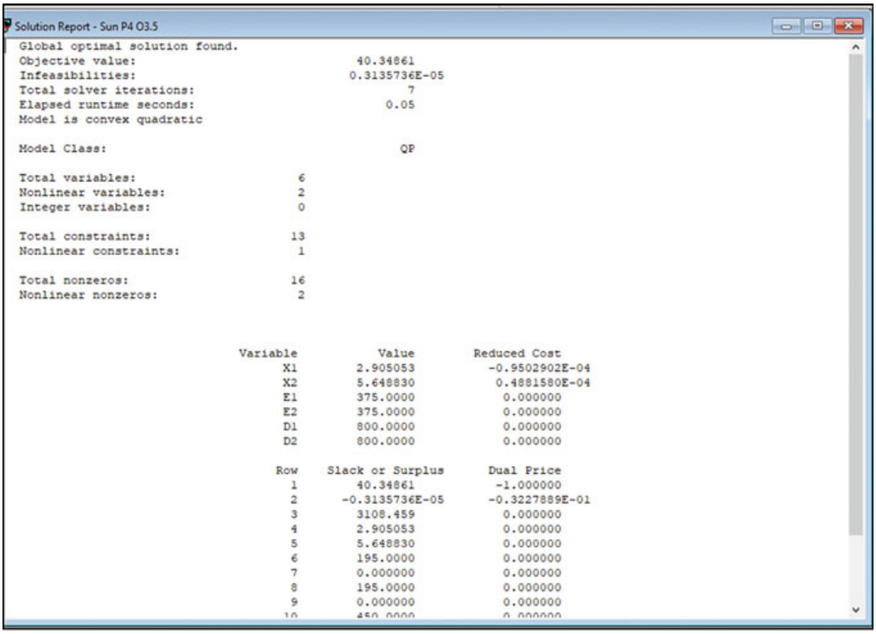
```
Global optimal solution found.
Objective value:                              0.2767323
Infeasibilities:                             0.1994863E-07
Total solver iterations:                             9
Elapsed runtime seconds:                          0.05
Model is convex quadratic

Model Class:                                        QP

Total variables:                     6
Nonlinear variables:                 2
Integer variables:                   0

Total constraints:                  13
Nonlinear constraints:               1

Total nonzeros:                     16
Nonlinear nonzeros:                  2


                              Variable           Value        Reduced Cost
                                    X1     0.1660206E-01     -0.3757818E-05
                                    X2       0.5257915        0.1186760E-06
                                    E1        150.0000         0.000000
                                    E2        150.0000         0.000000
                                    D1        1600.000         0.000000
                                    D2        1600.000         0.000000

                                   Row    Slack or Surplus      Dual Price
                                     1       0.2767323        -1.000000
                                     2     -0.1994863E-07     -0.1106929E-02
                                     3        343.7567         0.000000
                                     4     0.1660206E-01       0.000000
                                     5       0.5257915         0.000000
                                     6        150.0000         0.000000
                                     7        0.000000         0.000000
                                     8        150.0000         0.000000
                                     9        0.000000         0.000000
                                    10        650.0000         0.000000
                                    11        0.000000         0.000000
                                    12        650.0000         0.000000
                                    13        0.000000         0.000000
```

Fig. 6 Optimal solution for network lifetime in round 1 under 500 nodes



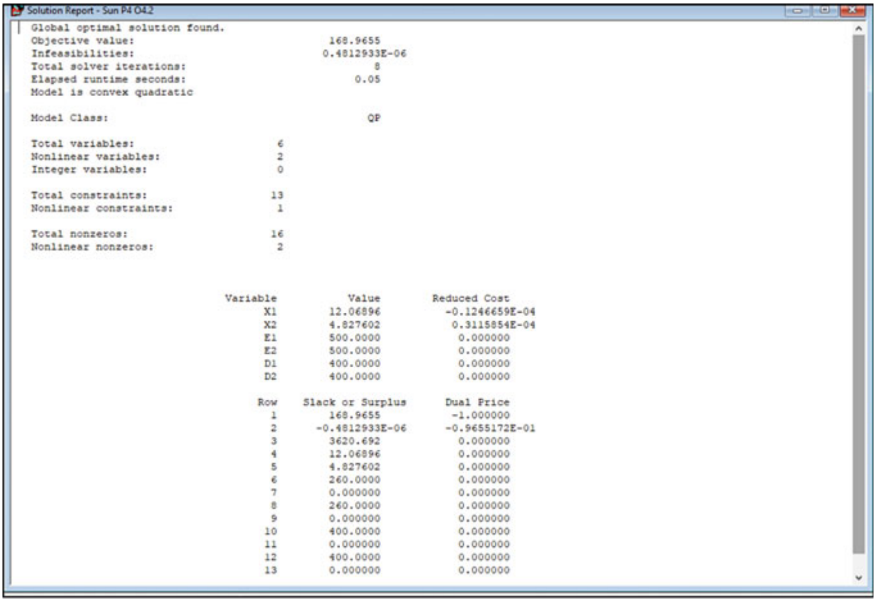
```
Global optimal solution found.
Objective value:                              8.330248
Infeasibilities:                              0.1245186E-05
Total solver iterations:                             8
Elapsed runtime seconds:                          0.05
Model is convex quadratic

Model Class:                                        QP

Total variables:                      6
Nonlinear variables:                  2
Integer variables:                    0

Total constraints:                   13
Nonlinear constraints:                1

Total nonzeros:                      16
Nonlinear nonzeros:                   2


                                Variable           Value        Reduced Cost
                                      X1       0.5553337       -0.3234236E-04
                                      X2        2.832288        0.6338086E-05
                                      E1        250.0000            0.000000
                                      E2        250.0000            0.000000
                                      D1        1200.000            0.000000
                                      D2        1200.000            0.000000

                                     Row    Slack or Surplus      Dual Price
                                       1        8.330248           -1.000000
                                       2     -0.1245186E-05     -0.1110700E-01
                                       3        2037.578            0.000000
                                       4       0.5553337            0.000000
                                       5        2.832288            0.000000
                                       6        150.0000            0.000000
                                       7        0.000000            0.000000
                                       8        150.0000            0.000000
                                       9        0.000000            0.000000
                                      10        700.0000            0.000000
                                      11        0.000000            0.000000
                                      12        700.0000            0.000000
                                      13        0.000000            0.000000
```

Fig. 7 Optimal solution for network lifetime in round 2 under 1500 nodes



```
Global optimal solution found.
Objective value:                              40.34861
Infeasibilities:                              0.3135736E-05
Total solver iterations:                             7
Elapsed runtime seconds:                          0.05
Model is convex quadratic

Model Class:                                        QP

Total variables:                      6
Nonlinear variables:                  2
Integer variables:                    0

Total constraints:                   13
Nonlinear constraints:                1

Total nonzeros:                      16
Nonlinear nonzeros:                   2


                                Variable           Value        Reduced Cost
                                      X1        2.905053       -0.9502902E-04
                                      X2        5.648830        0.4881580E-04
                                      E1        375.0000            0.000000
                                      E2        375.0000            0.000000
                                      D1        800.0000            0.000000
                                      D2        800.0000            0.000000

                                     Row    Slack or Surplus      Dual Price
                                       1        40.34861           -1.000000
                                       2     -0.3135736E-05     -0.3227889E-01
                                       3        3108.459            0.000000
                                       4        2.905053            0.000000
                                       5        5.648830            0.000000
                                       6        195.0000            0.000000
                                       7        0.000000            0.000000
                                       8        195.0000            0.000000
                                       9        0.000000            0.000000
                                      10        450.0000            0.000000
```

Fig. 8 Optimal solution for network lifetime in round 3 under 2500 nodes



```
Global optimal solution found.
Objective value:                              168.9655
Infeasibilities:                              0.4812933E-06
Total solver iterations:                             8
Elapsed runtime seconds:                          0.05
Model is convex quadratic

Model Class:                                        QP

Total variables:                      6
Nonlinear variables:                  2
Integer variables:                    0

Total constraints:                   13
Nonlinear constraints:                1

Total nonzeros:                      16
Nonlinear nonzeros:                   2



                                Variable           Value        Reduced Cost
                                      X1        12.06896       -0.1246659E-04
                                      X2        4.827602        0.3115854E-04
                                      E1        500.0000         0.000000
                                      E2        500.0000         0.000000
                                      D1        400.0000         0.000000
                                      D2        400.0000         0.000000

                                     Row    Slack or Surplus      Dual Price
                                       1        168.9655        -1.000000
                                       2     -0.4812933E-06    -0.9655172E-01
                                       3        3620.692         0.000000
                                       4        12.06896         0.000000
                                       5        4.827602         0.000000
                                       6        260.0000         0.000000
                                       7        0.000000         0.000000
                                       8        260.0000         0.000000
                                       9        0.000000         0.000000
                                      10        400.0000         0.000000
                                      11        0.000000         0.000000
                                      12        400.0000         0.000000
                                      13        0.000000         0.000000
```

Fig. 9 Optimal solution for network lifetime in round 4 under 3500 nodes

6 Conclusions

In this paper, network lifetime is optimized with the help of two basic parameters such as energy and distance. Nature of energy parameter is same as nature of network lifetime, but nature of distance parameter is different from the nature of network lifetime. So, the combination of energy and inverse of distance is used to form rule-based system for analysing route of the network. Fuzzy logic helps to estimate uncertainty of the network using membership function and mapped the imprecise network parameters efficiently. Decision-maker of the proposed model analyses the route of the network from source to destination nodes based on two efficiency terms as feasible and optimal. Feasible route indicates nearby best route which is the second choice of the route selection. Optimal route is the best choice for route selection from source node to the destination node. The optimal route is efficiently used to transmit the data packet within the WSN. This route helps to reduce energy consumption of the nodes and solve the ambiguity of the route selection in the network within several variations of the parameters.

References

1. Das SK, Samanta S, Dey N, Kumar R (2020) Design frameworks for wireless networks. Springer
2. Nguyen L, Nguyen HT (2020) Mobility based network lifetime in wireless sensor networks: a review. Comput Netw 107236. https://doi.org/https://doi.org/10.1016/j.comnet.2020.107236
3. Binh HTT, Dey N (eds) (2018) Soft computing in wireless sensor networks. CRC Press
4. Binh HTT, Nam NH (2018) Introduction to coverage optimization in wireless sensor networks. In: Soft computing in wireless sensor networks, pp 115–136. Chapman and Hall/CRC
5. Movassagh M, Aghdasi HS (2017) Game theory based node scheduling as a distributed solution for coverage control in wireless sensor networks. Eng Appl Artif Intell 65:137–146
6. Chen Z, Qiu Y, Liu J, Xu L (2011) Incentive mechanism for selfish nodes in wireless sensor networks based on evolutionary game. Comput Math Appl 62(9):3378–3388
7. Sun S, Chen N, Ran T, Xiao J, Tian T (2016) A Stackelberg game spectrum sharing scheme in cognitive radio-based heterogeneous wireless sensor networks. Signal Process 126:18–26
8. Yang W, Wang X, Song X, Yang Y, Patnaik S (2018) Design of intelligent transportation system supported by new generation wireless communication technology. In: Intelligent systems: concepts, methodologies, tools, and applications, pp 715–732. IGI Global
9. Jayakumar Loganathan J, Subbiah J (2020) Energy aware dynamic mode decision for cellular D2D communications by using integrated multi-criteria decision making model. Int J Ambient Comput Intell 11(3). (7 February 2020, IGI Global)
10. Shen J, Wang C, Wang A, Sun X, Moh S, Hung PC (2017) Organized topology based routing protocol in incompletely predictable ad-hoc networks. Comput Commun 99:107–118
11. Chatterjee S, Das S (2015) Ant colony optimization based enhanced dynamic source routing algorithm for mobile Ad-hoc network. Inf Sci 295:67–90
12. Fatemidokht H, Rafsanjani MK (2020) QMM-VANET: an efficient clustering algorithm based on QoS and monitoring of malicious vehicles in vehicular ad hoc networks. J Syst Softw 110561
13. Jat DS, Bishnoi LC, Nambahu S (2018) An Intelligent wireless QoS technology for big data video delivery in WLAN. Int J Ambient Comput Intell (IJACI) 9(4):1–14
14. Rasheed I, Banka H (2018, March) Query expansion in information retrieval for Urdu language. In: 2018 4th international conference on information retrieval and knowledge management (CAMP). IEEE, pp 1–6
15. Hao R, Yang H, Zhou Z (2019) Driving behavior evaluation model base on big data from internet of vehicles. Int J Ambient Comput Intell (IJACI) 10(4):78–95
16. Das SK, Tripathi S (2020) A nonlinear strategy management approach in software-defined Ad hoc network. In: Design frameworks for wireless networks. Springer, Singapore, pp 321–346
17. Singh P, Khosla A, Kumar A, Khosla M (2018) Optimized localization of target nodes using single mobile anchor node in wireless sensor network. AEU-Int J Electron Commun 91:55–65
18. Kotary DK, Nanda SJ (2020) Distributed robust data clustering in wireless sensor networks using diffusion moth flame optimization. Eng Appl Artif Intell 87:103342
19. Mohammed AS, Basha S, Asha PN, Venkatachalam K (2020) FCO–fuzzy constraints applied cluster optimization technique for wireless AdHoc networks. Comput Commun 154:501–508
20. Bera S, Chattopadhyay M, Dan PK (2018) A two-stage novel approach using centre ordering of vectors on agglomerative hierarchical clustering for manufacturing cell formation. Proc Inst Mech Eng Part B J Eng Manuf 232(14):2651–2662

Wireless Sensor Network Routing Protocols Using Machine Learning

Chaya Shivalingagowda, Hifzan Ahmad, P. V. Y. Jayasree, and Dinesh Kumar Sah

Abstract Routing is a predominant challenge in the field of WSNs because of insufficient power supply in each node. And low-transmission bandwidth required less memory space and handling limit. These sensors distributed randomly in nature and the environment, and each sensor nodes gather data from that environment for further analysis and additional processing and transmits the information and data to the base station. We discussed the different machine learning algorithms to develop routing protocols for the WSNs. These technologies have allowed the sensor to learn the experience data to make appropriate routing decisions and respond to changing the environment. We covered a wide range of machine learning (ML)-based routing protocols, such as distributed regression (DR), self-organizing map (SOM), and reinforcement learning (RL). This chapter affords a complete evaluation of the literature on the topic. The review has structured in such a way that suggests how network characteristics and necessities gradually viewed over time.

Keywords Computational intelligence · Distributed regression · Machine learning · Routing enhancement · Reinforcement learning · Routing protocols · Self-organizing map · Wireless sensor nodes

C. Shivalingagowda (✉)
Kalsekar Engineering College, New Panvel Mumbai and GITAM University, Vizag, India
e-mail: chaya.ravindra@gmail.com

H. Ahmad
Dr. A. P. J. Abdul Kalam Technical University (AKTU), Lucknow, India

P. V. Y. Jayasree
GITAM University, Vizag, India
e-mail: pvyjayasree@gitam.edu

D. K. Sah
Indian Institute of Technology (ISM), Dhanbad, India
e-mail: dksah.iitd@gmail.com

S. K. Das et al. (eds.), *Architectural Wireless Networks Solutions and Security Issues*, Lecture Notes in Networks and Systems 196,
https://doi.org/10.1007/978-981-16-0386-0_7

1 Introduction

WSN is one of the most assuring innovations in specific real-time applications due to its complexity, and cost-effectiveness also case of use [1]. The role of WSN is to track the area of interest and transmit the collected information to the base station for the post-analysis [2, 3]. WSN designers, in particular, need to fix fundamental issues identified with data accumulation, data replication, limitation, node grouping, energy-awareness routing, fault detection, or security. Wireless sensor network typically consists of a sensor, controller, and communication system. If the communication is wireless with each node, then it is called WSN. Machine learning is a technique of computer to learn and act like a human and improve their learning over time in an autonomous fashion, by collecting and processing a large amount of data through sensor nodes and feeding those data and collecting the information to form of observation and real-world interaction (Fig. 1).

Machine learning (ML) is a computer science and statistics area that encompasses a range of algorithms and methods that learn from datasets and can predict or support them. Machine learning (ML) developed as an artificial intelligence (AI) technique in the late 1950s. Over time, its emphasis changed, turning more to computationally viable and robust algorithms. These methods have widely used in different tasks in many applications, like classification, regression, and density prediction. Developing computational models for the learning process offers solutions and improves the present system. It also included IoT, cyber-physical systems(CPS), and machine-to-machine (M2M) integration [4].

Working : The first step in a wireless sensor network is to collect the data by the small-small wireless sensor nodes and collector sensor which is a sense the environmental condition and then the collector is collecting the data form the sensor after collecting the data is sent to the base station of the WSN after that data is sent to communication system and collecting the data then machine learning technique

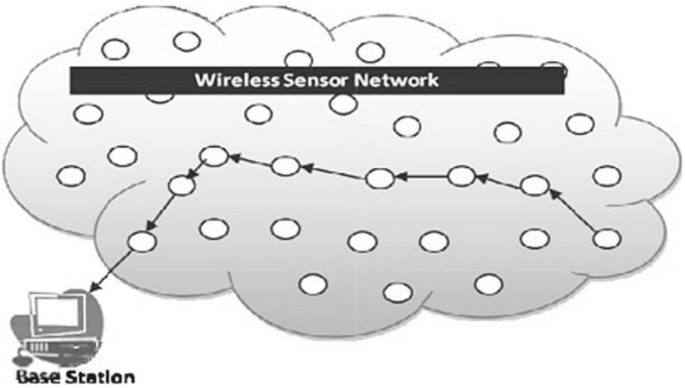


Fig. 1 Architecture of WSN [5]

and algorithm comes in the picture, the data which is collected from the sensor and collector arrange in proper format which is easily readable by the machine some preprocessing is performed on that data and noise is removed from the data after that data is divided into two section training and testing, the first step is to apply the ML algorithms on the training data and predict the category of and classification as output after that same technique is applying on testing data and compare these two data depended on their properties take a decision based on that data properties and applying different ML algorithm and take a proper decision which is applicable for WSNs system. In that process, a large amount of data transmitted, processed, and received by the base stations of WSN, which limit the computation because of the limited energy source and bandwidth of WSNs. The objective of ML is to lesson the data communicated and distributed on WSN. Machine learning implementation in routing plays an essential role in improving network performance. Sensor network dynamic behavior includes complex routing algorithms to optimize the performance of the system. The goal of this chapter is to include a detailed review of the use of machine learning techniques in sequence to enhance routing protocol output. WSNs have various other issues, such as centralized medium access, translation, optimum location, and distribution. Growing network lifespan by energy efficiency is a core problem in WSNs to execute these tasks energy costs. We assume that these dynamic distributed problems can be solved effectively and including routing and clustering. The early survey explains algorithms in their implementations. The fundamental explanation for this is that the memory and computing demands of machine learning methods are more significant than conventional WSN approaches. In this chapter, we wanted to provide a large selection of useful, up-to-date machine learning algorithms to compare their strengths and weaknesses. The rest of the chapter organized as follows. Section 2 presents the reader to the machine learning algorithm and types of machine learning. Section 3 explains the related work of machine learning techniques to WSN routing protocol. Section 3 specifies significant difficulties in routing layer and Sect. 4 outlines the application of ML in WSN and next section comparison of different ML techniques and CI techniques to evaluate the optimization achieved (Table 1; Fig. 2).

2 Machine Learning Techniques

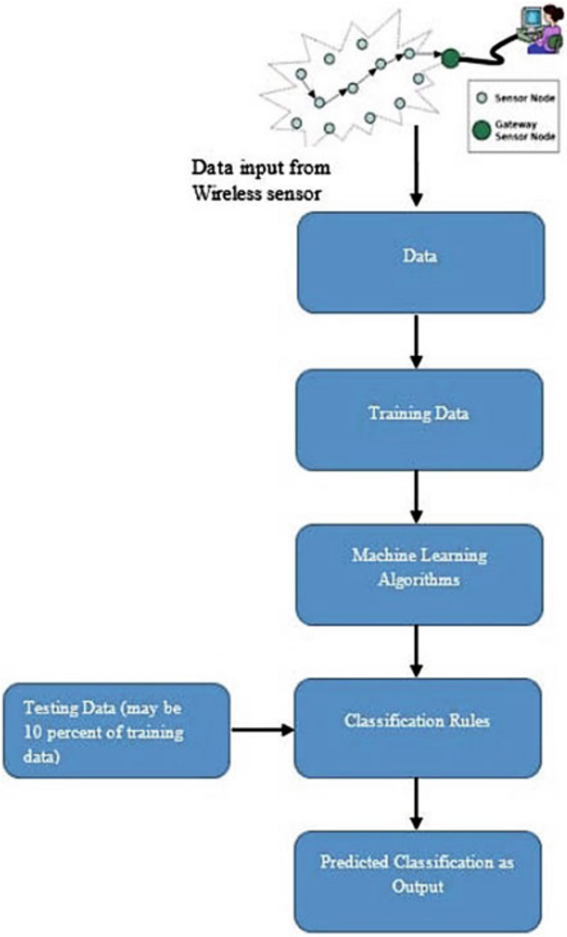
Machine learning is a method of analysis and prediction of the data to predict the feature outcome and build a model using different algorithms and techniques to predict the perfect output. Machine learning is a technique in which we give some input data and option some output on data and, based on that output, identify a relationship between them and patterns and trends to make a perfect decision wit minimum human interference. In this section, we are going to discuss some popular machine learning technique and their uses and advantages and disadvantages in wireless sensor network (WSN). Three types of machine learning are used to design wireless sensor network supervised learning, unsupervised learning, reinforcement

Table 1 Name and abbreviation used in the text

Abbreviation	Name
AI	Artificial intelligence
ACO	Ant colony optimization
ANN	Artificial neural network
ACD	Adaptive critic design
BS	Base station
CPS	Cyber-physical systems
CI	Computational intelligence
CS	Cuckoo search
CH	Cluster head
CS	Cuckoo search
DR	Distributed regression
DRQ	Distributed routing
FIS	Fuzzy inference method
FROMS	Feedback routing for optimizing multiple sinks
FL	Fuzzy logic
GA	Genetic algorithm
HIS	Human immune system
k-NN	*k*-nearest neighbor
M2 M	Machine to machine
ML	Machine learning
OWA	Ordered weighted averaging
PSO	Particle swarm optimization
Q-RC	Q-Routing with compression
RL	Reinforcement learning
SI	Swarm intelligence
SIR	Sensor intelligence routing
SOM	Self-organizing map
WSN	Wireless sensor network
ZEEP	Zone-based energy-efficient routing protocol

machine learning. The supervised learning describes a problem class that includes using a model to learn the mapping between example input and goal variable. There are two major categories of supervised learning: (a) classification & (b) regression. Unsupervised learning describes a series of problems concerning the use of a model for defining or extracting data relationships. Reinforcement learning represents a set of challenges where an agent works in an environment and needs to learn to use the input to work (Fig. 3).

Fig. 2 Machine learning in WSNs [28]


Sensor Node
Gateway Sensor Node
Data input from Wireless sensor
Data
Training Data
Machine Learning Algorithms
Testing Data (may be 10 percent of training data)
Classification Rules
Predicted Classification as Output

2.1 Supervised Learning

Supervised learning is one of the essential techniques in machine learning. In supervised learning, both input and output provided to the machine based on that it finds the relationship between the input and output variable and find the pattern between them when training the important system feature of supervised learning is to find the relationship between the input features and forecast objective output. Supervised learning solves the various problems in WSNs. Localization in a sensor node is defined as finding the geographical area to locate the sensor node of WSN. In simple WSNs, it required more time to find each location physically identifying the sensor nodes; it is an important task to handle such situations it required to reprogramming the architecture of WSNs. Machine learning algorithms easily classified between the anchor node in WSN and unknown nodes in the network. Mobile sensor nodes pro-

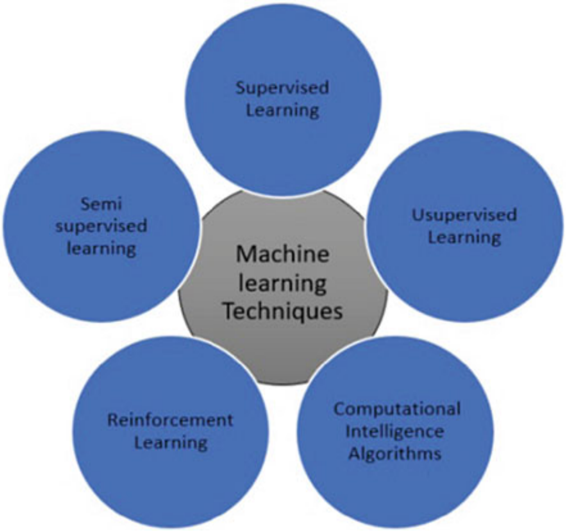


Fig. 3 Types of machine learning [31]

gressively and continuously change their situations/positions in WSNs, because of that to distinguish the exact location is increasingly satisfied and required more time, but it is quick with ML approaches. Using the ML algorithm in WSNs coverage and connectivity is increases find the number of sensor nodes connected can be found rapidly in WSN. This method, the machine model, is constructed using a specified training set (i.e., predefined data and labels). This model is employed to describe the scientific relationship between the parameters of input, output, and device. The use of these learning algorithms is to address numerous problems in WSNs, like identification and target tracking.

2.2 Logistic Regression

Logistic regression is a classification technique. Decision boundary (generally linear) derived based on probability interpretation results in a nonlinear optimization problem for parameter estimation. The goal of LR is to give a new data point and predict the class from which the data point is likely to have originated. The task of the classification is identifying a category that a new observation belongs to based on the previous data with known categories. When the number of categories is 2, it becomes a binary classification problem. Input features can be both qualitative and quantitative. If the inputs are qualitative, then there has to be a systematic way of converting them to quantities. Binary input like a “Yes” or “No” can be encoded as “1” and “0”.

I. **Linear classifier** is of two types if decision function is linear, binary classification can be performed depending on the side of the half-plane that the date falls in.

II. **Model as probabilities** Probability of a "Yes" or "No" gives a better understanding of the sample's membership to a particular category estimating that the binary outputs from the probabilities are straightforward through simple thresholding. Binary output for new samples can now easily be predicted using the following Eq. 1

$$p(x) = e(\beta 0 + \beta 1X)/1 + e(\beta 0 + \beta 1X) \tag{1}$$

If $\beta 0 + \beta 1X$ is non-negative then we get $p > 0.5$ and $Y = 1$ otherwise we get $p < 0.5$ and $Y = 0$. Decision boundary is the equation $\beta 0 + \beta 1X$. To prevent over-fitting, we need to penalize the coefficients. Regularization helps in building non-complex models that avoid capturing noise in the model due to over-fitting. The objective now becomes regularization that helps the model work better on test data due to the fact that over-fitting is minimized on training data.

2.3 K-NN Algorithm

The k-nearest neighbor algorithm uses both the label and unlabeled data used for classification and regression. It is a lazy learning algorithm where all computations are deferred until classification. It performs explicit generalization where the function is approximated. KNN does not show any operation on the training data. It is used when the nonlinear boundaries between the dataset; it is also used in the large dataset and also when the dataset is categorical. The KNN has used majorities of data points of the nearest location in the dataset. The parameter k is selected data and distance matrix to define the distance between the data point. The distance is measured by using different distance measuring techniques like Euclidean distance. These algorithms do not make any sense about underlying data distribution.

2.4 Unsupervised Learning

Unsupervised learning is a technique in which there is no label output data associated with the input. The first approach of unsupervised learning is to detect a pattern in a cluster. The main advantages of unsupervised learning in a wireless sensor network are to solve many problems like anomaly detection, routing, and data aggregation. Anomaly detection is improved using an unsupervised machine learning technique. It minimizes the complexity and communication problems; it also has advantages in fault detection attack. It detects outliers in WSNs in supervised methods using the past. Previous dataset is possible to adjust the parameter of sensor nodes dynamically. Routing is one of the most critical parameters in a wireless sensor network. There are many advantages by using ML-based routing like while using ML routing; it adopts changes in the environment without physically reprogramming and it also benefits ML in optimal routing low earning communication overhead and delay-aware. Data aggregation is a process of collecting the data to different sensor nodes called data aggregation; data aggregation affects the various parameters of the WSNs system like communication overhead units and power memory loss by using the ML unsupervised algorithm. It uses in dimension reduction of the data of the sensor node level. Therefore, it reduces communication overhead in the network; it performs this reduction in senor nodes to reduce the delay in the network nodes. It also used to reduce the dimension of the sensor node using environmental conditions. It is called unsupervised learning without a teacher. It is mostly related to the concept of using a series of gathered or clustered findings from a sample. This learning methodology differs from the supervised learning methodology in that there is no output vector, and the learning under this method does not require an input label.

2.5 K-Means Clustering

A technique to divide N observations data points into K clusters, where $K \leq N$. These observations, based on the group with the nearest mean, clustering can be considered one of the most straightforward and essential unsupervised learning techniques. Hence, as every other problem, it deals with finding the unlabeled data. A similar object belongs to the same group, and a different object belongs to the distinct group.

2.6 Reinforcement Learning

Reinforcement learning is a machine learning technique that is based on the trial-error principal to find the best possibilities in the search space. Reinforcement learning is learning continuously through the environment and gathers the data by sensor nodes to take appropriate action by the agent. It also provides excellent results because there is no need to store a large dataset. It uses the Q-learning technique where each agent gathers the data to the environment and generates a sequence of observation, and based on previous observation, it takes a proper decision. Therefore, reinforcement learning provides an outstanding result in WSN. It is helpful in the communication range control of sensor networks. Reinforcement learning basically consists of five-part like environment, state, reward, action, agent in WSNs. The environment consist of many small-small sensors which collect the data to the environment these data and information os pass to agent through the State State is playing the mediator role Between the environment and agent After collecting the information and data agent

Table 2 ML-based routing algorithms for WSNs

ML technique	Studies	Topology	Complexity	QoS	Environment	Mobility
ANN	[36]	Tree	High	No	Centralized	Static nodes
	[37]	Tree	Moderate	Yes	Distributed	Static nodes
	[38]	Tree	Moderate	No	Distributed	Mobile nodes
Deep learning	[39]	Hybrid	High	No	Centralized	Mobile nodes
SVM	[40]	Hybrid	Moderate	No	Distributed	Static nodes
Bayesian	[41]	Tree	Moderate	No	Distributed	Static nodes
	[42]	Hybrid	Low	No	Centralized	Static nodes
	[43]	Hybrid	Moderate	No	Centralized and Decentralized	Mobile nodes
k-means	[44]	Hybrid	Low	Yes	Distributed	Static nodes
	[45]	Tree	Moderate	No	Distributed	Static nodes
	[46]	Hybrid	Moderate	No	Centralized	Static nodes
SVD	[47]	Arbitrary	Moderate	No	Distributed	Static nodes

takes the proper decision and give the rewards to the environment depend on that rewards agent to take the action if the environment is changed rapidly then the agent is able to take the proper action it is also making changes depend on previous data but depend on the environmental situation therefore it is able to take proper decision instantly it also takes less time to make any changes is wireless sensor nodes and the accuracy is also good because of instant decision-making ability in reinforcement Learning technique it is using Q Learning techniques which is also advantages in dedicating the sensor nodes properly and also Q learning is provided good accuracy because of that properties is reinforcement learning one of the important ML Technique which is used in WSNs system. In a wireless sensor network, reinforcement learning works on the sensor node to interpret, process, and communicate data with its context. The agent must acquaint himself in this machine learning approach to take the necessary measures by making use of his own environmental familiarises. This technique for machine learning depends purely on two measures: trial-and-error check and delayed (Table 2; Fig. 4).

3 Related Work

The authors of [7] explored implementations of three standard machine learning algorithms at all communication levels in the WSNs (i.e., decision trees, neural networks, and reinforcement learning). In comparison, targeted surveys have also

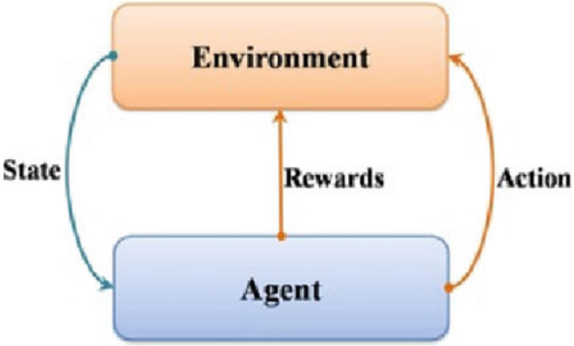


Fig. 4 Visualization of reinforcement learning [2]

published that focus on the use of machine learning, in particular, WSN challenges. For example, [8, 9] discussed the creation of effective outer detection methods, and some of these methods depend on machine learning principles. In [10], the author addresses techniques of artificial intelligence to tackle problems in WSNs, for example, routing, data collection and convergence, work management, optimum delivery, and localization. This computational intelligence is a subset of machine learning that concentrates on the biologically inspired method [11]. The zone-based energy-efficient routing (ZEEP) protocol was developed for mobile sensor networks in [12] using a fuzzy method. ZEEP uses the fuzzy inference method (FIS) to pick a range of CHs based on parameters like mobility, energy, and distance, including neural networks, fuzzy structures, and evolutionary algorithms. In [13], a routing protocol is based on deep learning presented, including the base station being an interface. It indicates the path which the base station has preserved, allocated, and retrieved. The suggested deep learning-based algorithm utilizes dynamic routing in a mobile sensor network. Firstly, the base station generates a set of simulated routing routes and determines the best path from them. For WSNs, an ACO-based routing algorithm has implemented in [14]. We consider various criteria, such as a node’s residual energy, transmission frequency, transmission direction, and the most precise route from the source nodes to the base station to determine the optimum routing. This algorithm achieves low energy loss and lengthens the network lifetime as we have seen many advantages in ML technique, but there are some limitations and disadvantages also using these machine learning algorithms. Machine learning algorithms are not able to produce an instant result and accurate prediction because of the ML algorithm is need to train first for training required a large amount of historical data. If the data size is large, then the more energy is consumed because of that energy limitation in WSNs and the high complexity of the ML algorithm. Also, the validation and prediction of the real-time data are a very difficult task in wireless sensor network because of the data which is an option from the sensor that is not in a proper format, and it also changes when the condition of the environment is also changing the size of the sensor which is important factor in WSNs [15]. When the size changes, then the amount of data also changes. Therefore, it is difficult in some situations to predict proper accuracy. Sometimes it is challenging to select a good ML technique for complex issues in WSNs.

4 Routing Layer

The routing problem applies to the fundamental question regarding moving a data packet from a node to another when there is no direct node connection in the network. This enigma also recognizes as multihop routing, which means that many intermediate nodes usually transfer the data packets to their destination. A routing protocol specifies the series of intermediate nodes needed to guarantee the packet distribution. There is a distinction between unicast and multicast routing protocols. The data packets routed from a single source to a unique destination comes under the unicast routing protocol. The data packets concurrently routed to multiple destinations which are known as the multicast routing protocol. These protocols remains a vast deal of literature on routing concerning WSNs and wireless ad hoc networks in general. The fundamental problem is to handle weak transmission connections, node crashes, and node stability and, most significantly, to allow effective use of resources. With the machine learning algorithm and computational intelligence, an attempt to produce energetic routing efficiency and increase the network lifetime is the primary concern of the author [7, 15].

5 Applications of Machine Learning in Routing

Machine learning methods can be used by approximating sensor data with algorithms to reduce interaction. In the area of WSNs, the main objective of learning strategies is to minimize the amount of energy used by sensor nodes by limiting the amount of contact within the node. Machine learning techniques are attractive in sensor networks for the following four reasons. There is generally an amount of consistency of data from the sensor network. Redundancy derives from the fact that relatively nearby sensors are likely to obtain identical measurements, and measurements taken at two different instants are strongly related as well. The exact measurement is rarely important, and generally, the observer accepts approximations. The reduced the accuracy, the more versatile the machine learning models and hence the communications gains that are possible. This makes it possible to design strategies for data collection, which exchange in energy precision. It associated with the involvement of software agents in an atmosphere to optimize the idea of cumulative prizes. The method elects a standard action and, at specific action, earns a credit. However, the right or successful move is not decided in advance. Therefore, the program carries out various actions and learns from its experience [13]. Applications of machine learning techniques in routing protocol are most commonly used in MANET and WSN distributed problems like routing, timing, medium access control, and network positioning.

6 Routing in WSN

Whenever a node is required to transmit a data packet, a nearby node may be chosen depending upon the values contained in its routing tables. Concerning a routing algorithm, the default action is to greedily select the neighboring node, including the best value of the routing table. The routing algorithms have specific guidelines for the ML that enables a sensor network to learn from past interactions, to implement appropriate routing decisions, and to respond to the changing environment–the estimation of the cost functions and the updating of the routing tables at execution time. ML reduces the difficulty of a standard routing issue by grouping it into more easy sub-routing problems. The nodes in each problem form by the formations of the graph, including only their immediate close neighbors, a low cost, reliable, and present-time routing, accomplished. To meet quality of service routing task specifications, use fairly basic computational methods and classifiers.

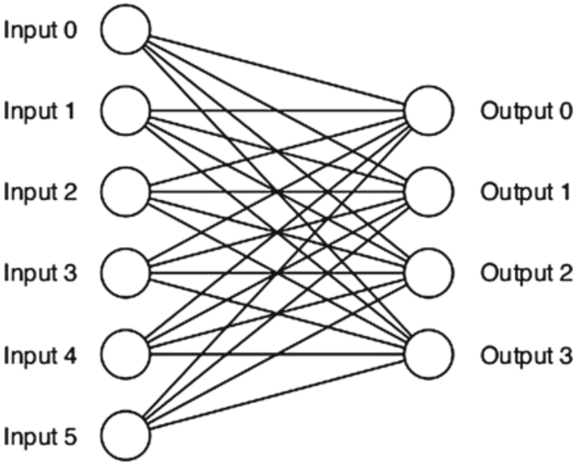
6.1 Distributed Regression Framework

This framework is an efficient and general framework in-network sensor data modeling. Within this context, the sensor network nodes cooperate to match each of their local measurement optimally with a global function. The algorithm is based on linear regression of the kernel, where this model takes the weighted sum of local dependent function [16]. Kernel functions map training observations to a certain function space to allow data analysis. The proposed system exploits the assumption that multiple sensor readings are strongly correlated. It would reduce the overhead for contact and identifies the sensor structure data. These findings act as an essential step in the creation of the distributed wireless network using different ML techniques and algorithms. The benefit of using these methods is very strong fitting performance low computing overheads.

6.2 ANN

Energetically effective and reliable artificial neural network routing scheme for WSNs called ELDC. In this methodology, the network remains focused on enormous data collection that includes almost all scenarios to make the system very useful and environmentally efficient. ELDC technique is exceptionally energy-efficient, capable of raising the lifetime of sensor nodes. An artificial neural network creates dominant threshold values for choosing the chief node of a system, and a cluster head depends upon the technique of backpropagation, including facilitates intelligent, efficient, and reliable system structure. ANN for the training of the protocol considers the appropriate parameters like remaining capacity, node distance from CH, distance from the border node, distance from the border node to BS, the traffic load on the specific link, and the likelihood of insurance status of the particular network [17] (Fig. 5).

Fig. 5 Architecture of ANN



6.3 Data Routing Using Self-Organizing Map (SOM)

The AI techniques in WSN are the self-organizing map that uses the concept concerning the self-organization concept. It is based on unsupervised learning to find efficient routing strategies, which are called sensor intelligence routing (SIR). SIR makes a minor improvement to the Dijkstra's algorithm to build the foundation of the node and the smallest paths from a base station to any node on the network. The second-layer neurons engage with each other during route learning to assign high weights within the learning chain. The weights of the winning neuron and its corresponding neurons are then modified to match the input patterns. The training process is a highly analytical method to construct neural networks [2]. SOM consists of two phases, including the learning phase and the execution phase. The learning phase arranges a two-dimensional map of the neurons, and it is done within a resourceful central station. The execution phase runs on network nodes. Each sensor node calculates the QoS of its links, which collects input samples and runs the algorithm of the winning neuron election. Each sensor node calculates the QoS of its connections, which gathers input samples and runs the algorithm of the winning neuron election. The complexity and the learning methods are the key challenges to implement this type of algorithm, whether the topology and structure of the network change [2, 17] (Fig. 6).

The use of AI technologies (e.g., neural networks) in WSNs declared to be a valuable method for optimizing network performance. The tremendous work of applying a SOM method on a sensor node implies that the usage of AI technologies will boost the efficiency of the WSNs.

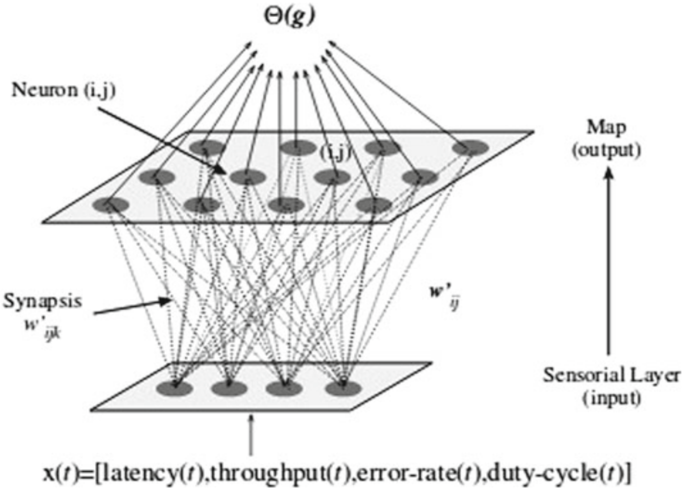


Fig. 6 SOM architecture [10]

6.4 Q-Learning

Q-learning is a simple and effective RL algorithm [18]. It does not require any environment model and can be used for online meaning feature learning. It is very easy to execute and has a good memory and resources needs optimum balance. In our multiple-sink example, each sensor node is an infinite learning agent, and actions are routing options for the next-hop(s) into a subset of sinks using separate neighbors. The following steps are followed to find Q-learning solutions agent states, action, Q-values, and updating *Q* values if not satisfied, reward function. A simple solution is to greedily choose the best (lowest) Q-value for the operation, the flexibility of the Q-learning methodology, and the prospect of enhancing current Q-learning protocols. There are different types of *Q* routing that are as follows.

1. DRQ-Routing distributed routing
2. Q-RC (Q-routing with compression)
3. AdaR-Yet another Q-learning-based routing
4. Q-PR: Q-probabilistic routing
5. RLGR geographic-based.

Some of Q learning machine learning techniques and protocols discussed here.

6.4.1 Routing Enhancement Using Reinforcement Learning (RL)

The reinforcement learning falls within a requirement of the algorithms of ML; meanwhile, an agent goes through encounters with the environment learning how to do it step by step producing some form of rewards. The primary purpose is to

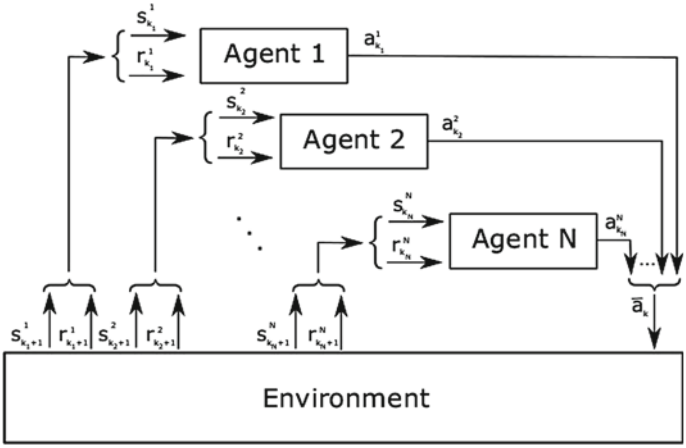


Fig. 7 Multiagent reinforcement learning schema

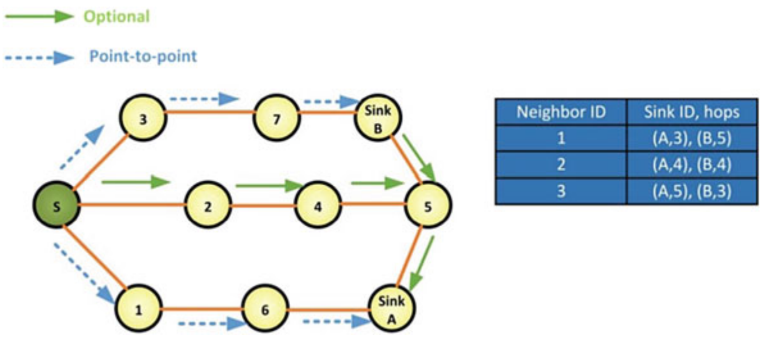
understand the environment through the method of optimizing incentives to identify appropriate policies. An RL agent knows what action by itself lead to optimality (Fig. 7).

A hierarchical algorithm is used in the enhanced learning model for opportunistic routing schemes. Figure 6 shows multicast routing; a node transfers the identical message to numerous recipients. MARL is appropriate for network distribution where multiple agents still exist. There are two types of MARL, namely single-hop and multihop [19].

6.4.2 Feedback Routing for Optimizing Multiple Sinks in WSNs with Reinforcement Learning (FROMS)

This protocol depends on Q-learning, and it operates on the mechanism of exchanging local knowledge of a node to other nearby nodes as input without creating any overhead network. The network used to implement this protocol is architectures in multisink form. The best thing about this routing system is to consider both the techniques of discovery and development to find an optimal route. The only extraction will, on the opposite, lead to a local optimum solution, unnecessary overhead exploration can prolong the time elapsed for the discovery of the path (Fig. 8).

Figure 7 shows a routing scenario in FORMS, which consists of one source node and two sink nodes. The solid line represents best-shared route, and the dashed line indicates the point to point routes. The exceptional commitment of this protocol that multiple sinks take into account in network architecture and by using multiple sink nodes significantly decreases network overheads. Even if any node fails, FORMS has a recovery mechanism. On the other hand, FORMS is vulnerable to node failure, and instability of sinks can lead to errors in the routing [20] (Table 3).



Neighbor ID	Sink ID, hops
1	(A,3), (B,5)
2	(A,4), (B,4)
3	(A,5), (B,3)

Fig. 8 Routing scenario in FORMS

Table 3 Routing algorithms for WSNs based on ML technique

Machine learning techniques	Topology	Routing protocol	Complexity	Qos	Environment
K-means	Hybrid/tree	Distributed regression framework	Low/moderate	No	Distributed
ANN	Tree	ELDC	High/moderate	No/Yes	Centralized/ distributed
Q-learning	Flat/multihop	Data routing using SOM	Low	No	Distributed
Q-learning	Flat/multihop	MARL	Moderate	Yes	Distributed
Q-learning	Flat/multihop	FORMS	High	No	Distributed

7 Computational Intelligence Algorithms

The principal drawback to routing algorithms focused on reinforcement learning is the restricted awareness of potential knowledge. The algorithms are thus not appropriate for extremely complicated conditions since it takes a long time to acquire the optimal paths. A study of improvement strategies that alter or encourage smart actions in diverse and dynamic environments lead to Computational Intelligence algorithms(CI) [33]. Such structures represent concepts that possess a potential for informing or adapting to new things, generalizing, abstracting, exploring, and associating. CI is illustrated because method descriptions and information resources that can individually position actual sensory statistical expertise have accurate and timely responses and endure high tolerance to faults. Computational intelligence frameworks are responsive to WSN's sophisticated design. The following subsections briefly explain several CI paradigms used in WSN clustering [31, 32].

7.1 Genetic Algorithm

Genetic is an adaptive heuristic search algorithm that models the genetic, biological evolution. It has shown to be a robust optimizer looking for a community of solutions and demonstrating versatility in addressing complex issues. Several NP-hard problems have been solved successfully. A key challenge in the GA problem is the translation of the problem into a series of chromosomes, while each of these chromosomes constitutes a solution to the problem. A fitness function measures the position of each chromosome. The main feature of GA is to find an optimization solution for the search route [21, 34]. To seek global optima, it retains a search frontier and solves problems of multicriterion optimization. However, GA's capability to describe rule-based, permutation-based, and proactive approaches to various pattern recognition and ML problems is a more precise benefit. GA-Routing Method for creating an aggregation tree that covers every sensor node, and it is used for homogeneous networks and small size networks. It enhances the lifetime of the node and improves node energy output. GA-EECP technique works to build clusters that are energy efficient. Proposes cluster size, relative size to the base station, energy flow, average cluster distance variance, and the number of communications as determining variables. It builds on Large Scale Networks [22].

7.2 Ant Colony Optimization

The ACO (Ant Colony Optimization) algorithm derives from the real action of ants that interact with each other via a medium designated as a pheromone [23]. During the walk, the ants lay a chemical signal on the space and detect the pheromone's current intensity for guidance. At first, no pheromone accumulated on the branches, and the ants do not know the length of the branches. However, if a shorter version is detected, pheromone will be given at a higher rate than the longer one. By contrast, pheromone modification is limited to a fixed area. Prematurity and instability can be easily avoided in this situation, but the convergence speed is slowed down. Second, at least, all of the connections have an identical pheromone, the ants roam with no warning, spontaneously. The list would remain inaccurate, and the other is misguided. Third, any mistake in route selection or pheromone updating will affect the final optimization in the ACO process. Some of the ACO based Routing protocols are BAR Protocol chooses the routing route by the allocation of probabilities. Many different paths are built in this protocol. It's ideal for applications where the priority is on the sink node, and ordinary nodes need to retain robust connectivity to the sink [24]. Another Protocol EEABR [24] proposes that Simulation tests show that it minimizes the contact load in various WSN situations and maximizes energy consumption, which stresses the increased lifespan of the network.

7.3 Particle Swarm Optimization

Swarm Intelligence (SI) is one of the emerging approaches for these heuristic computation challenges from the bio-inspired computation. James Kennedy and Russell Eberhart developed Particle Swarm Optimization (PSO) in 1995. It is a strong nonlinear stochastic optimization strategy focused on swarms mobility and intelligence. It is a trigger from bird or fish social behavior, where a group of birds is looking for food in an environment by chasing the closest bird to the food. It blends local search methods. Population-based stochastic optimization beginning with a randomly generated population group. They have fitness attributes for assessing their population, and with random processes, they refine the population and aim for the maximum. The PSO does vary from GA, though, in that there is no overlap and mutation [25]. Its particles do not die while their internal velocity modified. Finally, the structure of knowledge exchange in PSO is noticeably different. In a multidimensional space, each particle is viewed as a point and adjusts its location, determined by two components: the cognitive component and the social component arising from neighbor contact [35].

7.4 Artificial Immune System

The artificial immune system is primarily motivated by the natural or actual immune systems, and specific features of the artificial immune systems are modeled. The human immune system (HIS) has tremendous ability to suit sequences. It is used to discriminate among the alien cells that join the body (antigenic or non-self) and those that belong to the body (self). HIS combats the antigens and stores its structure. The networks depend on the idea that the immune system is using a typical network of interconnected B cells—a part of the adaptive immune system to recognize the antigen. The ability of the two B cells to interconnect is directly proportionate to the bond they participate in. A population of B cells holds two groups of subpopulations original population and the cloned population. If the excellent binding accomplished, the B cell is cloned and mutated, resulting in several antibodies. Once a new B cell generates, it connects to the closest B cells in the network. Unless the new cell cannot be inserted, the population is discarded. If all bindings fail, a B cell is produced utilizing the antigen as a template and then included in the artificial immune network, pattern recognition problems, classification, clustering, anomaly detection, and computer virus detection [25, 26].

7.5 Fuzzy Logic

A fuzzy logic (FL) is an arrangement of mathematics developed to describe theoretical human reasoning. Unlike the traditional set theory, which requires components to either be incorporated in a set or not. It can set average values based on semantic variables and rules of inference. In other words, in a fuzzy package, a particular entity is authorized to have an influenced membership, which is in the range [0,1]. There are two parts in fuzzy logic:

I. To determine the membership which corresponds to a given value of a linguistic attribute, a fuzzy membership function is set. The membership function may be programmed to represent the optimal goodness actions of a target in a versatile manner, depending on specific needs.
II. FL provides a floating aggregation operator, ordered weighted averaging (OWA), as an alternative to weighted numbers, to develop a multiobjective cost function.

FL was successfully implemented in optical image processing, pattern recognition, and control systems such as automotive motor controls, power systems, home appliances, elevators. However, FL is ideal for applying heuristics clustering and improving routing to accomplish various goals simultaneously. This algorithm generates non-optimal solutions, however, and fuzzy rules upon topology modifications tend to be relearned [27]. The FL protocols incorporate heuristics clustering or routing optimization to accomplish several entities simultaneously. FCH sees capacity, focus, and centrality as three linguistic factors to assess the probability of being the head of the cluster. In terms of the first description, it was confirmed to achieve a significant improvement in network lifespan. FMO is indeed excellent in this regard with many other well-recognized heuristics in Web routing. Fuzzy multiobjective routing characteristics applied to achieve several WSN routing goals at the same time.

7.6 Cuckoo Search

The suggested algorithm for multicast routing built on the CS algorithm, which relies upon the environment concerning the cuckoo species [27]. The algorithm takes into account three rules:

I. The bird lays one egg at a time, which is a solution, on a randomly selected host nest
II. The high-quality egg or the best-fit solution is the algorithm
III. The host bird leaves the poor quality alien egg when picking the best with a probability

Application of the newly developed fitness functions in the CS algorithm solution to automate multicast routing. Si and Pb (probability of egg discovery) have values between zero and 1. Eu is derived from a regular normal distribution with zero mean and standard deviation of unity. It can further be practiced as a Levy Flight distribution and holds a random step-length walk depends on the current position plus a chance of transition [28, 29].

7.7 Adaptive Critic Design

A dynamic optimization, a novel action-dependent adaptive critical design (ACD), is created. The suggested combination of an optimization-based particle swarm actor and a neural network critique illustrated by complex sleep scheduling of wireless sensor motes. The adaptive critical approach deters optimal control laws for a system by integrating two subsystems, an agent (dispensing feedback signals) and a critic (learning the ideal output index for each of the functions associated with the program) [30]. The ACD-determined integrated sleep pattern appears in high-quality data collection and enhanced energy efficiency. One possible way to this research to be expanded is to explore lightweight types of PSO and compact neuronal structures. Besides, there is room for investigating the use of the ACD in collective multinode sleep scheduling.

8 Conclusion

In this chapter, machine learning and computational intelligence technique improve WSN capability to respond to the complex behavior of the natural surroundings. Analysis of CI algorithms, which act as a reference for practicing CI algorithms for WSNs, is provided. An innovative CI methodology known as adaptive critical design (ACD) aims to provide realistic, optimal/suboptimal solutions to the problem of dispersed sensor scheduling has been discussed. In this chapter, a thorough analysis and current investigation of ML and CI techniques in routing has been presented.

References

1. Alsheikh MA et al (2014) Machine learning in wireless sensor networks: algorithms, strategies, and applications. IEEE Commun Surv Tutorials 16:4
2. Kumar DP, Amgoth T, Annavarapu CS (2019) Machine learning algorithms for wireless sensor networks: a survey. Inf Fusion 49:1–25
3. Rawat P et al (2014) Wireless sensor networks: a survey on recent developments and potential synergies. J Supercomput 68(1):1–48
4. Akyildiz IF et al (2002) Wireless sensor networks: a survey. Comput Netw 38(4):393–422
5. Guestrin C et al (2004) Distributed regression: an efficient framework for modeling sensor network data. In: 3rd international symposium on information processing in sensor networks. Berkeley, CA, USA, pp 1-10. https://doi.org/10.1109/IPSN.2004.238731

6. Alwakeel SS, Al-Nabhan NA (2012) A cooperative learning scheme for energy efficient routing in wireless sensor networks. In: 2012 11th international conference on machine learning and applications. vol 2. IEEE
7. Förster AM, Amy L (2011) Machine learning across the WSN layers. InTech
8. Zhang Y, Meratnia N, Havinga P (2010) Outlier detection techniques for wireless sensor networks: a survey. IEEE Commun Surv Tutorials 12(2):159–170
9. Hodge V, Austin J (2004) A survey of outlier detection methodologies. Artif Intell Rev 22(2):85–126
10. Kulkarni RV, Frster A, Venayagamoorthy GK, (2010) Computational intelligence in wireless sensor networks: a survey. IEEE Commun Surv Tutorials 13(1):68–96
11. Das S, Abraham A, Panigrahi B (2010) Computational intelligence: foundations, perspectives, and recent trends. Comput Intell Pattern Anal Biol Inf 1–37: https://doi.org/10.1002/9780470872352.ch1
12. Srivastava JR, Sudarshan TSB (2015) A genetic fuzzy system based optimized zone based energy efficient routing protocol for mobile sensor networks (OZEEP). Appl Soft Comput 37:863–886
13. Lee Y (2017) Classification of node degree based on deep learning and routing method applied for virtual route assignment. Ad Hoc Netw 58:70–85
14. Sun Y, Dong W, Chen Y (2017) An improved routing algorithm based on ant colony optimization in wireless sensor networks. IEEE Commun Lett 21(6):1317–1320
15. Förster A, Murphy AL (2010) Machine learning across the WSN layers. https://doi.org/10.5772/10516
16. Barbancho J, León C, Molina J, Barbancho A (2006) SIR: a new wireless sensor network routing protocol based on artificial intelligence. Adv Web Netw Technol Appl 3842:271–275
17. Barbancho J et al (2007) Using artificial intelligence in routing schemes for wireless networks. Comput Commun 30:2802–2811
18. Watkins CJCH (1989) Learning from Delayed Rewards. Ph.D. thesis, University of Cambridge, England
19. Arya A, Malik A, Garg R (2013) Reinforcement learning based routing protocols in WSNs: a survey. Int J Comput Sci Eng Technol 4:1401–1404
20. Habib MA, Arafat MY, Moh S. Routing protocols based on reinforcement learning for wireless sensor networks: a comparative study. J Adv Res Dyn Control Syst: 427-435
21. Kadam K, Navin S (2012) Application of machine learning (reinforcement learning) for routing in wireless sensor networks (WSNs). In: 2012 1st international symposium on physics and technology of sensors (ISPTS-1). IEEE
22. Munot H, Kulkarni PH (2016) Survey on computational intelligence based routing protocols in WSN. Int Res J Eng Technol 3:122–127
23. Gao W (2007) Study on immunized ant colony optimization. In: 3rd international conference on natural computation (ICNC 2007), vol 4. IEEE, pp 792–796. https://doi.org/10.1109/ICNC.2007.690
24. Camilo T, Carreto C, Silva JS, Boavida F, (2006) An energy-efficient ant-based routing algorithm for wireless sensor networks. In: Dorigo M, Gambardella LM, Birattari M, Martinoli A, Poli R, Stützle T (eds) Ant colony optimization and swarm intelligence. ANTS, (2006) Lecture notes in computer science, vol 4150. Springer, Berlin
25. Shah SH, Naseer K, Ali W, Jabbar S, Minhas AA (2011) Prolonging the network life time in WSN through computational intelligence. Lecture notes in engineering and computer science 2193
26. Solaiman B, Sheta A (2013) Computational intelligence for wireless sensor networks: applications and clustering algorithms. Int J Comput Appl 73(15):1–8
27. Guo W, Zhang W (2014) A survey on intelligent routing protocols in wireless sensor networks. J Netw Comput Appl 38:185–201
28. Babu DM, Ussenaiah M (2019) Cuckoo search and M-tree based multicast Ad hoc on-demand distance vector protocol for MANET. Int J Recent Technol Eng (IJRTE) 8(253). ISSN: 2277-3878

29. Yang X-S, Suash D (2010) Engineering optimisation by cuckoo search. arXiv preprint. arXiv:1005.2908
30. Kulkarni RV, Venayagamoorthy GK (2010) Adaptive critics for dynamic optimization. Neural Netw 23(5):587–591
31. Prajapati J, Jain SC (2018) Machine learning techniques and challenges in wireless sensor networks. In: 2018 2nd international conference on inventive communication and computational technologies (ICICCT). IEEE
32. Yang W, Wang X, Song X, Yang Y, Patnaik S (2018) Design of intelligent transportation system supported by new generation wireless communication technology. Intell Syst Concepts Methodol Tools Appl
33. Chandrakar P (2019) A secure remote user authentication protocol for healthcare monitoring using wireless medical sensor networks. Int J Ambient Comput Intell (IJACI) 10(1):96–116
34. Das SK, Samanta S, Dey N, Kumar R (2020) Design frameworks for wireless networks. Springer
35. De D, Mukherjee A, Das SK, Dey N (2020) Nature inspired computing for wireless sensor networks. Springer
36. Mehmood A, Lv Z, Lloret J, Umar MM (2017) An Artificial neural network based energy-efficient and robust routing scheme for pollution monitoring in WSNs. IEEE
37. Gharajeh MS, Khanmohammadi S (2016) Dynamic 3D fuzzy routing based on traffic probability in wireless sensor networks. IET
38. Srivastava JR, Sudarshan TSB (2015) A genetic fuzzy system based optimized zone based energy efficient routing protocol for mobile sensor networks (OZEEP). Elsevier
39. Lee Y (2017) Classification of node degree based on deep learning and routing method applied for virtual route assignment. Elsevier
40. Khan F, Memon S, Jokhio SH (2016) Support vector machine based energy aware routing in wireless sensor networks. IEEE
41. Jafarizadeh V, Keshavarzi A, Derikvand T (2017) Efficient cluster head selection using Naïve bayes classifier for wireless sensor networks. Springer
42. Liu Z, Zhang M, Cui J (2014) An adaptive data collection algorithm based on a Bayesian compressed sensing framework. Multidisciplinary2014
43. Kazemeyni F, Owe O, Johnsen EB, Balasingham I (2014) Formal Modeling and analysis of learning-based routing in mobile wireless sensor networks. Springer
44. El Mezouary R, Choukri A, Kobbane A, El Koutbi M (2016) An energy-aware clustering approach based on the K-means method for wireless sensor networks. Springer
45. Ray A, De D (2016) NEnergy efficient clustering protocol based on K-means (EECPK-means)-midpoint algorithm for enhanced network lifetime in wireless sensor network. IET
46. Jain B, Brar G, Malhotra J (2018) k-Means clustering algorithmic solution for low energy consumption for wireless sensor networks based on minimum mean distance from base station. Springer
47. Guo P, Cao J, Liu X (2017) Lossless in-network processing in WSNs for domain-specific monitoring applications. IEEE Trans Indus Inf 13(5):2130–2139

Distributed Traversal Based Fault Diagnosis for Wireless Sensor Network

Deepak Kumar, Rakesh Ranjan Swain, Biswa Ranjan Senapati, and Pabitra Mohan Khilar

Abstract Wireless Sensor Networks (WSNs) have become a new information collection and monitoring solution for the various application. Faults occurring to sensor nodes are prevalent due to the sensor device itself and the harsh environment, where the sensor nodes are deployed. To ensure the quality of service and to avoid further degradation of service, it is necessary for the WSN to be able to tolerant of the faulty nodes present in the network. The fault diagnosis techniques are classified based on the methods they employ to determine the faults. In this paper, we have proposed a traversal-based diagnosis algorithm that seeks to diagnose both permanent as well as intermittent fault in WSN. The proposed algorithm employs a special node called an anchor node to traverse the field. The traversal of the field is decided by a proposed traversal algorithm taking into consideration the length and breadth of the sensor field, and the transmission range of the nodes. The anchor node stops at defined positions in the deployment field where it executes the fault diagnosis algorithm taking into consideration the normal sensor nodes which are in its range. The diagnosis algorithm uses a timeout mechanism to identify hard faults and adjusted boxplot method to identify permanent and intermittent faults in the network. The adjusted boxplot method takes into consideration the skewness of the data generated by the nodes in the sensor field. The faulty sensor nodes are classified by using a Feed Forward Neural Net (FFNN) model with Gravitational Search (GS) learning algorithm. The proposed algorithm is implemented in the Omnet++ environment which shows very promising results. The performance parameters, such as detection accuracy, false

D. Kumar · B. R. Senapati · P. M. Khilar
National Institute of Technology Rourkela, Rourkela, India
e-mail: deepaksingh93j@gmail.com

B. R. Senapati
e-mail: biswa.rnjn@gmail.com

P. M. Khilar
e-mail: pmkhilar@nitrkl.ac.in

R. R. Swain (✉)
Department of CSE, ITER, Siksha 'O' Anusandhan (Deemed to be University), Bhubaneswar, India
e-mail: rakeshswain89@gmail.com

S. K. Das et al. (eds.), *Architectural Wireless Networks Solutions and Security Issues*, Lecture Notes in Networks and Systems 196,
https://doi.org/10.1007/978-981-16-0386-0_8

alarm rate, false positive rate, and energy consumption of the proposed algorithm show significant improvement over the existing algorithms.

Keywords Fault · Sensor nodes · Diagnosis · Skewness · Intermittent

1 Introduction

Wireless sensor networks (WSNs) have become a data collection and monitoring solutions for a large number of applications. WSN consists of automated sensor nodes, which are deployed to collect and disseminate the data to the sink node, where the appropriate application runs. Various types of sensors are there which supports various type of data collection. Sensor nodes can be deployed to sense, detect and report time-critical events and accordingly counter actions can be taken in a coordinated manner. WSNs have the capability to be used in applications like battlefield surveillance, environmental monitoring, detection of an intruder, monitoring of infrastructure and many more [1]. Due to the recent advancement in the field of the wireless sensor node, the availability of low-cost and low power consumption sensor nodes became possible. From military surveillance to health-care monitor, WSN is also applied to many new existing fields. They are used to monitor environmental conditions like temperature, pressure, pollutant, moisture etc [2]. Sensor nodes can be used to give early signals in case of environmental hazards. They can be implanted in the body of patients to monitor their health condition. Deployment of sensor nodes in the volcanic area can give us early signals, helping to avert the consequences of earthquakes and eruptions [3]. In the case of agriculture, it can be used to automate the process of watering the plants based on the sensor readings of the moisture of the field. Figure 1 shows deployment of sensor nodes in the area of interest. The nodes collect and disseminate the sensed data to the user application.

Reliability of sensor nodes is a prime concern in the case of safety-critical systems. Sensor nodes generally operate autonomously in very less attended and hostile environments and their deployment may last for some days to years. Thus, faults in WSNs are liable to occur. The fault is an incorrect state of the hardware or program as a consequence of the failure of a node's component [4]. Faults may range from simple crash faults to faults where nodes incorrectly or maliciously work [5]. As faults are inevitable, so it is very important to find which nodes are faulty and which are working correctly. Faults occurring in a sensor network can have severe consequences in terms of human life, impact on the environment or economic loss. For example, in the bridge monitoring application if faulty sensor nodes are not isolated then the application would not be able to take appropriate decisions. The sole purpose of employing the WSN to monitor the bridge will be defeated and will cause huge loss including human life. The erroneous output may result in an incorrect decision being made by the sensor network. So, in order to rectify this problem, we need to correctly identify the faulty nodes from that of the whole set of nodes in the sensor network.

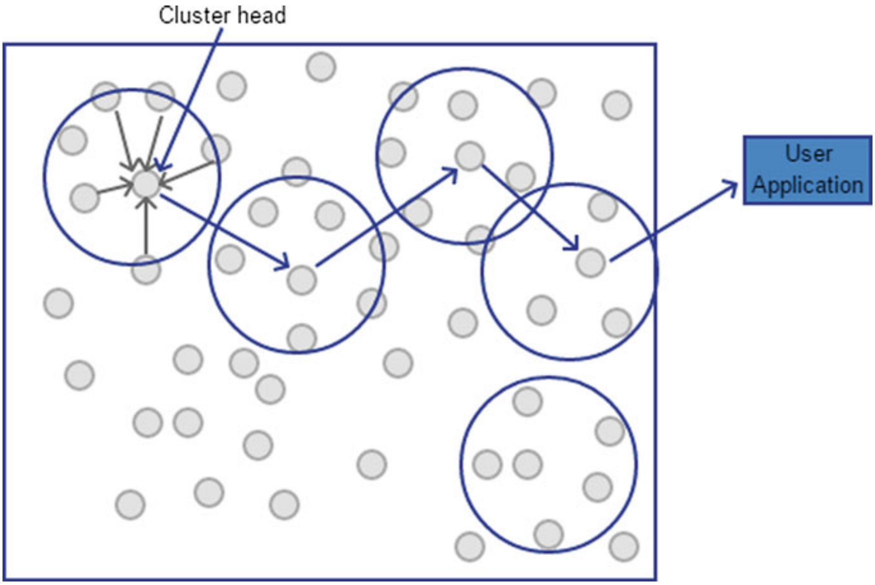


Fig. 1 Wireless sensor network monitoring a region

Faults can be classified based on the behavior of the faulty component of the node or duration of the fault [6]. Faults can be classified broadly into hard faults and soft faults. Hard faults are those kinds of fault in which the node is unable to communicate within the network. In case of a soft fault, the sensor node continues to work but with the altered behavior. Based on the duration of fault, faults can be distinguished into permanent, intermittent, and transient [6]. Permanent faults may be hardware or software fault which always produces errors. Intermittent faults are the one which shows an error for some instances of the time and works correctly for other instances. Transient faults are kinds of faults which occur momentarily due to some environmental conditions like lightening etc.

In this paper, we are concerned about the permanent fault and intermittent faults. One of the main issues in a wireless sensor network is energy constraint [7–11]. As the individual sensor nodes are not easily rechargeable, so it is highly beneficial to prolong the battery life of sensor nodes. Hence, the use of the fault detection algorithm must address this constraint and it should be energy efficient. We have used an anchor node whose task is to identify the faulty nodes in its vicinity. As the whole process of fault diagnosis is executed on the anchor node, the battery life of the sensor node is prolonged.

In this paper, we present a traversal-based fault diagnosis (DTFD) algorithm. Our major contributions are as follows:

i A fault diagnosis algorithm that works for both symmetric as well as skewed data.

ii A traversal based fault diagnosis algorithm, where a special node called an anchor node traverses the sensing region in an optimal way.

iii A protocol which is energy efficient and scalable in terms of energy requirement of WSN.

iv A protocol that utilizes the advantages of both centralized and distributed approaches.

The paper is organized in the following sections. Section 1 presents the introduction. Section 2 presents a brief description of related works. The system model of the proposed protocol is described in Sect. 3. Section 4 presents the proposed fault diagnosis protocol and the methodology of various phases. The analysis of the proposed algorithm is also presented here. Section 5 describes the fault classification phase using neural network model with Gravitational Search (GS) learning algorithm. The evaluation of the proposed protocol using simulation is presented in Sect. 6. Section 7 presents conclusion and future scope of this research.

2 Related Works

Many research work is being done in the field of the wireless sensor network to diagnose the network of faulty nodes. Fault diagnosis algorithms are broadly classified into two categories based on the participation of the sensor nodes in diagnosing the network i.e., (i) centralized approach and (ii) distributed approach. In a centralized approach, each sensor node transmits data to a common node which is comparatively powerful than other nodes in the network. Generally, the common node is a sink node or a base station. This node takes the responsibility of diagnosing the network of faulty nodes. The central node keeps track of all other nodes and thus is superior to other nodes in terms of memory and computational power. In a distributed approach, each node participates in the diagnosis process [12]. Each node tests other nodes in its vicinity and thus they are both tester as well as a testing node. In the distributed approach, a sensor node decides about the fault status of other nodes by monitoring the behavior of nodes locally. Thus, there is less congestion in the network as diagnosis related information are delivered locally and decisions are also taken locally. We discuss here some of the proposed work concerning centralized as well as distributed approaches.

2.1 Centralized Approaches

Staddon et al. [13] proposed an approach that allows the base station to learn the topology of the network. Once the base station knows about network topology can easily trace the failed nodes by executing divide and conquer strategy based on adaptive route update message. Koushanfar et al. [14] introduced a model-based testing technique which finds out incorrect computation faults. Ruiz et al. [15] proposed a management architecture based method for diagnosing faulty nodes, called MANNA. In this scheme, a manager is present which receives the energy level of each node in the network. For nodes which fail to reply to the query, the manager node looks up the energy map to check its residual energy. It deals with the crash faults. Kuo-Fend et al. [16] proposed a data-centric approach in which a source node selects two different disjointed paths to the sink node, to transmit the same data. If the received data are not identical, the source node sends another copy along with the third separate disjoint path. The sink identifies the faulty paths as well as faulty nodes by inspecting the contents of those packets.

Centralized approaches have lots of disadvantages and so the research has trended toward the distributed approaches. In a centralized approach, every sensor node conveys the diagnosis data to the central node, so it consumes huge bandwidth and therefore it cannot be recommended for large-scale WSNs. The most desirable property of the WSNs is that it should consume very less energy. But in order to dispatch data to the central node through multi-hop communication huge energy is depleted. Moreover, the detection latency of faulty nodes in the network for the centralized approach is high so they can not be used in real time. Due to these disadvantages, we moved towards distributed approaches [17, 18].

2.2 Distributed Approaches

Lee and Choi [19] proposed neighbor coordination based approach in which each sensor node performs a comparison between own sensed data with its neighbors at a particular instant of time. This process is repeated for a constant number of times and the result is locally stored, which is in the final step analyzed to calculate it's own fault status. Liang et al. [20] have proposed a weighted median based fault detection algorithm. This method exploits the spatial correlation of measurements by geographically distributed sensor nodes to identify soft faults. Nodes try to find the ratio of the difference between the own sensed data from the median of the neighboring data to the own sensed data. If this value is greater than a threshold, then the node is detected as faulty otherwise fault free. Ji et al. [21] extended the approach mentioned in [20]. They used the mean for finding out faulty node. Here nodes can find their fault status by comparing it's sensed data with that of the weighted average of their neighbor's sensed data. Panda and Khilar [22] used Neyman-Pearson method in a distributed environment to diagnose faulty nodes. Byzantine faults are detected using these methods. Mahapatro and Khilar [23] proposed an on-line cluster-based distributed fault diagnosis algorithm for hard and soft faults. Sahoo and Khilar [24] proposed an algorithm based on the comparison of sensed data of neighboring nodes and time-out mechanism. They diagnose the permanent and intermittent fault. Chen et al. [25] proposed majority neighbors voting based approach to diagnose soft faults in the network. Panda and Khilar [26] proposed one step based three sigma edit test fault diagnosis (DFD) approach. In this approach, every sensor node calculates it's probable fault status by using three sigma edit test on the sensed shared data. The probable sensed data of the node is found out by using the mean of the neighbor's reading. Sensor nodes then compare it's reading with that of the probable reading. Taking standard deviation into consideration to find out the outlyingness, the nodes whose result is in the range of 2–3 is considered to be fault-free, and greater than 3 is considered to be faulty. Panda and Khilar [27] proposed a distributed self-fault diagnosis (DSFD) algorithm for large scale wireless sensor networks using a modified three sigma edit test. This protocol uses the median of the sensed data to find out the fault status of itself. The message complexity is the order of constant. They diagnose the hard permanent fault as well as soft permanent faults. Swain et al. [28, 29] proposed a cluster-based heterogeneous fault diagnosis algorithm for wireless sensor network. The protocol is designed to detect and classify the faulty nodes (permanent, intermittent, and transient) in the network using a combination of statistical approach and soft computing approach. Similarly, the cluster based different neural network approaches are used for composite faulty nodes and links detection [30–37] in the WSNs.

3 System Model

This section describes the assumptions, network model, communication model, and fault model used in the proposed methodology. Table 1 shows the list of notations used in the proposed methodology.

3.1 Assumptions

i Anchor nodes are the special nodes which have somewhat more power than the normal sensor nodes. They are incorporated with the Global Positioning System(GPS). They are also used for localization of the sensor nodes in the network.

ii The sensing region is a rectangular field with length F_x and breadth F_y.

iii Sensor nodes are randomly distributed in the field. They have equal power in terms of transmission, battery, sensing etc. So they are homogeneous in nature.

iv Every node has a unique identification.

3.2 Network Model

Wireless sensor network is a network of spatially distributed autonomous sensor devices that cooperate within themselves to collect and disseminate data to the cen-

Table 1 List of notations used

Notation	Meaning
N	Total number of nodes in the network
F_x	Length of sensor field
F_y	Breadth of sensor field
$r_i(k)$	Sensed reading of v_i at time k
(v_{ii}, v_{ij})	coordinates of v_i sensor node
T_r	Transmission range
Neg	Number of neighbours of anchor node
Neg^x	Set of neighbour's sensed data at time x instance
N_a	Average degree of the network
NT	Neighbouring table stored at anchor node
F_i	Fault status of ith sensor node
T_{out}	Estimated transmission time of nodes in the network
(init, init)	Initial position of the anchor node
$(\text{curr}_x, \text{curr}_y)$	Current position of the anchor node
(x, y)	Next position of the anchor node
N_{curr}	Degree of the anchor node at current position

tral location, where it is used for specific applications. Wireless sensor Networks can be represented by using graph G(V,E), where $V = \{v_1, v_2, \ldots, v_n\}$ is the set of nodes and $E = \{e_1, e_2, \ldots, e_m\}$ is the set of edges. A specific edge in E is a pair of node that is $e_k = (v_i, v_j)$, where a link exist between nodes v_i and v_j. Edges exists between two nodes if and only if one is in the transmission range of another node. We can represent network by using adjacency matrix where each rows and columns represents nodes and a value = 1 at the confluence of row_i and column_j represent there exists a link between v_i and v_j and a value = 0 represents that there is no link between the specified nodes or here nodes are out of their transmission range. Adjacency matrix is a symmetric matrix meaning

$$(v_i, v_j) \in E \iff (v_j, v_i) \in E. \tag{1}$$

The nodes are haphazardly distributed in a rectangular field of length F_x and breadth F_y. The position of node is defined by (v_{ii}, v_{ij}), where v_{ii} and v_{ij} are the coordinates of node v_i in X and Y directions of the field. Given the positions of nodes v_i and v_j is (v_{ii}, v_{ij}) and (v_{ji}, v_{jj}) The distance between two nodes is calculated by using euclidean distance (euclid (x, y))

$$\text{euclid}(v_i, v_j) = \sqrt{(v_{ii} - v_{ji})^2 + (v_{ij} - v_{jj})^2}. \tag{2}$$

The neighbor nodes of a node is decided by its transmission range. The neighbors of a node are those nodes whose euclidean distance is less than it's transmission range(T_r).

$$\{(v_i, v_j) \in E \iff (v_j, v_i) \in E\} \iff \text{Euclid}(v_i, v_j) <= T_r \tag{3}$$

The proposed work seeks to implement a wireless network by randomly distributing 1024 nodes in a field of 120 × 120. The region has a special node called anchor node. The anchor node is the one which has the GPS capability. Anchor node works as a coordinator in its communication range. IEEE 802.15.4 protocol is used as the MAC protocol [38–40]. As the anchor node traverses the field, so it is required that the neighbor nodes are synchronized with it and send data whenever it receives a beacon signal. The anchor node sends a beacon signal which is used by the nodes in its vicinity to synchronize with that of anchor node. The anchor node moves in the defined path as shown in Fig. 2. The deep black color line shows the traversal path of the anchor node.

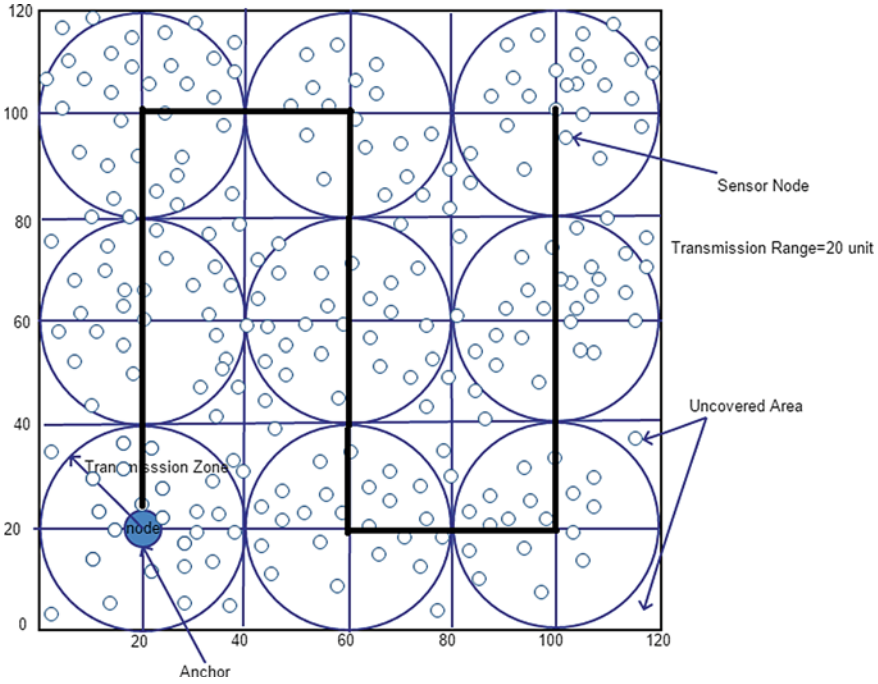


Fig. 2 Field containing anchor and non-anchor nodes

3.3 Fault Model

The proposed methodology diagnoses the permanent fault and intermittent fault in the network. Permanent faults are of two types such as hard permanent faults and soft permanent faults. In hard permanent fault the nodes do not respond to the request but in soft permanent fault nodes respond with sensed data but the reading is not correct. In the case of an intermittent fault, the nodes provide a correct result for some instances of time and wrong result for the rest. The readings of sensors are spatially and temporally correlated, since the reading observed at time instance is related to the readings at previous time instances and the readings observed at a geographic location is related to the readings observed at adjacent locations. The data $\{r_i(k)\}_{i=1}^{N}$ and $\{r_i(k)\}_{k=1}^{T}$ are to be analyzed in order to find out the faulty nodes in the network. Let the actual reading at time instance t be represented by a and the sensor node reading by node v_i at time t be represented by $r_i(t)$, then $r_i(t)$ can be written as:

$$r_i(t) = a + b + c + e_i(t), \quad \text{where} \quad i = 1, 2, \ldots, N. \tag{4}$$

here b is the spatial coherence factor, c is the temporal coherence factor and $e_i(t)$ is the presence of the error in the sensed data which is independent of the spatial and temporal coherence, since we already assumed it in the Eq. (4). Here we assume that a is the actual data for all the sensor node at time t but the error incorporated in each node is different. The reading $r_i(t)$ can be represented by a normal distribution with probability density function:

$$f(r_i(t)) = \frac{1}{\sqrt{2 * \pi * \sigma_i^2}} * \mathrm{e}^{\frac{-(r_i(t)-a)^2}{2*\sigma_i^2}} \tag{5}$$

For fault-free nodes, the variance σ_i^2 is very less but for soft faulty nodes, this is about 100 times of fault-free nodes. The factor b and c contribute very less to the variance σ_i^2.

4 Proposed Work

The Proposed algorithm has two parts one related to traversal of the anchor node and another related to the diagnosis of the faulty nodes in the network. The anchor node scans the whole sensor field based on the traversal algorithm. Anchor node stops at certain locations in the field defined by the size of the sensor field and transmission range of the sensor nodes in the network. This is explained in Algorithm 1. It gathers the data from its neighbors at that locations and then it executes diagnosis algorithm. The diagnosis algorithm identifies different kinds of faults like permanent hard faults, soft permanent faults, and intermittent faults.

4.1 Traversal Algorithm

The traversal algorithm seeks to scan the sensor field of $F_x \times F_y$ in such a way that the whole of the area is covered and at the same time energy consumption is less. Given the range of the node is T_r the anchor node is initially placed at (T_r, T_r) in X and Y direction respectively as specified in Fig. 3. The anchor node sends a beacon signal. Sensor nodes which are in anchor node's range of transmission send the sensed data to it. Anchor node executes Algorithm 2 to find out faulty nodes in its vicinity. Nodes which are faulty, isolated from the network. Now the node moves to the next location. Let x and y be the coordinates of the anchor node where the anchor node will move after the current position. Let the current position be represented by curr_x and curr_y in X and Y coordinates respectively. Then x, y can be given by Eqs. 6 and 7.

$$x = \text{curr}_x + 2 \times T_r \tag{6}$$

$$y = \text{curr}_y + 2 \times T_r \tag{7}$$

In Eqs. (6) and (7), two times transmission range is taken so as to optimize the area coverage [41]. This in turn also optimize energy consumption by not allowing

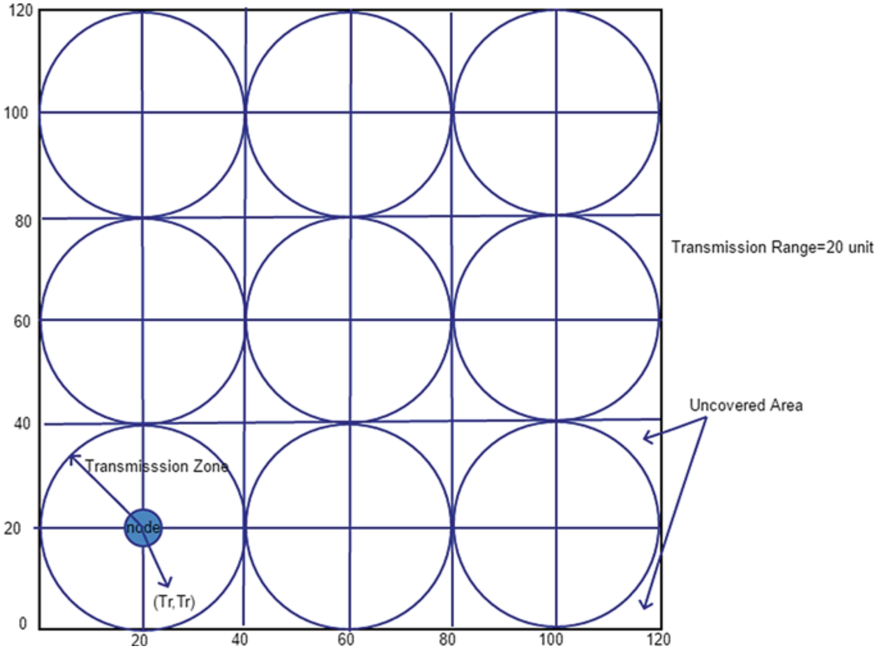


Fig. 3 The anchor node initially placed at (T_r, T_r)

to sense the area which has already been covered. The proposed traversal approach minimizes energy consumption, by optimizing the traversal area coverage.

This traversal scheme divides the rectangular field into $\left\lceil \frac{F_x}{T_r \times 2} \right\rceil \times \left\lceil \frac{F_y}{T_r \times 2} \right\rceil$ squares. Thus, each square formed is of side $= 2 \times T_r$. Each square bounds the circle that specifies the transmission zone inside the square by the anchor node, when it is present at the center of the square. The diameter of the circle inscribed inside the square is of $2 \times T_r$. The percentage of area covered by the anchor node is given in Eq. (8).

$$\frac{\left\lceil \frac{F_x}{T_r \times 2} \right\rceil \times \left\lceil \frac{F_y}{T_r \times 2} \right\rceil \times \pi \times T_r^2}{F_x \times F_y} \times 100\% \tag{8}$$

The Eq. (8) can be simplified to $25 \times \pi\%$, which is equal to 78.5%. As the area covered by this traversal algorithm is not 100%, so we propose another traversal algorithm which covers 100% area of the sensor field.

Considering the Sensor field of $F_x \times F_y$, we divide it into squares such that the whole area is covered. Earlier the circle was placed inside the square and therefore in each square, some of the area was not covered or only 78.5% of the area was covered. Now considering the range(T_r) to be constant and reaching till the vertex of the square, so that in a square whole area is covered as specified in Fig. 4. The square is circumscribed with diagonal $= 2 \times T_r$. Let the initial position of the anchor node in this scheme is represented by (init, init), which can be found out by using Pythagoras theorem and is given in Eq. (9).

$$\text{init} = \sqrt{T_r^2/2} \tag{9}$$

This traversal scheme shown in Fig. 4 divides the rectangular field into $\left\lceil \frac{F_x}{\text{init} \times 2} \right\rceil \times \left\lceil \frac{F_y}{\text{init} \times 2} \right\rceil$ squares. Each square thus formed is of side $= 2 \times \text{init}$ and whereas the radius of the circumcircle is T_r. Now the node moves to next location. Let x and y be the coordinates of the anchor node where the anchor node will move after the current position. Let the current position be represented by curr_x and curr_y in X and Y coordinates respectively, here it is initially (init,init) as specified in Fig. 4. Then x, y is given by Eqs. (10) and (11).

$$x = \text{curr}_x + 2 \times \text{init} \tag{10}$$

$$y = \text{curr}_y + 2 \times \text{init} \tag{11}$$

The detailed working of the proposed traversal algorithm is described in Algorithm 1.

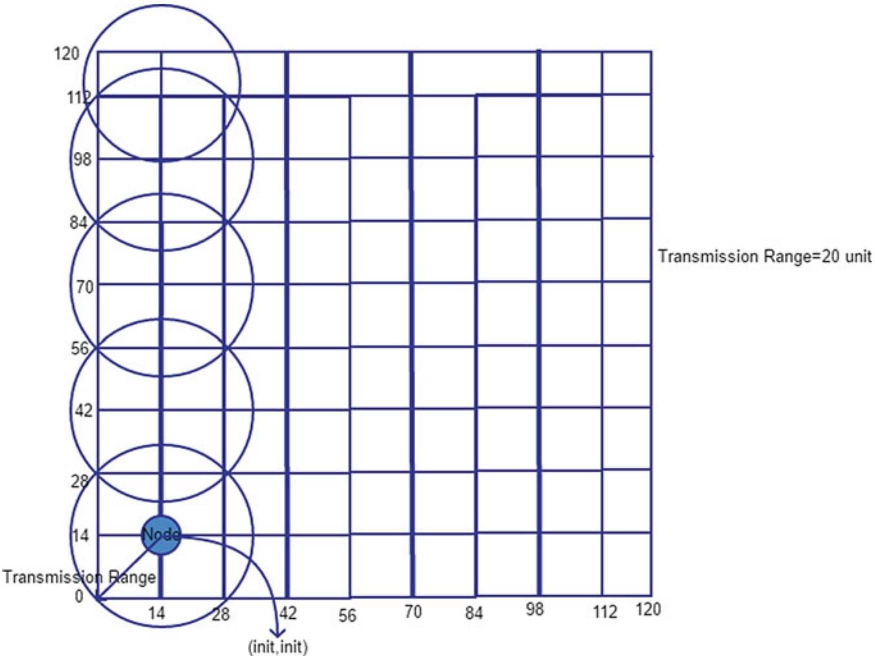


Fig. 4 The anchor node initially placed at (init,init)

4.2 *Diagnosis Algorithm*

Both the traversal algorithm and fault detection algorithm are executed simultaneously. The fault diagnosis phase comprises of two phases, (i) Hard fault detection and (ii) soft fault detection.

4.2.1 Hard Fault Detection

In the hard fault detection phase, the anchor node sends the beacon message at each of the locations. The sensor nodes after getting the beacon message transmit its sensed data to the anchor node. Those nodes which are unable to respond in timeout period (T_{out}) are considered to be faulty. The anchor node has a table, which stores the information regarding each non-anchor nodes in the network. Initially, it stores a 0 value, signifying a fault-free status of the non-anchor nodes. If the node is a hard fault, then it changes the fault status to 1 in the information table. IEEE 802.15.4 protocol is used as MAC protocol in which carrier sense multiple access with collision avoidance (CSMA-CA) mechanism is used for medium access by the nodes. Considering no collisions by the neighboring nodes while accessing the channel, the time taken (T) by a node to send the data to the anchor node is given by Eq. (12).

$$T = 2 \times pd + tr \tag{12}$$

where pd is the propagation delay, which is the time taken to transfer the packet from source to destination and tr is the transmission delay which is the time required to transmit a whole packet over a link. The other delays like processing delays and queuing delays are very less, so they are not considered here. Now as the non-anchor nodes can access the channel in any sequence due to contention among them, so we can say only about the cumulative time taken by all the neighbors of the anchor node for sending their data to it. The anchor node has Neg neighbors, so total time (Total) taken by all the neighbor nodes to transmit their data to the anchor node is given in Eq. (13).

$$\text{Total} = \text{Neg} \times T \tag{13}$$

Therefore, T_{out} time can be defined as $T_{\text{out}} = \text{Total}$. If a node does not respond in T_{out} time, then it is considered to be hard faulty.

Algorithm 1: Traversal algorithm

```
// $flag$:=0, when traversal is being done and becomes 1 when whole field is traversed.
// $(curr_x, curr_y)$: current position of the anchor node
// $minScaling$=: defines the constant value which describes the next position of the anchor node.
// $(F_x * F_y)$: area of the sensor field.
// $motion$: when the anchor is moving in positive Y direction, it is 0 otherwise it is 1.
if $(curr_y + 2 * minScaling) < F_y$ && $(motion == 0)$ then
    $curr_y \leftarrow curr_y + 2 * minScaling$;
else if $(curr_y + minScaling) < F_y$ && $(motion == 0)$ then
    $curr_y \leftarrow curr_y + minScaling$;
else if $(curr_y + minScaling) >= F_y$ && $(curr_x + 2 * minScaling) < F_x$ && $(motion == 0)$ then
    $curr_x \leftarrow curr_x + 2 * minScaling$;
    $motion = 1$;
else if $(curr_y + minScaling) >= F_y$ && $(curr_x + minScaling) < F_x$ && $(motion == 0)$ then
    $curr_x \leftarrow curr_x + minScaling$;
    $motion = 1$;
else if $(curr_y - 2 * minScaling) > 0$ && $(motion == 1)$ then
    $curr_y \leftarrow curr_y - 2 * minScaling$;
else if $(curr_y - minScaling) > 0$ && $(motion == 1)$ then
    $curr_y \leftarrow curr_y - minScaling$;
else if $(curr_y - minScaling) <= 0$ && $(curr_x + 2 * minScaling) < F_x$ && $(motion == 1)$ then
    $curr_x \leftarrow curr_x + 2 * minScaling$;
    $motion = 0$;
else if $(curr_y - minScaling) <= 0$ && $(curr_x + minScaling) < F_x$ && $(motion == 1)$ then
    $curr_x \leftarrow curr_x + minScaling$;
    $motion = 0$;
else
    $flag \leftarrow 1$;
return $(curr_x, curr_y)$.
```

4.2.2 Soft Fault Detection

In this phase, soft faults like soft permanent fault and intermittent faults are detected. Initially, after the completion of the hard fault detection phase, anchor node receives messages from its neighbors at time (t_1) and at time (t_2) instances. The anchor node detects the presence of faulty nodes by using statistical-based method i.e., the adjusted box plot introduced by Vanderviere and Huber [42]. This method does not assume anything regarding the distribution of the data. In the real world, data are almost skewed. This method can be applied to all distributional assumptions regarding the reading of the sensors.

Medcouple(MC):

It is a robust measure of skewness of univariate sample X from a continuous unimodal distribution F. let $X = \{x_1, x_2, \ldots, x_n\}$ is a set of data from a univariate continuous distribution F and x_m be the median of X. We define two sets X^L and X^R where

$$X^L = \left\{x_j | x_j <= x_m\right\} \tag{14}$$

$$X^R = \{x_i | x_i >= x_m\} \tag{15}$$

Medcouple(MC) can be defined as

$$MC = \text{med } g(x_i, x_j) \tag{16}$$

For $x_i \in X^R$ and $x_j \in X^L$ and $x_i \neq x_j$, we define the kernel function $g(x_i, x_j)$ as

$$g(x_i, x_j) = \frac{(x_i - x_m) - (x_m - x_j)}{x_i - x_j} \tag{17}$$

Equation (16) says that MC is the median of the set $\left\{g(x_i, x_j) | x_i \in X^R \text{ and } x_j \in X^L\right\}$. The kernel function is between -1 and 1.

Figure 5 shows the MC is negative for negatively skewed distributions, zero for symmetrical distributions, and positive for distributions that are right skewed. The interval of the adjusted boxplot is defined in Eq. (18).

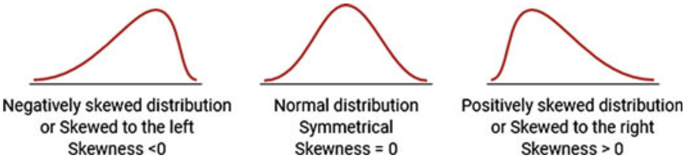


Fig. 5 Distribution and their skewness value

$$[L,U]=\begin{cases}\left[Q1-1.5\times exp(-3.5\times MC)\times IQR,\ Q3+1.5\times exp(4\times MC)\times IQR\right], & \text{if } MC>0\\ \left[Q1-3\times IQR,\ Q3+3\times IQR\right], & \text{if } MC=0\\ \left[Q1-1.5\times exp(-4\times MC)\times IQR,\ Q3+1.5\times exp(3.5\times MC)\times IQR\right], & \text{if } MC<0\end{cases} \tag{18}$$

where L and U are the lower and upper fence of the interval. $Q1$ and $Q3$ are first and third quartile, and IQR is the interquartile range. The neighboring nodes whose sensor reading in Eq. 18 fall out of range at t_1 and t_2 are considered to be soft permanent faulty node and those nodes whose reading fall out of range only at t_2 is considered to be an intermittent faulty node. The detailed working of this algorithm is summarized in Algorithm 2.

4.3 Analysis of DTFD Algorithm

4.3.1 Time Complexity

Every time the anchor node moves to a certain position in the field and carry out the fault diagnosis. At every location, it collects the data from the neighboring nodes. It collects data at two instance t_1 and t_2 and stores the data in Neg^{t_1} and Neg^{t_2}. The proposed work consists of two algorithms, (1) Traversal algorithm and (2) Diagnosis algorithm. Both the algorithms execute alternatively. The traversal algorithm is a constant time algorithm as every single step takes constant time, i.e. $\mathcal{O}(1)$. This algorithm will be called by main function $\left\lceil \frac{F_x}{\text{init}\times 2}\right\rceil \times \left\lceil \frac{F_y}{\text{init}\times 2}\right\rceil$ times. Thus the time complexity of the traversal algorithm is dependent on the area of the terrain where the sensors are deployed, which is equal to $\left\lceil \frac{F_x}{\text{init}\times 2}\right\rceil \times \left\lceil \frac{F_y}{\text{init}\times 2}\right\rceil \times$ constant.

Now coming to the diagnosis algorithm, this algorithm is executed alternatively with that of the traversal algorithm. So this algorithm is also called by main function $\left\lceil \frac{F_x}{\text{init}\times 2}\right\rceil \times \left\lceil \frac{F_y}{\text{init}\times 2}\right\rceil$ times. Now let us analyze each step of the diagnosis algorithm. Step 1 and 2 takes a constant amount of time. Step 4 takes $\mathcal{O}(N_{\text{curr}})$ time where N_{curr} is the degree of the anchor node at current position. Step 5 takes $\mathcal{O}(N_{\text{curr}} \log N_{curr})$, which is simply sorting the neighboring data. Step 6 to 12 is a constant time step. Again step 13–19 takes $\mathcal{O}(N_{\text{curr}})$. Step 20 finds out the medcouple of a set, in which left half and right half of the given data set is passed. This step takes $\mathcal{O}(N_{\text{curr}} \log N_{curr})$. Step 21 to 28 take $\mathcal{O}(1)$ time. Step 30 to 34 takes $\mathcal{O}(N_{\text{curr}})$ time. Finally step 38 to 40 takes $\mathcal{O}(N_{\text{curr}})$ time. Therefore we see that the maximum time being consumed by any single step is equal to $\mathcal{O}(N_{\text{curr}})$. As the diagnosis algorithm is called by anchor node at different positions in the field. Now we can take the value of N_{curr} to be N_a which is the average degree of the network. Therefore, in terms of the dimensions of the field and average degree of the network time complexity can be given as $\mathcal{O}(\left\lceil \frac{F_x}{\text{init}\times 2}\right\rceil \times \left\lceil \frac{F_y}{\text{init}\times 2}\right\rceil \times N_a)$.

```
Algorithm 2: Diagnosis Algorithm
    Input  : Reading $\{r_i(k)\}$
    Output: Fault status of sensor nodes $\mathrm{Neg}_i$
 1  Initialization Phase:
 2  Anchor node placed at $(init, init)$; Construct a table NT and set $F_i = 0$; Anchor node collects data from it's
    neighbours at $t_1$, $\mathrm{Neg}^{t}1$ and $t_2$, $\mathrm{Neg}^{t}2$;
 3  Diagnosis Phase:
 4  After $\mathrm{T_{out}}$ anchor node identifies hard fault by looking at data in $\mathrm{Neg}^{t}1$ & NT; Set $F_i = 1$
 5  sort($\mathrm{Neg}^{t1}$)                    //method for sorting in ascending order.
 6  md=med($\mathrm{Neg}^{t1}$)                    //method for finding out the median.
 7  if $|Neg^{t}1|\%2 == 0$ then
 8      $Q_1$ =median of $|\mathrm{Neg}^{t}1|/2$ smallest entries.
 9  else
10      $Q_3$ =median of $|\mathrm{Neg}^{t}1|/2$ largest entries.
11  end
12  $IQR = Q_3 - Q_1$
13  for $i$= $1 \ldots |Neg_x^{t1}|$ do
14      if $Neg^{t1}[i] \leq md_x$ then
15          $\mathrm{X}^L = \mathrm{X}^L \cup \mathrm{Neg}^{t}1\,[\mathrm{i}]$
16      else
17          $\mathrm{X}^R = \mathrm{X}^R \cup \mathrm{Neg}^{t}1\,[\mathrm{i}]$
18      end
19  end
20  mc = MC($\mathrm{X}^L$,$\mathrm{X}^R$,md)                    //method MC returns the medcouple.
21  if $mc > 0$ then
22      L = $\mathrm{Q}_1$ - 1.5 × exp(-3.5 × mc) × IQR
23      U = $\mathrm{Q}_3$ + 1.5 × exp(4 × mc) × IQR
24  else
25      if $mc < 0$ then
26          L = $\mathrm{Q}_1$ - 1.5 × exp(4 × mc) × IQR
27          U = $\mathrm{Q}_3$ + 1.5 × exp(-3.5 × mc) × IQR
28      else
29          L = $\mathrm{Q}_1$ - 3 × IQR
30          U = $\mathrm{Q}_3$ + 3 × IQR
31      end
32  end
33  $IFS = \emptyset$                    //IFS holds the node index.
34  for $i$=$1 \ldots |Neg^{t}1|$ do
35      if $(Neg^{t}1\,[i] < L \,||\, Neg^{t}1\,[i] > U)$ then
36          $\mathrm{FS}_i = 2$
37      else
38          $IFS = IFS \cup i$
39      end
40  end
41  repeat line 1 to 32 for $\mathrm{Neg}^{t}2$
42  for $i$=$1 \ldots |IFS|$ do
43      if $(Neg^{t}2\,[IFS[i]] < L \,||\, Neg^{t}2\,[IFS[i]] > U)$ then
44          $\mathrm{FS}_i = 3$
45      end
46  end
```

4.3.2 Message Complexity

Message complexity is defined as the number of messages exchanged over the network in order to find faulty nodes. Every node broadcasts a single message to the anchor node. Anchor node on the basis of the collected data from its neighbor finds out faulty nodes by applying the adjusted box-plot method. The proposed DTFD algorithm needs maximum N messages over the network, where N is the number of nodes in the network. Therefore message complexity of the DTFD is $\mathcal{O}(N)$.

4.3.3 Energy Consumption in the Network

Energy is consumed in the wireless sensor network for two purposes (i) Message transmission and (ii) Processing. Message transmission consumes much energy as compared to data processing. Therefore, in most of the related works, authors have only expressed the energy consumption in message transmission. Though little energy is consumed in processing but to some extent, the overall lifetime of the network depends on this factor.

In our proposed model, energy is consumed by two kinds of nodes [43]. Anchor node uses energy for sending, receiving messages as well as for its movement, whereas the non-anchor node uses energy only for sending and receiving messages. Energy is the biggest constraint in WSNs as nodes are deployed in the unattended environment. Energy consumption by anchor node is not of primary concern as anchor nodes does not have energy limitation and moreover, they are very few in the network. Here, we determine the energy requirement by both kinds of nodes.

4.3.4 Energy Consumption by Anchor Node

Energy is required for two purposes (i) Energy consumption for receiving messages (E_{rec}) and (ii) Energy required for moving along the predefined path (E_{move}).

The total energy consumption by the anchor node ($\text{TE}_{\text{anchor}}$) is given by Eq. (19).

$$TE_{\text{anchor}} = E_{\text{rec}} + E_{\text{move}}, \tag{19}$$

where E_{rec} is given in Eq. (20).

$$E_{\text{rec}} = N_{\text{r}} \times E_p, \tag{20}$$

where N_{r} is the avarage number of received messages and E_p is the energy required per packet. Now E_{p} is given by Eq. (21).

$$E_{\text{p}} = P_{\text{rec}} \times \frac{\text{packet size}}{\text{data rate}}, \tag{21}$$

where P_{rec} in Eq. (21) is the receiving power.

The energy required for the movement of the anchor node in its predefined path can be given by finding out the length of the path being traversed by it and energy required for unit movement along the path (E_{unit}).

$$E_{\text{move}} = lengthOfPath \times E_{\text{unit}} \tag{22}$$

4.3.5 Energy Consumption by Non-anchor Node

Energy consumption by the non-anchor node is of prime concern in our case as it affects network lifetime. Non-anchor nodes are static in the field. They send data to the anchor node when it comes in its range of communication. Therefore the energy required by the non-anchor node is given by Eq. (23).

$$E_{\text{send}} = N_{\text{t}} \times E_{\text{p}}, \tag{23}$$

where N_{t} represents an average number of messages transmitted and E_{p} represents energy required for sending each packet. E_{p} is given in the Eq. (24).

$$E_{\text{p}} = P_{\text{trans}} \times \frac{\text{packet size}}{\text{data rate}}, \tag{24}$$

where P_{trans} is the transmitting power.

4.3.6 Energy Consumption in Processing ($\boldsymbol{TE}_{\mathbf{P}}$)

Sorting is the single step in diagnosis that takes $\mathcal{O}(n\log(n))$ time. Therefore, the time complexity of DTFD is dominated by this step. Now let's compare DTFD with that of most optimized algorithm DSFD [27] in terms of energy consumption during processing. Both DSFD and DTFD has one common dominating step i.e., sorting of the neighborhood data. For an average degree of network N_a, the energy required for sorting by a single node in DSFD can be given by $\mathcal{O}(N_a \log(N_a))$. But, we can take it as a constant as network degree is constant. In DSFD each node performs computation, so total energy consumption for processing is given by in Eq. (25).

$$\text{TE}_{\text{P}} = N \times \mathcal{O}(1), \tag{25}$$

where N is the number of nodes in the network. Therefore, it is equal to $\mathcal{O}(N)$

Energy consumed in processing by the proposed algorithm is given by the fact that only anchor node does processing. Anchor node collects data and does sorting which is the main energy consuming step, it takes $\mathcal{O}(N_a \log(N_a))$ and N_a is constant. Therefore, this step takes constant time $\mathcal{O}(1)$. Anchor node moves to different number of locations in the field which can be given by $\left\lceil \frac{F_x}{\text{init}\times 2} \right\rceil \times \left\lceil \frac{F_y}{\text{init}\times 2} \right\rceil$. Therefore, energy consumed in processing is given by Eq. (26).

$$\text{TE}_{\text{P}} = \left\lceil \frac{F_x}{\text{init} \times 2} \right\rceil \times \left\lceil \frac{F_y}{\text{init} \times 2} \right\rceil \times \mathcal{O}(1), \tag{26}$$

which is much less than that of DSFD. Moreover, anchor node is a powerful node with extra processing capacity so, we can ignore this small amount of energy consumption, which does not affect the overall network lifetime. Therefore, energy consumed by fault diagnosis algorithm is negligible during processing.

4.3.7 Size of the Table at Anchor Node

The anchor node diagnoses the network of faulty nodes by executing DTFD and classify nodes as fault-free (0), hard fault (1), soft fault (2), and intermittent fault (3). Therefore, it requires two bits to store the fault status. There are N nodes in the network, to identify each nodes uniquely it requires $\lceil \log(N) \rceil$ bits. The total size of the table in bits is given in Eq. (27).

$$N(\lceil \log(N) \rceil + 2) \tag{27}$$

5 Fault Classification Phase Using Neural Net

In the fault classification phase, a feed forward neural net (FFNN) is used for faulty nodes classification. The neural net model could learn from the historical events of sensor node failures, then the acquired knowledge base is applied for the faulty node detection. The neural net follows 3 layers such as input, hidden, and output layer [44, 45]. The sensor values are feed in the input layer for the processing and the output of input layer are considered as input for the hidden layer. The hidden layer gives the input for output layer processing and the predicate output could be generated in the output layer with respect to the input instances.

Each sensor node n_i collects the sensor values of its corresponding neighbor sensor nodes. Let the sensor values $\overrightarrow{x_i} = (x_1, x_2, x_3, \ldots, x_n)$ are collected by a sensor node $n_i \in N$. The collected sensor values are the input patterns for the neural net training. The input patterns can be represent as $(\overrightarrow{x_i}, t)$, where $\overrightarrow{x_i}$ is the input vector of the sensor node $n_i \in N$ and t is the target class. The target class depends upon the failure types of sensor nodes. In the training dataset, the target classes with respect to the input patterns are known instances. Let n_{in} is the number of input layer instances, n_{hidden} is the number of hidden layer instances, and n_{out} is the number of output layer instances respectively.

Eq. (28) shows the input to hidden layer computation, where x_i is the input instances, v_i is the weight associated within input and hidden layer, and β_1 is the biased associated within input and hidden layer. Equation 29 shows the sigmoid function computation for activation of instances. The output of hidden layer h_{out} is represented in Eq. (30), where h_{in} is the hidden layer output. The output layer input y_{in} is calculated in Eq. (31), where h_i is the output of hidden layer, ω_i is the weight associated within hidden and output layer, and β_2 is the biased associated within hidden and output layer. The output of output layer y_{out} is computed in Eq. (32). For each instance of training dataset, the above computation is executed. Then, the mean squared error (MSE) is calculated in Eq. (33). The MSE is used to update the biases

β_1, β_2, and weights ν_i, ω_i of the neural net model. The weights and biases values are updated based on the Gravitational Search (GS) optimization learning algorithm. The Eq. (33) shows the MSE, where t_{ji} is the target output and o_{ji} is the actual output.

$$h_{\text{in}} = \beta_1 + \sum_{i=1}^{n_{\text{in}}} x_i \nu_i \tag{28}$$

$$\varphi\{x\} = \frac{1}{(1 + e^{-x})} \tag{29}$$

$$h_{\text{out}} = \varphi(h_{\text{in}}) \tag{30}$$

$$y_{\text{in}} = \beta_2 + \sum_{i=1}^{n_{\text{hidden}}} h_i \omega_i \tag{31}$$

$$y_{\text{out}} = \varphi(y_{\text{in}}) \tag{32}$$

$$\overrightarrow{E} = \frac{1}{2} \sum_{i=1}^{n_{\text{in}}} \sum_{j=1}^{n_{\text{out}}} (t_{ji} - o_{ji})^2 \tag{33}$$

The overview of neural net architecture is shown in Fig. 6. The input instances such as (x_i, t) is the input to the input layer, where x_i is the input instance and t is the target class with respect to different fault types. The sensor node $n_i \in N$ have input data, which is represented as $(x_1, x_2, x_3, \ldots, x_n, t)$ for the training process. The training process gives the information about the historical failure with respect to the input instances. For the testing process, the input instance should be represented as $(x_i, ?)$, where fault class of the input instance is unknown (?). So, after training process the actual fault type will be identified by the feed forward neural net model using Gravitational Search (GS) learning algorithm.

Gravitational Search (GS) algorithm was developed by Rashedi et al. in 2009 [46][47]. According to the gravitational force the masses are attracted towards each other. The lighter masses attracted towards the heavier masses and the heavier masses called as good solution. Algorithm 3 described the GS based learning for feed forward neural net model to classify the faulty sensor nodes in the network.

6 Performance Evaluation

The proposed algorithm has been evaluated and compared with the existing DFD [26] and DSFD [27] algorithms. Performance metrics such as detection accuracy (DA), false alarm rate (FAR), false positive rate (FPR), and energy consumption are compared. We have implemented the algorithms in Omnet++ environment [48]. Default simulation parameters are listed in Table 2.

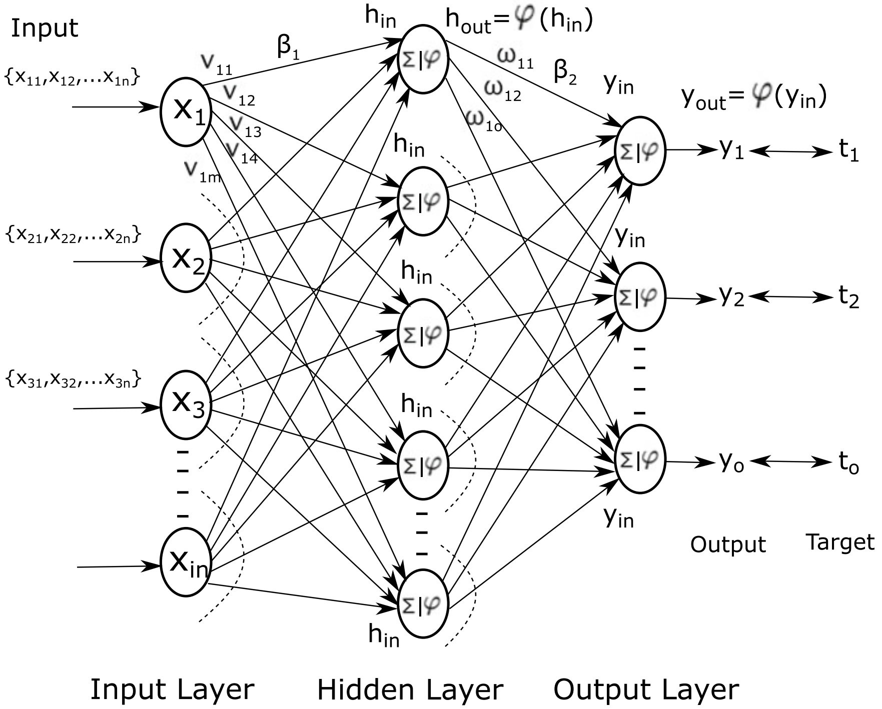


Fig. 6 Overview of neural net architecture

Table 2 Simulation parameters

Parameters	Value
Number of sensor node	1000
Deployment terrain	100×100 m^2
Radio range	5, 6, 8, 9, 10 m
Mac protocol	IEEE 802.15.4
Anchor node	1
Topology	Random network
Average degree	10, 15, 20, 25
Framework	Omnet++

Algorithm 3: GS Based Learning of NN

```
Initialize the parameters with masses and iterator=0;
while iterator ≤ Maximum iteration do
    iterator=iterator+1;
    Set Gravitational Force ($F$), Gravitational Mass ($M$), Velocity ($\vartheta$), Position ($\mathbb{P}$), Acceleration ($\alpha$), to 0;
    Set $\beta_i$ as random variable between [0, 1] interval;
    for each agent do
        Initialize weights ($\nu, \omega$) and biases ($\beta_1, \beta_2$) of FFNN;
    end
    Calculate fitness;
    if obtained fitness is better than gbest then
        Set gbest to obtained fitness;
    end
    Update Gravitational Mass ($M$) and Gravitational Constant ($G$);
    Calculate Gravitational Force ($F$) and Acceleration ($\alpha$);
    Update Velocity ($\vartheta$) and Position ($\mathbb{P}$);
    $\vartheta_i^k(t+1) = \beta_i \times \vartheta_i^k(t) + \alpha_i^k(t)$;
    $\mathbb{P}_i^k(t+1) = \mathbb{P}_i^k(t) + \vartheta_i^k(t+1)$;
end
```

6.1 Detection Accuracy (DA)

It is the ratio of the number of defective sensor nodes identified as faulty to the aggregate number of defective nodes present in the network. With the increase in fault probability the detection accuracy of faulty nodes decreases. When the average degree of the network is 20 and fault probability is 30%, the detection accuracy of the DTFD is close to 90%, which is shown in the graph plotted in Fig. 7. The DA of the proposed algorithm is greater than DFD and DSDF algorithms.

For the average degree of the network 25, with 30% fault probability, the detection accuracy of the DTFD is approximately 95%. The proposed algorithm performs better than the existing algorithms, which is shown in Fig. 8.

For the average degree of the network 30, with 30% fault probability, the detection accuracy of the DTFD is approximately 97%. This has been shown in Fig. 9. As the degree of the network increases the detection accuracy increases. With the increase in the average degree of the network, number of readings increases, which improves the detection accuracy performance. The reason behind the proposed algorithm better performance with respect to DA is that the statistical parameter used for finding out faulty nodes based on the data is less susceptible to extreme values (Table 3).

6.2 False Alarm Rate (FAR)

It is the ratio of the number of good sensor nodes identified as defective to the aggregate number of fault-free nodes present in the network. Figure 10 shows that at degree 20 and fault probability 30%, the FAR of DTFD is 2%, for DSFD it is around 2.8%, and for DFD, it is around 3.4%.

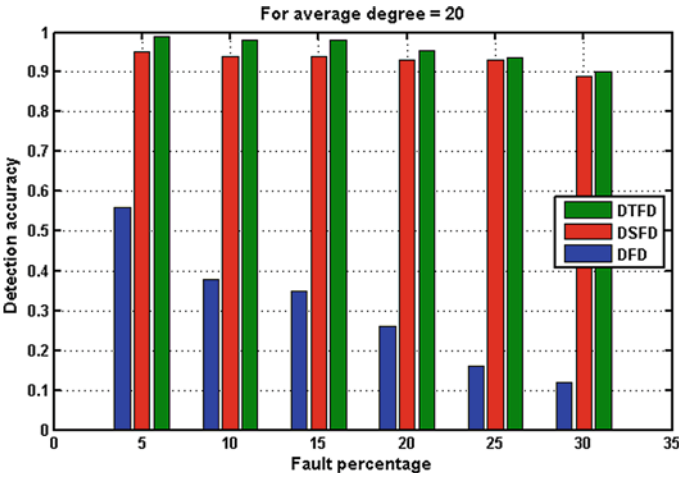


Fig. 7 Detection accuracy versus fault percentage

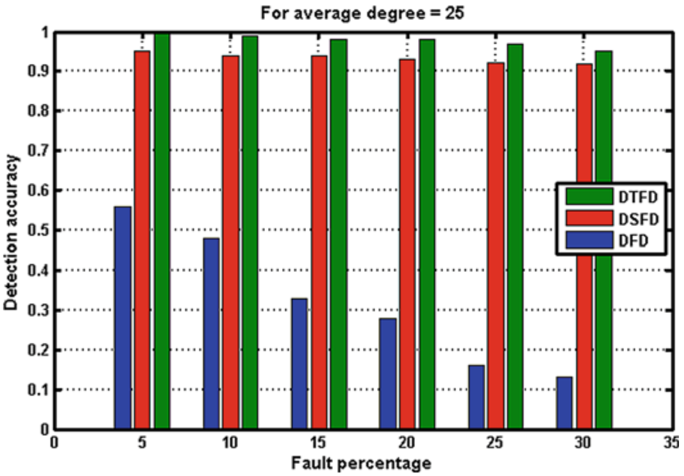


Fig. 8 Detection accuracy versus fault percentage

6.3 False Positive Rate (FPR)

It is the ratio of the number of defective nodes identified as fault-free to the total number of defective nodes present in the network. When the degree of the network is 30 and fault probability is 30%, FPR for DTFD is 3% and FPR of DSDF is 7%. When the degree of the network is 25 and fault probability is 30%, FPR of DTFD is 4% and for DSDF it is 8%. Figure 11 shows that at degree 20 and fault probability 30%, the FPR of DTFD is 10%, for DSFD it is around 13%, and for DFD, it is very

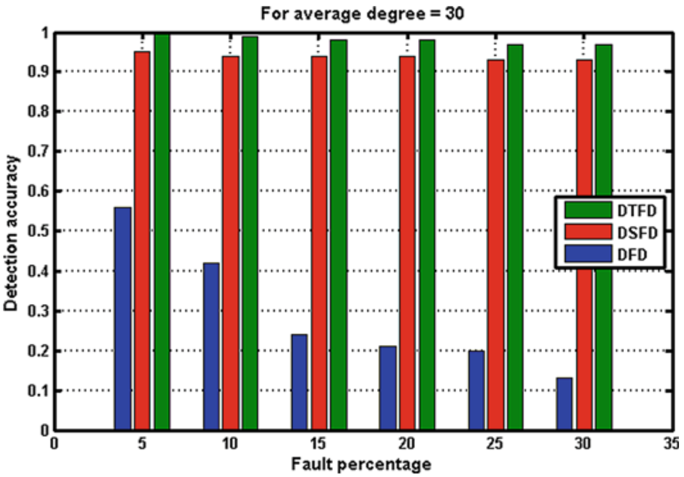


Fig. 9 Detection accuracy versus fault percentage

Table 3 Detection accuracy at degree = 30

Fault percentage (%)	DFD	DSFD	DTFD
5	0.56	0.94	1
10	0.42	0.95	0.99
15	0.24	0.94	0.98
20	0.21	0.94	0.98
25	0.20	0.93	0.97
30	0.13	0.93	0.97

high around 80%. As mean is affected by the presence of outlier data, FPR is very less for DFD.

The proposed algorithm does not make any distributional assumption. In case of symmetric distribution like normal distribution, mean = median = mode. The robust measure of skewness, i.e., medcouple (MC) is zero. In this case, the proposed algorithm sets the outlier detection boundary at equal distance from the median in both the directions. It behaves similarly to DSFD and detection accuracy is approximately the same for both the algorithms. When the distribution is negatively skewed, mean<median<mode. In this case, MC comes out to be negative. The median is closer to 3rd quartile as compared to 1st quartile. The span between the median and 1st quartile is more than the span between the median and 3rd quartile. The Proposed algorithm adjusts it's boundary accordingly. When the distribution is right-skewed, mean<median<mode. The MC comes out to be positive. Now, in this case, the median is closer to 1st quartile as compared to 3rd quartile. The proposed algorithm sets it's outlier detection boundary accordingly as explained in Algorithm 2. The interval of adjusted box-plot is right skewed in this case. Moreover, the proposed

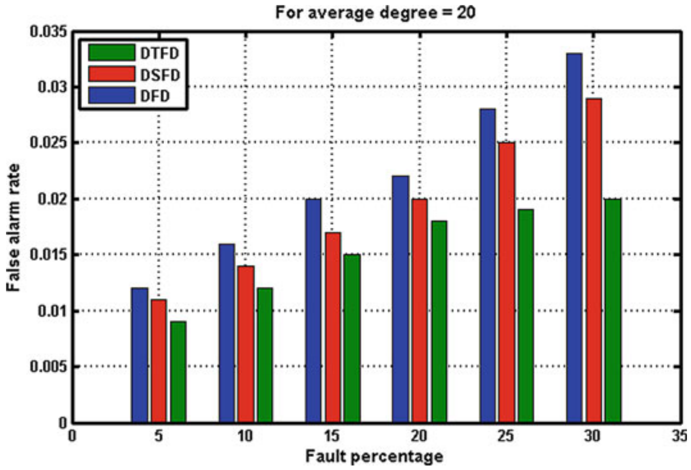


Fig. 10 False alarm rate versus fault percentage for DTFD, DSFD, and DFD algorithms

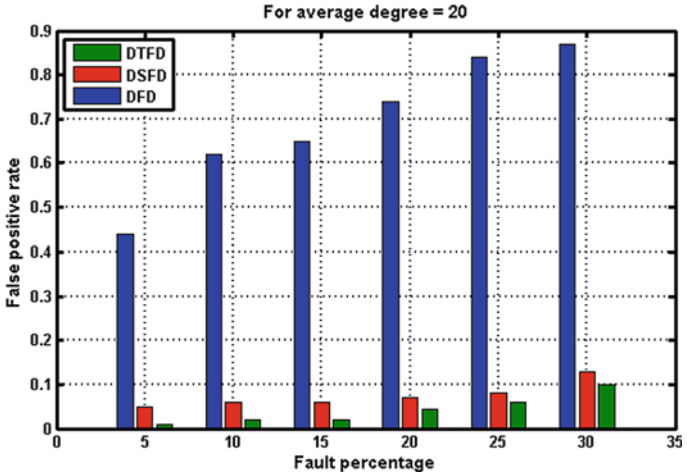


Fig. 11 False positive rate versus fault percentage

algorithm does not depend on mean and standard deviation, which are very prone to outlier data. When the data is skewed the proposed algorithm outperforms the existing algorithms in terms of detection accuracy, false alarm rate, and false positive rate.

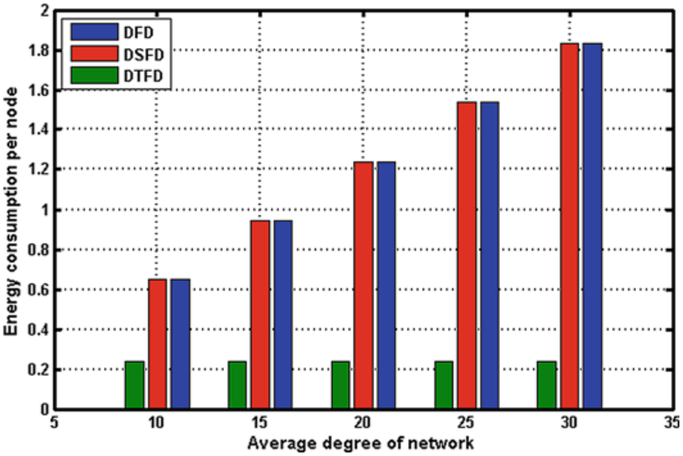


Fig. 12 Energy consumption per node

6.4 *Energy Consumption*

Energy Consumption of the proposed algorithm is calculated in terms of the number of messages transmitted and received in the network. The DTFD algorithm shows a very efficient result with respect to one of the most optimized algorithm DSFD. The energy consumption of the proposed algorithm is linear and it does not vary with the degree of the network whereas DSDF and DFD algorithm's energy consumption varies with the degree of the network. Figure 12 shows that the energy consumption of the proposed algorithm is constant.

In the proposed algorithm, each node broadcasts message at time instances t_1 and another message at time instance t_2. Therefore, the total number of messages broadcast in the network is $2 \times N$, where N is the number of nodes in the network. The total number of messages received in the network is dependent on the number of nodes in the network and traversal scheme being followed by the anchor node. As anchor node scans the field in an optimized way, there is a very less chance of a sensor node coming in the range of anchor node twice. So anchor node is able to receive the packet only once. Therefore, the total number of messages received in the network is approximately equal to the total number of messages being sent in the network, which is equal to $2 \times N$. As messages received in the proposed algorithm is much less than existing algorithms, the energy consumption of the network is very less. Moreover, the degree of the network does not influence the energy consumption of the proposed algorithm, so it is constant with respect to the degree of the network.

7 Conclusion

In this paper, we have proposed a traversal-based diagnosis protocol. The proposed traversal algorithm seeks to cover the whole field in an optimal way by considering the range of the nodes in the network. The anchor node having capabilities much more than that of normal sensor nodes in terms of computation, battery power etc., does all the diagnosis related computation. Sensor node deployed in the network sends messages to the anchor node when it comes in its neighborhood. The adjusted box-plot method is used to find out the faulty nodes in the network. This method employs parameters to find the outlyingness are robust to the presence of incorrect data produced by faulty sensor node. Whereas in other methods, the parameters deviate when a faulty node is present in the neighbor. This algorithm even takes care of the skewed data. In general, real-life data are skewed. The network lifetime increases as all the diagnosis related work are being done by the anchor node. The faulty sensor nodes are classified by using the Feed Forward Neural Network (FFNN) model with Gravitational Search (GS) optimization learning algorithm. The DA, FAR, FPR and other parameters of the DTFD outperforms the existing algorithms DFD and DSDF. The message complexity is $\mathcal{O}(N)$. Since very less number of messages are required to find faulty nodes, the algorithm is energy efficient.

In future, the proposed methodology needs to be implemented in real monitoring field. This can also be applied to the dynamic environment where all the sensor node moves thus changing its neighbors always. Moreover, apart from soft permanent fault and intermittent fault, other types of faults can also be found.

References

1. Huang P, Xiao L, Soltani S, Mutka MW, Xi N (2013) The evolution of mac protocols in wireless sensor networks: a survey. IEEE Commu Surv Tutorials 15(1):101–120
2. Senapati BR, Swain RR, Khilar PM (2020) Environmental monitoring under uncertainty using smart vehicular ad hoc network. In: Smart intelligent computing and applications. Springer, pp 229–238
3. Yick J, Mukherjee B, Ghosal D (2008) Wireless sensor network survey. Comput Netw 52(12):2292–2330
4. Koushanfar F, Potkonjak M, Sangiovanni-Vincentell A (2002) Fault tolerance techniques for wireless ad hoc sensor networks. In: Sensors, 2002; Proceedings of IEEE. vol 2. IEEE, pp 1491–1496
5. Yu M, Mokhtar H, Merabti M (2007) Fault management in wireless sensor networks. IEEE Wirel Commun 14(6)
6. Elhadef M, Boukerche A, Elkadiki H (2008) A distributed fault identification protocol for wireless and mobile ad hoc networks. J Parallel Distrib Comput 68(3):321–335
7. Das SK, Tripathi S (2019) Energy efficient routing formation algorithm for hybrid ad-hoc network: a geometric programming approach. Peer-to-Peer Netw Appl 12(1):102–128
8. Das SK, Tripathi S (2018) Adaptive and intelligent energy efficient routing for transparent heterogeneous ad-hoc network by fusion of game theory and linear programming. Appl Intell 48(7):1825–1845

9. Mazumdar N, Om H (2016) An energy efficient ga-based algorithm for clustering in wireless sensor networks. In: 2016 international conference on emerging trends in engineering, technology and science (ICETETS). IEEE, pp 1–7
10. Das SK, Tripathi S (2018) Intelligent energy-aware efficient routing for manet. Wirel Netw 24(4):1139–1159
11. Mazumdar N, Roy S, Nayak S (2018) A survey on clustering approaches for wireless sensor networks. In: 2018 2nd international conference on data science and business analytics (ICDSBA). IEEE, pp 236–240
12. Mazumdar N, Om H (2017) A distributed fault-tolerant multi-objective clustering algorithm for wireless sensor networks. In: Proceedings of the international conference on nano-electronics, circuits and communication systems. Springer, pp 125–137
13. Staddon J, Balfanz D, Durfee G (2002) Efficient tracing of failed nodes in sensor networks. In: Proceedings of the 1st ACM international workshop on wireless sensor networks and applications. ACM, pp 122–130
14. Koushanfar F, Potkonjak M, Sangiovanni-Vincentelli A (2003) On-line fault detection of sensor measurements. In: Sensors, 2003; Proceedings of IEEE. vol 2. IEEE, pp 974–979
15. Ruiz LB, Siqueira IG, Wong HC, Nogueira JMS, Loureiro AA et al (2004) Fault management in event-driven wireless sensor networks. In: Proceedings of the 7th ACM international symposium on Modeling, analysis and simulation of wireless and mobile systems. ACM, pp 149–156
16. Ssu KF, Chou CH, Jiau HC, Hu WT (2006) Detection and diagnosis of data inconsistency failures in wireless sensor networks. Comput Netw 50(9):1247–1260
17. Swain R, Dash T, Khilar P (2020) A lightweight approach to automated fault diagnosis in wireless sensor networks. IET Networks
18. Swain RR, Dash T, Khilar PM (2017) An effective graph-theoretic approach towards simultaneous detection of fault (s) and cut (s) in wireless sensor networks. Int J Commun Syst 30(13):e3273
19. Lee MH, Choi YH (2008) Fault detection of wireless sensor networks. Comput Commun 31(14):3469–3475
20. Gao JL, Xu YJ, Li XW (2007) Weighted-median based distributed fault detection for wireless sensor networks. Ruan Jian Xue Bao(J Softw) 18(5):1208–1217
21. Ji S, Yuan SF, Ma TH, Tan C (2010) Distributed fault detection for wireless sensor based on weighted average. In: 2010 2nd international conference on networks security, wireless communications and trusted computing. IEEE, pp 57–60
22. Panda M, Khilar PM (2015) Distributed byzantine fault detection technique in wireless sensor networks based on hypothesis testing. Comput Electr Eng 48:270–285
23. Mahapatro A, Khilar PM (2013) Online distributed fault diagnosis in wireless sensor networks. Wirel Pers Commun 71(3):1931–1960
24. Sahoo MN, Khilar PM (2014) Diagnosis of wireless sensor networks in presence of permanent and intermittent faults. Wirel Pers Commun 78(2):1571–1591
25. Chen J, Kher S, Somani A (2006) Distributed fault detection of wireless sensor networks. In: Proceedings of the 2006 workshop on dependability issues in wireless ad hoc networks and sensor networks. ACM, pp 65–72
26. Panda M, Khilar PM (2012) Distributed soft fault detection algorithm in wireless sensor networks using statistical test. In: 2012 2nd IEEE International Conference on Parallel distributed and grid computing (PDGC). IEEE, , pp 195–198
27. Panda M, Khilar PM (2015) Distributed self fault diagnosis algorithm for large scale wireless sensor networks using modified three sigma edit test. Ad Hoc Netw 25:170–184
28. Swain RR, Khilar PM, Bhoi SK (2018) Heterogeneous fault diagnosis for wireless sensor networks. Ad Hoc Netw 69:15–37
29. Swain RR, Khilar PM (2017) Soft fault diagnosis in wireless sensor networks using PSO based classification. In: Region 10 Conference, TENCON 2017-2017 IEEE. IEEE, pp 2456–2461
30. Swain RR, Khilar PM (2017) Composite fault diagnosis in wireless sensor networks using neural networks. Wirel Pers Commun 95(3):2507–2548

31. Swain RR, Khilar PM (2016) A fuzzy mlp approach for fault diagnosis in wireless sensor networks. In: 2016 IEEE Region 10 Conference (TENCON). IEEE, pp 3183–3188
32. Swain RR, Khilar PM, Dash T (2018) Fault diagnosis and its prediction in wireless sensor networks using regressional learning to achieve fault tolerance. Int J Commun Syst 31(14):e3769
33. Swain RR, Khilar PM, Dash T (2018) Multi-fault diagnosis in WSN using a hybrid metaheuristic trained neural network. Digital Commun Netw
34. Swain RR, Khilar PM, Dash T (2019) Neural network based automated detection of link failures in wireless sensor networks and extension to a study on the detection of disjoint nodes. J Ambient Intell Hum Comput 10(2):593–610
35. Swain RR, Dash T, Khilar PM (2019) A complete diagnosis of faulty sensor modules in a wireless sensor network. Ad Hoc Netw:101924
36. Swain RR, Khilar PM, Bhoi SK (2019) Underlying and persistence fault diagnosis in wireless sensor networks using majority neighbors co-ordination approach. Wirel Pers Commun: 1–36
37. Swain RR, Dash T, Khilar PM (2019) Investigation of rbf kernelized anfis for fault diagnosis in wireless sensor networks. In: Computational intelligence: theories, applications and future directionsm, vol 2. Springer, pp 253–264
38. Swain RR, Mishra S, Samal TK, Kabat MR (2017) An energy efficient advertisement based multichannel distributed mac protocol for wireless sensor networks (adv-mmac). Wirel Pers Commun 95(2):655–682
39. Swain RR, Mishra S, Samal TK, Kabat MR (2014) Adv-mmac: an advertisement based multichannel mac protocol for wireless sensor networks. In: 2014 international conference on contemporary computing and informatics (IC3I). IEEE, pp 347–352
40. Mishra S, Swain RR, Samal TK, Kabat MR (2015) Cs-atma: a hybrid single channel mac layer protocol for wireless sensor networks. In: Computational intelligence in data mining, vol 3. Springer, pp 271–279
41. Binh HTT, Hanh NT, Dey N et al (2018) Improved cuckoo search and chaotic flower pollination optimization algorithm for maximizing area coverage in wireless sensor networks. Neural Comput Appl 30(7):2305–2317
42. Hubert M, Vandervieren E (2008) An adjusted boxplot for skewed distributions. Comput Stat Data Anal 52(12):5186–5201
43. Rezazadeh J, Moradi M, Ismail AS, Dutkiewicz E (2015) Impact of static trajectories on localization in wireless sensor networks. Wirel Netw 21(3):809–827
44. Dash T, Nayak T, Swain RR (2015) Controlling wall following robot navigation based on gravitational search and feed forward neural network. In: Proceedings of the 2nd international conference on perception and machine intelligence. ACM, pp 196–200
45. Dash T, Sahu PK (2015) Gradient gravitational search: an efficient metaheuristic algorithm for global optimization. J Comput Chem 36(14):1060–1068
46. Rashedi E, Nezamabadi-Pour H, Saryazdi S (2009) Gsa: a gravitational search algorithm. Inf Sci 179(13):2232–2248
47. Sabri NM, Puteh M, Mahmood MR (2013) A review of gravitational search algorithm. Int J Adv Soft Comput Appl 5(3):1–39
48. Varga A, Hornig R (2008) An overview of the omnet++ simulation environment. In: Proceedings of the 1st international conference on Simulation tools and techniques for communications, networks and systems and workshops. p 60. ICST (Institute for computer sciences, social-informatics and and telecommunications engineering)

Fuzzy Q-Learning Based Controller for Cost and Energy Efficient Load Balancing in Cloud Data Center

**Subhra Priyadarshini Biswal, Satya Prakash Sahoo,
and Manas Ranjan Kabat**

Abstract The goal of cloud controller is to focus on continuous delivery of services to user on demand basis followed by "pay-per-use" model. Due to the increasing demand of cloud services, energy consumption on data center is increasing rapidly which lead to high operational cost. The harmful emission from this energy intensive data center affects our environment badly and cause climate change significantly. So as an alternative we have focused on onsite green power generation to reduce the harmful effects of greenhouse gases. In this paper, we proposed a fuzzy Q-learning based self-learning controller to optimize the load for specific data center. The proposed method also helps to reduce uncertainty and solve the congestion issue efficiently through fuzzy linguistic behavior and membership function. In this proposal, fuzzy output parameter considered as reward value which is used to learn and update the state for each data centre.

Keywords Cloud Computing · Q-learning · Data Center · Renewable Energy · Fuzzy Logic

1 Introduction

Cloud computing is an IT paradigm which provide various set of services to user on demand basis, that can be accessible at anywhere irrespective of time and place [1–4]. Here resources and services are always available on the internet and it can be released on demand basis without interacting with service provider. Resources are available to the user on pay per use model. User has to pay according to the resource usage by them like amount of space used, number of CPU cycle etc. Due to the increasing demands of cloud service, energy consumption of data center is increasing rapidly.

S. P. Biswal (✉)
School of Computer Science and Engineering, National Institute of Science and Technology (Autonomous), Institute Park, Pallur Hills, Berhampur, Odisha 761008, India

S. P. Sahoo · M. R. Kabat
Department of Computer Science and Engineering, Veer Surendra Sai University of Technology, Burla, Sambalpur, India

S. K. Das et al. (eds.), *Architectural Wireless Networks Solutions and Security Issues*, Lecture Notes in Networks and Systems 196,
https://doi.org/10.1007/978-981-16-0386-0_9

Data center are always deals with providing various set of services like web, e-mail, storage, processing etc. to user [5]. It always deals with delivering applications and services over the internet. For this, they requires higher amount of electricity. It will lead to high operational expenses. It adds significant impact on our environment by carbon emission.

Coal, petroleum and gas not only cause climate uncertainty through emissions of greenhouse gases, but also affect other economic, social and environmental area by adding up a dangerous negative balance sheet. This fossil fuel is not renewable. The harmful emission from this energy affects our environment badly and cause climate change significantly. So, the proposed method used fuzzy Q-learning technique to achieve the main goal. Fuzzy Q-learning is a meta-heuristic method which is used to optimize imprecise information efficiently and produce optimal result [6, 7]. There are several work are proposed based on meta-heuristic and fuzzy logic in the area of network based on optimization technique [8–11]. In the proposed method, renewable energy is derived from various natural processes like from sunlight, water, wind, biomass etc. From various study, it was verified that wind and solar energy has lower cost in comparison to other energy.

There are many approached has been proposed based on fuzzy inference system. In fuzzy system the rules are defined at design-time which leads to the following issues like user cannot describe any rule, user may only specify limited rules for some situations, users specify rules are not always effective and it may lead to uncertainty etc. There are several commonly used approaches are proposed like round robin, weighted round robin, ant colony optimization, particle swarm optimization, first come first serve, shortest job first etc. Each approach has its own advantages and disadvantages. But most of them have disadvantages like higher average waiting time, low throughput, requires detail task information, high response time. So to overcome that complexity here we proposed a fuzzy q-learning algorithm which is known as "a self-learning fuzzy cloud controller". It helps to run and update each state of the fuzzy logic based rules at runtime.

The main purposes of this paper are as follows.

(a) Design a self-learning cloud controller by combining Q-learning with fuzzy inference system.
(b) It gives importance on utilization of renewable energy sources and minimizes the consumption of brown energy.
(c) Deals with uncertainty caused by the incomplete knowledge.

The rest of the paper contain: Sect. 2 gives brief description cloud computing, Sect. 3 describes the related works done on load balancing problem, Sect. 4 describes about reinforcement learning. Section 5 contains proposed approach which consists of algorithm and fuzzy rule base. Section 6 was about experiments and results. Finally Sect. 7 contains the conclusion.

2 Cloud Computing

Cloud computing is an IT paradigm which enables user to consume computable resources (like storage, utility or application) on demand basis, that can be accessible at anywhere irrespective of time and place. Here resources and services are always available on the internet and it can be released on demand basis without interacting with service provider. Resources are available to the user on pay per use model. User has to pay according to the resource usage by them like amount of space used, number of CPU cycle etc.

2.1 Characteristics of Cloud Computing

(a) **On demand service**: Resources are provision on demand of the user without interaction of Cloud Service Provider. So it's an automated process.

(b) **Broad network access**: Provides platform independent access through heterogeneous client platform.

(c) **Resource pooling**: A pool of resources are provided by the cloud service provider for multiple user to fulfill their demand.

(d) **Rapid elasticity**: Resources can be rapidly scaled up or down on users demand basis. There are two types of scaling are there given in below

  (i) **Horizontal scaling**: it deals with lunching and provisioning additional resources.

  (ii) **Vertical scaling**: it involves with changing the virtual capacity of the server.

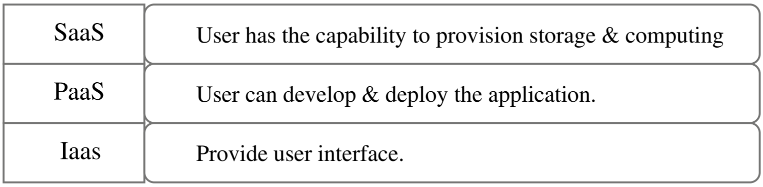
(e) **Measured service**: Cloud computing provides resources to the user as a measured service. It applies pay-per-use model. User has to pay based on the resource usage from the cloud.

(f) **Performance**: This may be scale up or down based on dynamic application workloads.

(g) **Security**: It improves security by centralization of data.

2.2 Cloud Models

There are various types of cloud models are available to the user like cloud service model and cloud deployment model. Cloud service model are categorize as IaaS (Infrastructure-as-a-service), PaaS (Platform-as-a-service), SaaS (Software-as-a-service) as shown in Fig. 1. Cloud deployment model are categorized as public cloud, private cloud, hybrid cloud and community cloud as shown in Fig. 1.



SaaS	User has the capability to provision storage & computing
PaaS	User can develop & deploy the application.
Iaas	Provide user interface.

Fig. 1 Cloud service model

2.2.1 Cloud Service Model

(a) **Infrastructure-as-a-Service**
It provides computing and storage resources to the users as virtual machine instances and virtual storage. User can deploy operating system and application of their choice. Cloud service provider maintains the underline infrastructure. Bill will be calculated in pay-per-use model. It means user has to pay for access on number of virtual machine hours and virtual storages. The example of I-a-a-S frame work is Amazon EC2.

(b) **Software-as-a-Service**

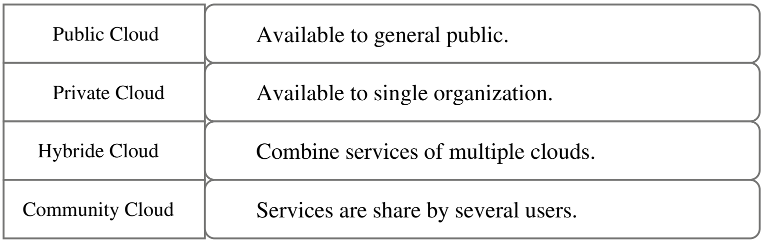
It provides user interface to the application itself. CSP maintains all. S-a-a-S applications are platform independent. Here the application software is provided to the user with in the cloud.

2.2.2 Cloud Deployment Model

(a) **Public Cloud**
Here cloud services are available to general public. The cloud resources are shared among every individuals, organizations, small and medium enterprises, government etc. The cloud services are provided by the third party provider. The example of public cloud are Amazon EC2 (Elastic Compute Cloud), Google App Engine etc.

(b) **Private Cloud**
Here cloud services are available for only a single organization. A single organization can access the resources from the cloud. So here security is more concern. Here the services are managed internally or by any third party or by the help of any single organization.

(c) **Hybride Cloud**
It combines the services of multiple clouds that may be of private or public. Hybride clouds are best suited for organizations that want their application to be secured applications and want to host their data on a private cloud. It is cost savings with hosting shared applications and data in public cloud.



Public Cloud	Available to general public.
Private Cloud	Available to single organization.
Hybride Cloud	Combine services of multiple clouds.
Community Cloud	Services are share by several users.

Fig. 2 Cloud deployment model

(d) **Community Cloud**
Here cloud services are shared by several organizations that have the same policy. It also supports load balancing techniques. The third party or any organization from the community will manage all the infrastructures (Fig. 2).

2.3 Load Balancing

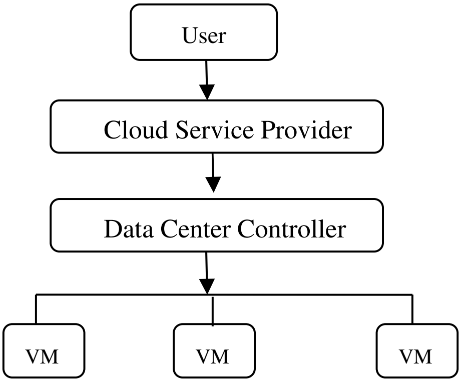
As we know the demand for cloud services increases day by day, as a result it will lead to load balancing as a major problem. When number of user request increases for more resources it becomes difficult for the server to execute their request in less time. So it may lead to lots of complexity with performance degradation. To solve this problem load balancing is efficiently required. This technique helps to distribute the workload among all available servers so that the user can get their resource in less time. Load balancing is an important concept which helps to increase the throughput. As a result, it helps to maximize the user satisfaction level. It helps to enhance the performance of the system by allocating virtual machine for the execution of user's request in less time. It also tries to minimize the user response time. The below figure will show how the workload will be distributed across all available servers to execute the users request as soon as possible (Fig. 3).

Generally there are so many load balancing algorithms are available. But most of them are either static or dynamic algorithms which are given in below. These algorithms are based on the existing status of the system.

(a) **Static algorithm**
These kinds of load balancing algorithm need earlier knowledge about detail task information to execute the task. It doesn't consider the present state of the system. This kind of algorithm may face the problem like performance degradation with lots of complexity which may lead to system failure. Some example of this algorithm is Round Robin, max–min load balancing, shortest job scheduling etc.

(b) **Dynamic algorithm**
These algorithms execute the work based on current state of the system. So that there will be no need of any prior knowledge to distribute the workload.

Fig. 3 Cloud load balance



They give better performance in comparison to static algorithm. The example of dynamic load balancing type of algorithm are fuzzy active monitoring, throttled load balancing etc.

3 Related Works

Uma Singhal [12] proposed a new fuzzy logic and GSO based load balancing mechanism for public cloud. Here the cloud was partitioned in to different groups. Each group will chose different load balancing strategy depending on the burstiness of workload. It consists of a cloud controller which will manage the entire load. Then it will forward to the load balancer. At last the most suitable virtual machine was selected based on the information provided by the memory and storage space usage etc. This algorithm has mainly focuses on the several parameters for load balance, i.e. burst detector, fuzzifier and load balancing approach. The fuzzifier is used to enhance the decision of choosing most suitable virtual machine. Here the author did not consider about any energy efficient concept. Pooyan Jamshidi [13] proposed a self-learning cloud controller. It automatically helps to compute the resource at runtime. The proposed algorithm is based on fuzzy inference system and Q-learning. It helps to adjust the rule at runtime. Here the parameter consider for further implementation are workload and response time. The output is change in number of virtual machine. The paper doesn't consider about any energy efficient concept related to the data center. Hamid Arabnejad [14] proposed comparison between different reinforcement learning algorithms. A self-adaptive fuzzy logic controller is combined with two reinforcement learning approaches (Fuzzy SARSA learning and Fuzzy Q-learning). Here both approaches are implemented and compared with their corresponding advantages and disadvantages. Pasha et al. [15] proposed round robin approach was proposed for virtual machine load balancing algorithm in cloud computing environment. It mainly focuses on execution of each and every request arrives at the data center from the user. It increases the throughput efficiently. The algorithm was based on Round Robin policy. Here time slot is allotted to every process. The workload will be assign to the virtual machine only when the virtual machine will be free. So when huge amount of task will be arrive at a time, then user has to wait until the next virtual machine will be available which the main disadvantage of the algorithm. Michael Dale [16] proposed comparative analysis of different types of renewable energy was proposed. Here the author has compared the costs of Photovoltaic, Solar Thermal, and Wind Electricity generation technologies. It describes the importance of renewable energy. The comparison is done basically on three types of renewable energy sources which are wind energy, photovoltaic solar power and concentrating solar power (CSP). The comparison is based on operational cost, capital cost and levelized cost of electricity. Capital cost means manufacture, energy requirement to process the material etc. Operational cost means energy requirements for proper maintenance of the system. For example washing solar systems, replacing worn parts, energy required to build the spare parts, operating the whole system, energy associate with the fuel cycle etc. Here they have describes different renewable technology sources developed in different country in year wise. By studying all these, they give a graph of capital cost of different renewable energy sources. From this they found that wind energy has lower cost in comparison to photovoltaic solar power and concentrating solar power (CSP). Sanyukta Raje [17] proposed different energy efficiency standard of the data center. Here the energy efficiency concept was based on the country India. The author focused on how the rising electricity cost, higher growth of Information Technology industries and higher utilization of fossil fuel in India will lead to high operational cost. It also gives harmful effects on our environment also. By using fossil fuel the carbon emission from the energy intensive data center are increases rapidly. It will leads to higher operational expenses. Due to the explosive growth of smart phone technology, social media apps, internet banking, e-commerce, multimedia etc. will lead to higher increasing demand for data center services. In India, the rising demand for data center reliability, efficiency will lead to some challenges like lake of technical awareness, energy efficient solutions etc. Suman Pandey [18] proposed a perspective study on cloud load balancing. Here the author has considered static, dynamic, genetic, decentralized aware based load balancing algorithms. The most common examples are round robin, shortest job first, max-mean, two phase load balancing, power aware load balancing, throttled, honey bee, active clustering algorithm etc. There are so many challenges are appear in cloud load balancing like virtual machine migration, automated service provisioning, energy management, stored data management. Pratibha Pandey [19] proposed fuzzy logic based job scheduling algorithm for cloud environment. Here this paper mainly focused on selection of virtual machines which are eligible to execute the task properly. The classification of task is done on the basis of quality of services. These QoS parameters are Completion time and Bandwidth. Bheda and Bhatt [20] has proposed an overview of all load balancing techniques in cloud computing environment. Here the author focuses on various load balancing algorithm available in cloud computing. They gives a comparison on the throughput value, overhead, fault tolerance, response time, recourse utilization, scalability, performance etc. For this comparison he considered load balancing algorithm like round robin, dynamic round robin, PLBA, active monitoring, FAMLB, throttled, active clustering load balancing algorithm etc. Er [21] has proposed online tuning of fuzzy inference system. Here the online tuning was done by using dynamic fuzzy Q-learning. Here an online self-learning algorithm was proposed. It mainly calculates the actions and Q-function in every iteration. If the temporal difference error value is higher than it will start adjusting the membership function and again calculate the action with Q-value. Ding et al. [22] proposed a method for task scheduling which is energy efficient. This is based on cloud computing technique. Basic key element is Q-learning which is used for reduce uncertainty and solve the related issue of cloud computing.

4 Reinforcement Learning

The most important features of artificial neural networks (ANN) are the ability to learn from the environment. The process of learning or training is required for making proper parameter adjustment. There are several types of learning process are available like supervised, unsupervised and reinforcement learning. In reinforcement learning process the exact information is not available properly. The learning based on this information is known as reinforcement learning. Here the aims is to find a proper actions, behavior or label which produce reward to maximize long term benefits.

Generally the RL doesn't have any prior knowledge about the environment. It only contains the information about valid set of actions with number of observations. By using repeatedly those actions it gains information about the corresponding environment. It will improve its policies automatically. The most common example of reinforcement learning is chess game. The block diagram of reinforcement learning is given in below (Fig. 4).

Here it performs like a closed loop. The output generated by the agent will affects the environment and similarly the environment will affects the agent. The environments will observe the current state of the system. Then the agent chooses the action from the controller policy depending upon the observations. It will affect to the environment. So, in this way one cycle will be completed and it will be repeated further to affects the next observations. After every cycle the agent will receives a reward value. Here the goal is to always choose the action with respect to the current state to

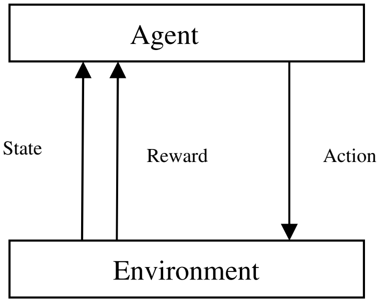


Fig. 4 Fuzzy inference system

maximize the reward value. Reward is defines that improvement of the performance after applying the suitable actions. As reinforcement learning doesn't contain any prior knowledge, exploration and exploitation is most important part. Here exploration means the knowledge gained by trying new things. Similarly exploitation me knowledge gained by gathering more information. The most common approaches of reinforcement learning are SARSA and Q-learning.

4.1 Elements of Reinforcement Learning

This subsection describe the basic concepts of reinforcement learning which consists of three basic parameters that described below and its algorithm shown in Algorithm 1 and its related notations are shown in Table 1.

(a) **Action**: The agent has to choose from set of actions to perform. By applying actions it will produce effects on the environment like state change.

(b) **State**: It is the original scenario where the problem and solutions are carried out.

(c) **Reward**: It was defined as the agent receives reward after taking some certain actions. Depending on the situation the reward value will be assigns to the system. If the system reach the goal, then the reward value will be higher else lower.

Algorithm 1 Pseudocode of Q-learning

```
Step1: Initialize Q ($s_t$, $a_t$)
Step 2: Repeat
Step 3: Select an action "a" for state s & execute it
Step 4: receive immediate reward "r"
Step 5: Evaluate the value of new state
$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha [ r_t + 1 + \gamma \max Q(s_{t+1}, a) - Q(s_t, a_t)]$
Step 6: Update value in q-table
s→s'
Step 7: End
```

Table 1 Notation description

Notation	Description
s	State
a	Action
a_t	action taken at time t
s_t	State at time t
r_{t+1}	Reward at time $t + 1$
α	Learning rate set in between [0, 1]
γ	Discount factor also lies in [0, 1]

5 Proposed Method

In order to make data center scheduling suitable for a cloud QoS aware, a number of requirements have to be met which are given in below:

(a) **Reliability**: The proposed system must be highly reliable. It has the ability to handle users request as many as possible.
(b) **Scalability**: The system must be capable of scaling itself when more number of users and applications are arises. It must be stable when the user's requirements change. The user must be access recourse from the system easily.
(c) **QoS and real time constraint**: The system must have advance QoS mechanism and policies to fulfill users demand.

5.1 System Model

This approach always focuses on onsite green power generation. In each time interval, the data center consumes energy from those renewable sources. The proposed approach is based on fuzzy q-learning where the arrival request will be forwarded to the most suitable data center. The selection of data center is done on the basis of information provided by the monitoring agent. Thus the workload will be done in minimum amount of time without any time complexity. Here the aim is to increase the process of load balancing with less time and maximize the utilization of renewable energy sources with lower electricity cost. Such that the throughput and the resource utilization will be maximized and the time complexity will be minimized.

This model consist of set of users transfer their request for the resource demand. The request will be received by the cloud service provider. Then it forwards the request to the cloud controller. Here the controller is based on fuzzy reinforcement learning. Here the monitoring agent will monitors different characteristics of the data center. It always deals with collecting the information and provides the data to both the fuzzy controller and knowledge learning components. For characteristics we have consider here workload, consumption of brown energy, utilization of renewable energy, and processor speed. By seeing the data provided by the monitoring agent, the controller chooses an action.

The fuzzy controller takes the observe data and generate the scaling actions. The learning components learn the appropriate rule and update the knowledge base continuously. After applying the action, system will generate the reward based on reward function. So reward means the performance improvement after applying the action. Then we observe the new state and reward. Here reward is calculated based on cost. Based on this information, we have to update the Q-table value. The important feature of Q-learning is that, it doesn't require any prior knowledge or detailed information about the system. Here if the fuzzy inference doesn't determine the scaling action, then the controller will randomly choose different actions and check whether it will produce the reward or not (Fig. 5).

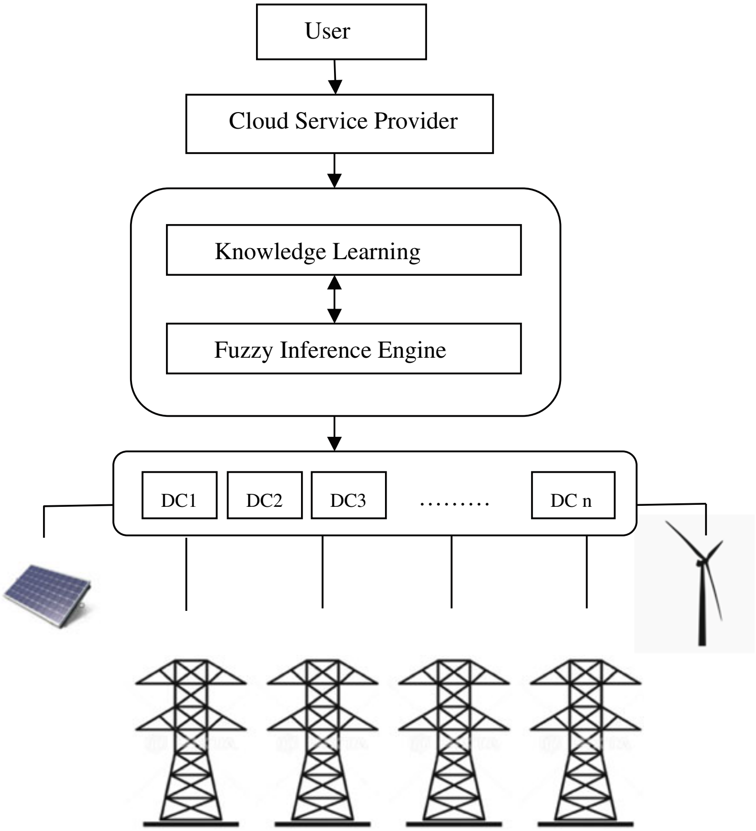


Fig. 5 System model

This model consists of two types of renewable energy sources like wind and solar energy. For wing energy, we have considered the wind turbine and for solar energy we have consider photo voltaic cell to convert it in to electrical energy. It mainly focuses on the utilization of renewable energy sources. The fuzzy rule base also consists of several parameters which help to redirect the request to the suitable datacenter. So that utilization of renewable energy sources will increase. As a result the effect of brown energy sources will be minimizes and it gives less bad impact on the environment. It also sees the cost of electricity at each data center based on the location. The data center which has less cost with maximum renewable energy sources will accept the request. Based on fuzzy q-learning every state of data center produces the reward. The data center which reward value is highest will accept the request from the user. Here the model consists of cloud controller, service provider, and set of users. For the onsite green power generation, here wind turbine and solar photo voltaic cells are considered. Here cloud controller is based on fuzzy logic.

5.2 *Fuzzy Q-Learning (FQL)*

In proposed approach, we used Q-learning as Reinforcement Learning approaches that we combine with the fuzzy controller. In this schema, a state '*s*' is modeled by four parameter (*Ux*, *Bx*, *Lx*, *Sx*) for which an RL approach looks for best action '*a*' to execute. The combination of the fuzzy logic controller with Q-learning, known as FQL, is explained in the following.

(a) **Initialize the *q*-values**: Here the *q*-table contains set of actions with state pair. At the time of learning process each value will be updated by considering reward value. So here we have to set the q table value to '0′ in initial stage.

(b) **Observe the current state**: After initialization of the *q*-values, the current state of each data center will be monitored. The monitoring agent will continuously monitored the characteristics of data center like utilization of renewable energy, consumption of brown energy, available processor speed and assigned load.

(c) **Select an action**: Here the control action is chosen by fuzzy logic controller. The fuzzy inference engine will observe each state value provided by the monitoring agent. Then it will process the value and generate the scaling actions. Here the scaling action means suitability of each data center. The fuzzy rule base only contains some rules defined by the user for some situations not for all. In that case the controller will automatically choose random actions and check whether it will generate the reward value or not.

(d) **Calculate the control action inferred by fuzzy controller**: It determines the output values produced by the fuzzy inference engine. Here we have considered Sugeno fuzzy inference engine. The output value is a constant value which lies in between $\{-2, -1, 0, 1, 2\}$. The action can be any number of finite set of numbers. But for the simplicity we have consider here 5 possible actions depending on our problem.

(e) **Calculate the reward value**: The controller receives the current values of state and actions. Here the reward value is calculated based on cost value. If the corresponding data center gives lower cost then it receives the reward otherwise the reward value will be zero.

(f) **Calculate the value of new state**: After calculating the action value, it will calculate the value of new state i.e. $V\left(s'\right)$.

$$V\left(s'\right) = Q(st, at) + \alpha[rt + 1 + \gamma \max Q(st + 1, a) - Q(st, at)]$$

(g) **Update the *q*-values**: After calculating the value of new state, it will update the corresponding state value in the *q*-table.

Algorithm 2 Fuzzy Q-learning

```
Step1: Initializes the q-values to zero
Step2: Observes the current state s
```

```
Step3: Repeat
Step4: Choose partial action 'a' from state s
Step5: Computes the action 'a' inferred by fuzzy controller
Step6: Receives reward
Step7: Apply the action and observe the new state s'
$V(s') = Q(st, at) + \alpha [ rt + 1 + \gamma \max Q (st + 1, a) - Q (st, at)]$
Step8: Updates the q-values
$s \rightarrow s'$
Step9: End
```

5.3 *Fuzzy Rule Base*

We have considered here 12 fuzzy rules for the fuzzy rule base. All the above fuzzy rules are in "if-else" format. Here we have considered four input variables for fuzzy inference which are given in Table 2.

(a) **R_X: Utilization of renewable energy sources**

It defines the ratio of number of renewable energy sources used from the total number of renewable energy sources. It can be defined as

$$R_x = \frac{\sum R(x)}{R_{\text{total}}} \tag{1}$$

Here $\sum R(x)$ is the total number of renewable energy sources used and Rtotal is the total renewable energy available.

Table 2 Fuzzy rules to identify efficient data center

R_x	B_x	P_x	L_x	Suitability(Scaling action)
Low	Low	High	Low	Very high(2)
Low	Low	Low	High	Mid (0)
Low	High	High	Low	High (1)
Low	High	Low	High	Mid (0)
Mid	Low	High	Low	High (1)
Mid	Low	Low	High	Mid (0)
Mid	High	High	Low	Mid (0)
Mid	High	Low	High	Low (−1)
High	Low	High	Low	Mid (0)
High	Low	Low	High	Low (−1)
High	High	High	Low	Mid (0)
High	High	Low	High	Very low (−2)

(b) **B_X: Consumption of brown energy sources**
It is defined as the amount of brown energy consumed from with respect to total available renewable energy.

$$B_x = \frac{\Sigma B(x)}{R_{\text{total}}} \tag{2}$$

Here $\Sigma B(x)$ is the total amount of brown energy used.

(c) P_X: Processor speed
(d) L_X: Assigned load at the data center.

The output value produced by fuzzy inference is defined in terms of "suitability". But for fuzzy q-learning, we have considered as scaling action. Here we have taken some constant value in between $\{-2, -1, 0, 1, -1\}$ for easier evaluation. It can be any finite number.

For example- IF *Ux* is low, *Bx* is low, *Lx* is low and *SX* is high THEN "sa = + 2".

5.4 Reward Calculation

After finding proper action given by fuzzy inference system, the next job is to apply the action and evaluate the next state. Then it will checks that whether this action produce the reward or not. The reward calculation is based on cost.

$$\begin{aligned}\text{Reward} = \mathbf{0},\ &\text{if cost} > 0.5\\ 1,\ &\text{if cost} < 0.5\end{aligned}$$

Here cost is calculated by amount of brown energy consumed by the data center * electricity price per each unit.

$$\text{i.e. Cost } (F_x) = B_x * \text{ electricity price per each unit} \tag{3}$$

5.5 New State Evaluation

The monitoring agent will monitor the state of each data center. If the current state scaling action is greater than 0, the new state will be calculated as the addition of both current state with the corresponding action value. If scaling action is less than 0, the new state will be same as the previous state.

$$\text{New state} = x + u,\ \text{if}\, u > 0\,\&\, x > 1$$

$$x, \text{ if } u <= 0 \,\&\, x <= 1$$

Here u = action, x = current state.

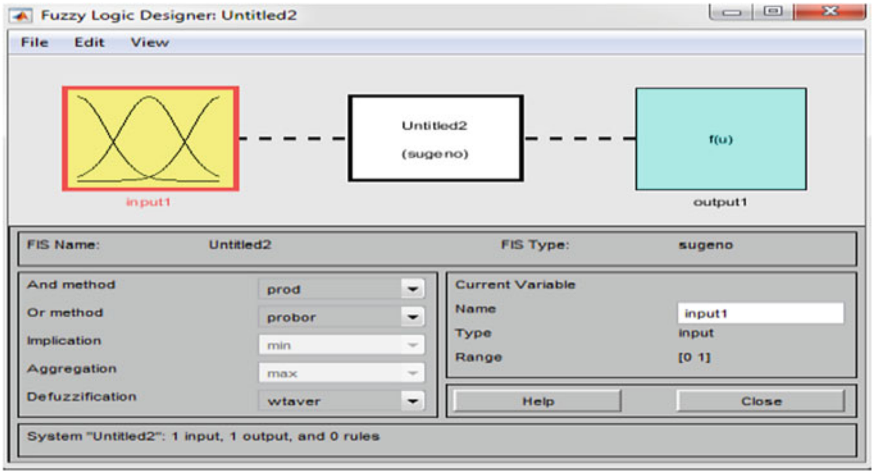
6 Result and Discussions

For the implementation of fuzzy rules, we used "MATLAB". It provides a fuzzy tool box system. The tool box contains the list of inputs, output and types of fuzzy inference engine. Here we take four inputs i.e. R_X, B_X, P_X, L_X and one output is the suitability of each data center in form of scaling actions. For every input and output, we have used triangular membership function. The name of the inference engine is fuzzy load balance.

The inference engine is based on "fuzzy Sugeno inference engine". It is the most popular inference engine as it has less complexity. In fuzzy inference engine all IF–THEN rules are defined by the fuzzy set. For different input, we have taken different membership function to define whether it is in the range of high, medium, low etc. The below figure will show the design of our fuzzy inference system (Fig. 6).

6.1 Fuzzy Sugeno Inference Engine

It is mostly suitable for mathematical analysis. Sugeno inference always gives output that is either constant or a linear (weighted) mathematical expression. Sugeno-type



FIS Name: Untitled2

FIS Type: sugeno

And method: prod

Or method: probor

Implication: min

Aggregation: max

Defuzzification: wtaver

Current Variable

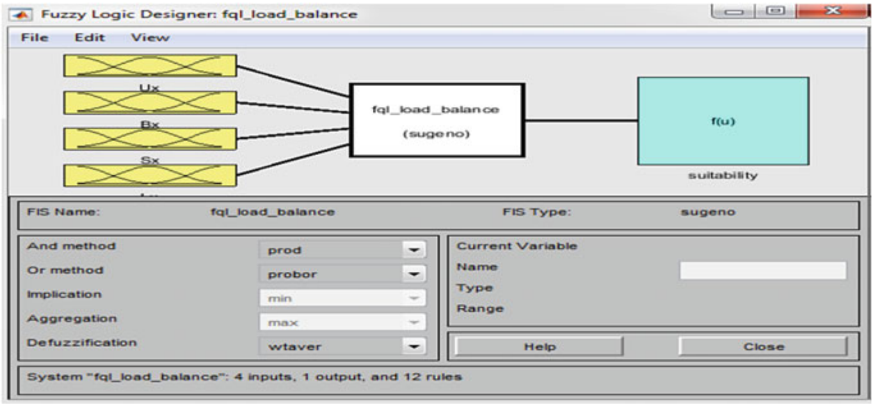
Name: input1

Type: input

Range: [0 1]

System "Untitled2": 1 input, 1 output, and 0 rules

Fig. 6 Fuzzy logic toolbox



FIS Name: fql_load_balance

FIS Type: sugeno

And method: prod

Or method: probor

Implication: min

Aggregation: max

Defuzzification: wtaver

Current Variable

Name

Type

Range

System "fql_load_balance": 4 inputs, 1 output, and 12 rules

Fig. 7 Fuzzy Sugeno Inference

FIS uses weighted average to compute the crisp output. This is mostly helpful in optimization and adaptive techniques. Sugeno FIS has some advantage over Mamdani FIS which are given in below:

(1) It is computationally more efficient.
(2) It works well with optimization and adaptive techniques.
(3) It is so convenient to mathematical analysis.

Example- If A is X_1 and B is X_2 then $C = ax_1 + bx_2 + c$ (linear expression) where a, b and c are constants (Fig. 7).

6.1.1 Membership Function for Input and Output

For the design of inference engine we have used mat lab toolbox which is given in below. Here the membership functions are in the form of triangular shape function.

This function is defined by a lower limit x, an upper limit y, and a value m, where $x < m < y$. It is easier to evaluate the membership value in compare to other membership function.

(a) Utilization of Renewable Energy
This input variable is denoted by U_x. It defines the ratio of the number of renewable energy sources used to the total number of renewable energy sources. Here the value of membership functions lies in between [0, 1]. This membership function has divided in to three types i.e. low, mid, high (Fig. 8).

(b) Consumption of Brown Energy
It is denoted by Bx. It defines how much brown energy sources will consumed for the electricity. Here the values of membership functions lie in between [0, 1]. This membership function has divided in to two types i.e. low and high (Fig. 9).

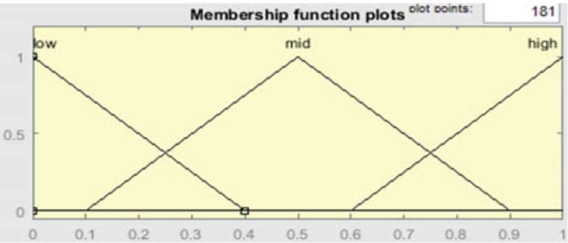


Fig. 8 Membership function of utilization of renewable energy

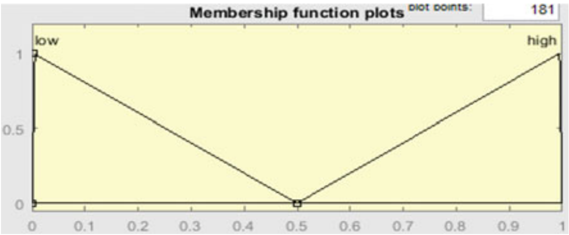


Fig. 9 Membership function of consumption of brown energy

(c) Processor Speed

This input variable is denoted by *Sx*. It defines the speed of the processor. Here the values of membership functions lie in between [0, 1]. This membership function has divided in to three types i.e. low, mid, high (Fig. 10).

(d) Assigned Load

This input variable is denoted by *Lx*. It defines the assigned load of the given data center. Here the values of membership functions lie in between [0, 1]. This membership function has divided in to two types i.e. low, high (Fig. 11).

(e) Suitability of the Data Center (Scaling Actions)

This input variable is denoted by "suitability". It is the output function produced by the fuzzy inference system. Here the values of membership functions lie in between [−2, − 1, 0, 1, 2]. This membership function has divided in to four types i.e. very low, low, mid, high, very high. If the output values lies in between [−2, − 1, 0, 1, 2], then it will be considered as in given format (Fig. 12).

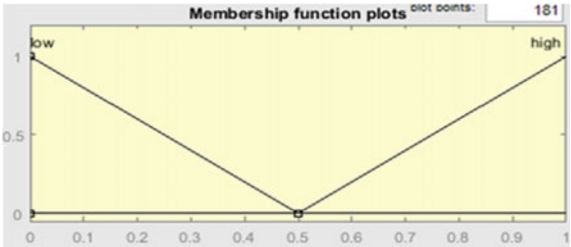


Fig. 10 Membership function of processor speed

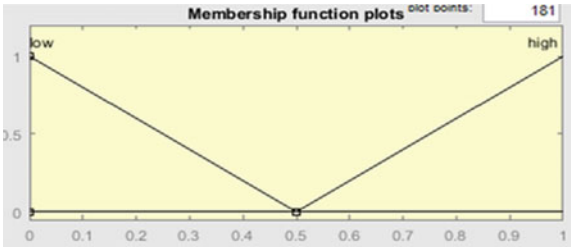
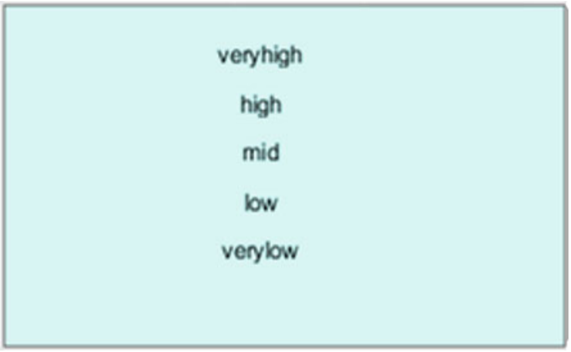


Fig. 11 Membership function of assigned load



veryhigh

high

mid

low

verylow

Fig. 12 Linguistic variables for output parameter

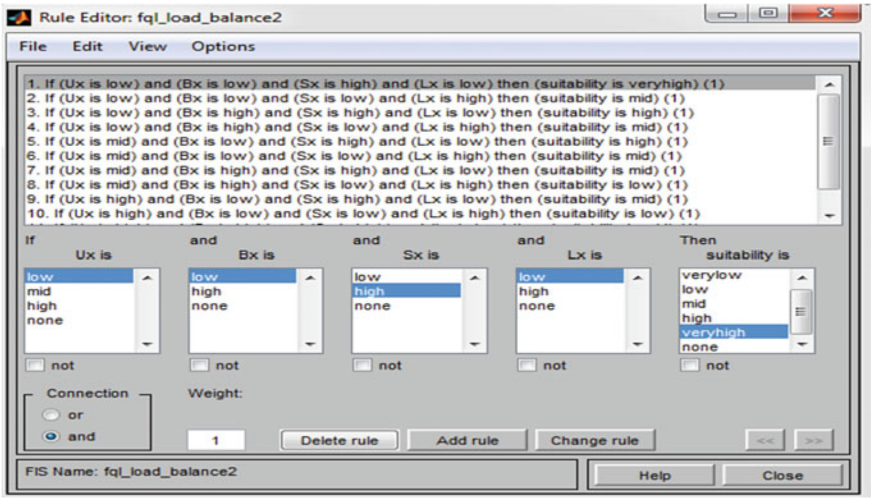
[−2] – Very low
[−1] – Low
[0] – Mid
[1] – High
[2] – Very high

6.1.2 Rule Base

The rule base consists of 12 rules by using the input variables to determine the output value or the scaling actions (Fig. 13).

6.1.3 Output of Fuzzy Inference Engine

We have taken an example to produce the output. The values of every parameter are given in the figure. The data center which consists of higher suitability value will accept request from the user for execution (Fig. 14).



1. If (Ux is low) and (Bx is low) and (Sx is high) and (Lx is low) then (suitability is veryhigh) (1)
2. If (Ux is low) and (Bx is low) and (Sx is low) and (Lx is high) then (suitability is mid) (1)
3. If (Ux is low) and (Bx is high) and (Sx is high) and (Lx is low) then (suitability is high) (1)
4. If (Ux is low) and (Bx is high) and (Sx is low) and (Lx is high) then (suitability is mid) (1)
5. If (Ux is mid) and (Bx is low) and (Sx is high) and (Lx is low) then (suitability is high) (1)
6. If (Ux is mid) and (Bx is low) and (Sx is low) and (Lx is high) then (suitability is mid) (1)
7. If (Ux is mid) and (Bx is high) and (Sx is high) and (Lx is low) then (suitability is mid) (1)
8. If (Ux is mid) and (Bx is high) and (Sx is low) and (Lx is high) then (suitability is low) (1)
9. If (Ux is high) and (Bx is low) and (Sx is high) and (Lx is low) then (suitability is mid) (1)
10. If (Ux is high) and (Bx is low) and (Sx is low) and (Lx is high) then (suitability is low) (1)

Ux is: [low] mid high none

Bx is: [low] high none

Sx is: low [high] none

Lx is: [low] high none

suitability is: verylow low mid high [veryhigh] none

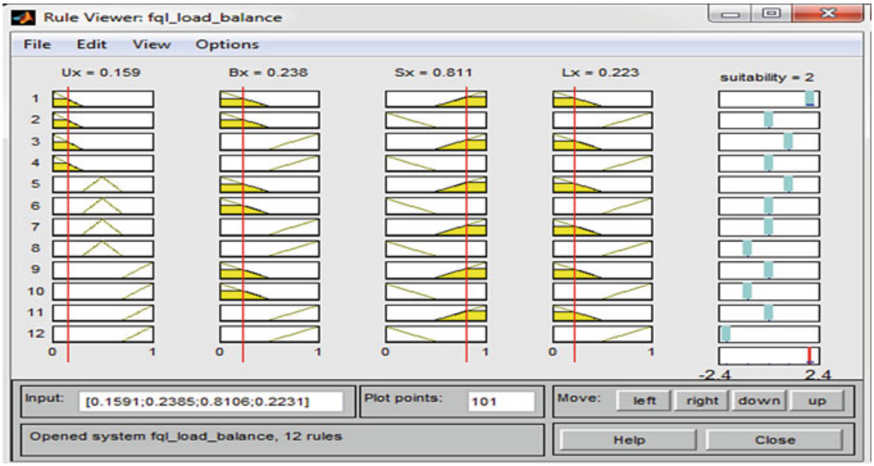
□ not □ not □ not □ not □ not

Connection: ○ or ● and

Weight: 1

FIS Name: fql_load_balance2

Fig. 13 Rule base for the proposed method



Input: [0.1591;0.2385;0.8106;0.2231]

Plot points: 101

Opened system fql_load_balance, 12 rules

Fig. 14 Validation of fuzzy inference system

6.2 Result Comparison of Fuzzy and Fuzzy Q-Learning

In Table 3, comparison is done in between fuzzy and fuzzy Q-learning. The inference system is based on Mamdani and Sugeno inference engine.

From the below comparison fuzzy q-learning gives batter result in compare to simple fuzzy inference system (Fig. 15).

Table 3 Comparison between fuzzy and fuzzy Q-learning

	FIS	Input	MFs	Reward calculation
Fuzzy	Mamdani	RX	3	N/A
		BX	2	
		SX	2	
		LX	2	
		FX	3	
FQL	Sugeno	RX	3	Based on cost
		BX	2	
		PX	2	
		LX	2	

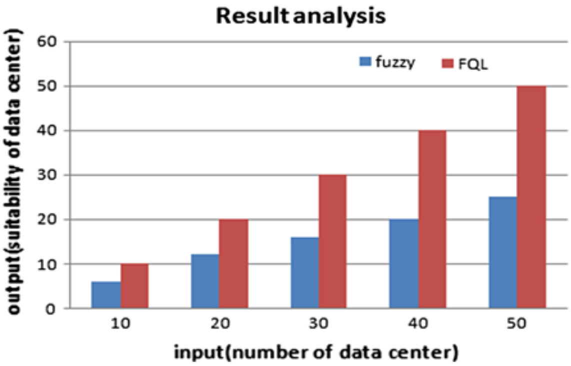


Fig. 15 Result analysis of fuzzy and FQL

6.2.1 Comparison of Cost with Respect to Utilization of Renewable Energy

Here the comparison is done on the basis of cost value. Cost value is calculated based on consumption of brown energy from each data center as given above on Eq. (3) (Fig. 16).

As a result if the consumption of brown energy will increases then the cost of that corresponding data center will be also increased. So that it is important to choose the alternate of brown energy which is known as renewable energy sources.

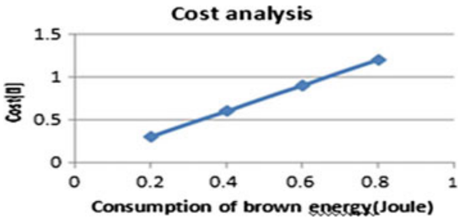


Fig. 16 Result analysis of Fuzzy and FQL

7 Conclusion

Fuzzy q-learning based algorithm for knowledge evolution is proposed here. This is based on fuzzy inference and Q-learning. This approach provides a non-linear mapping from the inputs like processor utilization of renewable energy, consumption of brown energy, electricity cost, speed of the processor, assigned load on data center to an output showing the appropriateness of the data center for the request redirection. This can be done by using some fuzzy rule which make this approach simple with out of any complexity. The Q-learning approach helps to learn each rule and update it in knowledge base at runtime. In future work we will try to implement this using cloud analyst software.

References

1. Khayer A, Talukder MS, Bao Y, Hossain MN (2020) Cloud computing adoption and its impact on SMEs' performance for cloud supported operations: a dual-stage analytical approach. Technol Soc 60:101225
2. Gill SS, Tuli S, Xu M, Singh I, Singh KV, Lindsay D, Tuli S, Smirnova D, Singh M, Jain U, Pervaiz H (2019) Transformative effects of IoT, Blockchain and artificial intelligence on cloud computing: evolution, vision, trends and open challenges. Internet of Things: 100118
3. Lin HC, Kuo YC, Liu MY (2020) A health informatics transformation model based on intelligent cloud computing–exemplified by type 2 diabetes mellitus with related cardiovascular diseases. Comput Methods Programs Biomed 191:105409
4. Kim T, Min H, Choi E, Jung J (2020) Optimal job partitioning and allocation for vehicular cloud computing. Future Gener Comput Syst 108:82–96
5. Toosi AN, Buyya R (2015) A fuzzy logic-based controller for cost and energy efficient load balancing in geo-distributed data centers. In 2015 IEEE/ACM 8th international conference on utility and cloud computing (UCC). IEEE, pp 186–194
6. Dey N (ed) (2017) Advancements in applied metaheuristic computing. IGI Global
7. Dey N, Ashour AS (2016) Antenna design and direction of arrival estimation in meta-heuristic paradigm: a review. Int J Serv Sci Manage Eng Technol (IJSSMET) 7(3):1–18
8. Das SK, Tripathi S (2019) Energy efficient routing formation algorithm for hybrid ad-hoc network: a geometric programming approach. Peer-To-Peer Netw Appl 12(1):102–128
9. Das SK, Tripathi S (2018) Adaptive and intelligent energy efficient routing for transparent heterogeneous ad-hoc network by fusion of game theory and linear programming. Appl Intell 48(7):1825–1845
10. Das SK, Tripathi S (2018) Intelligent energy-aware efficient routing for MANET. Wirel Netw 24(4):1139–1159
11. Chatterjee S, Sarkar S, Dey N, Ashour AS, Sen S, Hassanien AE (2017) Application of cuckoo search in water quality prediction using artificial neural network. Int J Comput Intell Stud 6(2–3):229–244
12. Singhal U, Jain S (2014) A new fuzzy logic and GSO based load balancing mechanism for public cloud. Int J Grid Distrib Comput 7(5):97–110
13. Jamshidi P, Sharifloo AM, Pahl C, Metzger A, Estrada G (2015) Self-learning cloud controllers: fuzzy q-learning for knowledge evolution. In: 2015 international conference on cloud and autonomic computing (ICCAC). IEEE, pp 208–211
14. Arabnejad H, Pahl C, Jamshidi P, Estrada G (2017) A comparison of reinforcement learning techniques for fuzzy cloud auto-scaling. In: 2017 17th IEEE/ACM international symposium on cluster, cloud and grid computing (CCGRID). IEEE, pp 64–73

15. Pasha N, Agarwal A, Rastogi R (2014) Round robin approach for VM load balancing algorithm in cloud computing environment. Int J 4(5):34–39
16. Dale M (2013) A comparative analysis of energy costs of photovoltaic, solar thermal, and wind electricity generation technologies. Appl Sci 3(2):325–337
17. Raje S, Maan H, Ganguly S, Singh T, Jayaram N, Ghatikar G, Greenberg S, Kumar S, Sartor D (2015) Data center energy efficiency standards in India. In: Proceedings of the 2015 ACM 6th international conference on future energy systems. ACM, pp 233–240
18. Pandey S (2017) Cloud load balancing: a perspective study. Int J Eng Comput Sci 6(6)
19. Pandey P, Singh S (2017) Fuzzy logic based job scheduling algorithm in cloud environment. Comput Model NEW Technol 21(3):25–30
20. Bheda H, Bhatt H (2015) An overview of load balancing techniques in cloud computing environments. Int J Eng Comput Sci 4:9874–9881
21. Er MJ, Deng C (2004) Online tuning of fuzzy inference systems using dy-namic fuzzy Q-learning. IEEE Trans Syst Man Cybern B (Cybern) 34(3):1478–1489
22. Ding D, Fan X, Zhao Y, Kang K, Yin Q, Zeng J (2020) Q-learning based dynamic task scheduling for energy-efficient cloud computing. Future Gener Comput Syst 108:361–371

Modelling of Aggregation Systems

Localization Techniques Using Machine Learning Algorithms

Chandrika Dadhirao and RaviSankar Sangam

Abstract Wireless sensor networks monitor environments that amendment apace over time. This dynamic behavior of the networks is either caused by external factors or initiated by the system itself. Machine learning techniques help us to work with extreme conditions and assist in avoiding the redesign of the network. The prominent feature of training the machine or network itself to modify according to such kinds of environments is being introduced in the sensor networks using machine learning techniques. However, the performance of the sensor networks has many constraints like energy efficiency, information measure or bandwidth, etc. Localization of nodes is one of the major issues that have to be worked on, as proper placement of nodes solves above-mentioned performance issues. The sensors in wireless networks gather knowledge regarding the objects they are to be sensed by which machine learning algorithms conjointly evoke several sensible solutions for localization of nodes that maximize resource utilization and prolong the lifetime of the network. The machine learning algorithms are categorized into three categories, namely supervised learning, unsupervised learning and reinforcement learning algorithms. As localization is the method of deciding the geographic coordinates of network's nodes and its relevant components as position awareness of sensing element of every sensor nodes plays a vital role in network communication for further process. In this chapter, we are going to focus on how the localization issue in wireless sensor networks can be solved using the three categorized machine learning algorithms.

Keywords Wireless sensor networks · Localization · Machine learning algorithms · Supervised learning · Unsupervised learning · Reinforcement learning

C. Dadhirao · R. Sangam (✉)
SCOPE, VIT-AP University, Vellore Institute of Technology -AP University, Amaravathi, India

S. K. Das et al. (eds.), *Architectural Wireless Networks Solutions and Security Issues*, Lecture Notes in Networks and Systems 196,
https://doi.org/10.1007/978-981-16-0386-0_10

1 Introduction

Machine learning(ML) is raised out of artificial intelligence (AI). Humans are intelligent species on earth, they learn from past experiences and act accordingly. Introducing this concept in machine learning, the machine to learn from past experiences and act accordingly is known as machine learning. Here, when related to computer previous experiences are termed as data, we are introducing the intelligence in the computer is termed as AI. Applying AI, we tend to needed to create higher and intelligent machines. However, aside from a few minor tasks like finding the shortest path between purpose A and B, we tend to were unable to program a lot of complicated and perpetually evolving issues. There was a realization that the sole thanks to being able to accomplish this task were to let machine learn from itself. This is similar to the scenario compared with a child learning itself. Therefore, machine learning was developed as a replacement capability for computers. And currently, machine learning is the gift in such a significant amount of segments of technology that we tend to do not even realize it. Machine learning is an inspiration to be told from examples and knowledge, while not being expressly programmed. Rather than writing code, you feed information to the generic formula, and it builds logic supported the info given [1]. On the other side, wireless sensor networks(WSNs) are extraordinarily popular and are ubiquitous at present, and they have gained popularity over a decade now. These wireless sensor networks became the key to the formation of the Internet of things. They both collaboratively work in making a country digitization. WSN, a big element of comprehensive computing that is presently being employed on an oversized scale to supervise period environmental standing further on stimulate the gathered results for future analysis. The main focus of the research in WSNs is the deployment of nodes in the network to maximize its life and minimize energy consumption in the system during communication. Sensors are mainly utilized under extreme energy constraints, i.e., human intervene highly impossible. To overcome the above scenario, creating a new wireless sensor node [2] is incredibly a tough task and involves a range of different parameters of accessibility to the required application which includes various ranges, transmitter/transceiver type, target technology, components, collective memory, storage space, power, lifetime, security and safety, quantum capability, inter/intra-communication technology, energy and resources, etc. So, we can train the sensor networks to act accordingly to the environment or according to the need by introducing machine learning algorithms(MLA) into wireless sensor networks [3–5]. Figure 1 shows the sensor network structure layout with a sensing field, communication lines, nodes and base station sensor network.

1.1 Problem Formulation

In wireless sensor networks, node deployment and node localization are the primary issue, and it is the root cause for most of the challenges/issues related in wireless

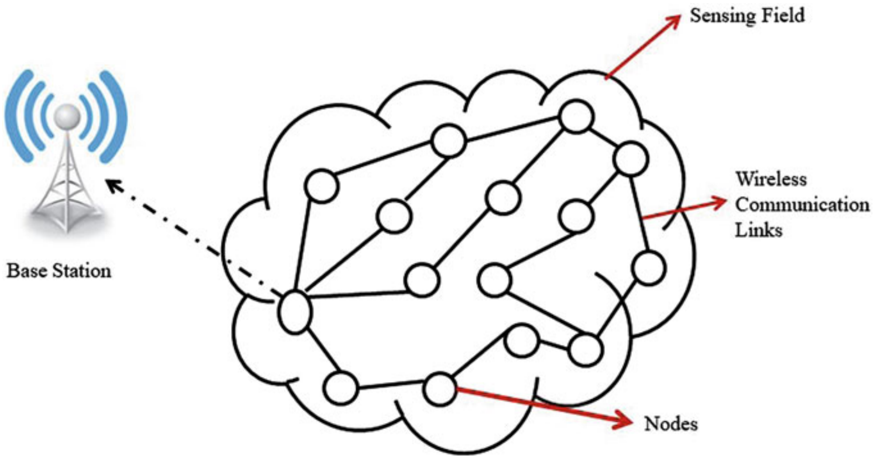


Fig. 1 Wireless sensor network layout

sensor networks. The sensor's position may be known or not. To get the correct details about the target area, we should accurately define the location of each node in the network. The actual localization problem can be resolved using two phases. Phase one is called the assessment/estimation phase and phase two is the testing phase. In three simple steps, this two-phase localization can be developed.

To address the localization problem, let us assume a randomly node deployment of network consists of number of predefined position nodes expressed in two-dimensional coordinates and a set of sensors with unknown locations N. The localization problem can be addressed using machine learning algorithms. Due to the increasingly widespread presence of sensitive sensors on WSNs, an overall performance location technique is insufficient for all applications. In recent years, therefore, there have been some significant advances in WSN localization techniques. Steps in Localization

1. Estimating distance between the nodes by time arrival of the signal and strength of the message.
2. Apply localization algorithms for optimizing the distance between the nodes.
3. Based on the result in step 2, get the position of the node which needs to be identified.

Figure 2 address the localization process. Table 1 represents the abbreviations used in this chapter and Table 2 presents the symbols and notations used in this chapter.

This localization is mainly concerned with minimizing the error rates of the actual node position and the position calculated when randomly deployed [6].

The rest of the chapter is organized as follows. In Sect. 2, we prioritize the role of machine learning (ML) in networking of wireless sensors. Followed by the benefits and detriments of implementing machine learning algorithms in sensor networks in Sects. 3 and 4. Then, the categorization of MLA based on localization issue in wire-

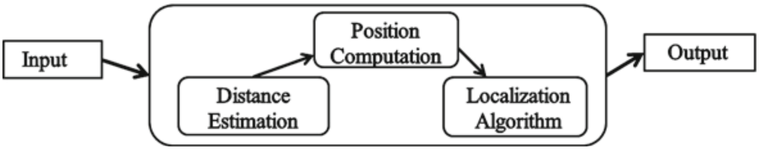


Fig. 2 Generalized localization process

Table 1 Abbreviations used in chapter

Abbreviation	Full name
ML	Machine learning
AI	Artificial intelligence
WSNs	Wireless sensor networks
MLA	Machine learning algorithms
SMC	Sequential Monte Carlo
DFL	Device-free localization
MBCS	Multitask Bayesian compressive sensing
NN	Neural networks
TOA	Time of arrival
RSSI	Received signal strength indication
TDOA	Time difference of arrival
DE	Distance estimation
PSO	Particle swarm optimization
ANN	Artificial neural networks
LNSM	Log normal shadow model
FFANN	Feed forward artificial neural network
DT	Decision trees
EB-MAC	Event-based medium access control
DGPR	Distributing the Gaussian process regression
GPR	Gaussian process regression
SVM	Support vector machines
LSVM	Localization based on support vector machine
THMSO	Two-hop mass-spring optimization
KNN	K-nearest neighbor
GRT	Geographical routing box
KBT	KNN boundary box
SOM	Self-organizing map
PCA	Principal component analysis
MB	Mobile beacons
lKNN	l distant K-nearest neighbors
GPS	Global position system

Table 2 Symbols and notations used in chapter

Symbol	Description
X	Independent or input variable
Y	Dependent or output variable
ϵ	Possible random error
f	Function that makes relation between X and Y
ϕ	Probability distribution function to learn uncertain concepts
T	Collected data
$(\phi\|T)$	Posterior probability of the parameter ϕ given the observation T
$(T\|\phi)$	It is the likelihood of the observation T given the parameter ϕ
S_t	Initial state
S_{t+1}	New state
A_t	Action taken
α	Learning rate
$\mathrm{b}(S_t, A_t)$	Achieved reward

less sensor networks(WSNs) is mentioned in Sect. 5. In Sect. 6, the first category of MLA, i.e., supervised learning is addressed and its related algorithms are discussed. Followed by the second category the unsupervised learning is discussed along with the relevant algorithms in Sect. 7. The third categorization of MLA which is the reinforcement learning is mentioned in Sect. 8. Finally, we draw conclusion and we give future scope in Sect. 9.

1.2 Motivation and Contributions

Researchers are working on the concept of deployment of sensor in networks applicable globally for any application and localization of sensors to identify aggregate data from a particular location without any congestion, corruption, or redundancy. Machine learning is powerful tool for sensors that help to calibrate and correct sensors when connected to other sensors measuring environmental variables. WSN monitor dynamic environments that change rapidly over time. Sensor networks often utilize ML techniques to eliminate unnecessary redesign while identifying locations of sensors or during deployment of sensor nodes in the sensing area, etc. Machine learning also inspires many practical solutions to maximize resource utilization along with prolonging the survival of nodes in the network, as machine learning algorithms originated from various fields like mathematics, neurosciences, statistics, including computer science.

Our contribution in this chapter is that we gathered a few of the machine learning algorithms as categorized into three major categories, namely supervised learning, unsupervised learning, and reinforcement learning. We given insights on few of the techniques in each group which are more suitable for WSN domain. Finally, we concluded the chapter with few ideas for future scope.

2 The Role of Machine Learning in WSN

The traditional WSN approaches are programmed that create the networks firm to retort dynamically. Techniques for machine learning can be used by reacting to overcome such eventualities. The method of self-learning from the experiences and actions, while not human intervention or re-programming is machine learning. Due to its size, efficiency and simply deployability, the WSN is responsible for monitoring, gathering sure data and transferring it to the bottom of the station for a wide range of sensor applications in the post-knowledge analysis range. The WSN is one of the most promising technologies in every field. WSN has an overlay large number of sensor systems. So managing such an oversized range of nodes need scalable and efficient algorithms. In most scenarios, sensor networks adopt ML techniques to remove the need for unnecessary redesign. Applying ML algorithms to WSN, it invites you to use various sensitive solutions to maximize the use of resources and extend the network life [7].

3 Benefits of Machine Learning in Wireless Sensor Networks

There are many benefits in implementing machine learning algorithms in wireless sensor networks. Few of them are listed below with an explanation.

(a) If we adapt machine learning techniques/algorithms in wireless sensor networks, where we train the network to monitor dynamically in environments that change over time, i.e., for example, environmental change can be erosion in soil or by turbulence in sea, and many more situations similar to it can adopt such changes dynamically and operate itself efficiently without any human intervention.

(b) WSN is mainly used in cases where human intervention is not possible, and these sensor networks operate on behalf of human to identify the situations and gather the data from such environmental locations.

(c) Machine learning algorithms can be used for maximum data coverage by the sensors. As WSN applications cowl minimum knowledge thanks to limited detector capability or hardware resources of the sensors.

(d) Big data, cloud computing, Internet of things, cyber systems, and communication between machinery with an intelligentsia motivation to support decisions along with the control of free networks have made a significant contribution to ML development in WSN. So, ML plays a prominent role to extract various levels of distinct abstractions necessary to perform tasks of artificial intelligence with very little human involvement [8, 9].

4 Detriments of Machine Learning in Wireless Sensor Networks

However, when using machine learning techniques in wireless sensor networks, a few disadvantages and limitations should be considered.

(a) Multiple ML methods are costly to compute. It can have adverse effects when used on wireless sensor networks, depending on how much the measurements are performed, leading to higher power/energy consumption.
(b) ML better perform as the amount of data you have increased. Because WSNs are commonly used in volatile settings, we cannot be sure of the data we enter (except in supervised learning).
(c) Training the nodes in the network with a large number of samples may not be sufficient, and it cannot react automatically if any unusual thing has been sensed, i.e., it may lead to reasonably tiny error bounds.
(d) At times it may land up with a resource crisis when training the network with high computational units to manage developed systems with centralization to perform the learning task [9].

5 Categorization of Machine Learning Algorithms Based on Localization

The primary purpose of localization is used to identify the geographical location/placement of a sensor node. When we deploy the nodes randomly in a field/ environment, we are unaware of its location. When nodes are used randomly in an environment there is a chance of change in the climate dynamically, during those changes machine learning algorithms [10] help to improve the location accuracy. The existing machine learning algorithms can be categorized into three types based on the mode of the intended network structure of sensors. Here, we are primarily focusing on localization issue in wireless sensor networks, and we are categorizing the machine learning algorithms based on localization [11]. There are machine learning algorithms that are categorized into two-three classes, namely supervised learning, unsupervised learning, and reinforcement learning as shown in Fig. 3.

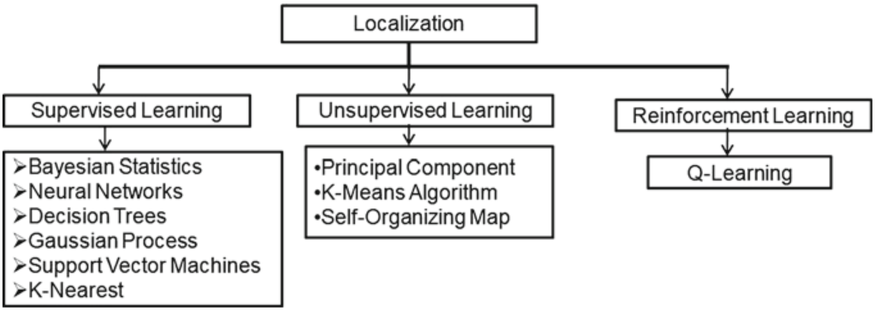


Fig. 3 Classification of machine learning techniques in wireless sensor techniques

6 Supervised Learning

In supervised learning, the data sets with related labels(inputs) [12] are collected and the relationships are found with the system, while training. At the end of the training method, an activity from associate grade input can be carried out with the best estimation of output y, important tasks of supervised algorithms are to develop a model that reflects links between input options and predicted objective outputs. During supervised learning, the machine learning algorithms consider two phases, i.e., training phase and prediction phase [13]. The data set which is being used represents the learned relationship between input, output, and system parameters. In this, the data set is divided into two categories, namely regression-based supervised learning and classification-based supervised learning.

A regression-based supervised learning is used when the resultant variable or output is some real or continuous value. Some value (Y) should be predicted based on a certain number of features (X) represented in Eq. (1). The variables (X) are either continuous or quantitative in the regression model to predict precise results (Y) with minimal errors.

$$Y = f(X) + \epsilon \tag{1}$$

where Y is a dependent output variables (output), X indicates an independent input variable (input) and represents the potential random error by the f is the function making a relationship between X and Y [10].

A classification-based supervised learning attempts to draw certain conclusions (Y) from the observed values (X). Given one or more inputs (X) in a classification model will try to predict the value of one or more outcomes (Y). This classification of supervised algorithm is based on four basic ideas they are namely logic-based, instance-based, perceptron-based and statistical-based classification which can be classified into six popular algorithms.

(a) Bayesian statistics
(b) Neural networks(NN)
(c) Decision trees(DT)

(d) Gaussian process
(e) Support vector machines(SVM)
(f) K-nearest neighbor (KNN).

6.1 Bayesian Statistics

Bayesian methods adapt the distribution of probability to learn certain concepts effectively and without over-fitting. It is logical thinking desires a comparatively tiny set of coaching samples [14] in distinction to machine learning algorithms. Bayesian strategies use chance distribution operate to find out unsure ideas (e.g., ϕ) while not over-fitting with efficiency. The algorithmic rule applies the present data, i.e., collected knowledge abbreviated as T to update previous beliefs into posterior beliefs $p(\phi|T) \infty p(\phi)\ p(T|\phi)$, wherever $p(\phi|T)$ is that the posterior chance of the parameter ϕ given the observation T, and $p(T|\phi i)$ is that the opportunity of the observation T given the setting ϕ. One application of Bayesian logical thinking in WSNs is to assess event consistency (ϕ) exploitation incomplete knowledge sets (T) by investigation previous knowledge concerning the atmosphere. But, such applied math data demand restricts the full usage of Bayesian algorithms in the field of WSNs. The key issue is to use the current to update prior assumptions in background assumptions where the subsequent probabilities of the parameter given the observation are that the setting is likely to be considered. The evaluation of the consistency of events with integrated data sets using prior environmental knowledge is used to access Bayesian inference in WSNs.

In [15], the authors Nguyen et al. dealt with the multi-source location of WSN and proposed for this issue a statistical sampler solution based on new application of the Monte Carlo sequence (SMCs) with unknown quantified source data obtained in the fusion center by various sensors from anonymous wireless channels. The experimentation results show that the proposed algorithm is best when compared with classical methods. The authors [16] addressed the localization of the node, and the issue is resolved by proposing a refinement of the Bayesian algorithm and referred to as the progressive correction. The particle filtering technique is used for node localization in progressive correction using a small number of parameters, but for a large number of parameters, this technique is not implemented. So a generalization procedure is proposed to significantly find more accurate node localization. The numerical simulations demonstrated that the algorithm proposed is giving more accurate results for node localization with the moderate expense of computation

Device-free localization (DFL) techniques were proposed by authors [17] to estimate target locations by analyzing their shadowing impacts in the area of interest on radio signals. They proposed a DFL technique-based compressive sensing to use the diversity of the frequency of subcarrier information in fine grains. In this method, it is a sparse recovery problem to build dictionaries from several channels on the saddle surface model with a multi-target DFL. To estimate the location vector, a multitask Bayesian compressive sensing (MBCS) framework develops an iterative

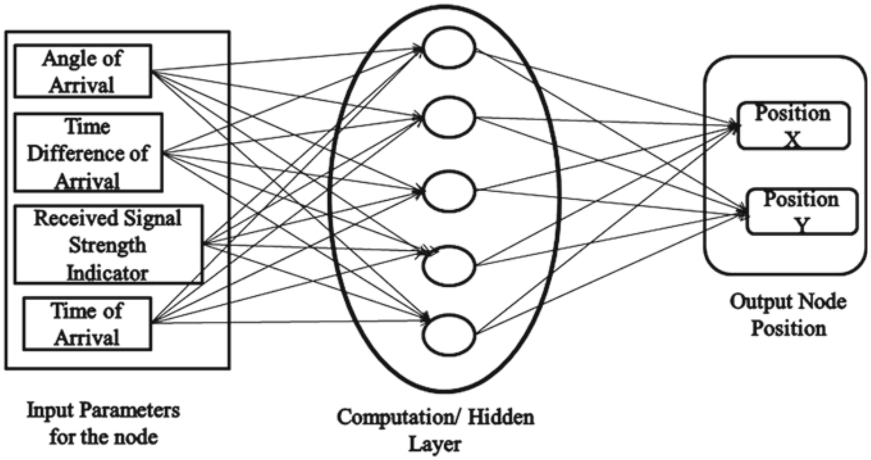

Angle of Arrival
Time Difference of Arrival
Received Signal Strength Indicator
Time of Arrival
Input Parameters for the node
Computation/ Hidden Layer
Position X
Position Y
Output Node Position

Fig. 4 Position estimation using supervised learning

location vector estimation algorithm. In comparison with CS-based multi-target DFL approaches, the superiority of its work is demonstrated through the simulation results.

6.2 *Neural Networks*

This algorithmic learning rule cascades chains of decision units, i.e., perceptron's or radial basis functions and is employed with efficiency to spot nonlinear and sophisticated features. The appliance of neural networks in WSNs in distributed manners continues to be not thus pervasive because it wants high procedure power to be told the network weights and depends on senior management overhead. In distinction, at the centralized manner, the NNs learn multiple outputs and call boundaries without delay that resolves many networks challenges victimization the identical model. Neural networks are known for various inputs and multiple outputs along with uncountable hidden layers for processing between inputs and outputs. The application of neural networks, for instance, sensor node localization downside is taken into account in WSNs. The node is located at the angle spread and distance measurement of anchor node signal received, which means arrival time (TOA) with received signal strength indicator(RSSI) received as well as arrival time distinction(TDOA) as illustrated in Fig. 4. The localization considers in three layers. The input layer is responsible for considering distance or angle estimation between the nodes. The former hidden layer it can be called as single layer or multiple layers for computing position calculation of each single node. The final stage is the output layer in which localization of whole network.

In [17], the authors identified localization of nodes could not be determined using traditional mechanisms due to hardware restrictions of the nodes. Later, they identified soft-computing techniques like neural networks could be used as a solution for localization problem. They suggested the new technique of the neural network by minimizing the number of neurons from hidden neural network layers by an algorithm of particle swarm optimization. They proved that the unique algorithm has a lower rate of localization errors [18] and a lower storage requirement than the existing analog methods. The authors in [19] suggested a problem of range-free localization with artificial neural networks and introduced an algorithm named "the anisotropic signal attenuation robust localization" that uses distance estimation (DE) approach to effectively derive the gap in closed from in order to attain anisotropic signal attenuation for the location of the node. The simulation results prove that the proposed algorithm presents a range-free localization both in accuracy and robustness. In [20], the authors identified a solution for localization problem for application-based which uses both techniques particle swarm optimization (PSO)-based in advanced version and artificial neural networks for the application indoor and outdoor application tracking. There are two approaches to the proposed method. The first approach is based on a proposed hybrid particle swarm optimization and artificial neural networks (PSO-ANN) algorithm to improve the distance estimation between the nodes of accurate node localization by using the feed forward neural network type and the Levenberg–Marquardt training algorithm. The first approach is based on a log normal shadow models(LNSM) for canal propagation and the next strategy. The authors concluded their research in this paper by saying "there was a mean absolute error of 0.022 and 0.208 m for both the outdoor and indoor environments." Indoor environments were investigated by the density of anchor nodes for the precision of the location. In the paper by Payal et al. [21] proposed the creation of a fast coverage and low costs neural network feed forward artificial neural networks (FFANN) to develop the wireless sensor networks (WSN) location framework. In order to build a cost-effective locale framework, this FFANN method has shown conclusively conjugation grade-based sensor nodes.

6.3 Decision Trees

Decision trees (DT) are a type of supervised ML classification method focused on if-then rules to improve readability. This algorithm predicts labels of data by iterating sample data (input) through a learning tree. In this processing, a comparison of feature property is made with decision conditions to reach a specific category (output) based on the decision condition [13]. DT offers an easy but efficient method to identify WSN connection reliability by defining a few critical features such as loss rate, corruption rate, mean failure time (MTTF), and mean restore time (MTTR). DT works with linearly separable results, however, and the process of building optimal learning trees is NP-complete [9]. Merhi et al. [22] developed a method for WSNs for acoustic target localization. Exacting locations of targets are determined using one of the two ways in decision trees is the time difference of arrival (TDOA) metric using a spatial correlation in the decision tree. They proposed the design of the protocol "event-based medium access control" (EB-MAC) to establish an acoustic localization in WSNs.

6.4 Gaussian Process

This Gaussian process is the solution in the selection optimum sensor locations to achieve resistance to node failures in the network and model ambiguity. Krause et al. [23] proposed an optimized solution algorithm called lazy learning algorithm for placement of sensors based on the application for which it is being used. One exciting feature of this solution is the development of an investigation phenomenon. Lazy learning algorithms store samples of training and delay the main workload until the request for classification is received.

A distributed node motion protocol for location in networks of wireless sensor systems was developed by the authors [24]. This approach is used to predict optimal locations of motive nodes based on their movements by distributing the Gaussian process regression (DGPR). It overcomes the disadvantage of traditional Gaussian process regression (GPR) algorithms, where *N* is the number for sample size, with O(N3)'s computational complexity. The algorithm proposed formed the solution to overcome the computational complexity by adopting a sparse Gauss process regression algorithm. In this process, each node will independently execute the regression algorithm using only local neighbor's spatial time information for locating nodes.

6.5 Support Vector Machines

This method is used in each unfeasible sensor in the wireless sensor network to self-place the node in every device. The localization based on support vector machines (LSVM) method for locating the nodes in WSNs has been proposed by Tran and Nguyen [25]. LSVM adopts a number of decision metrics, including connectivity information and indicators to achieve its design goals and to produce suitable training data. Although LSVM provides a distributed localization quickly and efficaciously, its performance in training samples is still sensitive to outliers. The nonlinear SVM 2D visualization is shown in Fig. 5.

In [26], the author proposed a new range-free localization algorithm based on polar coordinates sensor nodes to solve the locational problem within wireless sensor networks by support vector machine (SVM). With the WSN field boundaries, each sensor node can be located in one of the endless networks by dividing into a certain number of polar grids. Then, the center of the resident polar grid is calculated as the sensor node location. Furthermore, the authors suggested a new algorithm to enhance node accuracy. In THMSO, both neighborhood information and northern

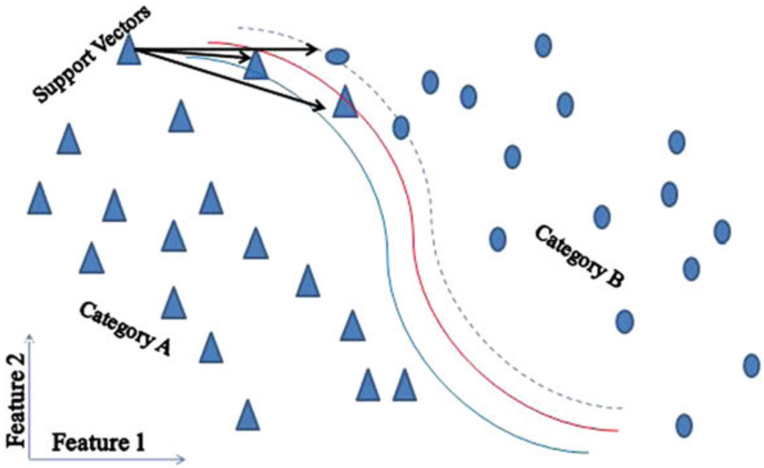


Fig. 5 Nonlinear support vector machine. *Source* Pandey, "Localization Adopting Machine Learning Techniques in Wireless Sensor Networks", 2018

node information is used as refinement to locate the sensor node. The algorithm proposed is the THMSO two-hop mass-spring optimization (THMSO). The findings show that the algorithm proposed improves better than existing methods of localization. When the detection area is too wide, the author [27] proposed a solution that requires each sensor node to be classified several times to locate SVMs, which means that placement time is too long which hampers good SVM performance. For the similarity measure, the proposed quick-SVM uses the minimum spanning by dividing the support vectors into groups according to the minimized functions. A linear combination of " determining factor" and "adjusting factor," based on a similarity of classification speed, is replaced by each group of support vectors. Vector support machines provide the most popular options for resolving no convex free improvement problems for neuropathic networks. The malicious behavior of sensing element nodes, security, and location should be used in the context of the WSN for intrusion detection or police work. It is possible to reveal in knowledge with SVM the spatiotemporal correlations [28].

6.6 *K-Nearest Neighbor*

K-nearest neighbor is a query processing algorithm. This query is applied to the classified data and generates output values for the adjacent data samples as labels. Several functions are available to determine the closest node set. *K*-nearest neighbor requires a high computer capacity, as it is calculated on the basis of simple connected points. In this article [29], the authors proposed a solution for the novel spatial query question in mobile sensor networks, l distant K-nearest neighbors (l-KNN). The consequence of the question implies well-scattered objects closest to the interest field. The l-KNN method can be used in most KNN applications, whether we want the KNN result to be well-distributed or narrowly protected. l-KNN divides the search space into several track-sectors where all sides are equal or greater than the distance limit. By choosing Q-nodes in alternating track-sectors, we ensured l distances between any two Q-nodes. To keep the gap tight, we changed the track-sectors' central angles and radius.

7 Unsupervised Learning

In unsupervised learning, there is no output (unlabeled) related to the inputs; even the model try and extract the relationships from the information. Unsupervised learning approach used as classifying the set of comparable patterns into clusters, spatiality reduction, and anomaly detection from the info. The main contributions of unsupervised learning in WSNs are to tackle different problems like property downside. In this, the output vector is not provided. Its primary goal is to classify the simple sets to different clusters or groups by investigating the similarity between the input samples. It tries to extract the relationships among the data associated with the input. Classification based on unsupervised learning: The classification of supervised algorithm is based on four basic ideas they are, namely logic-based, instance-based, perceptron-based, and statistical-based classification.

(a) Self-organizing map (SOP)
(b) Principal component analysis (PCA)
(c) K-means clustering.

7.1 Self-Organizing Map

The WSN-based self-organizing maps (SOM) solution consisting of thousands of nodes was introduced by Paladina et al. [30]. The proposed solution involves the execution of each node with a simple SOM algorithm which considers three layers which have one input layer and two output layers. Here, the input layer has spatial coordinates of anchor nodes which are 8 in number. An unknown node surrounds these eight anchor nodes. After training the hidden nodes which are spatially coordinated in a 2D space by the output layer. To find the absolute locations from the nodes is difficult from traditional methods, as a solution in [31] proposed a localization algorithm based on node connectivity information and the SOM algorithm. This algorithm works well for networks with limited resources. The proposed algorithm can be termed as a centralized algorithm as each nodes information is transmitted to the central processing unit to design an adjacent matrix for identification of node location. In [32], Lee the author proposed a scheme for node localization. In this, the author presented a new way of localizing nodes without the need for the anchor nodes. This scheme also uses the SOM algorithm process efficiently without any restriction on the number of nodes. From [31, 32], we can analyze that both the authors used the SOM algorithm but in different ways. By the analysis, we can say that [31] is more efficient over [32] because it minimizes the overhead of node transmission by eliminating the need for the central unit.

7.2 Principal Component Analysis

It is essentially a technique to compress information by reducing its dimension by extracting vital info exhibited from collected information set and rework it into a brand new orthogonal variable known as principal elements. It is a dimensionality reduction with the multivariate method for compression of data, aiming to extract relevant priority-based information from gathered data in the form of orthogonal variables called principal components. The technique is used for the acoustic location of the underwater and the detection of abnormalities in wireless networks. The 2D visualization of the principal component analysis is shown in Fig. 6.

In [33], the authors proposed a scheme using probabilistic pattern recognition in eigenspace of PCA for underwater localization. Based on the proposed system, the information can be easily obtained by probabilistic pattern recognition of projected features in PCA space. Experimental results have shown that the proposed underwater localization scheme is efficient and accurate when compared with existing techniques.

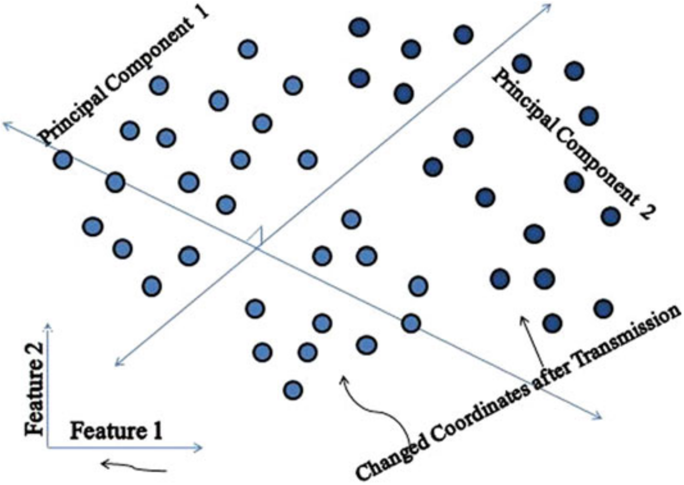


Fig. 6 PCA 2D visualization. *Source* Pandey, "Localization Adopting Machine Learning Techniques in Wireless Sensor Networks", 2018

7.3 K-Means Clustering

The k-means algorithm groups information into entirely different categories referred to as clusters. The unsupervised learning formula is widely adopted in a bunch of detector node as its implementation is smooth and has a linear process complexness. The authors [34] addressed cost-effectively to measure cost-effective K-nearest neighbor queries in a 3D sensor network using intelligent mobile data collectors.3D plane rotation algorithm that maps selected sensor nodes on different planes to a reference plane and a novel neighbor selection algorithm based on node distance and signal-to-noise parameters. We have implemented GlomoSim's 3D-KNN algorithm and validated the cost efficiency of the proposed algorithm through comprehensive performance evaluation over well-defined device parameters.

8 Reinforcement Learning

Reinforcement learning allows a sensing element node to find out its surroundings by interacting it. The agent learns to require the most effective actions that maximize its long edges mistreatment its expertise.

8.1 Q-Learning

A widely known rule, Q-learning [35] is a form of reinforcement learning technique is explained in. As illustrated in Fig. 7, associate degree agent updates its achieved edges no inheritable due to the action taken at a given state regularly.

The entire advantages awarded called the Q-value of performing arts associate degree action (A_t) at a given state St is calculable as shown in equation (2).

$$Q(S_{t+1}, A_{t+1}) = Q(S_t, A_t) + \alpha(b(S_t, A_t) - Q(S_t, A_t) \tag{2}$$

This rule is often applied only in a highly distributed design, such as a network of wireless sensing elements, when each node takes action to maximize its length. It is essential to note the extensive use of Q-learning and efficiency in WSN routing

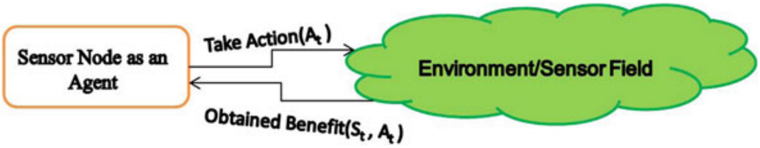


Fig. 7 Visualization of the Q-learning method

downside. where α (S_t, A_t) denotes the immediate reward of acting A_t a given state S_t and is the learning rate that determines how fast learning occurs (usually set to value between 0 and 1) [13].

Li et al. [36] stated that the Q-learning rules does not represent the various positions of the MB, and also the rule target is to hide all the sensors within the monitored space (i.e., all the sensors ought to hear a location update message from the MB at some stages). The entire operation is conducted within a mobile phone, which can save hidden node resources. However, the whole system can fail in case of mobile beacon defects as a centralized technique.

9 Conclusion and Future Work

In several applications of WSNs, localization of the physical/geographical location of a node is outlined. The sensor node that is placed in an excessive field without knowing its position and no other infrastructure on the market is available to track its status. The location of the sensing element, however, may be a vital task in this unique situation. This task is familiar by means that of manual assignment or geographical position system (GPS). The position of device nodes within the surroundings will support an amendment by dynamically due to some external causes. To handle such situations, machine learning algorithms can be needed to avoid the stress and strain of re-programming or reconfiguring the network and improve the accuracy of the location of nodes in wireless sensor networks.

Machine learning offers a range of techniques to strengthen the power and dynamic behavior of wireless sensor networks. Although machine learning techniques are applied to several applications in WSNs, several problems are still open, and additional analysis efforts must be put into execution to solve many issues in WSNs. Furthermore, process intelligence paradigms like neural networks and neuro-fuzzy strategies, swarm intelligence algorithms like ant colony optimization, and evolutionary formulas like the competitive imperialist algorithm might also be applied to enhance the performance of networks, and soft-computing techniques can also be used to solve the challenges of wireless sensor networks. Moreover, numerous problems are still open for future analysis like developing lightweight and distributed message passing techniques, stratified agglomeration patterns, online learning algorithms, and adopting machine learning is additionally in resource management drawback of wireless sensor networks.

References

1. Bonaccorso G (2017) Machine learning algorithms. Packt Publishing Ltd
2. Chowdhury TJ, Elkin C, Devabhaktuni V, Rawat DB, Oluoch J (2016) Advances on localization techniques for wireless sensor networks: a survey. Comput Netw 110:284–305

3. Singh N, Rautela K (2016) Literature survey on wireless sensor network. Int J Eng Comput Sci 5(8)
4. Wang J, Ghosh RK, Das SK (2010) A survey on sensor localization. J Control Theor Appl 8(1):2–11
5. Thanh Binh Huynh Thi, Nilanjan Dey (2018) Soft computing in wireless sensor networks. CRC Press, Boca Raton
6. Binh HT, Hanh NT, Nghia ND, Dey N et al (2020) Metaheuristics for maximization of obstacles constrained area coverage in heterogeneous wireless sensor networks. Appl Soft Comput 86:105939
7. Bera S, Das SK, Karati A (2020) Intelligent routing in wireless sensor network based on african buffalo optimization. In: Nature inspired computing for wireless sensor networks. Springer, pp 119–142
8. Saad E, Elhosseini M, Haikal AY (2018) Recent achievements in sensor localization algorithms. Alexandria Eng J 57(4):4219–4228
9. Alsheikh MA, Lin S, Niyato D, Tan HP (2014) Machine learning in wireless sensor networks: algorithms, strategies, and applications. IEEE Commun Surv Tutorials 16(4):1996–2018
10. Kumar DP, Amgoth T, Annavarapu CS (2019) Machine learning algorithms for wireless sensor networks: a survey. Inf Fusion 49:1–25
11. De D, Mukherjee A, Das SK, Dey N (2020) Wireless sensor network: applications, challenges, and algorithms. In: Nature inspired computing for wireless sensor networks. Springer, pp 1–18
12. Bhatti G (2018) Machine learning based localization in large-scale wireless sensor networks. Sensors 18(12):4179
13. Pandey S (2018) Localization adopting machine learning techniques in wireless sensor networks
14. Nguyen TL, Septier F, Rajaona H, Peters GW, Nevat I, Delignon Y (2015) A bayesian perspective on multiple source localization in wireless sensor networks. IEEE Trans Signal Process 64(7):1684–1699
15. Morelande MR, Moran B, Brazil M (2008) Bayesian node localisation in wireless sensor networks. In: 2008 IEEE international conference on acoustics, speech and signal processing. IEEE, pp 2545–2548
16. Box GE, Tiao GC (2011) Bayesian inference in statistical analysis, vol 40. Wiley, London
17. Guo Y, Yu D, Li N (2018) Exploiting fine-grained subcarrier information for device-free localization in wireless sensor networks. Sensors 18(9):3110
18. Banihashemian SS, Adibnia F, Sarram MA (2018) A new range-free and storage-efficient localization algorithm using neural networks in wireless sensor networks. Wirel Pers Commun 98(1):1547–1568
19. El Assaf A, Zaidi S, Affes S, Kandil N (2016) Robust anns-based wsn localization in the presence of anisotropic signal attenuation. IEEE Wirel Commun Lett 5(5):504–507
20. Gharghan SK, Nordin R, Ismail M, Abd Ali J (2015) Accurate wireless sensor localization technique based on hybrid pso-ann algorithm for indoor and outdoor track cycling. IEEE Sens J 16(2):529–541
21. Payal A, Rai CS, Reddy BV (2014) Artificial neural networks for developing localization framework in wireless sensor networks. In: 2014 international conference on data mining and intelligent computing (ICDMIC). IEEE, pp 1–6
22. Merhi Z, Elgamel M, Bayoumi M (2009) A lightweight collaborative fault tolerant target localization system for wireless sensor networks. IEEE Trans Mob Comput 8(12):1690–1704
23. Krause A, Singh A, Guestrin C (2008) Near-optimal sensor placements in gaussian processes: theory, efficient algorithms and empirical studies. J Mach Learn Res 9(2):235–284
24. Gu D, Hu H (2012) Spatial gaussian process regression with mobile sensor networks. IEEE Trans Neural Netw Learn Syst 23(8):1279–1290
25. Tran DA, Nguyen T (2008) Localization in wireless sensor networks based on support vector machines. IEEE Trans Parallel Distrib Syst 19(7):981–994
26. Wang Z, Zhang H, Lu T, Sun Y, Liu X (2018) A new range-free localisation in wireless sensor networks using support vector machine. Int J Electron 105(2):244–261

27. Zhu F, Wei J (2017) Localization algorithm for large scale wireless sensor networks based on fast-svm. Wirel Pers Commun 95(3):1859–1875
28. Lu CH, Fu LC (2009) Robust location-aware activity recognition using wireless sensor network in an attentive home. IEEE Trans Autom Sci Eng 6(4):598–609
29. Han Y, Park K, Hong J, Ulamin N, Lee YK (2015) Distance-constraint k-nearest neighbor searching in mobile sensor networks. Sensors 15(8):18209–18228
30. Paladina L, Paone M, Iellamo G, Puliafito A (2007) Self organizing maps for distributed localization in wireless sensor networks. In: 2007 12th IEEE symposium on computers and communications. IEEE, pp 1113–1118
31. Giorgetti G, Gupta SK, Manes G (2007) Wireless localization using self-organizing maps. In: Proceedings of the 6th international conference on Information processing in sensor networks, pp 293–302
32. Hu J, Lee G (2008) Distributed localization of wireless sensor networks using self-organizing maps. In: 2008 IEEE international conference on multisensor fusion and integration for intelligent systems. IEEE, pp 284–289
33. Lee KC, Ou JS, Huang MC (2009) Underwater acoustic localization by principal components analyses based probabilistic approach. Appl Acoust 70(9):1168–1174
34. Jayaraman PP, Zaslavsky A, Delsing J (2010) Intelligent processing of k-nearest neighbors queries using mobile data collectors in a location aware 3d wireless sensor network. In: International conference on industrial, engineering and other applications of applied intelligent systems. Springer, pp 260–270
35. Vijayakumar V (2019) Application of machine learning in wireless sensor network. Springer, Cham, pp 1–7
36. Li S, Kong X, Lowe D (2012) Dynamic path determination of mobile beacons employing reinforcement learning for wireless sensor localization. In: 2012 26th international conference on advanced information networking and applications workshops. IEEE, pp 760–765

Vehicular Delay Tolerant Network Based Communication Using Machine Learning Classifiers

Amit Kumar Singh and Rajendra Pamula

Abstract In this intelligent era, vehicles are exploited for a different mobile sensor activity. Vehicular delay tolerant network is the application of delay tolerant network. Nowadays, when conventional network does not work or fails in the emergency situation, vehicular delay tolerant networks provide solutions. Vehicular delay tolerant networks is very useful for solving many problems such as sensor-based applications, intelligent traffic, weather forecasting and many more delay tolerant services like campus information services etc. Many more delay adaptive services to save infrastructure-based network load, and these type of networks are very successful. For efficient routing strategy, the efficient selection of vehicular relay node is very important. So in this chapter, we have proposed "vehicular delay tolerant network-based communication using machine learning classifiers." First, we have analyzed which machine learning classifier is the best solution for our problem. We have used machine learning classifiers for filtering efficient vehicular nodes, so that packets can be delivered from source to destination.

Keywords Vehicular delay tolerant network · Routing strategy · Machine learning · Classifier · Relay selection

1 Introduction

In this chapter, we have considered network scenario in which group of moving vehicular delay tolerant nodes(VDTN) and fixed nodes operate as a relay for transferring data opportunistically, when they come in the range of each other. The considered network is delay tolerant network (DTN) [1, 2] and the considered mobile vehicles

A. K. Singh (✉) · R. Pamula
IIT (ISM) Dhanbad, Jharkhand, India
e-mail: amit.jgec@gmail.com

R. Pamula
e-mail: rajendra@iitism.ac.in

S. K. Das et al. (eds.), *Architectural Wireless Networks Solutions and Security Issues*, Lecture Notes in Networks and Systems 196,
https://doi.org/10.1007/978-981-16-0386-0_11

operate as vehicular delay tolerant network (VDTN) nodes, the network topology continuously changes and disconnection depends upon other factors also.

DTN is a research area concentrated on network, wherein interface availability might be habitually disconnected and the continuous path to a destination is absent. These types of networks generally applicable to interplanetary communications, very provincial territories, emergency situations, and different type of mobile ad hoc networks(MANETs). The DTN architecture [3] includes bundle [4] layer provides many of the facilities to the architecture. Bundle protocol is an important aspect of DTN architecture; on the basis of this layer, network layer between the application layer and transport layer, datalink layer, and physical layers are designed. The bundle layer works as a medium layer between different types of network. Bundle layer provides a store, carry, and forward dependent paradigm, so that when disconnection occurs packets can be stored and then carried to the destination and forwarded once connectivity restores. The packets can be forwarded to a destination even continuous path does not exists with the help of relay nodes moving between source and destination. In addition to the above concept, it provides the way to differentiate the lower level of strategy, so that different protocols can use transport layer to physical layer according to the situation. The bundle layer is also responsible for binding and storage and carrying of the packets till a suitable transfer time found, and also responsible for transfer of bundle custody, and packet routing also.

Vehicular delay tolerant network (VDTN) is an emerging application area of DTN which uses the same paradigm of store, carry, and forward mechanism as DTN. In real life, VDTN is very prominent to solve many problems, e.g., traffic jam problem, awareness in public, and provide connectivity in disconnected area. With the help of machine learning, VDTN can provide a better solution to the problem of disconnectivity in emergency situation also. In real life, generally, conventional network stops working or it gets down because of some emergency situations or natural disaster. And disconnection occurs between any two point in the network or end-to-end network does not exists. In this case, VDTN is very reliable solution to the problem. For example, in connection between any to road side units(RSU) disconnects any point of time for a certain time period, then VDTN can provide solution and fills the gap of disconnectivity by relaying packets with the help of vehicular nodes. But for this, it is very important to have efficient routing strategy. An efficient routing strategy can only designed when an nodes are efficient enough to deliver the packets to the destination. So these problems motivated us to proposed our strategy.

VDTN is an easy alternative to conventional network when conventional networks fails to work. Conventional network depends upon continuous end-to-end connectivity nut VDTN can work where network is dis-continuous. In Fig. 1, there is a disconnention between Point 1 and Point 2. These two points may me road side units (RSUs) or any source or destination. Suppose Point 1 wants to send packet to Point 2 but there is no end-to-end connectivity between these points. So in the absence of end-to-end network, VDTN nodes are forwarding packets from Point 1 to Point 2.

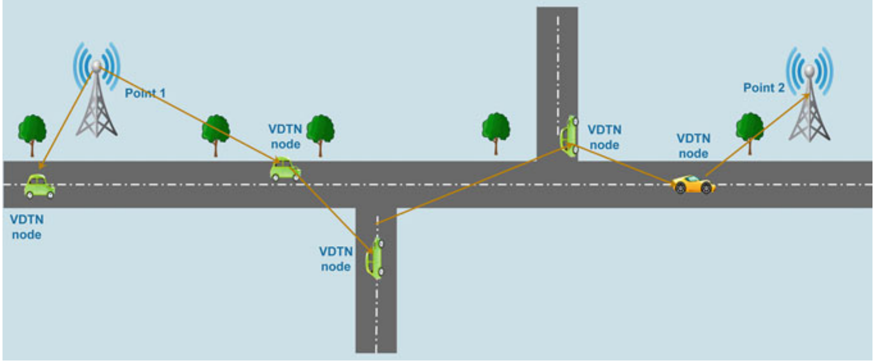


Fig. 1 Vehicular Delay Tolerant Network

1.1 Motivation and Contribution

The motivation of this chapter centers around the advancement of a machine learning-based routing protocol for VDTN. Many routing protocols have been produced for both opportunistic and DTN. RAPID [5], PRoPHET [6], Spray-and-Wait [7], and DTLSR [8] are among the most notable for sharp organizing situations. Inside the domain of interplanetary organizing, one of the most well-known routing protocols is contact graph routing [3, 9] or CGR, which utilizes the known contact times and separations of booked system advantages for deciding a useful course. Routing protocol expected for defer open-minded systems and specifically space networks must address a few issues. Flight equipment is frequently restricted in terms of handling ability and memory assets, so the calculation must be productive and not utilize over the top protocols or require much information stockpiling. In deep space, there might be a vast spread postponement between arranging nodes, so the exchanging of network status information becomes exorbitant and may not mirror the present condition of the system. In expansion, correspondence joins are frequently uneven, with the goal that an enormous measure of information might be sent from a network node. However, there might be constrained transmission capacity to get affirmations or then again different status data.

2 Literature Review

VDTN is one type of wireless networks which rapidly usage in real-life applications due to its flexible and efficient nature. Wireless network has several variation based on it frameworks which is describe in [10, 11]. This book contains several frameworks such as optimization, security and privacy, localization, and network lifetime enhancement. This book provides the basic frameworks ideas of the users and new researchers. In [12], the authors proposed nature-inspired-based methods in WSN. This book contains several popular nature-inspired algorithms that helps to the readers and researchers both for designing as well as innovating any algorithm.

There is a lot of successful standard research work available in the field of opportunistic routing for DTNs [7, 13–17]. Many are in light of the fundamental technique for scourge directing in which nodes experiencing each other exchange whatever messages the other does not as of now have. Right now, different hubs experience one another, numerous duplicates of the message are spread all through the system. Except if the message lapses previously, the goal is reached, or it is in certainty routed to a few hubs that are inaccessible inside the system by some other node, the packet will be conveyed. The main disadvantage of this method is that the various duplicates of a similar packet in the system expend hub information stockpiling and packet moving time pointlessly. Most shrewd steering conventions center on an approach to decide how to pick the best hubs to duplicate packets, which means nodes that have the best possibility of reaching the destination node. PRoPHET is one such strategies which utilizes the history of delivery records of nodes to decide the probability that a node will reach the destination. PRoPHET depends on human versatility designs, and also the perception that countless contact openings between the two nodes will follow a regular pattern [6]. Packets are recreated and sent to nodes present in its n that have a high likelihood of conveying it to its goal.

PRoPHET decides this probability dependent on a conveyance consistency metric. Every node keeps up a vector of conveyance probability of all nodes experienced and trades this data with different nodes during an underlying contact stage. The conveyance consistency is determined at whatever point two hubs are in contact. Nodes those are regularly in contact will have a higher conveyance consistency, and as such, the calculation picks up the two nodes as the favored way. There is very few words that apply machine learning methods to directing in DTNs [18–20]. Choice of tree-based classifiers tried to boost directing choices in case of epidemic routing by creating a group of nodes utilizing a quality vector and a determined grouping name [21] The technique used in [19] additionally utilizes locales, a time-sensitive list, and message goal as characteristics for a Bayesian classifier. These two strategies used to put away network traffic. Yang [22] designed an intelligent system for transportation system in wireless network. This is based on an existing transportation system based on process structured system. Finally, it helps to enhance network capabilities and services of the network. It also helps to user function and usages in the network and network metrics properly to maintain the network. Chandrakar [23] designed an authentication system for the users in wireless network. This is basically based on healthcare system and used for medical purpose. This proposal is used for sensing patient body information and send to the doctor for treatment and diagnosis purpose. It also helps in user authentication, privacy and data security purpose, so that efficient result comes from the diagnosis system. In WSN, data is gathered from multiple homogeneous or heterogeneous sources because, and real-life data is connected with different IoT, IoV, or cloud environment. So, it is difficult to keep the natures of the data in same structure. Information retrieval is very important part in modern research areas which indicates collect information that are stored in unstructured form based on multiple local languages and process it in particular pattern after observing. Hao et al. [24] designed an evaluation system for big data analysis. This data is based on IoV where it means Internet of vehicle. This proposal is based on K-means algorithm that is used here as a clustering. In this work, different behaviors of the driving are involved for controlling vehicle. Finally, it helps in reducing fuel consumption and helps in transportation globally. Shen et al. [25] designed a predictable-based routing method for ad hoc network. In this work, topology is organized by the helps to static and dynamic topology distribution with the help of not completely predictable method. Here, incomplete predictable technique initiated by anti-pheromone system. Finally, it achieves energy efficiency and node utilization both for enhancing the network lifetime. Chatterjee and Das [26] designed an ACO-based routing technique for MANET. The main aim of this routing method for enhance the QoS by increase ratio of packet delivery ratio and decreasing network delay by using ant. This method uses DSR routing as a base routing protocol. The basic route packet like RREQ and RREP is used here "request ant" and "reply ant" packet for managing the network. Finally, it determined level of pheromone for each route to decide optimal route of the network. Fatemidokht and Rafsanjani [27] designed an anomaly detection method for VANET based on clustering approach. In this paper, VANET contains some malicious nodes that act as several vehicles in the area of transportation. The nodes in this work disconnected and organize frequently in term of changes of topology. The clustering method in this work used for decision of gateway selection, proper neighbor selection, and also cluster head selection. Finally, it helps in packet delivery ratio and reduction of end-to-end delay. Lakshmanaprabu et al. [28] designed ACO-based routing technique for smart city in the context of VANET. This work basically, proposed influence of big data technology in term of smart city. Basically, these type of cities provide more secure travel and effectiveness lifestyle of the people. The big data technology of this work helps to enhance the traffic system of the cities. And it also helps to manage cluster of the nodes. Kadono et al. [29] proposed routing method for ad hoc network. This routing is based on GPSs system. The proposed method uses ACO technique for handling network path of the ad hoc network. In this network, ant works as an agent that handle and compromise several network parameters and maintain parameters changeability. In this work, due to GPS system, ant behaves as intelligent agent uses their pheromone intelligently. Finally, it helps in packet delivery ratio. Vinoba and Vijayaraj [30] designed a topology control-based method for ad hoc network. The proposed method is the fusion of ACO and Bayesian reasoning technique. The combination of both is used to manage dynamic position of the network and control residual energy of the nodes. Finally, it outperforms by several network metrics based on several parameters. Bello-Salau et al. [31] designed a routing method for VANET based on an optimization technique. This technique is based on GA optimization where several parameters are used to design constraints along with objective function. The names of the parameters are loss of path, link frequency, residual energy, and strength of the signal. The aim of this method is to reduce anomaly and help to reduce noise in the network. Robinson et al. [32] proposed routing method for MANET that is used to reduce overhead of the network. This routing method is based on knowledge information of the neighbor that is consists of two popular packets such as RREQ and RREP for requesting and replying. This neighbor knowledge helps to reduce to traffic of the network and overhead of the network, enhances broadcasting mechanism of the MANET. Wang et al. [33] designed ACO-based routing technique for MANET. In this network, bandwidth and energy both are crucial parameters, so here nature-inspired swarm intelligence technique is used for controlling these two parameters. It manages by the fusion of zone based routing and concept of ACO optimization. During this fusion, some feature is extracted from two popular routing protocols such as DSR and ZRP routing protocols. Giagkos and Wilson [34] designed a routing method for WANET using swarm intelligence nature-inspired technique. The main agent of this routing method is insect that behave intelligently to solve the solution. This is a multipath routing technique that work fusion of some intelligent behavior such as scalable, efficient, and adaptive nature. Finally, it outperforms under several network conditions. Misra et al. [35] designed energy aware routing protocol for WANET based on swarm intelligence. This is based on ACO technique that helps to optimize the network efficiently in term of battery energy. In this work, several energy consumption situations are controlled such as idle mode, sleep mode, and helps to prolong the battery life. Finally, it reduces the energy consumption and enhances the network lifetime. Bitam and Mellouk [36] designed a multicast system for routing in VANET. This is based on bee colony optimization which is used here to solve NP complete issue in VANET. This routing protocol is designed by the fusion of network simulator with C++ language. It helps to enhance the QoS of the network by optimizing some metrics such as bandwidth, jitter, network delay, and cost. These metrics considered as objective function for solving the network issues. Rosati et al. [37] designed an algorithm for ad hoc network. This algorithm is based on ant optimization where base routing protocol is used as AODV. This is based on distributed system. In this method, used heuristic approach helps to increase network lifetime and decrease the complexity of the node expenses. Finally, it helps to produce optimal solution that is suitable for the network. Kumar et al. [38] designed a technique for route determination in WSN based on several constraints. The basic key element of this method is ACO for optimizing the path between source node and sink node. In this work, the sensor node collect information from the environment and process it path finding. Finally, validate it in terms of network metrics and help to reduce delay and enhance the network lifetime.

3 System Model and Proposed Protocol

3.1 System Model

Our plan endeavors to take care of the VDTN routing issue as a machine learning problem. We picked up this technique for a few reasons. The arrangement ought to be versatile to new entering nodes in the concerned network, an assortment of the conditions, working in possibly unique time network and leaving the network. Furthermore, we needed an answer which may be ready to decide examples of interruption or examples of communication traffic which is not quickly self-evident. The network of machine learning can utilize information got from the network condition to decide such examples.

Very few types of techniques of machine learning can be considered for routing problems. The machine learning strategies are divided into supervised and unsupervised. The strategy with supervise learning is having huge number of associated information which is utilized for learners. The researchers create rules from the dataset to characterize new occasions of the information dependent on the preparation set. Information in the preparation set is marked, and the student makes forecasts that might be right or mistaken. Similarly, unsupervised learning strategy is utilized to show the information and learn increasingly about relationships inside the informational index. One more class of learning which is important is reinforcement learning which allow student to settles on choices at first in an experimentation strategy. Choices that bring about a positive result to gain the student a prize. Right now, researchers figure out what is acceptable and terrible choices in a given case.

Reinforcement type of learning is recommended for directional conventions in a few works [17–19]. There are, in any case, a few disadvantages on account of VDTN. One is that capacity must be resolved to empower the learners to get rewards. In VDTN, the objective is, for the most part, to limit conveyance time and boost conveyance likelihood. Utilizing time as a measurement in VDTN may prompt uncertain outcomes since postpone time may change on organize conditions that are out of the control of the learner (spread postponements between nodes, for instance) or, on the other hand, postponements may happen as a result of poor directing decisions. Conveyance consistency is a decent pointer of steering achievement, nonetheless by and large, in VDTN, it might be obscure at the source node side if the packet was in certainty conveyed. Conventions such as TCP/LTP can guarantee dependable conveyance; however, since these depend on affirmations or retransmission demands from the goal, there might be impressive postponement before this is known at the source node relying upon the separation between nodes and the information rate. It might be favored as far as speed, what's more, productivity to send information in straightforward datagrams (UDP/LTP green portions). Inside bundle protocol, conveyance receipts and care move can be mentioned at the group layer, yet once more it is constraining to expect that these instruments will consistently be utilized. It very well may be restrictive to expect that affirmations what's more, receipts will be engendered back to the sender and much work has been done in the VDTN people group to attempt to address the downsides of having possibly long full circle times to send a message and get an affirmation back. Along these lines, we might want to maintain a strategic distance from a convention that depends on affirmation, conveyance receipts or status parcels, especially the case in which the practicality of getting such input is critical to the exhibition of the calculation. On account of support learning, if the research depends on positive criticism from the destination to settle on better choices, this input could come at a significant time later and result in a progression of lackluster showing.

3.2 Proposed Strategy

Our motive is to focus on the learning so that nodes can learn to find the best possible path between source and destination. The objective is basically to decide the future system state-dependent on the historical backdrop of the system. We have considered few factors which can impact which highway a group of nodes should be taken; furthermore, and these factors can be utilized as a input for the algorithm made for learning the paths. The mentioned factors are future and current topologies of the network (all the information about source node, relay node, and destination node), duration of contact in interface range, data transfer rate, buffer capacity, and location. Every highlighted factors will change with time; however, we have implemented that a few and all will follow the same periodicity. We have considered this duration of time the age and it will be separate every age into time cuts. This is basically the idea as one hour, which can be separated into 60 min. All the nodes will have the same movement model for packet creating in the specified significant time.

3.2.1 Working Steps of the Protocol

- Passing by nodes come in the interface of source node Point 1 in Fig. 1.
- Source node uses classification of VDTN nodes.
- VDTN nodes are classified as performing and non-performing.
- All the classifiers are checked one by one for efficiency.
- Efficient classifier is selected.

We chose three notable classifiers (decision tree, Naive Bayes, and K-nearest neighbors) to figure out which can give the best execution. The classifiers are straightforward and naturally fit the portrayed issue. Nodes follow a given an example all through the age (people heading to work each day simultaneously); thus, it is sensible to expect that they keep on following this design in whenever section (commute home simultaneously also and it will return back at a comparative time the following day). The input in our classifier depends on a trait vector X comprising of the time file in the age, source node, destination node, and if the message was conveyed or not (1 or then again 0). The name information Y, or yield of the classifier, is the arrangement of nodes that the message sends to relay. On the off chance that the message has visited node I, at that point, the bit in the position inset; it is zero in any case.

The technique depicted partitions the routing by classification issue into n number of separate issues, with one classifier can produce a twofold yield showing a node is or is not an individual from the arrangement of hubs along a given course. It can be considered a multi-level classifiers, precisely the binary relevance technique (BR) [39]. It is one of the least complicated strategies for multi-level issues and uses an issue change approach. A separate characterization issue with different yields is changed into various grouping issues. Much work in the multi-mark arrangement has been centered around deciding the connections between marks to improve arrangement exactness.

Classifier chains [40] (CC) likewise change the multi-mark issue into a lot of individual twofold characterizations; in any case, the property space for each model is reached out with the twofold mark pertinent relationships of every single past classifier, shaping a chain. CC takes the relationship between recently grouped marks into account when performing grouping of the following mark. A poor decision of request can contrarily affect execution.

The above-mentioned approach utilizes past effective ways taken to attempt to foresee an acceptable way in the following age. Be that as it may, one of the features of machine learning also, characterization is to consider various beforehand watched credits to give a general likelihood for guaranteed result. Extra highlights, for example, area, support limit, also, information rate may improve execution by considering potential postponements brought about by moderate connections or unreasonable queueing times. We initially start to assess diverse node properties by considering the node area. Our methodology is to utilize the notable K-mean clustering algorithm [41] to decide districts in which hubs much of the time visit, which would then be able to be utilized as a credit to our classifier, much like the area code utilized in [16]. Instead of essentially isolating the zone into equivalent parcels, the K-mean clustering algorithm will give an information-driven way to deal with gathering node areas.

3.3 Performance Evaluation

To evaluate the multi-label classification problem, four notable multi-label predictions parameter are used generally. The two important related parameters for multi-label prediction are Hamming loss, furthermore, zero-one misfortune [42]. Hamming loss ascertains the part of marks that are erroneously arranged. That is:

$$L_{\mathrm{H}}(y, h(x))) = \frac{1}{m}\sum_{i=1}^{m}[[y_i \neq h_i(x)]] \tag{1}$$

In Eq. (1), $L_{\mathrm{H}}(y, h(x))$ is the Hamming loss(Hl) function, in the same equation y is the arrangement of m measured marks for a provided occurrence and $h(x)$ is the output of the classifier (the m number of predicted labels). The articulation $[X]$ assesses to 1 if X is valid and 0 in any case. The Hl is rather than zero-one loss, which considers the whole forecast is mistaken if any mark in the expectation is inaccurate, as appearing in Eq. (2):

$$L_{\mathrm{s}}(y, h(x))) = [y \neq (h(x))] \tag{2}$$

Hl is an increasingly permissive parameter that scores based on singular marks. In both Hl and zero-one loss values inclining toward zero demonstrate great execution, while values inclining toward one show a higher level of misclassification.

In both the engineered portability situations and in the real-world follows, decision tree-based classifiers played out the best over each of the four measurements. The base classifier chose had a lot more prominent effect on execution than either the basic steering strategy used to create the preparation information or the policy for multi-label classification. Result shows the exactness of ordering singular hubs dependent on past information recorded from steering with a few routing algorithms (PRoPHET and Epidemic). Orders are finished utilizing Naive Bayes, K-nearest neighbor, and Decision Tree classifiers. Results show the different measurements for the multi-level order approach. And from the results, it appears that decision tree-based characterization is a promising technique as a fundamental classifier, and this appeared to have the most noteworthy effect on execution, considerably more so than the multi-label approach utilized (chain classifiers, independent classifiers, gathering, and label powerset). Our results show the normal results from this arrangement of testing, looking at epidemic strategy, a multi-label classification is the best for all scenario. We have used Python for our simulation and scikit-learn module of Python to simulate the classifier. Scikit-learn is machine module for Python used for various classification, regression, and clustering. The results from Python simulation are plotted with the help of Matplotlib package of Python. The plotted results are shown in Figs. 2, 3, 4 and 5.

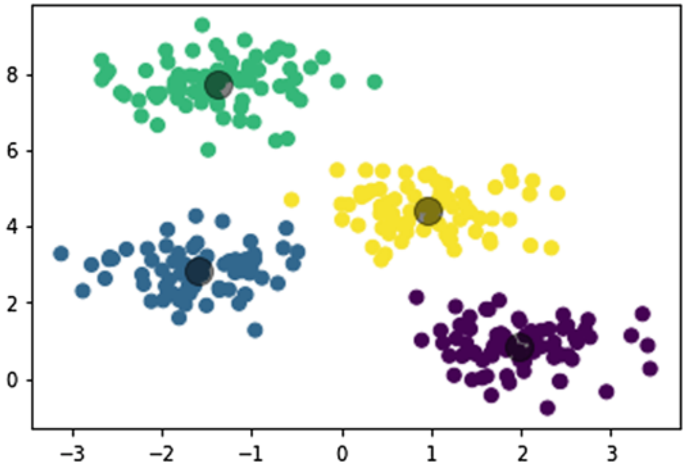


Fig. 2 Regions in nodes located by K-means clustering

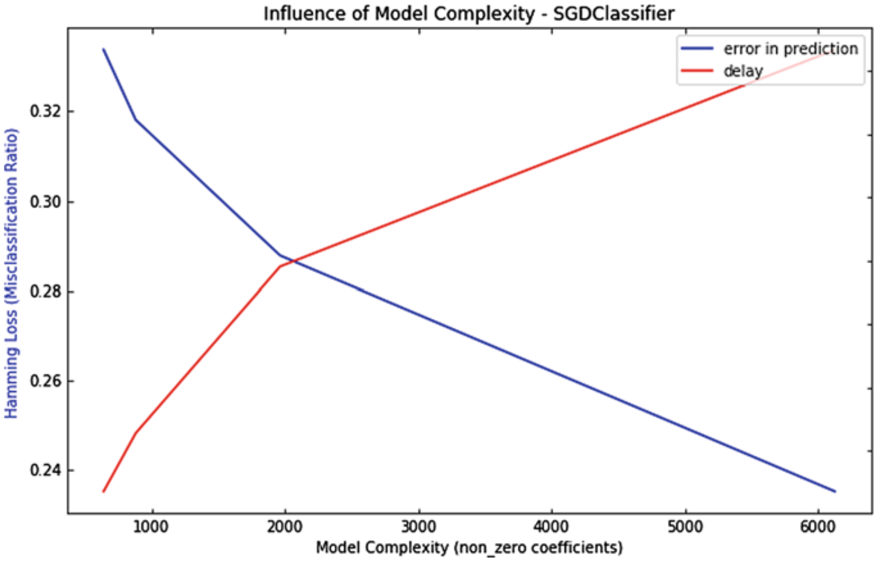


Fig. 3 Hamming loss

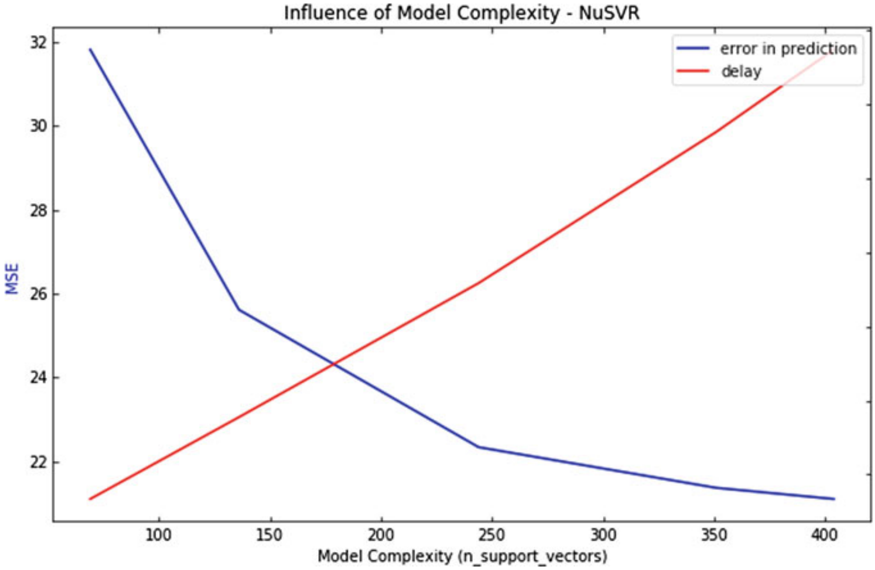


Fig. 4 Model complexity

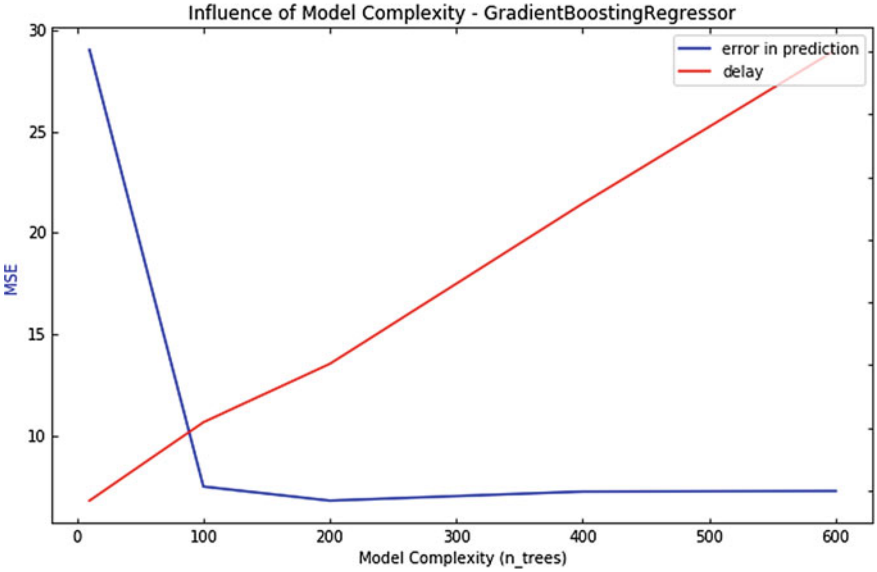


Fig. 5 Complexity

4 Conclusion

We have proposed the VDTN routing strategy “vehicular delay tolerant network-based communication using machine learning classifiers.” And our results propose that classification technique is a suitable strategy to anticipate and classify the traffic and decide the in all probability, and VDTN nodes to be experienced in a given way which can be utilized to settle on increasingly educated steering choices, diminishing overhead in pestilence-based steering draws near. The regular time and asset expending undertaking of preparing a learning calculation should be possible in a disconnected way, with information put away from a prior time. When the learning model has been produced, it very well may be traded to nodes in the communication scenario which basically need to play out the forecast figuring, which is substantially less escalated than preparing. This strategy has experimented as restricted to learning progressively since numerous order calculations include a preparation stage, trailed by model approval previously they can be utilized to make forecasts on new occurrences of information. In any case, it is conceivable in certain calculations to refresh the model as new information shows up, with the end goal that the model is continually adjusted. We leave this methodology for future work.

References

1. Singh AK, Pamula R (2018) IRS: Incentive based routing strategy for socially aware delay tolerant networks. In: 2018 5th international conference on signal processing and integrated networks (SPIN). IEEE, pp 343–347
2. Singh AK, Bera T, Pamula R (2018) PRCP: Packet replication control based prophet routing strategy for delay tolerant network. In: 2018 4th international conference on recent advances in information technology (RAIT). IEEE, pp 1–5
3. Araniti G, Bezirgiannidis N, Birrane E, Bisio I, Burleigh S, Caini C, Feldmann M, Marchese M, Segui J, Suzuki K (2015) Contact graph routing in dtn space networks: overview, enhancements and performance. IEEE Commun Mag 53(3):38–46
4. Cerf V, Burleigh S, Hooke A, Torgerson L, Durst R, Scott K, Fall K, Weiss H (2007) Delay-tolerant networking architecture, Keith Scott
5. Balasubramanian A, Levine BN, Venkataramani A (2009) Replication routing in dtns: a resource allocation approach. IEEE/ACM Trans Netw 18(2):596–609
6. Doria A, Lindgren A (2007) Probabilistic routing protocol for intermittently connected networks
7. Spyropoulos T, Psounis K, Raghavendra CS (2005) Spray and wait: an efficient routing scheme for intermittently connected mobile networks. In: Proceedings of the 2005 ACM SIGCOMM workshop on delay-tolerant networking, pp 252–259
8. Demmer M, Fall K (2007) DTLSR: delay tolerant routing for developing regions. In: Proceedings of the 2007 workshop on Networked systems for developing regions, pp 1–6
9. Jat DS, Bishnoi LC, Nambahu S (2018) An intelligent wireless qos technology for big data video delivery in wlan. Int J Ambient Comput Intell (IJACI) 9(4):1–14
10. Das SK, Samanta S, Dey N, Kumar R (2020) Design frameworks for wireless networks. Springer, Berlin
11. Mazumdar N, Roy S, Nayak S (2018) A survey on clustering approaches for wireless sensor networks. In: 2018 2nd international conference on data science and business analytics (ICDSBA). IEEE, pp 236–240
12. De D, Mukherjee A, Das SK, Dey N (2020) Nature inspired computing for wireless sensor networks
13. Mundur P, Seligman M (2008) Delay tolerant network routing: beyond epidemic routing. In: 2008 3rd international symposium on wireless pervasive computing. IEEE, pp 550–553
14. Rodolfi M (2015) DTN discovery and routing: from space applications to terrestrial networks. Ph.D. thesis
15. Burgess J, Gallagher B, Jensen DD, Levine BN et al (2006) Maxprop: routing for vehicle-based disruption-tolerant networks. In: Infocom, vol 6. Barcelona, Spain, pp 1–11
16. Dudukovich R, Raible DE (2016) Transmission scheduling and routing algorithms for delay tolerant networks. In: 34th AIAA international communications satellite systems conference, p 5753
17. Ahmed S, Kanhere SS (2010) A bayesian routing framework for delay tolerant networks. In: 2010 IEEE wireless communication and networking conference. IEEE, pp 1–6
18. Portugal-Poma LP, Marcondes CA, Senger H, Arantes L (2014) Applying machine learning to reduce overhead in dtn vehicular networks. In: 2014 Brazilian symposium on computer networks and distributed systems. IEEE, pp 94–102
19. Dudukovich R, Hylton A, Papachristou C (2017) A machine learning concept for dtn routing. In: 2017 IEEE international conference on wireless for space and extreme environments (WiSEE). IEEE, pp 110–115
20. Boyan JA, Littman ML (1994) Packet routing in dynamically changing networks: a reinforcement learning approach. In: Advances in neural information processing systems, pp 671–678
21. Valadarsky A, Schapira M, Shahaf D, Tamar A (2017) A machine learning approach to routing. arXiv preprint arXiv:1708.03074

22. Yang W, Wang X, Song X, Yang Y, Patnaik S (2018) Design of intelligent transportation system supported by new generation wireless communication technology. In: Intelligent systems: concepts, methodologies, tools, and applications. IGI Global, pp 715–732
23. Chandrakar P (2019) A secure remote user authentication protocol for healthcare monitoring using wireless medical sensor networks. Int J Ambient Comput Intell (IJACI) 10(1):96–116
24. Hao R, Yang H, Zhou Z (2019) Driving behavior evaluation model base on big data from internet of vehicles. Int J Ambient Comput Intell (IJACI) 10(4):78–95
25. Shen J, Wang C, Wang A, Sun X, Moh S, Hung PC (2017) Organized topology based routing protocol in incompletely predictable ad-hoc networks. Comput Commun 99:107–118
26. Chatterjee S, Das S (2015) Ant colony optimization based enhanced dynamic source routing algorithm for mobile ad-hoc network. Inf Sci 295:67–90
27. Fatemidokht H, Rafsanjani MK (2020) QMM-VANET: an efficient clustering algorithm based on qos and monitoring of malicious vehicles in vehicular ad hoc networks. J Syst Softw 165:110561
28. Lakshmanaprabu SK, Shankar K, Rani SS, Abdulhay E, Arunkumar N, Ramirez G, Uthayakumar J (2019) An effect of big data technology with ant colony optimization based routing in vehicular ad hoc networks: towards smart cities. J Clean Prod 217:584–593
29. Kadono D, Izumi T, Ooshita F, Kakugawa H, Masuzawa T (2010) An ant colony optimization routing based on robustness for ad hoc networks with gpss. Ad Hoc Netw 8(1):63–76
30. Vinoba R, Vijayaraj M (2020) Novel control topology with obstacle detection using rdpso-gba in mobile ad-hoc network. Comput Commun 60:847–857
31. Bello-Salau H, Aibinu AM, Wang Z, Onumanyi AJ, Onwuka EN, Dukiya JJ (2019) An optimized routing algorithm for vehicle ad-hoc networks. Eng Sci Technol Int J 22(3):754–766
32. Robinson YH, Krishnan RS, Julie EG, Kumar R, Thong PH et al (2019) Neighbor knowledge-based rebroadcast algorithm for minimizing the routing overhead in mobile ad-hoc networks. Ad Hoc Netw 93:101896
33. Wang J, Osagie E, Thulasiraman P, Thulasiram RK (2009) HOPNET: a hybrid ant colony optimization routing algorithm for mobile ad hoc network. Ad Hoc Netw 7(4):690–705
34. Giagkos A, Wilson MS (2014) Beeip-a swarm intelligence based routing for wireless ad hoc networks. Inf Sci 265:23–35
35. Misra S, Dhurandher SK, Obaidat MS, Gupta P, Verma K, Narula P (2010) An ant swarm-inspired energy-aware routing protocol for wireless ad-hoc networks. J Syst Softw 83(11):2188–2199
36. Bitam S, Mellouk A (2013) Bee life-based multi constraints multicast routing optimization for vehicular ad hoc networks. J Netw Comput Appl 36(3):981–991
37. Rosati L, Berioli M, Reali G (2008) On ant routing algorithms in ad hoc networks with critical connectivity. Ad Hoc Netw 6(6):827–859
38. Kumar P, Amgoth T, Annavarapu CSR (2018) Aco-based mobile sink path determination for wireless sensor networks under non-uniform data constraints. Appl Soft Comput 69:528–540
39. Tsoumakas G, Katakis I (2007) Multi-label classification: an overview. Int J Data Warehouse Min (IJDWM) 3(3):1–13
40. Simon GJ, Kumar V, Li PW (2011) A simple statistical model and association rule filtering for classification. In: Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining, pp 823–831
41. MacQueen J et al (1967) Some methods for classification and analysis of multivariate observations. In: Proceedings of the 5th Berkeley symposium on mathematical statistics and probability, vol 1. Oakland, CA, USA, pp 281–297
42. Dembczyński K, Waegeman W, Cheng W, Hüllermeier E (2010) Regret analysis for performance metrics in multi-label classification: the case of hamming and subset zero-one loss. In: Joint European conference on machine learning and knowledge discovery in databases. Springer, pp 280–295

Applications of Big Data and Internet of Things in Power System

Ramesh Chandra Goswami, Hiren Joshi, Sunil Gautam, and Hari Om

Abstract In recent years, Internet of things (IoT) technology is the fastest growing technology which connects physical device or sensors to Internet. IoT devices collect the information from object's then store or transfer information over the Internet without help of any manual involvement and with the help of embedded technology. The big data play a vital role in IoT because it is a process of a huge amount of information on real-time basis. This chapter highlights the use of big data and IoT for the power systems. IoT can be used in various areas of power system such as metering, transformer monitoring, prediction of demand and planning for future consumption. The main objective of this chapter to make a clear understanding of the use of big data and IoT in the power system and how it will improve customer service and social welfare.

Keywords Internet of things · Power system · Smart grid · Big data

1 Introduction

All of us know that electric energy is an essential element for today's life. The automation is at the top level in current society. Everything in human life is either automated or in pipeline of automation. The automation require electric energy, similarly industrial activity of any country, society also require electric energy. In similar way, whatever we do since morning to the evening, practically consume electric energy; either water heating in winter season or toaster in daily routine needs

R. C. Goswami (✉) · S. Gautam
Department of Engineering and Physical Science, Institute of Advanced Research, Gandhinagar, India

H. Joshi
Department of Computer Science, Gujarat University, Ahmedabad, India
e-mail: hdjoshi@gujaratuniversity.ac.in

H. Om
Department of Computer Science and Engineering, Indian Institute of Technology (ISM), Dhanbad, India

S. K. Das et al. (eds.), *Architectural Wireless Networks Solutions and Security Issues*, Lecture Notes in Networks and Systems 196,
https://doi.org/10.1007/978-981-16-0386-0_12

electrical energy. In offices, without electric energy no computer can work for all time. After generation of electric energy, it is transmitted as load may not be near to the power generation center. Thus, transmission should be cost effective as in this activity cost is an important factor. The transmission takes much time also because the coal or raw material needs not to be available in the vicinity of load. There is variation in consumption pattern throughout the year; so there is a need of storage of electric energy. The storage of energy increases its cost.

Unfortunately, energy cannot be stored in large volume. So, the requirement of energy; it will produce. A system that deals with the business of electricity generation along with its transmission and distribution is called a power system. It is one of the most complex and largest human-made systems, which provides a very important service to the human society. As much the air is required for human life, in the same way, the electrical energy is just like an air for industrial system.

The early power grid systems contained number of interconnected synchronous alternate current grids [1]. In these systems, the flow of electric power is in only one direction from an electric service station to consumers. In the first step for power generation where a large number of power plants are involved to generate electrical energy, the majority of them use carbon and uranium-based fuels. In the second step, i.e., power transmission, the generated electricity is sent from power generation centers to remotely located places using the power transmission lines. The third step, i.e., power distribution, it is distributed to end users by reducing voltage. The power consumption of per capita plays a vital role for the growth of country. The electricity consumption of increase when the industrial development is fast and it goes reduced when industrial revolution is either completed or about to complete. The electricity demand is growing quickly according to the International Energy Agency, it is predicted that it will be increased by two-third by 2035[2]. It is committed by the European Union to reduce 20% their energy consumption by 2020. The continuous increasing demand of electricity is putting pressure on old power setup. Apart from this also creating congestion problem and reduce power quality. The current power grid lost reliability due to inefficient monitoring, fault analysis and automation techniques. This put a wrong impact for solutions point of view. Smart grid is the solution of such issues where name itself indicate an intelligent power infrastructure. This technology enhances the world power system to efficient, secure and reliable. These are happen by use of information and communication networks. In smart grid systems, the real-time monitoring is very useful to solve the various difficult situations in power system.

In the nation building, power management plays a vital role. There are various countries involve on the project working to conserve energy because there is rapid increase in energy consumption and the energy is not sufficient to meet the rising demand. In last century, hydroelectric power was the major source of energy is now being replaced by nuclear energy. But nuclear plants have several security issues. The power problem can be managed by the efficient utilization of energy and incorporating efficient method of power management system based on IoT [3]. In the current scenario, the power system is facing different types of challenges from power generation to transmission and distribution, and due to this reason, various technological innovations happen through the digital system [3]. Smart grid is a major step in this direction which is producing large amount of data. The power system can take benefits from the large data such it can find the consumption pattern which is varying throughout the year, i.e., it can be found when the demand is high, low or medium by utilizing the data with the help of various data analysis technique of big data. There are some factors affecting the performance of existing grid.

i. *Supply shortfalls of electricity:* There is an energy gap between what we produce and what we expected to deliver.
ii. *Need of reducing losses:* When the electricity is transmitted there is loss of electricity between the production unit and the consumption unit. If this loss is greater it will create an imbalance between the demand and supply.
iii. *Management of peak demand:* The consumption of electricity is not constant throughout the year. Some part of the year there is less demand and some part have highest demand.
iv. *Increasing demand of electricity:* As the per capita consumption of energy is increasing so the demand of electricity is increasing every year.

Internet of things is simply "a network of objects connected to the Internet that is capable of collecting and exchanging data." It contains two main parts: Internet for connectivity and things mean objects or devices and it plays a vital role in smart grid.

The primary contribution of this chapter explains the power grid is to make the grid more efficient such as reducing overall system losses and losses on transmission systems.

2 Motivation

In a traditional system, it is very difficult to predict the price of electricity due to the uncertain and nonlinear pattern. Big data facilitates to simplify difficult patterns of data and forecasts accurately. There are several automated extraction feature process of advanced learning method machine and deep learning which are able to extract important and hidden pattern of data very effectively. There are some other factors which are also work as driving force are out dated infrastructure, improving system efficiency, big data for better management, balancing supply–demand, stability and safety. The deficiency of global energy directly affects the growth of a nation and also the environment due to greenhouse gases. The requirement of energy is growing exponentially worldwide. Outdated network infrastructure and climate, etc., result inefficient and unstable electric system are the basic reason for innovative thinking in the direction of smart grid.

3 Related Work

In smart grids, the big data techniques can be helpful finding the pattern of energy cost and its consumption. Further, they can also help design the program for predictive supply and demand, a fundamental prerequisite for balance of demand–supply. Various efforts have been done which overview the facility of smart meters and advanced metering infrastructure in smart grids. For example, authors discuss advantages and challenges of integrating smart meters and also focus on smart meters [4]. Sun et al. [5] discussed various smart energy meters such as electricity, gas and heat meters and also discuss benefits and applications. Researchers [6] discuss the structure for smart grids from all perspective such as operation, transmission, distribution and customer perspective from security and privacy point of view. They also discuss about advanced metering infrastructure. Authors in [7] review the communication technologies which are utilizing in smart metering systems and network deployment schemes with reference to regional context. Paper [8] focuses on smart meters functionalities with data analysis point of view such as collection speed, volume of data and complexity. Paper [9] analyzes the IoT-based framework for managing the need of contingencies for the future. The concept of smart cities is the latest one in the growth of IoT technology. Smart grid is latest attraction between the engineers and researches in information technology and electric power transmission areas [10]. Author describes [11] that smart grid is the enhancement of the combination several technologies such as communication technology and electrical information technology. In paper [12], the author focuses on cloud-based IoT structure that combines the need of cloud computing and edge computing. Authors in [13] propose a scheme for lightweight data report which uses mechanism based on the hash tree and data integrity verification for protecting user's privacy and data integrity of data. Paper [5] discusses about the smart energy meters for electricity, heat and gas along with their benefits. Paper [14] studies the use of the smart meters including such as advanced metering infrastructure and communication technology at medium voltage and low voltage levels.

Author in [15] states that the information which is stored and processed is used for tracking and observing. The per capita consumption of energy is too low in some developed countries like India as compared to developed countries. So the per capita consumption of energy is likely will grow in developing countries with growth in economy which put more pressure on energy demand**.**

4 Big Data and Its Characteristics

The big data was defined as early as 2001. It can easily be understood that amount of data which is large in size such that it is difficult to store, manage and process in proper way [16]. It is not an absolute concept but it is relative concept. There is no uniformity among the definition of big data. If we compare big data with traditional dataset big

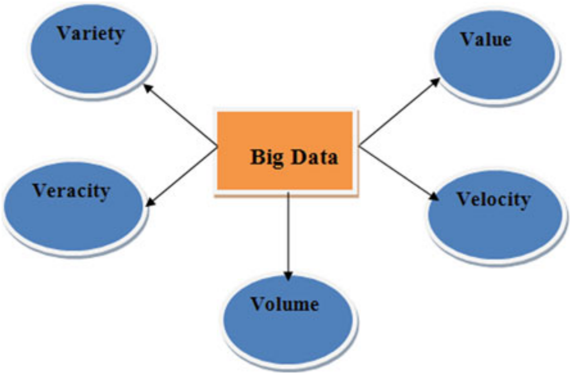


Fig. 1 Big data characteristics

data include a lot of unstructured type of data which required real-time analysis. It also provides a very proper platform for searching new values which help to gain deep knowledge of value available in the data and also gives challenges for properly managing the datasets. The services related to big data of the Internet companies are growing very fast. Real-time big data which is generated on the Internet have several challenges such as evaluation, storage, analysis and transmission [17]. Facebook generates 4 petabytes of data per day. The basic characteristics of big data [18] include *volume, velocity, variety*, *veracity* and *value,* as shown in Fig. 1. The *volume* refers to the amount of data being generated and its understanding is important for making the data-based decisions, and the *velocity* refers to the rate at which the data is produced. It is helpful to understand and design appropriate techniques to process the data. The *variety* can be defined as the structural heterogeneity in a data, which can be of three types: *structured, semi-structured* and *unstructured* data. The *structured* data is organized in table structure such as relational databases and the *semi-structured* data also has an organization, but no table structure. This type of data can more easily be read and manipulated, for example, XML files and Web pages. The unstructured data has no organizing structure and the big data technologies provide different ways to add structure, for example, a combination of images, videos, text files, etc. The *veracity* refers to the trustworthiness of data in terms of its quality and accuracy and the *value* refers to how much data is useful and meaningful.

5 Smart Grid

A smart grid is a technology that permits bi-directional communication between the utility and its customers, and sensing along the transmission lines. In a smart grid, there are two flows: electricity and information, which flow in different directions. The smart grids monitor the customer preferences, their appliance as well as each activity at power plants. So, in near future, the new technology will change the

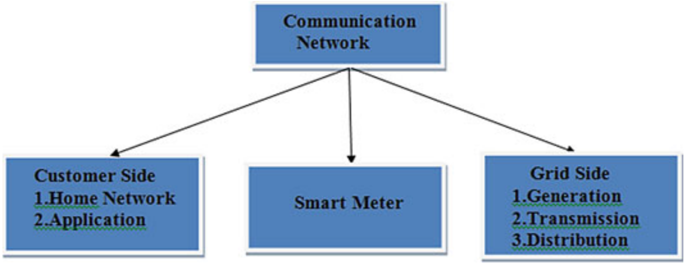


Fig. 2 Smart grid architecture

whole direction of power systems. The bi-direction flow of electricity and information is controlled by the demand. One important aspect of smart grid is distributed generations that play a significant role in new area of power grid.

5.1 Smart Grid Architecture

The smart grid incorporates the sources of renewable energy into the system and manages power consumption, its production, and its flow along with data among the power generating systems, distributing system and consuming centers [19]. The power generating systems consist of energy sources such as wind, nuclear, hydropower and renewable and they directly communicate and coordinate with distributing systems. The power distributing systems comprises a network connecting with the consuming centers with the electricity grid and transmits the data using advanced metering infrastructure. The power-consuming centers comprise the users of electricity residential as well as industrial units. To optimize the services, it is very important to supervise the consumption and production. Figure 2 shows the basic architecture of a typical smart grid.

5.2 Smart Grid Infrastructure

The infrastructure of a smart grid comprises of five main layers that interact with each other and exchange data [19]. These layers include component layer, communication layer, information layer, functional layer and business layer as discussed below [19].

(i) **Component layer**
This layer primarily contains the physical devices which are responsible of getting information, function and communication means from other layers.

(ii) **Communication layer**
This layer employs different techniques and protocols to exchange the data between the components of grid.

(iii) **Information layer**
The information layer contains the details about the data model and communication systems used to exchange the information.

(iv) **Function layer**
The function layer provides the details about the logical functions. The logical functions or applications are independent of the physical architecture.

(v) **Business layer**
This layer provides the details about the business models and regulatory requirements being deployed in the grid.

The various smart grid layers interact with each other to ensure energy management.

5.3 Source of Data in Smart Grid

The smart grid implementation generates large amount of data due to smart meter and sensors which are installed in the smart grid networks. Data is collected from power generation by plants to power consumption by customers. Such data has readings from smart meters, economic situation, and weather condition of a certain region. Smart meter sends the energy consumption after a certain intervals so large number of meters installed generates huge amount of data [20]. Due to this reason, apart from the energy management, the smart grids are required to able to data management which is of high velocity, storage capacity and also data analytics.

6 Architecture of IoT

There is no consensus about the single architecture for IoT, which is agreed unanimously and several researchers have discussed different architectures. The basic architecture of IoT consists of six layers: perception layer, network layer, application layer, transport layer, processing layer and business layer [21].

(i) **Perception layer**
This layer contains the physical devices such as sensors that collects information by sensing about their environment. It identifies the smart objects available in the surrounding and also physical parameters by sensing.

(ii) **Network layer**
This layer is basically used to connect the smart things, network devices, servers and to transmit and process the sensed data.

(iii) **Application layer**
This layer is used to deliver the application-specific services to users and contains information about various applications that can use the IoT, e.g., smart health, smart homes, etc.

(iv) **Transport layer**
It sends the collected data from sensor devices to processing layer and from processing layer to the perception layer using various networks such as wireless, LAN, RFID, NFC, etc.

(v) **Processing layer**
This layer manages and processes the data coming from transport layer. It provides various services by employing several technologies like databases, cloud computing, etc.

(vi) **Business layer**
This layer manages the whole IoT system, including applications, business and profit models and also the user privacy.

7 Application in Power System

The smart grid provides intelligent transmission and distribution networks to deliver electricity. It improves the electric system's reliability, security and efficiency by using bi-directional communication of consumption data and dynamic optimization of electric-system operations, maintenance and planning. The smart grid consists of various resources, applications and technologies. The resources comprise the devices, which affect the energy supply, load, or grid conditions, information about infrastructure, networks, end-user systems and distributed energy resources. The applications comprise the operational strategies that create the value using resources. The technologies comprise elements of smart grid to provide resources and applications [22].

In this system, customers will know how much electricity he/she used at any moment of time without waiting the monthly statement which is generated at end of the month and this is possible by smart meters. In other words, the customer can analyze the consumption pattern and by using it they can reduce their monthly bill. So we can say customers have control over their electricity which will overall reduce the electricity consumption. It is very expensive to build power plant to supply occasional high demand of electricity. Apart from this, some economical approach can also be applied to reduced electricity such as automatic turn off devices which are no longer in use, these are possible only by using the smart grid and its related technologies. The smart grids have billions of smart objects or things such as sensors, smart meters and smart appliances.

All the implementation is sensor-based. Sensor detects the environment and takes the decision. For example, electricity consumption is changed according to season. In each season, humidity and temperature are the major variable factor when humidity and temperature increase the air conditioner consume more electricity but when humidity reduced and temperature increased the air conditioner consume less electricity [22].

The important functionality of smart grid is the efficient transmission of electricity quicker restoration of electricity after the disturbances power, reduced operation and management of cost for utilities and finally lower power cost for consumers.

In smart grid, each step of power network contains intelligence and bi-directional communication capabilities to monitor and control of the grid from anywhere. For example, the smart homes have all the devices smart including smart meters. The power generators, electric transmission and distribution networks contain sensors and actuators. Apart from this, the smart grid contains a balance in real time between the energy generation and its consumption by monitoring and controlling the power chain. Factors associated with the performance of existing grid are:

i. Increasing demand of electricity
ii. Management of electricity at peak duration
iii. Shortfall in supply of electricity
iv. Integrating of renewable energy systems.

In smart grid system, the real-time data is collected after that transmitted through the smart meters that gives information related to energy to customer as well as utility company. Smart grid requires reliable as well as real-time monitoring to provide the quick solution to users at the time of natural turbulence. So, smart sensing and monitoring potential are required. There is an important term, called as advanced metering infrastructure, that consists other things such as set of smart meters, communication module, data collectors, LAN, WAN, network management system, outage management system, meter data management systems, etc. [23]. So, it is a collective term used to explain the complete infrastructure from smart meter to two-directional communication to control center equipment and other applications that help gathering and transferring the energy usage information in regular activity.

7.1 *Smart Grid Communication*

The smart grid communication deals with communication among various components of the smart grid, which has the following components.

i. Smart appliance at home
ii. Smart meter
iii. Gateways
iv. Data aggregator units
v. Meter data management system.

The smart meters are installed in smart homes including appliances. The data aggregator units (DAUs) basically collect the data from different appliances of the homes, which along with their smart meters are connected to DAUs and then meter data management system (MDMS). Figure 3 shows the flow of data among various components.

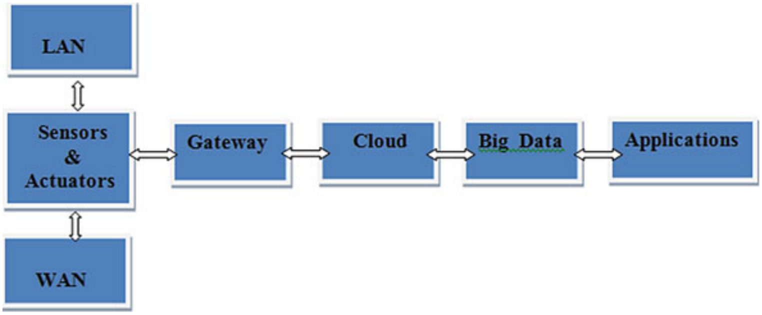


Fig. 3 Data flow in various components

7.1.1 Smart Meters and Gateways

The gateway connects very closely to the smart meters the smart meters forward the energy consumption information from the appliances located at homes to the gateways and also forward the billing amount and the control information from the gateways to the home appliances these, i.e., gateways act as a link between the smart meters and the data aggregation units. So, we can say that gateway helps in two-way communication.

The smart objects facilitate good management and an efficient balancing of demand and consumption. In this context, smart meter is a key technological element of IoT. A smart meter gives facility of reading, recording and monitoring of user consumption in systematic manner such as at certain time period and daily basis. Smart meters also provide bi-directional real-time communication between the utility system and meters, which allow the utility to collect interval data, time-based demand data, outage management, interruption/ restoration, monitoring the service quality, network analysis of distribution, distribution planning, peak time demand, reduction of demand, customer billing, etc. In smart meters, the microcontroller works as a fundamental part where the majority of data processing occurs [24]. So, the activity which is controlled by the microcontroller in smart meters consists of power management, reading of smartcard, calculations, detection of tamper and data management.

7.1.2 Data Aggregator Units and Meter Data Management System

The data aggregator basically perform the task such as energy request of a certain area, forward energy consumption information to the central co-coordinator which is meter data management system and also maintain a buffer to queue the energy consumption information of the customers. Basically, the meter data management system has the following functionality.

i. Acts as a central coordinator for smart grid communications
ii. It is handled by the energy service providers
iii. Part of operation center
iv. Decide the price per unit energy one needs pay.

7.1.3 Reliability of Power and Smart Meter

Power reliability is associated with the electric interruptions that occur in a power system. A power system is said to be reliable when the consumers are getting power continuous without any interruptions. There are various factors which are affecting the reliability of power system. Nearly 80% reliability of power-related issues is happening in the distribution part of power system [25]. In traditional system, the details of interruptions were difficult to know, but in the current system to find out issue of power system is very easy due to data analysis and measuring the seriousness of interruption [26].

7.2 Requirement of IoT and Big Data in Smart Grid System

The primary purpose of big data analysis in smart grid is to find the trends of electricity consumption and its cost. It helps to develop a program for predictive demand and supply, an important basic factor for balancing demand and supply. Big data and smart grid studied for detection of anomaly in power system and proper positioning of computing unit for communicating data to smart grid. Author in [27] has studied about price predictions.

Smart grid-based IoT technology is generally used in long distance communication of data from the utility center to end user. For this, we require a better wireless technology in order to reduce the complexity related to long range transfer. The information that flows has two parts.

i. Flow between all smart meters that are connected in a hub with IoT-enabled devices
ii. Flow between the control centers and utility providers.

Figure 4 shows the smart grid and IoT layer relationship.

8 Big Data Analytics

In current industrial phase, the data analytics play a crucial role that incorporates the information and communication technology. A new additional phase is added to the conventional network by using smart meter.

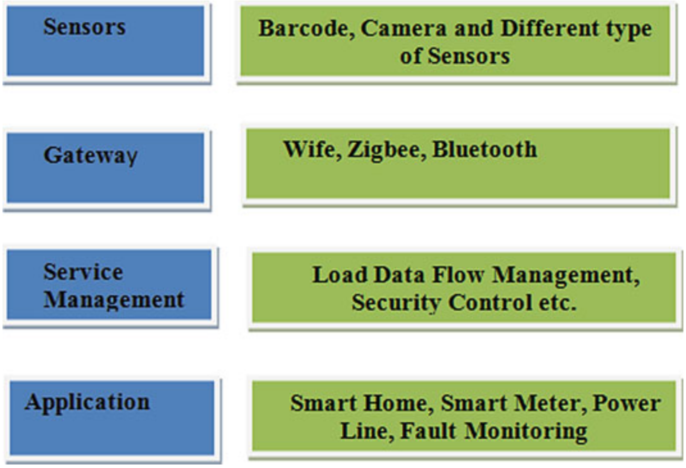

Sensors
Barcode, Camera and Different type of Sensors
Gateway
Wife, Zigbee, Bluetooth
Service Management
Load Data Flow Management, Security Control etc.
Application
Smart Home, Smart Meter, Power Line, Fault Monitoring

Fig. 4 IoT layers and smart grid

The industry generally treats data in the order of terabytes or petabyte and beyond as big data. Majority of this data consists of information generated by Web-based sources such as social networks like Facebook and video sharing sites like YouTube. In big data, this is also known as unstructured data, that is, not in fixed format such as spreadsheet or the kind that can be easily stored in a traditional system [28]. In other words collectively the volume of data being generated has come to be termed big data and analytics that include a wide range of faculties from basic data mining to advanced machine learning is known as big data analytics, i.e., performing analysis on large-scale datasets in order of tens or hundreds of gigabytes to petabytes can be termed big data analytics.

The main aim of using big data analysis is to draw useful information or value from the data. This value can be extracted from the collected data by performing analysis on the data as shown in Fig. 5 [29]. Consumers and utilities may make informed decisions based on the resulting value.

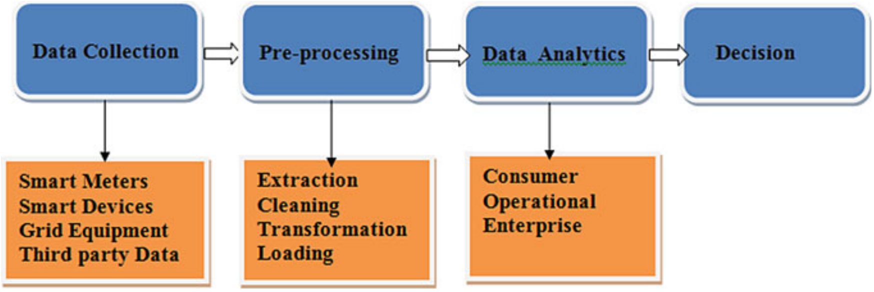


Fig. 5 Extracting value using big data analytics

(i) **Data collection**
In this phase collection of data take place on which analytics will be performed. The collection of data can be from various sources such as smart meters and smart devices.

(ii) **Data pre-processing**
The data collected from various sources are extracted. These data values are then cleaned to remove error values [30]. After that the data is transformed to the repository's format from where data are loaded into a repository. After this phase, various data analytics techniques are applied on the pre-processed data to extract valuable information based on which some action/decisions can be made.

(iii) **Analytics**

a. Consumer analytics
b. Operational analytics
c. Enterprise analytic.

9 Smart Grid Based on IoT Security, Requirements, Issues and Challenges

The smart grid involves thousands of devices; so, the security is one of the important factors. Here, we investigate security issue, challenges and requirement [31].

9.1 Smart Grid Security Requirements

The security services considered for the Internet of things-based smart grid are as follows.

(i) **Authentication**
It is the capability to verify the identity of any communicating device or object in the smart grid system. For example, to generate the bill of a particular customer, it necessary to authenticate smart meter [32].

(ii) **Data integrity**
It refers to that the data is not modified by anyone.

(iii) **Confidentiality**
It ensures that details available only to the main recipients. For example, the consumption of a user's will be known only to the energy provider and smart grid operator apart from these no third party should know the information [32].

(iv) **Availability**
It ensures that resources and information are only available to those who need them.

9.2 Security Challenges in Smart Grid

The smart grid extensive infrastructure provides significantly better awareness along with smooth command and control, which is important to demand–supply, electricity storage and also to automate distribution and transportation. Just like any other complex system, the smart grid also has some vulnerabilities and challenges, which arise due to the integration of several technologies. Some challenges are listed as follows [32].

(i) **Trust**
Trust can be defined as confidence in a system that a user accesses correct data produced by the desired device and also believes that the data is not modified.

(ii) **Communication and device security**
The smart grids use the Internet technologies without adequate security and reliability planning. The traditional communication contains the devices installed with physical access controls in locked buildings, and the two-way meters deployed can be accessed by consumers and adversaries. So, there should be automatic meter reading (AMR) environments in such cases.

(iii) **User Privacy**
The major security concerns are availability, integrity and confidentiality. The smart grid incorporates smart metering and load management. The user privacy is an important issue. The pattern of electricity consumption can lead to not only the amount of energy a customer uses but it also indicates when he is at home or work. It can also be analyzed to some extent that when one is at home whether he is sleeping or watching television. Further, it can be found what types of appliances and devices are present in home.

(iv) **Scalability**
In smart grid, it is difficult to make security solutions scalable, for example, authentication management due to very huge number of objects and smart devices and broad area coverage.

(v) **Constrained Resources**
Smart devices are resource-constrained device so it is very difficult to implement the classical security solution. So, security solution for these devices should be specially designed.

(vi) **Heterogeneity**
The inconsistency in devices of smart grids creates problems in having end-to-end communication secure, which is a major difficult issue.

9.3 Smart Grid Security Issues

In the environment of smart grid, several millions of devices and its related infrastructure are connected to each other's which expose the smart grid to security vulnerabilities. On the other hand, the cloud computing enables applications to be virtualized, i.e., sharing the cloud platform with millions of users creates another set security concerns. Apart from this smart grid must be highly scalable and accessible in real-time application where lower latency is a huge challenge. The important security issues are given as follows.

(i) **Data sensing and communications security**
In smart grid, the sensing and communication of data may be target of intrusion, malicious activity or any other kinds of threats. The used public network could be vulnerability for various threats. In other words, we can say that when the data is traveled between two smart object then intruder can modify the exchanged data, which creates a major problem to service providers as well as customers.

(ii) **Authorization**
It gives permission that only an authenticated person or object allowed to carry out some tasks or access some resources.

(iii) **Privacy Issue**
The smart meters in day-to-day life collect and transfer huge amount of data to consumers, and the service providers contain the private user information that can be sometimes used against the consumer, devices, etc. But, in smart grid, there are some intelligent devices involved in management of both demand and supply of the electricity. These devices may act as entry point of attack. So security can be considered as severe issue in smart grid. Authors in [33] describe the techniques of privacy preserving, focus mainly on data aggregation. Due to this, authors decide to focus techniques of data aggregation which are used for preserving privacy of smart grids. Based on analysis, major challenges such as secure cryptographic algorithm, limitation related with hardware and signal processing without considering some security requirements such as authentication, integrity, access control and privacy smart grids cannot be broadly deployed.

10 Future Scope

The researchers consider smart grid is next generation grid, which supply bi-directional flow of information and electricity, with more power, reliability, security, and efficiency of generation, transmission and distribution of power. As smart grid continues to develop, the realization of a reliable and stable system is necessary. Further, it also provides real-time information, lower cost of operation and electricity. The scope of future scope is the researcher will overcome the failure of smart grid the save the time.

11 Conclusion

The importance of electricity is a well-known thing. The demand of electricity is increasing with rapid speed. Smart grid is an initiative in this direction which plays a significant role from generation, transmission and distribution of power. Internet of things is the developing technology which facilitates the communication between two objects or devices irrespective of location across the globe. In the large-scale deployment and adoption of the smart grid, it takes maximum use of IoT. Thus, we can say that the power management system based on IoT is an effective method for power management. Big data also plays an important role to understand various aspect of power consumption, seasonal trend of consumption of electricity consumption and self-healing. Thus, in short, we can say that in power management through smart grid IoT and big data technology plays a key role.

References

1. Collier SE (2017) The emerging enernet: convergence of the smart grid with the Internet of Things. IEEE Ind Appl Mag 2:12–16
2. How will global energy markets evolve to 2035? (2013) World energy outlook factsheet. International Energy Agency
3. Gielen D, Boshell F, Saygin D, Bazilian MD, Wagner N, Gorini R (2019) The role of renewable energy in the global energy transformation. Energy Strategy Rev 24:38–50
4. Depuru SSSR, Wang L, Devabhaktuni V (2011) Smart meters for power grid: challenges, issues, advantages and status. Renew Sustain Energy Rev 15(6):2736–2742
5. Sun Q et al (2016) A comprehensive review of smart energy meters in intelligent energy networks. IEEE Internet Things J 3(4):464–479
6. Al-Turjman F (2019) 5G-enabled devices and smart-spaces in social-IoT: an overview. Elsevier Future Gener Comput Syst 92(1):732–744
7. Pillai R, Thukral H (2017) Next generation smart metering: IP metering. CIRED Open Access Proc J 2017(1):2827–2829
8. Pawar S, Momin BF (2017) Smart electricity meter data analytics: a brief review. In: IEEE region 10 symposium (TENSYMP), pp 1–5
9. Ciavarella S, Joo JY, Silvestri S (2016) Managing contingencies in smart grids via the internet of things. IEEE Trans Smart Grid
10. Fouda MM, Fadlullah ZM, Kato N, Lu R, Shen X (2011) A lightweight message authentication scheme for smart grid communications. IEEE Trans Smart Grid 2(4):675–685
11. Locke G, Gallagher PD (2010) NIST framework and roadmap for smart grid interoperability standards, release 1.0. National Inst Stand Technol 33
12. Meloni A, Pegoraro PA, Atzori L, Benigni A, Sulis S (2018) Cloud-based IoT solution for state estimation in smart grids: exploiting virtualization and edge-intelligence technologies. Comput Netw 130:156–165
13. Bao H, Chen L (2016) A lightweight privacy-preserving scheme with data integrity for smart grid communications. Concurrency Comput Pract Exp 28(4):1094–1110
14. Kabalci E (2016) Emerging smart metering trends and integration at MV-LV level. In: International smart grid workshop and certificate program (ISGWCP), pp 1–9
15. Binh HTT, Hanh NT, Nghia ND, Dey N (2020) Metaheuristics for maximization of obstacles constrained area coverage in heterogeneous wireless sensor networks. Appl Soft Comput 86. ASOC 105939

16. Kaisler S, Amnour F, Alberto J (2012) Big data: issues and challenges moving forward. In: 46th IEEE international conference on system science, Wailea, Maui, HI, USA
17. Jat DS, Bishnoi LC, Nambahu S (2018) An intelligent wireless QoS technology for big data video delivery in WLAN. Int J Ambient Comput Intell (IJACI) 9(4):1–14
18. Hassan MK, El Desouky AI, Elghamrawy SM, Sarhan AM (2019) Big data challenges and opportunities in healthcare informatics and smart hospitals. In: Security in smart cities: models, applications, and challenges. Springer, pp 3–26
19. Daki H, El Hannani A, Aqqal A, Haidine A, Dahbi A (2017) Big Data management in smart grid: concepts, requirements and implementation. J Big Data 4(1):1–19
20. Avancini DB, Rodrigues JJ, Martins SG, Rabêlo RA, Al-Muhtadi J, Solic P (2019) Energy meters evolution in smart grids: a review. J Clean Prod 217:702–715
21. Sethi P, Sarangi SR (2017) Internet of things: architectures, protocols, and applications. J Electr Comput Eng
22. Ghosal A, Conti M (2019) Key management systems for smart grid advanced metering infrastructure: a survey. IEEE Commun Surv Tutorials 21(3):2831–2848
23. Ferrag MA, Ahmim A (2017) Security solutions and applied cryptography in smart grid communications. IGI Global, USA
24. Khan MW, Zeeshan M (2019) QoS-based dynamic channel selection algorithm for cognitive radio based smart grid communication network. Ad Hoc Netw 87:61–75
25. Billinton R, Allan RN (1996) Reliability evaluation of power systems. Springer, Boston
26. Kuhi K, Korbe K, Koppel O, Palu I (2016) Calculating power distribution system reliability indexes from Smart Meter data. In: 2016 IEEE international energy conference (ENERGYCON), pp 1–5
27. Statistical office of the European Union (2016) Shedding light on energy in the EU—a guided tour of energy statistics. Technical report
28. Ahmed M, Choudhury S, Al-Turjman F (2019) Big data analytics for intelligent Internet of Things. In: Artificial intelligence in IoT. Springer, pp 107–127
29. Munshi AA, Yasser ARM (2017) Big data framework for analytics in smart grids. Electric Power Syst Res 151:369–380
30. Hao R, Yang H, Zhou Z (2019) Driving behaviour evaluation model base on big data from internet of vehicles. Int J Ambient Comput Intell (IJACI) 10(4):78–95
31. Bekara C (2014) Security issues and challenges for the IoT-based smart grid. In: FNC/MobiSPC, pp 532–537
32. Bagri D, Rathore SK (2019) A review paper on existing security algorithm to find out issues for achieving the confidentiality and integrity of metering data in smart grid. Available at SSRN 3446660
33. Erkin Z, Troncoso-pastoriza JR, Lagendijk RL, Perez-Gonzalez F (2013) Privacy-preserving data aggregation in smart metering systems: an overview. IEEE Signal Process Mag 30(2):75–86

Analysis of Network Parameters for Network Lifetime in WSN: A Fuzzy Quadratic Programming Approach

Manoj Kumar Mandal, Arun Prasad Burnwal, Abhishek Kumar, Divya Mishra, and Nikhil Saxena

Abstract Wireless sensor network (WSN) is a collection of sensor nodes that are attached with base station (BS) and sink node to achieve a specific purpose. The main purpose of the WSN is sensing environmental parameters such as energy, temperature, and humidity. There are several parameters of the WSN that changes time to time and frequently based on the operation. Each sensor node contains limited capacity of battery that is insufficient during any operation and fails to send the data packet to the BS. So, there is need of some modeling using some intelligent technique. In this paper, a fuzzy quadratic programming (FQP) is used to optimize network parameters efficiently. FQP is the fusion of fuzzy logic and quadratic programming. Fuzzy logic is a multi-values logic which is used to reduce uncertainty and estimate imprecise parameters efficiently. Quadratic programming is a nonlinear programming based on second order of mathematical polynomial for reducing the main objective. The combination of both helps to analyze conflicting network parameters and decide the optimal objective value along with constraints. The proposed method is validated in LINGO optimization software in terms of several rounds to predict the optimal solution.

Keywords Wireless sensor network · Fuzzy set theory · Quadratic programming · Nonlinear optimization · Membership value

M. K. Mandal (✉)
Department of Mathematics, Jharkhand Rai University, Ranchi 835222, India

A. P. Burnwal
Department of Mathematics, GGSESTC, Bokaro, Jharkhand 827013, India

A. Kumar
Department of Electronics and Communication Engineering, Swami Vivekananda Subharti University, Meerut 250005, India

D. Mishra
Department of Computer Science Engineering, Swami Vivekananda Subharti University, Meerut 250005, India

N. Saxena
University of Cincinnati, Cincinnati, OH 45221, USA

S. K. Das et al. (eds.), *Architectural Wireless Networks Solutions and Security Issues*, Lecture Notes in Networks and Systems 196,
https://doi.org/10.1007/978-981-16-0386-0_13

1 Introduction

Wireless sensor network (WSN) is a collection of sensor nodes that provide the services to the users and customer with the help of sensing system [1–3]. It consists of base station (BS) and some sink node. The purpose of sink node is to receive data and information from the source node. The purpose of BS to collection data and information from the multiple sensor nodes and analyze it for performing operation. Figure 1 shows WSN network that consists of several information such as sensor nodes, BS, user, and computer. In this figure, within range all sensor nodes are deployed, this range is connected with BS for analyzing and validating data packets. BS station is further connected with a computer that helps to store sensed information for predicting and analyzing. This computer is connected with user that work as an administrator that works fully with sensor nodes and BS for managing several applications as follows.

(a) Military application
(b) Entertainment application
(c) Business and marketing
(d) Educational
(e) Disaster management.

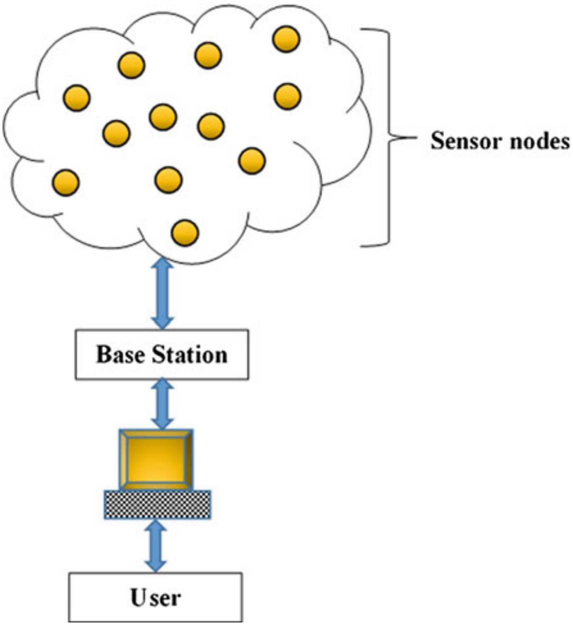


Fig. 1 Wireless sensor network

The WSN have several applications in terms of efficient working principles, but it has also some limitations such as battery issue of the sensor nodes, low communication ranges within different hop nodes, speed between sensor nodes, etc. These several types limitations are arise due to its variation of network parameters. In this paper, a fuzzy quadratic programming (FQP) [4] is used to optimize network parameters efficiently. FQP is the fusion of fuzzy logic and quadratic programming. Fuzzy logic is a soft computing technique which produces soft results that indicate approximation results [5]. It is a multi-values logic which is used to reduce uncertainty and estimate imprecise parameters efficiently. Quadratic programming is a nonlinear programming based on second order of mathematical polynomial for reducing the main objective. The combination of both helps to analyze conflicting network parameters and decide the optimal objective value along with constraints.

The rest of the paper is divided as follows. Section 2 highlights some existing work related to some routing techniques. Section 3 shows details analysis parts of the proposed method. Section 4 shows simulation and analysis part of the proposed method. Finally, in Sect. 5, conclude the paper.

2 Related Works

In several years, various works are proposed in the context of ad-hoc network and WSN. Some works are discussed in this section as follows. Mandhare et al. [6] designed a meta-heuristic-based routing protocol for MANET. The purpose of this routing is to enhance QoS of the network. The proposed method reduces the issue of non-deterministic NP hard issue. The key technique is used in this method that is cuckoo search method. The method is used in the AODV routing technique with the help of RREQ and RREP packets for finding the shortest path. Finally, it compares with some nature-inspired techniques such as PSO, ACO, and simple AODV, enhances the matrices scalability and mobility, and reduces congestion of the network. Phoemphon et al. [7] proposed a hybrid method for WSN using PSO based. The proposed method is based on localization technique. In this work, two basic parameters are considered such as hop-count and distance for evaluating localization system. In WSN, all parameters are based on approximation technique. The proposed method uses this localization system using PSO technique. The basic parameters of the PSO and its tuning parameters considered with the help of network parameters, and finally, it helps to enhance the network lifetime. Tripathi and Das [8] proposed five input parameters based intelligence routing using multiple criteria of ad-hoc network. This is based on soft set method which mixed by extended fuzzy set, i.e., intuitionistic fuzzy set and two techniques of the multi-criteria decision system. Each input parameter is mapped into the soft set in terms of three elements such true membership value, false membership value, and between both which is known as hesitation membership value. Finally, it helps to resolve the uncertainty of the network efficient and derive optimal route of the network. Sun et al. [9] proposed an optimization technique for handling attack in WSN. This is based on PSO optimization along with multiple objectives of the network. It minimizes the energy consumption of the nodes and maximizes the load balancing of the network. The natures of both objectives are nonlinear formulation, and nature of the PSO optimization is binary. Finally, it helps to dynamically maintain all the objectives and their constraints, maintains the convergence of the system, and finally helps to enhance the network lifetime. Cao et al. [10] designed a PSO based distributed optimization technique for WSN. It is based on the deployment system of the network. In this system, nodes are heterogeneous types and divided into two types as relay nodes and simple nodes. This system helps to maximize coverage of the network and maximize the network lifetime with the help of relay nodes. In this paper, nodes are deployed in the 3D environment system, and finally, it helps to reduce time of the passing message and cost of the network. Yang et al. [11] designed an intelligent system for transportation system in wireless network. This is based on an existing transportation system based on process structured system. Finally, it helps to enhance network capabilities and services of the network. It also helps user function and usages in the network and network metrics properly to maintain the network. Loganathan and Subbiah [12] designed an energy-based communication system for device-to-device communication in the network. It is based on multi-criteria decision-making system where multiple criteria are involved for integrating the network metrics efficiently. Finally, it helps to enhance the network lifetime and helps in communication system. Hu et al. [13] designed an algorithm for WSN for cooperative maintenance of the nodes using PSO. This method is based on multiple sink nodes of the network for recovery of the route. This routing technique is used to reduce overhead of the network that rose in the communication system by using two basic parameters such as delay and overhead of the network. The tuning parameters of the network help to enhance the global optimization of the PSO, so that accuracy convergence of the network is increases. Elhabyan and Yagoub [14] designed a PSO based optimization technique for WSN. This is based on clustering technique in the network. In this work, the nature-inspired technique PSO mixed with linear programming boosts cluster head for identifying its work within the network. This cluster head helps to collect information from all the sensor nodes and send it to the base station for analyzing and aggregating information to the source to the destination node. Finally, it outperforms the network metrics in terms of delivery ration and scalability. Hayes and Ali [15] proposed a routing protocol named as RASeR for mobile sensor networks. In this paper, topology changing issue solves with the help of global time division multiple accesses using fixed nodes. Singh et al. [16] designed a distributed routing algorithm for WANET. The proposed work is based on two existing routing protocols such as DSDV and AODV where performance of these protocols is analyzed based on pause time of the nodes mobility. Chandrakar [17] designed an authentication system for the users in wireless network. This is basically based on healthcare system and used for medical purpose. This proposal is used for sensing patient body information and sending it to the doctor for treatment and diagnosis purpose. It also helps in user authentication, privacy, and data security purposes, so that efficient result comes from the diagnosis system. Jat et al. [18] designed an intelligent technique for QoS in WLAN. This proposal is based on video delivery system. This is based on multimedia application for video data processing and analyzing. The data analyzed here is based on real-time data that generated by the Internet. It also helps in video data transmission, storage, evaluating, and broadcasting. Sattari et al. [19] used an ACO inspired algorithm for routing in VANET. In this work, several traffics are controlled and managed for finding feasible paths of the network. It is also used for reliable transmission. Finally, a cellular-ant-based algorithm is designed to find an optimal solution. Abdallah [20] proposed a smart partial flooding routing algorithm for ad-hoc network. In this paper, two 3D geographical routing algorithms are used for maintaining two things like overhead and flooding in the network. In WSN, data is gathered from multiple homogeneous or heterogeneous sources because real-life data is connected with different IoT, IoV, or cloud environment. So, it is difficult to keep the natures of the data in same structure. Information retrieval [21] is very important part in modern research areas which indicates collecting information that are stored in unstructured form based on multiple local languages and process it in particular pattern after observing. Hao et al. [22] designed an evaluation system for big data analysis. This data is based on IoV where it means Internet of vehicle. This proposal is based on K-means algorithm that is used here as a clustering. In this work, different behaviors of the driving are involved for controlling vehicle. Finally, it helps in reducing fuel consumption and helps in transportation globally. Singh et al. [23] designed a method for pattern adopting in ad-hoc network. The proposed method is based on security system. It illustrated different attacks of the ad-hoc in terms of passive and active attacks. It also highlights process to overcome traditional issues and constraints. In the above literature, all authors did not analyze the network parameters with respect to network lifetime. In this paper, this is proposed based on mathematical optimization.

3 Proposed Method

In this section, the proposed method is illustrated briefly with the help of mathematical formulation and analysis. Let N is the set of wireless sensor nodes that are deployed in the workable area as shown in Eq. (1).

$$N = \{n_i\} \quad \text{where} \quad i = 1, 2, 3, \ldots, k \tag{1}$$

In this equation, k varies based on user requirement and deployment. WSN consists of several parameters for transmitting data packet such as “energy consumption,” “hop-count,” “distance,” “delay,” and “overhead,”. In this paper, there parameters are considered as “energy consumption,” “delay,” and “overhead.” These parameters affect the network lifetime of the WSN. If energy consumption increases, then network lifetime decreases; if delay increases, then network lifetime decreases, and if overhead increases, then network lifetime decreases. Hence, nature of these three parameters is contradictory with network lifetime. Table 1 shows several assumptions of the network with parameters that help to formulate the network lifetime statistic.

Table 1 Assumptions and statistics of the network

Node (N)	Energy consumption (E)	Delay (D)	Overhead (O)
n_1	e_1	d_1	o_1
n_2	e_2	d_2	o_2
n_3	e_3	d_3	o_3
…	…	…	…
…	…	…	…
…	…	…	…
…	…	…	…
…	…	…	…
n_k	e_k	d_k	o_k

Table 2 Dataset for round 1 where number of node, i.e., k_1 is 5

Energy consumption	Delay	Overhead
$e_1 = 5$	$d_1 = 2$	$o_1 = 1$
$e_2 = 1$	$d_2 = 4$	$o_2 = 3$
$e_3 = 3$	$d_3 = 2$	$o_3 = 4$
$e_4 = 1$	$d_4 = 2$	$o_4 = 4$
$e_5 = 3$	$d_5 = 2$	$o_5 = 2$

In this paper, nature of the considered parameters is conflicting with network lifetime, so here objective is to minimize these parameters based on satisfied constraints. So, objective function and their constraints are shown in Eq. (2). In this equation, x_1, x_2, and x_3 are decision variables for three parameters such as energy consumption, delay, and overhead. Summation of three needs to be minimized, and its constraints contain considered parameters and statistic of the network. The symbol 'k' is the variation of nodes that vary several time based on the deployment as k_1, k_2, k_3, …, k_n. Here, difference of each 'k' varies by five nodes. So, correspondence equation for different rounds is shown in Eqs. (3)–(6). The correspondence datasets are given in Tables 2, 3, 4, 5, and 6.

Table 3 Dataset for round 1 where number of node, i.e., k_2 is 10

Energy consumption	Delay	Overhead
$e_1 = 11$	$d_1 = 4$	$o_1 = 2$
$e_2 = 10$	$d_2 = 4$	$o_2 = 5$
$e_3 = 5$	$d_3 = 8$	$o_3 = 4$
$e_4 = 6$	$d_4 = 2$	$o_4 = 8$
$e_5 = 3$	$d_5 = 7$	$o_5 = 2$

Table 4 Dataset for round 1 where number of node, i.e., k_3 is 15

Energy consumption	Delay	Overhead
$e_1 = 10$	$d_1 = 7$	$o_1 = 8$
$e_2 = 16$	$d_2 = 2$	$o_2 = 4$
$e_3 = 4$	$d_3 = 5$	$o_3 = 6$
$e_4 = 9$	$d_4 = 6$	$o_4 = 2$
$e_5 = 11$	$d_5 = 6$	$o_5 = 5$

Table 5 Dataset for round 1 where number of node, i.e., k_4 is 20

Energy consumption	Delay	Overhead
$e_1 = 12$	$d_1 = 11$	$o_1 = 9$
$e_2 = 11$	$d_2 = 5$	$o_2 = 3$
$e_3 = 9$	$d_3 = 8$	$o_3 = 7$
$e_4 = 6$	$d_4 = 2$	$o_4 = 8$
$e_5 = 3$	$d_5 = 7$	$o_5 = 2$

Table 6 Dataset for round 1 where number of node, i.e., k_5 is 25

Energy consumption	Delay	Overhead
$e_1 = 14$	$d_1 = 11$	$o_1 = 10$
$e_2 = 10$	$d_2 = 14$	$o_2 = 15$
$e_3 = 5$	$d_3 = 8$	$o_3 = 4$
$e_4 = 6$	$d_4 = 2$	$o_4 = 8$
$e_5 = 7$	$d_5 = 7$	$o_5 = 3$

$$\begin{aligned}
&\text{Minimize :} && Z_1 = x_1 + x_2 + x_3 \\
&\text{Subject to constraints :} && e_1x_1 + d_1x_2 + o_1x_3 \geq k_1 \\
& && e_2x_1 + d_2x_2 + o_2x_3 \geq k_1 \\
& && e_3x_1 + d_3x_2 + o_3x_3 \geq k_1 \\
& && e_4x_1 + d_4x_2 + o_4x_3 \geq k_1 \\
& && e_5x_1 + d_5x_2 + o_5x_3 \geq k_1
\end{aligned} \tag{2}$$

$$\begin{aligned}
&\text{Minimize :} && Z_2 = x_1 + x_2 + x_3 \\
&\text{Subject to constraints :} && e_1x_1 + d_1x_2 + o_1x_3 \geq k_2 \\
& && e_2x_1 + d_2x_2 + o_2x_3 \geq k_2 \\
& && e_3x_1 + d_3x_2 + o_3x_3 \geq k_2 \\
& && e_4x_1 + d_4x_2 + o_4x_3 \geq k_2 \\
& && e_5x_1 + d_5x_2 + o_5x_3 \geq k_2
\end{aligned} \tag{3}$$

$$\begin{aligned} &\text{Minimize :} && Z_3 = x_1 + x_2 + x_3 \\ &\text{Subject to constraints :} && e_1x_1 + d_1x_2 + o_1x_3 \geq k_3 \\ &&& e_2x_1 + d_2x_2 + o_2x_3 \geq k_3 \\ &&& e_3x_1 + d_3x_2 + o_3x_3 \geq k_3 \\ &&& e_4x_1 + d_4x_2 + o_4x_3 \geq k_3 \\ &&& e_5x_1 + d_5x_2 + o_5x_3 \geq k_3 \end{aligned} \tag{4}$$

$$\begin{aligned} &\text{Minimize :} && Z_4 = x_1 + x_2 + x_3 \\ &\text{Subject to constraints :} && e_1x_1 + d_1x_2 + o_1x_3 \geq k_4 \\ &&& e_2x_1 + d_2x_2 + o_2x_3 \geq k_4 \\ &&& e_3x_1 + d_3x_2 + o_3x_3 \geq k_4 \\ &&& e_4x_1 + d_4x_2 + o_4x_3 \geq k_4 \\ &&& e_5x_1 + d_5x_2 + o_5x_3 \geq k_4 \end{aligned} \tag{5}$$

$$\begin{aligned} &\text{Minimize :} && Z_5 = x_1 + x_2 + x_3 \\ &\text{Subject to constraints :} && e_1x_1 + d_1x_2 + o_1x_3 \geq k_5 \\ &&& e_2x_1 + d_2x_2 + o_2x_3 \geq k_5 \\ &&& e_3x_1 + d_3x_2 + o_3x_3 \geq k_5 \\ &&& e_4x_1 + d_4x_2 + o_4x_3 \geq k_5 \\ &&& e_5x_1 + d_5x_2 + o_5x_3 \geq k_5 \end{aligned} \tag{6}$$

Equations (2)–(6) are in the form of linear where both objective functions and constraints are linear for three parameters such as energy consumption, delay, and overhead for five rounds such as k_1 to k_5, here k_1 to k_5 are varied based on difference of five. Dataset in Tables 2, 3, 4, 5, and 6 has shown several constraint values of the objective values. Quadratic programming is more efficient than linear programming in term of estimation and efficiency. It is an optimization technique like meta-heuristic [24] technique for optimizing several problems based on objective function and their constraints. So, Eqs. (2)–(6) are converted into the form of quadratic and shown in Eqs. (7)–(11).

$$\begin{aligned} &\text{Minimize:} && Z_1 = x_1^2 + x_2^2 + x_3^2 \\ &\text{Subject to constraints:} && e_1x_1 + d_1x_2 + o_1x_3 \geq k_1 \\ &&& e_2x_1 + d_2x_2 + o_2x_3 \geq k_1 \\ &&& e_3x_1 + d_3x_2 + o_3x_3 \geq k_1 \\ &&& e_4x_1 + d_4x_2 + o_4x_3 \geq k_1 \\ &&& e_5x_1 + d_5x_2 + o_5x_3 \geq k_1 \end{aligned} \tag{7}$$

$$\begin{aligned} &\text{Minimize:} && Z_2 = x_1^2 + x_2^2 + x_3^2 \\ &\text{Subject to constraints:} && e_1x_1 + d_1x_2 + o_1x_3 \geq k_2 \\ &&& e_2x_1 + d_2x_2 + o_2x_3 \geq k_2 \\ &&& e_3x_1 + d_3x_2 + o_3x_3 \geq k_2 \\ &&& e_4x_1 + d_4x_2 + o_4x_3 \geq k_2 \\ &&& e_5x_1 + d_5x_2 + o_5x_3 \geq k_2 \end{aligned} \tag{8}$$

$$\begin{aligned} &\text{Minimize:} && Z_3 = x_1^2 + x_2^2 + x_3^2 \\ &\text{Subject to constraints:} && e_1x_1 + d_1x_2 + o_1x_3 \geq k_3 \\ &&& e_2x_1 + d_2x_2 + o_2x_3 \geq k_3 \\ &&& e_3x_1 + d_3x_2 + o_3x_3 \geq k_3 \\ &&& e_4x_1 + d_4x_2 + o_4x_3 \geq k_3 \\ &&& e_5x_1 + d_5x_2 + o_5x_3 \geq k_3 \end{aligned} \tag{9}$$

$$\begin{aligned} &\text{Minimize:} && Z_4 = x_1^2 + x_2^2 + x_3^2 \\ &\text{Subject to constraints:} && e_1x_1 + d_1x_2 + o_1x_3 \geq k_4 \\ &&& e_2x_1 + d_2x_2 + o_2x_3 \geq k_4 \\ &&& e_3x_1 + d_3x_2 + o_3x_3 \geq k_4 \\ &&& e_4x_1 + d_4x_2 + o_4x_3 \geq k_4 \\ &&& e_5x_1 + d_5x_2 + o_5x_3 \geq k_4 \end{aligned} \tag{10}$$

$$\begin{aligned} &\text{Minimize:} && Z_5 = x_1^2 + x_2^2 + x_3^2 \\ &\text{Subject to constraints:} && e_1x_1 + d_1x_2 + o_1x_3 \geq k_5 \\ &&& e_2x_1 + d_2x_2 + o_2x_3 \geq k_5 \\ &&& e_3x_1 + d_3x_2 + o_3x_3 \geq k_5 \\ &&& e_4x_1 + d_4x_2 + o_4x_3 \geq k_5 \\ &&& e_5x_1 + d_5x_2 + o_5x_3 \geq k_5 \end{aligned} \tag{11}$$

Equations (7)–(11) are in the form of quadratic where the nature of the objective functions is nonlinear, but nature of the constraints is in linear. So, the combination of both is nonlinear for three same parameters and same k_i with difference five.

4 Simulation and Analysis

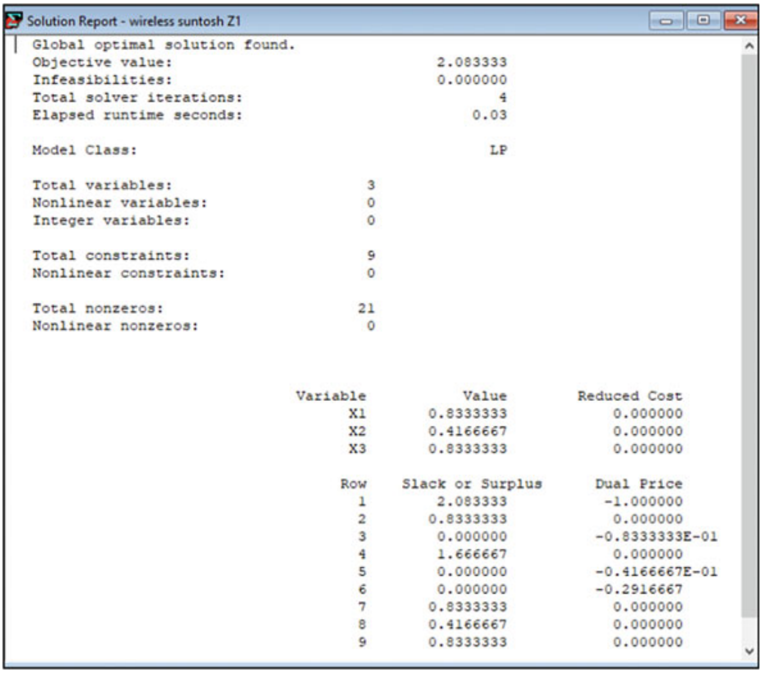
The proposed method is simulated and verified in LINGO optimization software which is used to optimize linear and nonlinear both formulations. In this paper, total linear objective functions used are five, and nonlinear objective functions used are five, so total objective function is ten. Each objective function contains five linear constraints. So, here, total constraints are 5 × 10, i.e., 50, 25 for linear and 25 for quadratic formulation which is nonlinear formulation. In this simulation, minimum node is 5, maximum node is 25, and in each iteration or round, node varies by 5. The objectives are to minimize three network parameters such as energy consumption, delay, and overhead. Total simulation parameters are given in Table 7.

Figures 2, 3, 4, 5, and 6 show linear formulation of three objective function such as (x_1 to x_3) such as energy consumption, delay, and overhead for five rounds (Z_1 to Z_5) such as 5, 10, 15, 20, and 25. The values of different rounds depicted as values of the objective functions and values of the decision variables such as $Z_1 = 2.08333$, $x_1 = 0.833333$, $x_2 = 0.4166667$, and $x_3 = 0.833333$ in round 1 by 5 nodes; $Z_2 = 2.146018$, $x_1 = 0.3982301$, $x_2 = 1.061947$, and $x_3 = 0.6858407$ in round 2 by 10 nodes; $Z_3 = 2.904412$, $x_1 = 0.5147059$, $x_2 = 1.397059$, and $x_3 = 0.9926471$ in round 3 by 15 nodes; $Z_4 = 4.257426$, $x_1 = 0.3465347$, $x_2 = 2.227723$, and $x_3 = 1.683168$ in round 4 by 20 nodes; and $Z_5 = 4.51807$, $x_1 = 1.506024$, $x_2 = 1.35542$, and $x_3 = 1.656627$ in round 5 by 25 nodes. In each round, it is observed that when number of nodes increases, optimized objective value also increases.

Figures 7, 8, 9, 10, and 11 show formulation of quadratic programming based on three objective functions such as (x_1 to x_3) such as energy consumption, delay, and overhead for five rounds (Z_1 to Z_5) such as 5, 10, 15, 20, and 25. The values of different rounds depicted as values of the objective functions and values of the decision variables such as $Z_1 = 1.515152$, $x_1 = 0.7575742$, $x_2 = 0.6060645$, and $x_3 = 0.7575743$ in round 1 by 5 nodes; $Z_2 = 1.689190$, $x_1 = 0.6079219$, $x_2 = 1.013556$, and $x_3 = 0.5406711$ in round 2 by nodes 10; $Z_3 = 2.922078$, $x_1 = 0.7790705$, $x_2 = 0.9739913$, and $x_3 = 1.168960$ in round 3 by nodes 15; $Z_4 = 6.756759$, $x_1 = 1.216193$, $x_2 = 2.027032$, and $x_3 = 1.081099$ in round 4 by nodes 20; $Z_5 = 6.849669$, $x_1 = 1.509525$, $x_2 = 1.354422$, and $x_3 = 1.654251$ in round 5 by nodes 25. In each round, it is observed that when number of nodes increases, optimized objective value also increases; but in quadratic programming, values are more optimal than linear programming.

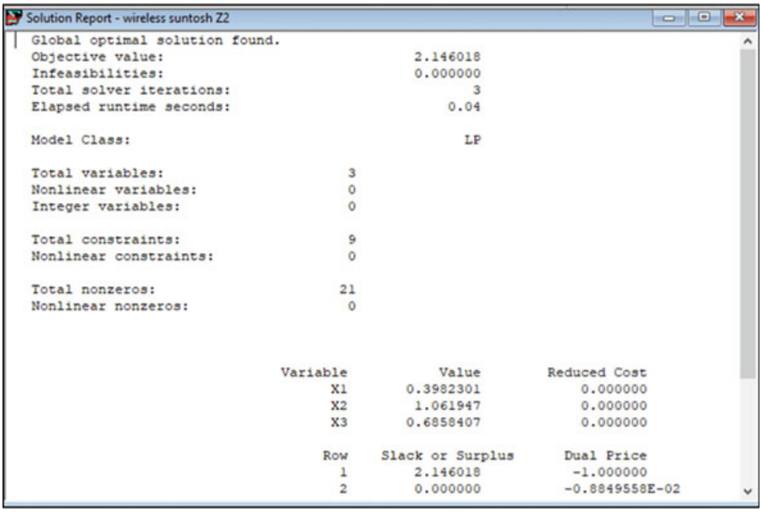
Table 7 Simulation parameters

Parameter	Description
Software	LINGO
Linear objective function	5
Nonlinear objective function	5
Constraints	50
Number of nodes	Minimum 5, maximum 25
Number of rounds	5
Network parameters	Energy consumption, delay, overhead



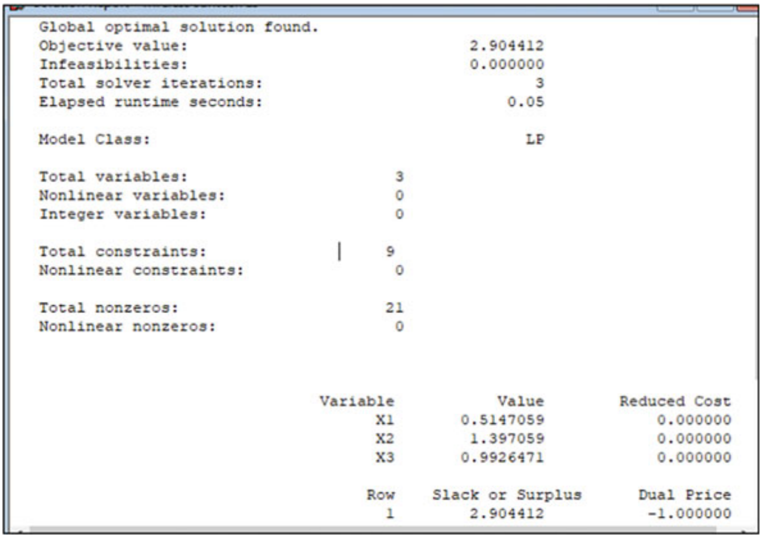
```
Global optimal solution found.
Objective value:                              2.083333
Infeasibilities:                              0.000000
Total solver iterations:                             4
Elapsed runtime seconds:                          0.03

Model Class:                                        LP

Total variables:                     3
Nonlinear variables:                 0
Integer variables:                   0

Total constraints:                   9
Nonlinear constraints:               0

Total nonzeros:                     21
Nonlinear nonzeros:                  0



                              Variable           Value        Reduced Cost
                                    X1       0.8333333            0.000000
                                    X2       0.4166667            0.000000
                                    X3       0.8333333            0.000000

                                   Row    Slack or Surplus      Dual Price
                                     1        2.083333           -1.000000
                                     2       0.8333333            0.000000
                                     3        0.000000       -0.8333333E-01
                                     4        1.666667            0.000000
                                     5        0.000000       -0.4166667E-01
                                     6        0.000000          -0.2916667
                                     7       0.8333333            0.000000
                                     8       0.4166667            0.000000
                                     9       0.8333333            0.000000
```

Fig. 2 Linear formulation based on node 5



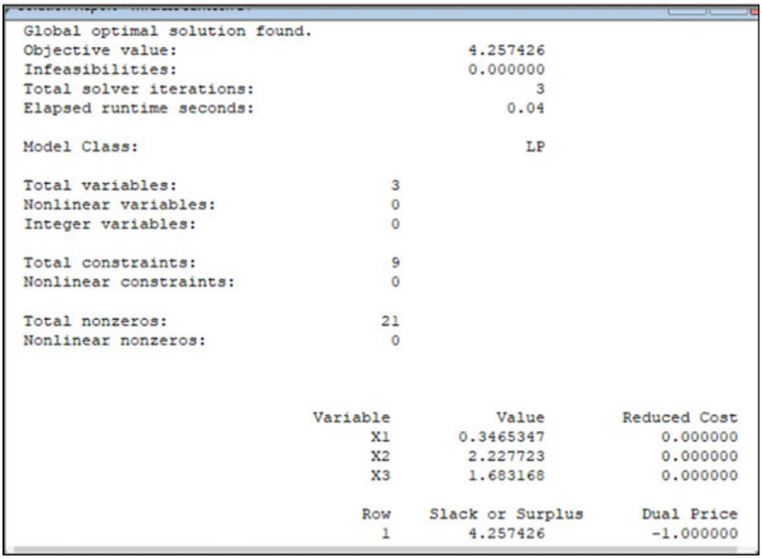
```
Global optimal solution found.
Objective value:                              2.146018
Infeasibilities:                              0.000000
Total solver iterations:                             3
Elapsed runtime seconds:                          0.04

Model Class:                                        LP

Total variables:                     3
Nonlinear variables:                 0
Integer variables:                   0

Total constraints:                   9
Nonlinear constraints:               0

Total nonzeros:                     21
Nonlinear nonzeros:                  0



                              Variable           Value        Reduced Cost
                                    X1       0.3982301            0.000000
                                    X2        1.061947            0.000000
                                    X3       0.6858407            0.000000

                                   Row    Slack or Surplus      Dual Price
                                     1        2.146018           -1.000000
                                     2        0.000000       -0.8849558E-02
```

Fig. 3 Linear formulation based on node 10



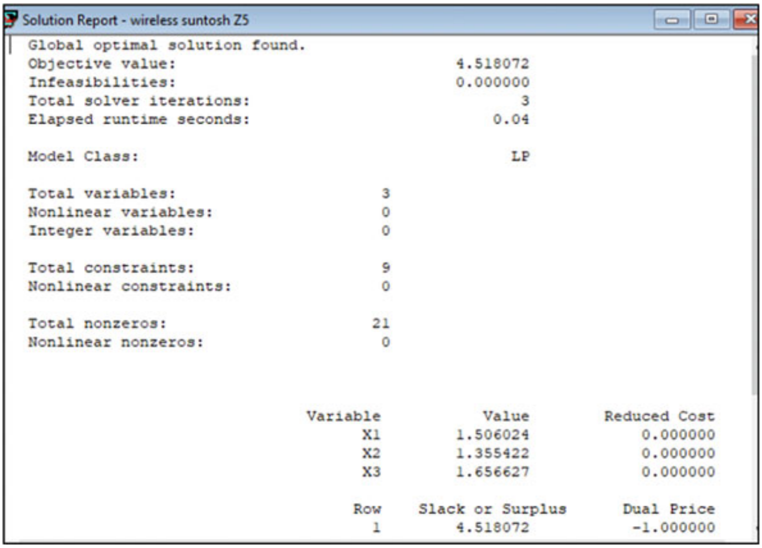
```
Global optimal solution found.
Objective value:                              2.904412
Infeasibilities:                              0.000000
Total solver iterations:                             3
Elapsed runtime seconds:                          0.05

Model Class:                                        LP

Total variables:                      3
Nonlinear variables:                  0
Integer variables:                    0

Total constraints:                    9
Nonlinear constraints:                0

Total nonzeros:                      21
Nonlinear nonzeros:                   0


                          Variable           Value        Reduced Cost
                                X1       0.5147059            0.000000
                                X2        1.397059            0.000000
                                X3       0.9926471            0.000000

                               Row    Slack or Surplus      Dual Price
                                 1        2.904412           -1.000000
```

Fig. 4 Linear formulation based on node 15



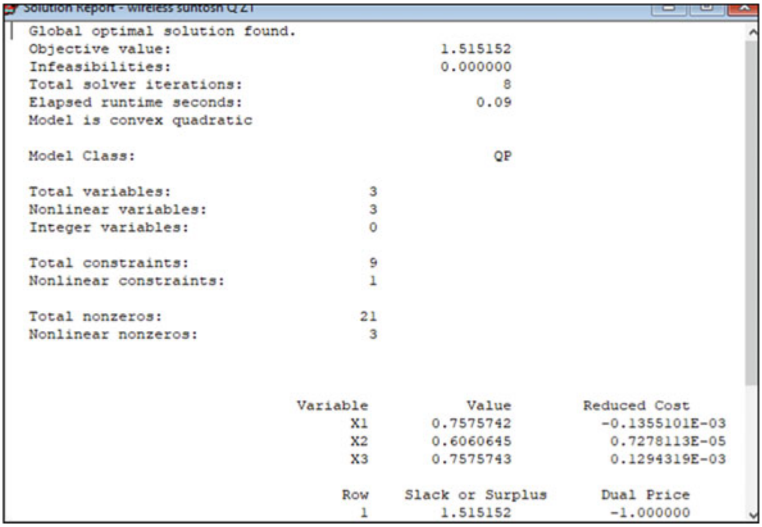
```
Global optimal solution found.
Objective value:                              4.257426
Infeasibilities:                              0.000000
Total solver iterations:                             3
Elapsed runtime seconds:                          0.04

Model Class:                                        LP

Total variables:                      3
Nonlinear variables:                  0
Integer variables:                    0

Total constraints:                    9
Nonlinear constraints:                0

Total nonzeros:                      21
Nonlinear nonzeros:                   0


                          Variable           Value        Reduced Cost
                                X1       0.3465347            0.000000
                                X2        2.227723            0.000000
                                X3        1.683168            0.000000

                               Row    Slack or Surplus      Dual Price
                                 1        4.257426           -1.000000
```

Fig. 5 Linear formulation based on node 20



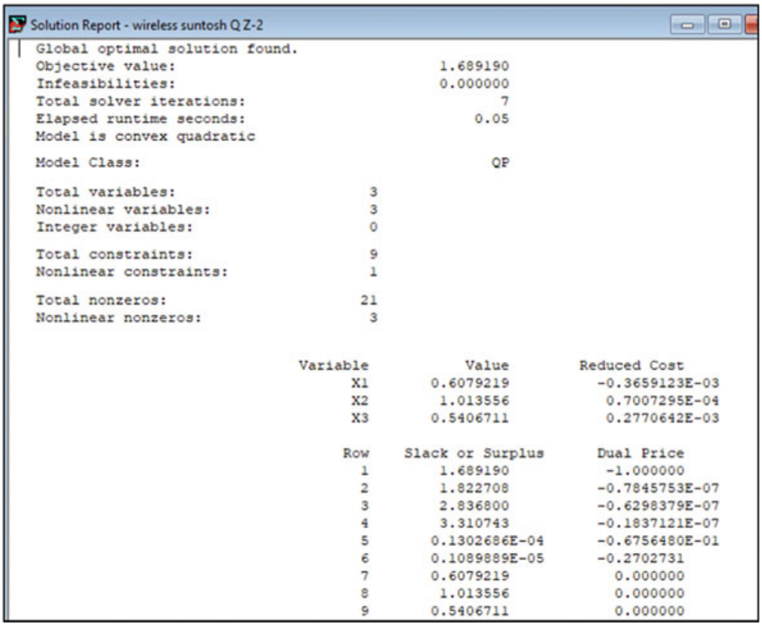
```
Global optimal solution found.
Objective value:                              4.518072
Infeasibilities:                              0.000000
Total solver iterations:                             3
Elapsed runtime seconds:                          0.04

Model Class:                                        LP

Total variables:                      3
Nonlinear variables:                  0
Integer variables:                    0

Total constraints:                    9
Nonlinear constraints:                0

Total nonzeros:                      21
Nonlinear nonzeros:                   0



                              Variable           Value        Reduced Cost
                                    X1        1.506024            0.000000
                                    X2        1.355422            0.000000
                                    X3        1.656627            0.000000

                                   Row    Slack or Surplus      Dual Price
                                     1        4.518072           -1.000000
```

Fig. 6 Linear formulation based on node 25



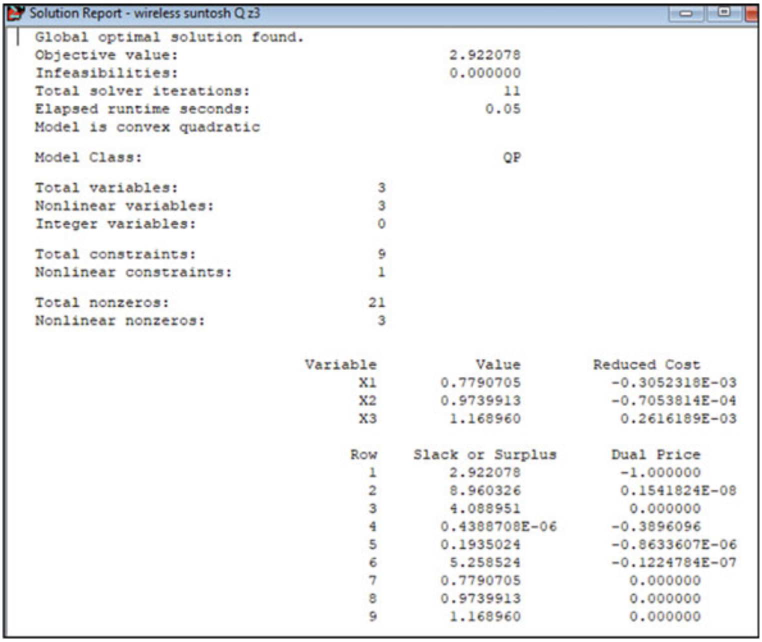
```
Global optimal solution found.
Objective value:                              1.515152
Infeasibilities:                              0.000000
Total solver iterations:                             8
Elapsed runtime seconds:                          0.09
Model is convex quadratic

Model Class:                                        QP

Total variables:                      3
Nonlinear variables:                  3
Integer variables:                    0

Total constraints:                    9
Nonlinear constraints:                1

Total nonzeros:                      21
Nonlinear nonzeros:                   3



                              Variable           Value        Reduced Cost
                                    X1       0.7575742       -0.1355101E-03
                                    X2       0.6060645        0.7278113E-05
                                    X3       0.7575743        0.1294319E-03

                                   Row    Slack or Surplus      Dual Price
                                     1        1.515152           -1.000000
```

Fig. 7 Quadratic formulation based on node 5



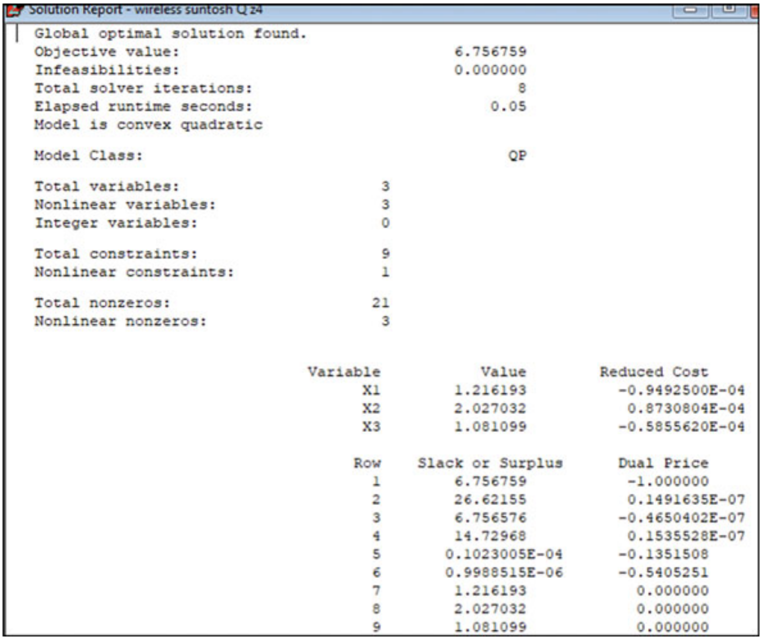
```
Global optimal solution found.
Objective value:                              1.689190
Infeasibilities:                              0.000000
Total solver iterations:                             7
Elapsed runtime seconds:                          0.05
Model is convex quadratic

Model Class:                                        QP

Total variables:                     3
Nonlinear variables:                 3
Integer variables:                   0

Total constraints:                   9
Nonlinear constraints:               1

Total nonzeros:                     21
Nonlinear nonzeros:                  3

                              Variable           Value        Reduced Cost
                                    X1       0.6079219       -0.3659123E-03
                                    X2        1.013556        0.7007295E-04
                                    X3       0.5406711        0.2770642E-03

                                   Row    Slack or Surplus      Dual Price
                                     1        1.689190          -1.000000
                                     2        1.822708       -0.7845753E-07
                                     3        2.836800       -0.6298379E-07
                                     4        3.310743       -0.1837121E-07
                                     5       0.1302686E-04   -0.6756480E-01
                                     6       0.1089889E-05   -0.2702731
                                     7       0.6079219          0.000000
                                     8        1.013556          0.000000
                                     9       0.5406711          0.000000
```

Fig. 8 Quadratic formulation based on node 10



```
Global optimal solution found.
Objective value:                              2.922078
Infeasibilities:                              0.000000
Total solver iterations:                            11
Elapsed runtime seconds:                          0.05
Model is convex quadratic

Model Class:                                        QP

Total variables:                     3
Nonlinear variables:                 3
Integer variables:                   0

Total constraints:                   9
Nonlinear constraints:               1

Total nonzeros:                     21
Nonlinear nonzeros:                  3

                              Variable           Value        Reduced Cost
                                    X1       0.7790705       -0.3052318E-03
                                    X2       0.9739913       -0.7053814E-04
                                    X3        1.168960        0.2616189E-03

                                   Row    Slack or Surplus      Dual Price
                                     1        2.922078          -1.000000
                                     2        8.960326        0.1541824E-08
                                     3        4.088951          0.000000
                                     4       0.4388708E-06     -0.3896096
                                     5       0.1935024       -0.8633607E-06
                                     6        5.258524       -0.1224784E-07
                                     7       0.7790705          0.000000
                                     8       0.9739913          0.000000
                                     9        1.168960          0.000000
```

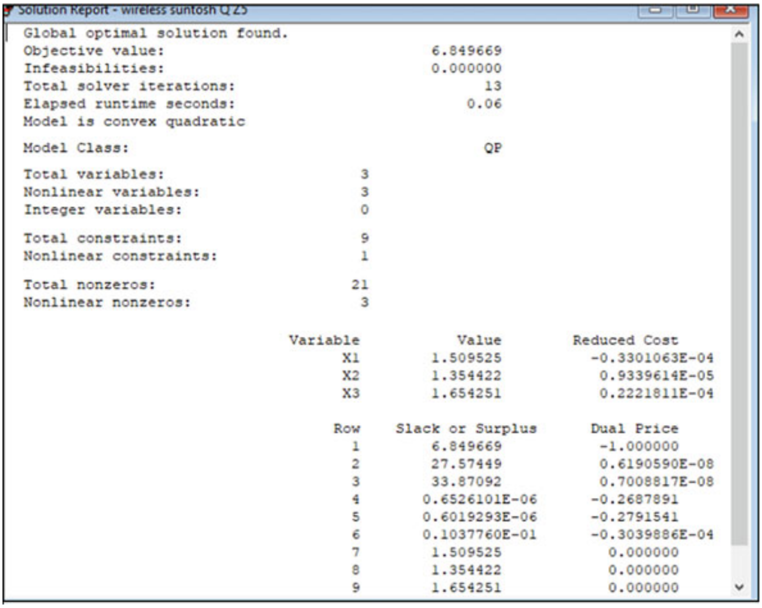
Fig. 9 Quadratic formulation based on node 15



Solution Report - wireless suntosh Q z4

```
Global optimal solution found.
Objective value:                              6.756759
Infeasibilities:                              0.000000
Total solver iterations:                             8
Elapsed runtime seconds:                          0.05
Model is convex quadratic

Model Class:                                        QP

Total variables:                      3
Nonlinear variables:                  3
Integer variables:                    0

Total constraints:                    9
Nonlinear constraints:                1

Total nonzeros:                      21
Nonlinear nonzeros:                   3


                              Variable           Value        Reduced Cost
                                    X1        1.216193       -0.9492500E-04
                                    X2        2.027032        0.8730804E-04
                                    X3        1.081099       -0.5855620E-04

                                   Row    Slack or Surplus      Dual Price
                                     1        6.756759        -1.000000
                                     2        26.62155         0.1491635E-07
                                     3        6.756576        -0.4650402E-07
                                     4        14.72968         0.1535528E-07
                                     5       0.1023005E-04    -0.1351508
                                     6       0.9988515E-06    -0.5405251
                                     7        1.216193         0.000000
                                     8        2.027032         0.000000
                                     9        1.081099         0.000000
```

Fig. 10 Quadratic formulation based on node 20



Solution Report - wireless suntosh Q Z5

```
Global optimal solution found.
Objective value:                              6.849669
Infeasibilities:                              0.000000
Total solver iterations:                            13
Elapsed runtime seconds:                          0.06
Model is convex quadratic

Model Class:                                        QP

Total variables:                      3
Nonlinear variables:                  3
Integer variables:                    0

Total constraints:                    9
Nonlinear constraints:                1

Total nonzeros:                      21
Nonlinear nonzeros:                   3

                              Variable           Value        Reduced Cost
                                    X1        1.509525       -0.3301063E-04
                                    X2        1.354422        0.9339614E-05
                                    X3        1.654251        0.2221811E-04

                                   Row    Slack or Surplus      Dual Price
                                     1        6.849669        -1.000000
                                     2        27.57449         0.6190590E-08
                                     3        33.87092         0.7008817E-08
                                     4       0.6526101E-06    -0.2687891
                                     5       0.6019293E-06    -0.2791541
                                     6       0.1037760E-01    -0.3039886E-04
                                     7        1.509525         0.000000
                                     8        1.354422         0.000000
                                     9        1.654251         0.000000
```

Fig. 11 Quadratic formulation based on node 25

5 Conclusion

This paper is based on the fusion of linear programming and quadratic programming with the help of three basic network parameters such as energy, delay, and overhead. The paper is formulated in two ways, first is based on linear programming, and second is quadratic programming. In each formulation, objective function is considered with the help of some constraints. Finally, it is observed that in each iteration, data is optimized based on increase of number of nodes in the network. In this proposed work, both linear and quadratic programmings are compared and validated that quadratic programming formulation is better than linear programming formulation.

References

1. Li H, Lin Z (2017) Study on location of wireless sensor network node in forest environment. Procedia Comput Sci 107:697–704
2. Curry RM, Smith JC (2016) A survey of optimization algorithms for wireless sensor network lifetime maximization. Comput Ind Eng 101:145–166
3. Tekin N, Gungor VC (2020) The impact of error control schemes on lifetime of energy harvesting wireless sensor networks in industrial environments. Comput Standard Interf 70:103417
4. Kheirfam B, Verdegay JL (2012) Strict sensitivity analysis in fuzzy quadratic programming. Fuzzy Sets Syst 198:99–111
5. Das SK, Kumar A, Das B, Burnwal AP (2013) On soft computing techniques in various areas. Comput Sci Inf Technol 3:59
6. Mandhare VV, Thool VR, Manthalkar RR (2016) QoS Routing enhancement using metaheuristic approach in mobile ad-hoc network. Comput Netw 110:180–191
7. Phoemphon S, So-In C, Leelathakul N (2020) A hybrid localization model using node segmentation and improved particle swarm optimization with obstacle-awareness for wireless sensor networks. Expert Syst Appl 143:113044
8. Das SK, Tripathi S (2018) Intelligent energy-aware efficient routing for MANET. Wirel Netw 24(4):1139–1159
9. Sun Z, Liu Y, Tao L (2018) Attack localization task allocation in wireless sensor networks based on multi-objective binary particle swarm optimization. J Netw Comput Appl 112:29–40
10. Cao B, Zhao J, Lv Z, Liu X, Kang X, Yang S (2018) Deployment optimization for 3D industrial wireless sensor networks based on particle swarm optimizers with distributed parallelism. J Netw Comput Appl 103:225–238
11. Yang W, Wang X, Song X, Yang Y, Patnaik S (2018) Design of intelligent transportation system supported by new generation wireless communication technology. In: Intelligent systems: concepts, methodologies, tools, and applications. IGI Global, pp 715–732
12. Jayakumar Loganathan J, Subbiah J (2020) Energy aware dynamic mode decision for cellular D2D communications by using integrated multi-criteria decision making model. Int J Ambient Comput Intell 11(3). (7 February 2020, IGI Global)
13. Hu YF, Ding YS, Ren LH, Hao KR, Han H (2015) An endocrine cooperative particle swarm optimization algorithm for routing recovery problem of wireless sensor networks with multiple mobile sinks. Inf Sci 300:100–113
14. Elhabyan RS, Yagoub MC (2015) Two-tier particle swarm optimization protocol for clustering and routing in wireless sensor network. J Netw Comput Appl 52:116–128
15. Hayes T, Ali FH (2016) Robust Ad-hoc sensor Routing (RASeR) protocol for mobile wireless sensor networks. Ad Hoc Netw 50:128–144

16. Singh J, Bhardwaj M, Sharma A (2015) Experimental analysis of distributed routing algorithms in ad hoc mobile networks. Procedia Comput Sci 57:1411–1416
17. Chandrakar P (2019) A secure remote user authentication protocol for healthcare monitoring using wireless medical sensor networks. Int J Ambient Comput Intell (IJACI) 10(1):96–116
18. Jat DS, Bishnoi LC, Nambahu S (2018) An intelligent wireless QoS technology for big data video delivery in WLAN. Int J Ambient Comput Intell (IJACI) 9(4):1–14
19. Sattari MRJ, Malakooti H, Jalooli A, Noor RM (2014) A dynamic vehicular traffic control using ant colony and traffic light optimization. In: Advances in systems science. Springer, Cham, pp 57–66
20. Abdallah AE (2016) Smart partial flooding routing algorithms for 3D ad hoc net-works. Procedia Comput Sci 94:264–271
21. Rasheed I, Banka H (2018) Query expansion in information retrieval for Urdu language. In: 2018 4th international conference on information retrieval and knowledge management (CAMP). IEEE, pp 1–6
22. Hao R, Yang H, Zhou Z (2019) Driving behavior evaluation model base on big data from internet of vehicles. Int J Ambient Comput Intell (IJACI) 10(4):78–95
23. Singh J, Singh A, Shree R (2011) An assessment of frequently adopted unsecure patterns in mobile ad hoc network: requirement and security management perspective. Int J Comput Appl 24(9):0975–8887
24. Singh J, Banka H, Verma AK (2019) Locating critical failure surface using meta-heuristic approaches: a comparative assessment. Arab J Geosci 12(9):307

Analysing of Troubleshooting Techniques

IDS Detection Based on Optimization Based on WI-CS and GNN Algorithm in SCADA Network

S. Shitharth, N. Satheesh, B. Praveen Kumar, and K. Sangeetha

Abstract Industry control systems (ICS) are considered as one of the inevitable systems in this contemporary smart world. In that supervisory control and data acquisition (SCADA) is the centralized system that control the entire grid. When a system is considered to be a whole and sole control, obviously an uncompromised security would be the prime. By having that as a major concern, a lot of research is being done on IDS security. In spite of that it has several cons including increased fake positive and fake negative rates, which will invariably lead to a larger chaos. To get rid of these problems, a weighted-intrusion based cuckoo search (WI-CS) and graded neural network (GNN) methods are proposed in this chapter. The key purpose of this chapter is to identify and categorize the anomalies in a SCADA system through data optimization. At initial stage, the collected real-time SCADA dataset is given as input. Then, by using the aforementioned proposed machine learning algorithms, these data are clustered and optimized. Later to find, the type of intrusion will remain as a further challenge and for that we propose HNA-AA algorithm. The investigational results estimate the efficiency of the system by considering sensitivity, false detection rate, precision, recall, Jaccard, accuracy, dice and specificity.

Keywords SCADA · Cukoo search · Neural network · Intrusion detection system · Clustering · Feature optimization

S. Shitharth (✉)
Department of CSE, Vardhaman College of Engineering, Hyderabad 501218, India

N. Satheesh
Department of CSE, St. Martin’s Engineering College, Hyderabad 500041, India

B. P. Kumar
Department of EEE, Bharat Institute of Engineering and Technology, Hyderabad 501510, India

K. Sangeetha
Department of CSE, Sri Satya Sai University of Technology and Medical Sciences, Sehore, Madhya Pradesh 466001, India

S. K. Das et al. (eds.), *Architectural Wireless Networks Solutions and Security Issues*, Lecture Notes in Networks and Systems 196,
https://doi.org/10.1007/978-981-16-0386-0_14

1 Introduction

SCADA collects information from the remote systems which is a part of ICS [1, 2]. It is an automation system that is used to receive the data from various sensors and components located at remote sites and transmit the data at a master site for monitoring purposes. SCADA systems are mainly used to both control and monitor various industrial applications. The significant and most essential duty is to detect intrusions in SCADA network. In real life, every country is moving toward a digital and smart environment. Smart grid is one of the prime setups in making a smart city. Marinating the smart grid security is always a vital area to look out and that is been done by SCADA systems. SCADA is primarily used to monitor electric power generation, transmission and distribution in smart grid systems. Along with that water sewage system, traffic signal controlling, mass transit, manufacturing, building facilities and environment are the other prime real-time uses of a SCADA system [3]. Intrusion detection is the process used to detect and monitor such malicious activity on a network. Monitoring the network manually is quite impossible. The foremost target of IDS [4–6] is to defend the availability, integrity and confidentiality of the system by digging out the malicious nodes, intrusions and destructive attacks. The IDS achieves the subsequent processes: Both user and system behavior were observed and analyzed. Then, the network configurations and its vulnerabilities are audited and also assess the reliability of critical organization and data files. The research work by Santosh and Sachin elaborated on [7] methodologies like software defining network (SDN) along with linear programming in ad hoc network are used to overcome uncertainties and imprecisions. These stats again have a larger impact while calculating confusion matrix results in any network security. Based on calculating those results and matching of well-known attacks, activity patterns are analyzed along with the irregular activities of the nodes in the system are also analyzed. The Modus operandi of anomaly detection in SCADA system is shown in Fig. 1.

Intrusion detection systems (IDS) are used to observing network and system activities to detect malicious activities. It can be categories into network-based intrusion detection system (NIDs) and host-based intrusion detection system (HIDs) [8, 9]. For protecting host computer networks and monitoring networks, the host-based IDS [10, 11] are used largely and also to spot the burglar attack activities, malicious actions nodes in the system and application anomalies. It classifies and prevents the risky incidents to investigate and observe the actions happening over the host system. The network-based IDS [12, 13] openly examines the traffic in the network system. For supervising the intrusive actions, based on its invariable changing state, it collects system's data initially. Then, finding and identifying the intrusions in the SCADA network becomes a significant and tedious task. Perhaps, it has its own disadvantages like complex installation and frequent network overheads. The major issue with the existing approaches is its high fake alarm rates and classifying unknown attacks and its feature absence in the intrusion attack library. To get rid of these issues, this paper proposes an enhanced IDS system and an optimization techniques using machine learning algorithm. The prime motto of this work is to pick the finest attributes for

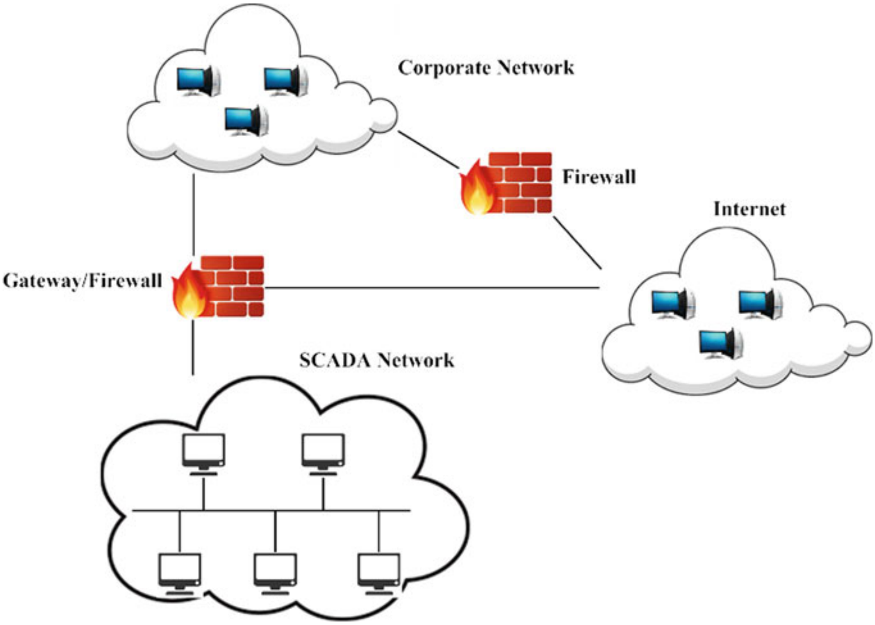


Fig. 1 Anomaly detection in SCADA network

optimization using WI-CS algorithm. In order to categorize the attacks or intrusions accurately in a SCADA system, a GNN categorization method is developed. The grading or hierarchical method is used in order to calculate the performance of the developed system. Moreover, this research work uses the real-time power system dataset which has been subjected to intrusions. The other sections of this chapter are structured as follows: Sect. 2 is more of literature survey that takes hold of previous research methods done in SCADA IDS. Section 3 is all about the full description of the developed WI-CS and GNN method, and the Sect. 4 assesses the enactment of the developed methods with already available methods for claiming the improved performance. The section V concludes in a way how this work is carried forward for future enhancements and concluded.

2 Related Works

This related work is more of a literature survey that explains previous research works done by various authors in SCADA IDS. A novel IDS idea given by Almalawi et al. [14], which identified only the usual and dangerous criteria that degrade the system efficiency. The prime focus of this paper is followed by:

- State recognition

- Extracting own decision statements on detection
- Decrease of high false positive rate
- Compute to assess criticality.

The major drawback of this proposed method is that the work did not dig deep into the constancy modifications in the network. Mitchell and Chen [15] proposed some major takeaways such as:

- To create a logical methodology for counteracting the attack
- To identify and find out the dissimilar failures types in industry control systems (IPs).

The major drawback of this research is, CPS system's solidity is been questioned during the intrusion times. Based on telemetry study, an approach to find out anomalies in SCADA system was developed by Ponomarev and Atkison [16]. In this paper, using master–slave process, the traffic is well divided between sender and receiver. The telemetry-based IDS supervised all packets in the ICS system, so it detected the irregularity in traffic. To spot the control-related attacks in SCADA system, a semantic analysis framework was designed by Lin et al. [17]. In this research, with a power flow analysis, the control results are optimized. A regular expression (RegEx) was proposed by Liu and Torng [18]. The supplies contain nondeterministic finite automata (NFA) size and deterministic finite automata (DFA) speed which is made more scalable and adaptable. Moreover, the condition copying and evolution copying were arrested using the projected automata mode. Marchang et al. [19] developed a mobile ad hoc network (MANET) to decrease the period of lively moment without compromising their efficiency. Here, a behavioral observance of nodes is calculated at a specific check point to reduce the energy consumption of the nodes in the network. Dirichlet-based detection scheme (DDOA) was introduced by Li et al. [20] mainly to identify the outliers residing in the smart grid. They done it by using the real-time IEEE power system database and by using a software tool called power world simulator. Ambusaidi et al. [21] proposed a method using square vector to find out the outliers in SCADA. They used KDD and NSL dataset for feature classification and optimization.

Hasan et al. [22] proposed a trust-based monitoring system for incoming and outgoing traffic. The main objective is to get rid of excessive capital and operational cost used by the nodes and he has done it by calculating trust values if every node in the system. Yang et al. [23] proposed a research methodology for the attacks having dissimilarity and parameter variation among themselves. He used behavioral approach of the nodes and try to whitelist the genuine nodes. Based on this, the author can separate the inliers and outliers from the network. A three-layer IDS for providing protection to SCADA was developed by Samdarshi et al. [24]. Based on partition system and dividing the IoT devices based on their router gateway, it is the major takeaway from this research. The core intent of this research is to use various cybersecurity applications use the trust system placement scheme. Sayegh et al. [25] proposed a way for IDS based on patterns and traffic of the incoming nodes in the network. This research includes the below five stages:

- Intrusion whiffing
- Attribute extraction
- Training phase
- Testing phase
- Anomaly detection.

Now, the relationships among the packet were anticipated to recognize its authentication. The problem of attack isolation and intrusion detection was investigated by Amin et al. [26]. The objectives of this research are as follows:

- It provides solutions for anomaly detection
- It considers the sensor data for attribute classification.

Maglaras et al. [27] proposed a vector classification which is used to improve the efficiency by a numerical method. In addition by using clustering methodology, it classifies the SCADA alarm into low, medium and high. The major problem of this technique is the need to reduce the fake alarm rate. Yang et al. [28] proposed an IDS to track malicious interventions in the SCADA network. The objective of this work was based on the signature-based rules and it detects the unknown attacks and the major benefit of this work was to detect the outliers accurately. Yasakethu and Jiang [29] proposed a methodology using Markov model that finds out the intruders. It also uses Markov matrix plot for data representation. This is again based on a set of pre-defined rules and if the nodes violate that it would be blacklisted. Finally, the confusion matrix results how the improved the true positive rate is and a reduced fake positive rate. A relation-based intrusion detection (RID) system was proposed by Wang et al. [30] to find a SCADA network's false data injection attacks. The major claims made in this research are of:

- The system state of RID is observed; the inconsistent situation is detected and the compromised origins were inferred.
- The intrusion detection model is used to attain real-time exposure on resource-constrained machines.

A model-based IDS was proposed by Goldenberg and Wool [31] which was extremely responsive that flags the irregularities with the help of the Modbus scheme. In the SCADA system, for finding the integrity attacks, an unsupervised anomaly-based IDS was developed by Almalawi et al. [32]. The major ideas of this research were as given below:

- SCADA system's steady and conflicting states were identified.
- From the identified states, it took out the proximity of finding a policy.

Here, based on the finest inconsistency, threshold conflicting annotations were divided from the steady annotations. A survey of several detection methods for finding the irregularities in the system was done by Ahmed et al. [33]. This survey contains the below groups: organization, numerical, information theory and groups. Besides, it also discusses the challenges in research while collecting dataset that is used for detecting system intrusions. Santhosh and Sachin [34] have detailed about multi-criteria decision-making (MCDM) technique which works based on a ranking methodology based on preferences. The work offers a decisive route or the mobile nodes to travel by combining the aforementioned MCDM technique with intuitionistic fuzzy soft set (IFSS).

3 Proposed Method

This part of the chapter discusses the detailed explanation of the proposed optimization technique for anomaly detection. The ultimate target of this research is to correctly spot the intrusions in a system with the help of extracted features using a learning methodology. WI-CS and GNA-NN are the two types of algorithms proposed in this research. In the first stage, the real-time data derived from the Mississippi University SCADA test bed is fed into the system, in which the features and the groups are arranged. After that, the features are extracted using the projected WI-CS-based optimization technique. Then, it selects the finest attributes and training data which are given as input to classification. In this phase, the HNN-based categorization algorithm is introduced to categorize the attack types. Figure 2 shows the flow graph of the research method and it consists of the below stages:

- Initialization of dataset
- Arrangement of attribute
- Initialization of cluster
- Optimization of feature and selection of attribute
- Classification.

In this analysis, the SCADA test bed dataset is fed as the input to evaluate the performance of the proposed IDS, which is formulated based on the PRC technique. A unique dataset feature contains 37 power system events that accompanied 15 different sets, which are used to fabricate the dataset. These features are further classified on an event basis like a natural event, attack event and no events. To get more in detail, these conditional scenarios are furthermore subdivided into other categories in a random manner such as three-class, multi-class datasets and binary. The integration of a huge number of sensors is available in the SCADA system. As a result, the memory intricacy of the system is amplified. Based on the response from all sensors, feature selection is executed to resolve the above-mentioned problem. This enhances the efficiency and, decreases the storage intricacy. The conditions are created, derived from every attribute, because of that one can use a vast data sources. With the aim of finding variations among parameters, GNN structure is developed. Huge attributes are used for accurate detection throughout the progression of classification.

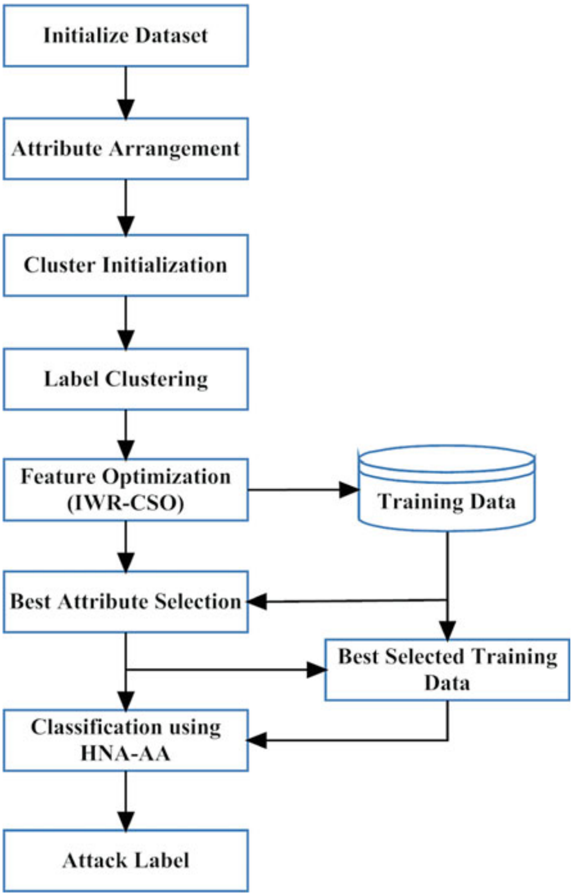


Fig. 2 Flow graph of the research method

3.1 Optimization of Feature

Here, all the extracted features are scanned for any possible intrusion of an anomaly. The main motto of our work is to select the appropriate attributes for optimization. It finds outs and senses the irregular node behavior based on preferred attributes. The developed WI-CS is an optimization-based technique which is used to cluster and organize the node attributes. Generally, fitness and a cost function comprised to form input. Optimization is the course of regulating the inputs and uniqueness of a machine to get the utmost outcome. Here, we use cuckoo search algorithm which is inspired by the brood parasitism of cuckoo birds. The cuckoos laying their eggs in the nests of other host birds. If a host bird discovers the eggs in the nests are not their own, it will throw away these alien eggs away. Else it will abandon its nest and build a new nest somewhere. Based on this observation, the cuckoo search algorithm is described by the following rules:

- Every cuckoo lays solely 1 egg at a time and the eggs are exactly set in a nest (randomly selected).
- The nest having better quality eggs which are carried onto the next phase.
- The number of nest is fixed and the quality of nest is static and is not alterable.

In this research, huge attributes are used to develop a SCADA organization. Therefore, it is necessary to spot whether the attributes are grouped or not. If the attributes are grouped, the precise results can be achieved; or else, this directs to the misclassification rate. Then, the cost function is predicted based on the difference that caused by the anomalies compared with legal nodes. Also, based on the weight, the probability of the particle is predicted. Then, the fitness value falls somewhere between anomaly and non-anomaly node ranges. These calculated values would be considered for cuckoo's egg laying probability [35]. Then, the finest attribute is selected based on previous and current values. This repeats for every iteration. The finest attribute is picked once the fitness average rate becomes higher than the calculated cost estimated value. In the proposed algorithm, input feature matrix T, cuckoo particles P, initial radius r and cost C_{st} are initialized. The cost rate is predictable by the length and radius of searching particles.

```
Algorithm I –Weighted Intrusion based Cuckoo Search (WI-CS)
Input:     T Feature matrix;
Output:    ST Select Feature;
Step 1:    Initialize cuckoo particles and cost rate;
           $P = \{T_1, T_2 \ldots T_N\}$ // Cuckoo Particles;
           Cst = 0; //Initial cost value;
           r=1; //Initial Radius;
Step 2:    Estimate cost value as,
           $Cst = \{P(1), (g \times h)\}$
           Where,
           $g = 1 + \frac{r}{(N-1) \times \sum P_i'}$        // i = 1, 2 … N;
           r – Radius of searching;
           N – Length of particles;
           $h = 1 - \sqrt{\frac{P_1}{g}}$
Step 3:    Co-ordinates of particles;
           $x = P(Fitness, 1)$;
           $y = P(Fitness, 2)$;
```

```
Step 4:   Objective function;
          $O = \begin{Bmatrix} P_i & if\ x \ \leq y \\ 0 & else \end{Bmatrix}$ // Objective function;
Step 5:   Update radius;
          $r_1 = d_1 + \frac{(V_{0\ to\ r} \times (d_2 - d_1))}{r}$
          Where,
          $d = Min\ (Cst) \pm \left(\alpha \times \left(Max(Cst)\right) - \left(Min(Cst)\right)\right)$
Step 6:   Reproduce and Update Cuckoo particles;
          $for\ i = 1\ to\ M$ // 'M' – Number of iteration
          $if\ (Cst_i < Cst_{i-1})$
                $C_{head} = P(indx)$ // Cluster head selection;
                Where, $indx = \begin{cases} 1, if\ \left(P \times e^{-\beta N}\right) < 0 \\ 0, \qquad\qquad else \end{cases}$
      $X_{update}(i) = x(i-1) + \left(\left(Rand^{-\frac{1}{\alpha}}\right) * cos(Cst * 2 * pi)\right)$
      $Y_{update}(n) = y(i-1) + \left(\left(Rand^{-\frac{1}{\alpha}}\right) * cos(Cst * 2 * pi)\right)$
      $Y_{update}(n) = y(i-1) + \left(\left(Rand^{-\frac{1}{\alpha}}\right) * cos(Cst * 2 * pi)\right)$
        $P(m) = \left(1 - \frac{i-1}{(M-1)^{\frac{1}{\mu}}}\right)$ //Probability of laying eggs;
        $if\ P(Cst) < P(m)$ //Mutation
          $X_{Mutation}(i) = x(i-1) + \left(P(m) \times \left(Max(x) - Min(x)\right)\right)$
          $Y_{Mutation}(n) = y(i-1) + \left(P(m) \times \left(Max(y) - Min(y)\right)\right)$
                    $Cst_{mutation} = \{P(1), (g \times h)\}$
        End if;
             Update radius r;
      End i loop;
Step 7:  $\boldsymbol{ST = T(Cst > Average(Cst))}$
```

3.2 Classification Using HNN

By using the projected HNN method, the attacking and non-attacking labels are properly categorized only next to optimize the features. In the proposed method, the inputs given are the label Lb, select training set Str and testing feature Sv. At the start, the preferred testing feature, φ, directionality D, and iteration(theta) are initialized as zero. The unsystematic values like Wxy and Wxh are initialized with the range of testing Sv and also training Str. The NN includes three dissimilar kinds of layers contains hidden, input and output layers. Normally, the final result can be obtained from the output layer. Here, net 1 and net 2 indicated the input–output layer, respectively. Then, the exponential is calculated for the layers *H1* and *Y*. Then, the temporary distance is calculated. The temporary distance s_i is obtained only after getting the calculated load value of φ that is upgraded by including the present value. After that, for the differential features, the directionality *D* is estimated, *D* and theta values are upgraded. The association *r* among the attribute sets are estimated, in case, it is larger than the *D* value, Lb(i) label is allocated as the classified label (CL).

Algorithm II – Hierarchical Neural Network (HNN)

```
Input: Str as choosen training set; Sv as choosen testing feature and Lb as Label;
Output: Classified Label CL;
Step 1:  Initialization,
         $SV = \frac{SV}{norm\ (SV)}$;
         $Theta = 0, D = 0, \varphi = 0$;
Step 2:  $for\ (i = 1\ to\ Row_{size}\ (Str))$
Step 3:          $Wxh = Rand(STr), Why = Rand(SV)$;
Step 4:          $net1 = SV \times Wxh - Theta$;
Step 5:          $H1 = \frac{e^{(net1)} - e^{(-net1)}}{e^{(net1)} + e^{(-net1)}}$;
Step 6:          $net2 = H1 \times Why - Theta$;
Step 7:          $Y = \frac{e^{(net2)} - e^{(-net2)}}{e^{(net2)} + e^{(net2)}}$;
Step 8:          $s_i = \sqrt{\left(\frac{\sum((Str(n,:) - Y)^2)}{Y}\right)}$;
Step 9:          $\Delta Y = (1 + Y) \times (1 - Y) \times (STr_i - Y)$
Step 10:         $\varphi = \varphi + s_i + Theta$;
Step 11:     $D = (\varphi * \Delta Y + s_i) * \varphi$; // Directionality for differential features;
Step 12:         $Theta = Theta * D$;
Step 13:         $\varphi = \varphi + Theta$;
Step 14:     $r = \frac{\sum_{j=1}^{N}(Y_i - \bar{Y}) \times (s_i - \bar{s})}{\sqrt{\sum_{j=1}^{N}(Y_i - \bar{Y})^2 \times \sum_{j=1}^{N}(s_i - \bar{s})^2}}$;// Correlation between the feature sets;
Step 15:     If $(r > min(D))$ // Feature verification condition;
Step 16:         $CL = Lb\ (i)$; // Classified label;
Step 17:     End if;
Step 18:   end i loop;
```

Table 1 Parameter dataset with and without intrusion

Factor		Dataset			
		OLIRD	OLORD	MLIRD	MLORD
Humidity factor	Without intrusion	44–52	32–65	42–55	44–75
	With intrusion	45–90	50–87	50–92	57–94
Temperature factor	Without intrusion	25–30	24–36	28–32	25–35
	With intrusion	25–55	26–38	25–55	28–50

4 Performance Investigation

This part of the chapter gives the outcome using detection rate and false alarm rate for both proposed as well as existing algorithms. Here, in the presented research, with the help of the Ns-2 simulator, one dataset is formed by designing the system organization with a hundred nodes. After that, MATLAB tool is used to detect the intrusion by the implemented dataset. For validating the proposed method, two scenarios were taken, i.e., with and also without attacks. The compared datasets [13] in this section are as follows, one-leap indoor real data (OLIRD), multi-leap indoor real data (MLIRD), one-leap outdoor real data (OLORD) and multi-leap outdoor real data (MLORD).

4.1 Description of Dataset

Table 1 illustrates the dataset formed and utilized in this proposed work. Temperature and humidity are some of the parameters considered for analysis. Those above-mentioned factors point out the significant states from various densities in two-dimensional (2D) space. It shows the various levels of discriminations which is near to the low, medium and large-risk conditions. Besides, considering without and with attack, the factors are evaluated. Now, the lowest and largest values of temperature and humidity factors are calculated. For example, 44–52 in OLIRD dataset, 44 is the lowest and 52 is the largest humidity level. The dataset including low and high values are tabulated below:

4.2 Metrics for Accuracy

In the proposed method by using detection and false positive rate, the accuracy metrics is measured for intrusion detection. Efficient data-driven clustering (EDDC) [36] and proposed WI-CS with GNN techniques the detection rate and fake positive rate are given in Table 2.

Table 2 Detection rate and false positive rate

Dataset	Detection rate (%)		False positive (%)	
	EDDC	WI-CS with GNN	EDDC	WI-CS with GNN
OLIRD	100	100	0	0
OLORD	96.88	97.57	1.94	1.14
MLIRD	100	100	0.31	0.18
MLORD	92.98	94.28	0.04	0

$$\text{Detection Rate} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{1}$$

where

TP	True positive
TN	True negative
FP	Fake positive
FN	Fake negative

The fake positive rate is calculated as follows:

$$\text{False Positive Rate} = \frac{\text{FP}}{\text{FP} + \text{TN}} \tag{2}$$

where TP specifies the count of perfectly identified critical states, similarly FN specifies the count of the present but not identified critical states, FP specifies the wrongly flagged normal states as critical, and furthermore, TN specifies the count of properly identified normal states. Hence, it is observed from the analysis when compared to the other methods, the developed method gives the finest results for the datasets of four different cases.

4.3 Accuracy

$$\text{Accuracy} = \frac{\text{Number of critical states}}{\text{Total number of states}} \tag{3}$$

For different datasets, the accuracy values are shown in Table 3 for proposed as well as existing methods. For all datasets, high accuracy is obtained for the proposed method when compared to the existing method.

Table 3 Accuracy

Dataset	Accuracy (%)	
	Existing	Proposed
OLIRD	100	100
OLORD	97	98
MLIRD	100	100
MLORD	93	95

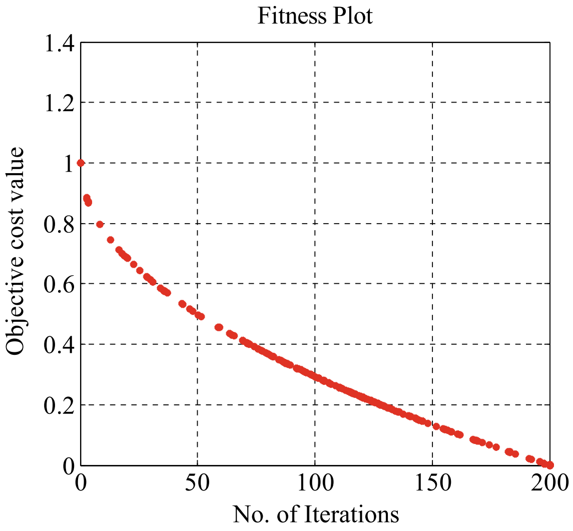


Fig. 3 Number of iterations versus objective cost value

4.4 Fitness Plot

Concerning the number of iterations, Fig. 3 shows the objective cost value. An increase in the numbers of iterations (0–200) will decrease the cost value as in progress. In the proposed method, the cost is reduced for intrusion detection which can be concluded from the above analysis.

4.5 Performance Rate

The incorporated WI-CS with SVM does not give the improved performance results, because of the restrictions of SVM but the proposed WI-CS is a competent optimization method. The major disadvantage of SVM is that it digs out an additional quantity of irrelevant features. Lacking of WI-CS, the time complexity is increased because of the whole data is necessary for analysis. During the normal attack detection and abnormal attack detection, it may guide to high misclassification rate. To overcome

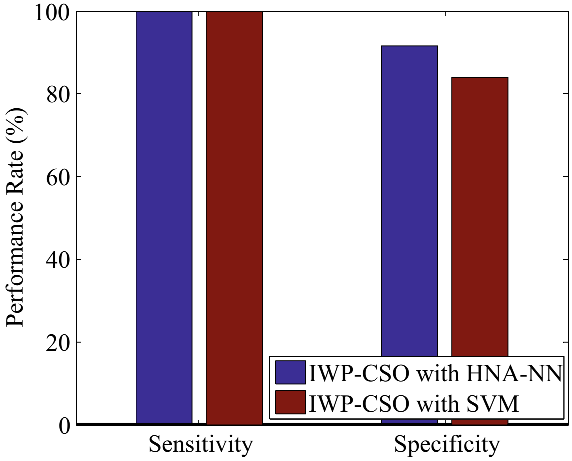


Fig. 4 Sensitivity and specificity

all these, WI-CS is created to obtain the optimal features. Similarly, to rise above the difficulty of SVM, the GNN method is collaborated with the WI-CS method. In general, to calculate the efficiency of the system, two major values are considered. One is sensitivity and the other one is specificity. It can be calculated as given below:

$$\text{Sensitivity} = \frac{\text{TP}}{(\text{TP} + \text{FN})} = \frac{\text{Number of true positive assessments}}{\text{Number of all positive assessments}} \tag{4}$$

Likewise, specificity is calculated as given below:

$$\text{Specificity} = \frac{\text{TN}}{(\text{TN} + \text{FP})} = \frac{\text{Number of true negative assessment}}{\text{Number of all negative assessment}} \tag{5}$$

These values of IWP-CSO/WI-CS with SVM and proposed IWP-CSO/WI-CS with GNN techniques are shown in Fig. 4. When compared to other methods, the proposed method gives a high-performance rate which is obtained from the above analysis.

4.6 Jaccard and Dice Coefficients

The similarity between data can be measured mainly by using the Jaccard and dice coefficients. It is given as the fraction between the sizes of the intersection of a couple of set to the union of two sets. These coefficients are used in the paper [37, 38] for comparison of different algorithms like random forest, *K*-means, etc. The metrics can be calculated as given below:

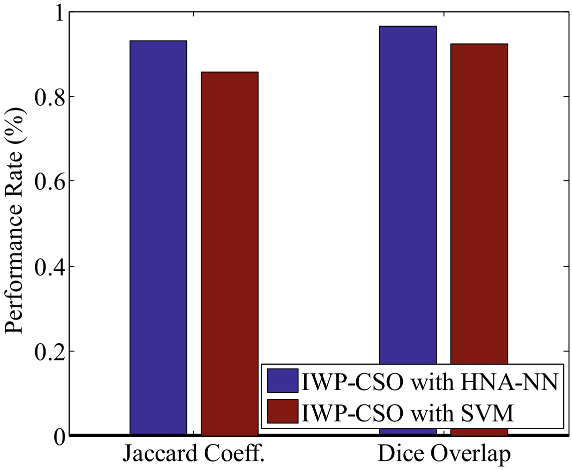


Fig. 5 Jaccard and dice coefficients

$$\text{Jaccard} = \frac{|X \cap Y|}{|X| + |Y| - |X \cap Y|} \tag{6}$$

where X and Y specify a couple of datasets. Likewise, dice is furthermore a similarity measuring method that discovers the similarity between information. It can be calculated as given below:

$$\text{Dice} = 2\frac{|X \cap Y|}{|X| + |Y|} \tag{7}$$

Jaccard and dice of WI-CS with SVM and proposed IWP-CSO / WI-CS with HNN method are shown in Fig. 5. The high-performance results are obtained in the proposed method when compared to the existing method.

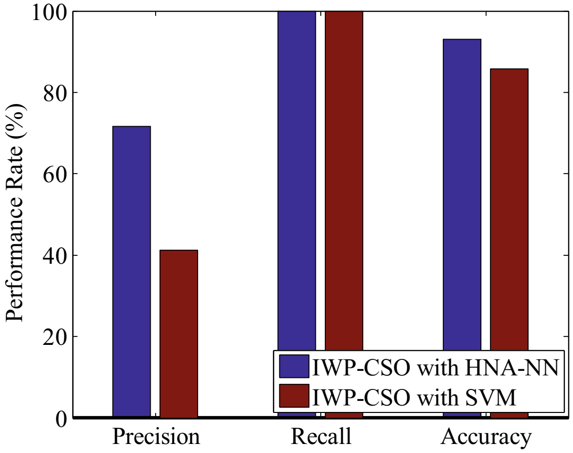
4.7 Precision, Recall and Accuracy

Figure 6 shows the performance rate of IWP-CSO with SVM and proposed IWP-CSO/WI-CS with HNN methods. These three rates are highly considered or the performance criteria. Precision and recall are calculated mainly based on true positive detection rate, whereas accuracy majorly depends on sensitivity and specificity values as mentioned below:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{8}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{9}$$

Fig. 6 Performance rate



$$\text{Accuracy} = \frac{(\text{TN} + \text{TP})}{(\text{TN} + \text{TP} + \text{FN} + \text{FP})} = \frac{\text{Number of true correct assessment}}{\text{Number of all assessment}} \tag{10}$$

It is very clear that the proposed method gives a very high precision, recall, as well as accuracy measures than the existing methods.

4.8 False Detection Rate

It is very important to get rid of genuine nodes which are misclassified as anomalies. To find out that we calculate false detection rate (FDR) as given below:

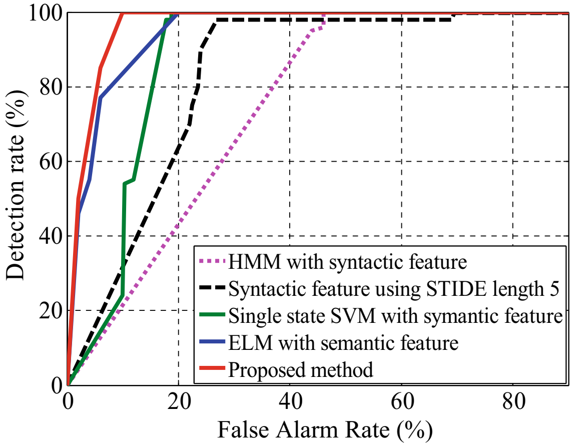
$$\text{False Detection Rate} = \frac{\text{number of honest users misidentified}}{\text{number of nodes identified as attackers}} \tag{11}$$

The false detection rate of an existing method as well as the proposed method is shown in Fig. 7, in which the x-axis corresponds to the alarming rate (in %) and the y-axis corresponds to the detection rate (in %). The proposed method gives a very high detection rate than the existing methods which is obtained from the analysis.

5 Conclusion and Future Work

Nowadays, even SCADA and ICS systems are more likely to be using a wireless sensor network. Even in the medical field, wireless sensor system became so inevitable nowadays. This study [39] deals with a detailed review of how ECG monitoring model works and about the evaluation of zig bee medical sensor networks in

Fig. 7 Detection rate



clinical networks. This shows really how much impact WSN has made in health care sectors. Hence, this chapter's security proposal is not just enough for the traditional SCADA network but also for SCADA that is using WSN. This chapter presents an improved WI-CS and GNN techniques for a SCADA system to filter the anomalies in it. The major motto of this work is to find out the outliers in this network and to classify the unknown attacks in the SCADA system. In this, the real-time dataset of the SCADA network is fed as input, and appropriate parameters are chosen. Then, in the training dataset, the attributes are selected which are optimized by employing the proposed WI-CS technique. After that, the filtering the finest attributes they are optimized by the GNN classification algorithm that guesses the intruder and also non-intruder label. The investigational results estimate the performance of the proposed method considering sensitivity, false detection rate, precision, recall, Jaccard, accuracy, dice and specificity. Furthermore, SIRD, SORD, MIRD and MORD which are the four different datasets are considered in this research for showing the improved performance of the introduced system. From this study, when compared to the other algorithms, the developed WI-CS through HNN algorithm gives better results.

References

1. Aghajanzadeh N, Keshavarz-Haddad A (2015) A concise model to evaluate security of SCADA systems based on security standards. Int J Comput Appl 111
2. Shahzad A et al (2015) A SCADA intermediate simulation platform to enhance the system security. In: 2015 17th international conference on advanced communication technology (ICACT), pp 368–373
3. Debashis De et al (2020) Wireless sensor network: applications, challenges, and algorithms, Springer tracts in nature-inspired. Springer, pp 1–18
4. Wei H et al (2015) SOM-based intrusion detection for SCADA systems. In: Electronics and electrical engineering: proceedings of the 2014 Asia-Pacific electronics and electrical engineering conference (EEEC 2014), 27–28 Dec 2014, Shanghai, China, p 57

5. Mcquillan JL, Lloyd CA (2016) SCADA intrusion detection systems. ed: US Patent 20,160,094,578
6. Liao H-J et al (2013) Intrusion detection system: a comprehensive review. J Netw Comput Appl 36:16–24
7. Das SK, Tripathi S (2019) A nonlinear strategy management approach in software-defined ad hoc network. In: Lecture notes in networks and system, pp 321–346
8. Manikandan ST (2014) Removal of selective black hole attack in MANET by AODV protocol. Int J Innov Res Sci Eng Technol 3(3):2372–2377
9. Shitharth S, Winston DP (2015) An appraisal on security challenges and countermeasures in smart grid. Int J Appl Eng Res 10(20):16591–16597
10. Ou C-M (2012) Host-based intrusion detection systems adapted from agent-based artificial immune systems. Neurocomputing 88:78–86
11. Shitharth S, Winston DP (2015) A comparative analysis between two countermeasure techniques to detect DDoS with sniffers in a SCADA network. Procedia Technol 21:179–186
12. Koc L et al (2012) A network intrusion detection system based on a Hidden Naïve Bayes multiclass classifier. Expert Syst Appl 39:13492–13500
13. Selvarajan S, Shaik M, Ameerjohn S, Kannan S (2019) Mining of intrusion attack in SCADA network using clustering and genetically seeded flora based optimal classification algorithm. Inf Secur IET 14(1):1–11
14. Almalawi A et al (2016) An efficient data-driven clustering technique to detect attacks in SCADA systems. IEEE Trans Inf Foren Secur 11:893–906
15. Mitchell R, Chen R (2016) Modeling and analysis of attacks and counter defense mechanisms for cyber physical systems. IEEE Trans Reliab 65:350–358
16. Ponomarev S, Atkison T (2016) Industrial control system network intrusion detection by telemetry analysis. IEEE Trans Dependable Secur Comput 13:252–260
17. Lin H et al (2016) Runtime semantic security analysis to detect and mitigate control-related attacks in power grids. IEEE Trans Smart Grid 1–1
18. Liu AX, Torng E (2016) Overlay automata and algorithms for fast and scalable regular expression matching. IEEE/ACM Trans Netw 1–16
19. Marchang N et al (2016) A Novel approach for efficient usage of intrusion detection system in mobile Ad Hoc networks. IEEE Trans Veh Technol 1–1
20. Li B et al (2016) DDOA: a dirichlet-based detection scheme for opportunistic attacks in smart grid cyber-physical system. IEEE Trans Inf Foren Secur 1–1
21. Ambusaidi M et al (2016) Building an intrusion detection system using a filter-based feature selection algorithm. IEEE Trans Comput 1–13
22. Hasan MM, Mouftah HT (2016) Optimal trust system placement in smart grid SCADA networks. IEEE Access 4:2907–2919
23. Yang Y et al (2014) Multiattribute SCADA-specific intrusion detection system for power networks. IEEE Trans Power Delivery 29:1092–1102
24. Samdarshi R et al (2015) A triple layer intrusion detection system for SCADA security of electric utility. In: 2015 annual IEEE India conference (INDICON), pp 1–5
25. Sayegh N et al (2014) SCADA intrusion detection system based on temporal behavior of frequent patterns. In: MELECON 2014–2014 17th IEEE mediterranean electrotechnical conference, pp 432–438
26. Amin S et al (2013) Cyber security of water SCADA systems—part II: attack detection using enhanced hydrodynamic models. IEEE Trans Control Syst Technol 21:1679–1693
27. Maglaras LA et al (2014) Integrated OCSVM mechanism for intrusion detection in SCADA systems. Electron Lett 50:1935–1936
28. Yang Y et al (2013) Intrusion detection system for IEC 60870–5–104 based SCADA networks. In: 2013 IEEE Power Energy Soc Gener Meet: 1–5
29. Yasakethu S, Jiang J (2013) Intrusion detection via machine learning for SCADA system protection. In: Proceedings of the 1st international symposium on ICS and SCADA cyber security research 2013, pp 101–105

30. Wang Y et al (2014) SRID: state relation based intrusion detection for false data injection attacks in SCADA. In: European symposium on research in computer security, pp 401–418
31. Goldenberg N, Wool A (2013) Accurate modeling of Modbus/TCP for intrusion detection in SCADA systems. Int J Crit Infrastruct Prot 6:63–75
32. Almalawi A et al (2014) An unsupervised anomaly-based detection approach for integrity attacks on SCADA systems. Comput Secur 46:94–110
33. Ahmed M et al (2016) A survey of network anomaly detection techniques. J Netw Comput Appl 60:19–31
34. Das SK, Tripathi S (2018) Intelligent energy-aware efficient routing for MANET. Wirel Netw 24(4):1139–1159
35. De D, Mukherjee A, Das SK, Dey N (2020) Nature-inspired computing for wireless sensor networks, Springer tracts in nature-inspired. Springer, pp 1–341
36. Suthaharan S et al (2010) Labelled data collection for anomaly detection in wireless sensor networks. In: Intelligent sensors, sensor networks and information processing (ISSNIP), 2010 6th international conference on, pp 269–274
37. Selvarajan S, Shaik M, Ameerjohn S, Kannan S (2019) Integrated probability relevancy classification (IPRC) for IDS in SCADA', design framework for wireless network. Lect Notes Netw Syst 82(1):41–64
38. Das SK, Samanta S, Dey N, Kumar R (2019) Design frameworks for wireless networks. Lecture notes in networks and system. Springer, Singapore, pp 1–439
39. Dey N et al (2017) Developing residential wireless sensor networks for ECG healthcare monitoring. IEEE Trans Consum Electron 63(4):442–449

Performance Analysis of MANET Under Grayhole Attack Using AODV Protocol

Samiran Gupta and Harsh Nath Jha

Abstract Mobile ad hoc network (MANET) has been a challenging field with its foremost criteria like heterogeneity of nodes, dynamic topology, energy constraint and security over the years. MANETs are globally popular for their cost-effectiveness ease of access and configuration. However, MANETs are vulnerable to many types of attacks like Blackhole, Wormhole, Grayhole, etc., which makes MANETs pretty much risky to rely upon when scaling up on a large scale. Under mobile ad hoc networks, all the transmission between the mobile nodes occurs wirelessly. Due to the infrastructure-less, self-organizing and dynamic nature of the nodes, it is an arduous task to enforce any security solutions against these kinds of vulnerabilities. Ad hoc on-demand vector (AODV), a supremely significant route-on-demand routing protocol for MANET, relies on the routing table at each intermediate node location. In this paper, we are mainly analyzing the performance of a MANET under Grayhole attack as per AODV routing protocol using NS-2 simulation environment.

Keywords Mobile ad hoc network · Grayhole attack · Wireless networks · Ad hoc on-demand vector · Smart node · Dynamic routing protocols · Throughput · Quality of service

S. Gupta (✉)
Department of Computer Science and Engineering, Asansol Engineering College, Asansol, West Bengal 713305, India
e-mail: samiran.bappa@gmail.com

H. N. Jha
Department of Information Technology, Asansol Engineering College, Asansol, West Bengal 713305, India
e-mail: ind.harshit@gmail.com

S. K. Das et al. (eds.), *Architectural Wireless Networks Solutions and Security Issues*, Lecture Notes in Networks and Systems 196,
https://doi.org/10.1007/978-981-16-0386-0_15

1 Introduction

Mobile ad hoc network (MANET) [1] is constituted of dynamically self-orienting mobile nodes, making it an infrastructure-less model of network design. These nodes may function as servers as well as clients, as required, demolishing the demand of a dedicated server or router in the network [2]. This provides autonomity to the system, boosting its performance. These nodes have the ability to create a suitable path for the communication channel to form and function. However, MANETs are not a good choice if seen from the point of view of security and integrity of data. The absence of a dedicated server or router may also produce serious security breaches as there is nearly no authentication or encryption available.

MANET as a network arrangement is also much cost effective than the conventional ones, although they cannot be scaled up to a large scale, as despite having undeniably excellent features, have never been a preferable choice as the data in the communication channels are left exposed to tons of security threats and other limitations that are associated with MANETs.

Figure 1 presents a pictorial portrayal of the arrangement of the nodes in the network and their configuration to form a bigger communication channel. We have a total of 7 nodes participating in the channel with no dedicated router or a similar central medium to act as a bridge to connect the nodes. The individual nodes have one or more transceivers between them [3]. The application of MANETs is not as popular in small or medium-sized business or personal/home networks, as compared to a conventional router-driven setup.

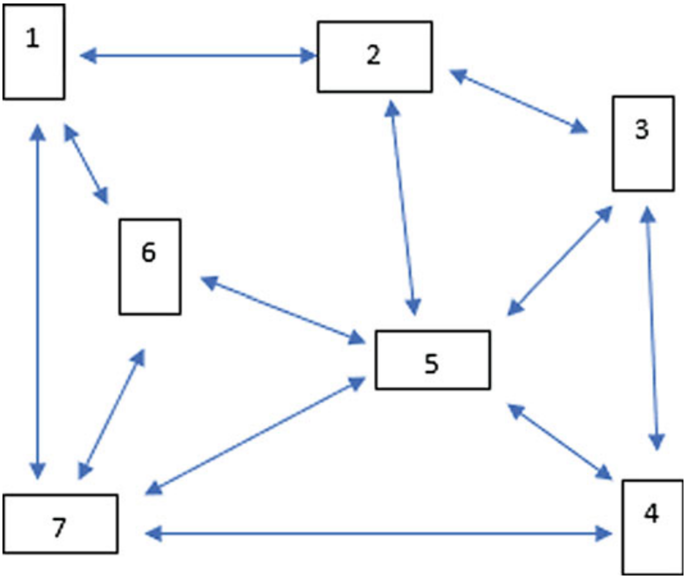


Fig. 1 Overview of MANET

Although MANETs have tons of advantages to count on [4], the flip side of the coin has some serious demerits of MANETs, which are mostly security oriented. These types of networks are mostly prone to Grayhole attack, especially when operating under the AODV protocol.

2 Literature Study

We are presenting an elaborative study on MANET and its characteristics (Sect. 3), its overview (Sect. 3.1) and its working principle and the AODV routing protocol (Sect. 3.2). In Sect. 4, we have presented the problem statement, i.e., about the risks that MANETs are exposed to. Section 5 gives an account on Grayhole attacks and its functional mechanism. For generating a real-like attack on a dummy network arrangement have used NS-2 (network simulation environment), via Linux Fedora. We have discussed the attack environment in Sect. 6, while Sect. 7 contains results and comprehension of our simulation.

We have tried to stress on the risks that MANETs come packaged with, which will facilitate us and also other researchers to come up with better solutions against this kind of attacks and be able to fix those issues, resulting in making MANET a safer communication environment than before.

In this section, we discuss some related and underlying research works by different researchers/authors in the field of wireless networks and Grayhole attacks. Over time, a considerable number of researchers have shared their ideas, findings and conclusions in this subject and also suggested several defense techniques to detect and diffuse Grayhole attacks on ad hoc networks based on intrusion detection systems (IDS) and related mechanisms.

Gupta [5] et al. discuss about sensor networks with regards to load-based routing schemes. Through their work on MANET, Jha et al. [6] shed light on the loss of performance and other security breaches associated with Wormhole attacks. Many authors have worked on Grayhole being launched on ad hoc wireless networks, which itself highlights its severity. Sharma [7] has done a survey on Grayhole attacks on MANETs, which makes it crystal clear that Grayhole attacks can prove deadly in terms of compromising with the network. Dhaka et al. [8] proposed a method to detect Grayhole attacks and Blackhole attacks in MANETs. Later on, Aarti et al. [9] and Mittal [10] have proposed an enhanced multipath approach to deal with the threat of Grayhole attacks on MANETs.

Researchers in this field have made noteworthy breakthrough in this area, but unfortunately, we are yet to have a high-accuracy defense system against Grayhole attack. With regards to the scope for development that we have in this area, we are properly visualizing the attack scenario in a detailed fashion through this paper for (we researchers) being able to develop an enhanced fighting mechanism against Grayhole attacks.

3 Brief Study of Mobile Ad hoc Networks:

i. Dynamic topology: MANET's multi-hop network topology is capable of sudden and spontaneous reorganization in both unidirectional and bi-directional routing architecture.
ii. Cost effective: Being hardware-less and peer-to-peer in nature, MANETs are considerably cheaper for small to medium level business as well as residential networks.
iii. Power supply constraint: Battery led (or similar energy source) power supply being incorporated here is not a reliable or promising source of power per se, and it is the reason why the mobile nodes in the network have light weighted features, low power and storage capacity.
iv. Autonomous/self-configuring: The prime feature of MANETs is the ability of its components (participating nodes) to re-role themselves into routers and hosts themselves.
v. Mediocre throughout: As MANET is a wireless form network arrangement, it struggles against factors like noise, multi-access, interference condition, etc., which dramatically reduces its productivity based on efficiency, throughput and reliability.
vi. Lack of data security: Being infrastructure-less by design, MANETs have no dedicated routers, because of which a standard host configuration or firewall rule-set cannot be enforced. This gives rise to potential threat to the data present in the channel [11] as well as the quality of service (QoS) [12] of the network.

3.1 MANET—Highlights

Wireless ad hoc networks are fairly popular with its users at a mass level. Being a 'plug and play' kind of network setup, MANETs do not require a dedicated router. Although MANET is an awesome mode of network, but it also has some flaws attached with it when implemented at a large scale. Below are some of the forward most ins and outs of MANETs:

i. Infrastructure-less mode of design.
ii. No central administration.
iii. Human intervention independent, as each node can re-purpose themselves as a router or host as needed.
iv. Vulnerable to security threats.
v. Intercommunication interferences causes poor throughput.
vi. Cost effective.

3.2 MANET—Working Principle

MANETs are mostly developed using a table-driven network protocol. AODV protocol [13] is one of the foremost protocols in this matter which enables its nodes to be follow a dynamic, self-configuring and multi-hop routing method. This proves to be a key element in route maintenance. Maintaining routes with inactive nodes are not required because of the dynamic re-routing in AODV.

If there are 5 nodes in a channel and only three of them are participating in an active communication and the remaining two are merely present in the network, then the working nodes need not preserve a route with them. To promote optimal load balancing, AODV supports real-time re-routing and re-orientation of the nodes and avoid any disruption in the channel.

Each node has a specific range till which it can establish communication. This is much similar to a scenario of a classroom where a student from the first bench wants to pass a notebook to their friend at the last bench. Here, the notebook will be passed to the recipient student via many students acting as intermediate sender. If the destination node in MANET is unreachable from the sender, then the nodes use a similar strategy of sending it via multiple intermediate senders. This process is known as multi-hopping in AODV routing premises.

These nodes are designed to be able to re-design the network topology as a response to a security breach, when detected. Once a malicious activity is reported in any node, it is denied permission to perform any action in the communication channel. Again, since this whole process may require some time and until then some sensitive data might already have been compromised; hence, it cannot be accepted as a fail-safe mechanism.

AODV strictly follows a request-reply technique to verify the authenticity of the participants in the network. It contains a few message type definitions such as route requests (RREQs), route replies (RREPs), route errors (RERRs) and acknowledgment (ACK). For every transfer of a data packet, the source generates a route request (RREQs) toward the recipient node and the receiving node replies with an acknowledgment (ACK) of receiving the data in order to prove its authenticity. In case if this process fails, a breach is assumed to have taken place and it leads to broadcasting an error message (RERRs), which immediately suspends all transactions until the node is verified.

AODV routing involves of a couple of episodes:

i. Discovery: Discover new paths using RREQ and RREP.
ii. Maintenance: Report an error when found, using RERR.

AODV protocol maintains a separate routing table per node. Each node's route table contains information about the distance to other nodes in the channel, which is measured in terms of hop-counts. The route table contains the following details gathered while the route discovery phase:

i. Source/previous node
ii. Next node/hop
iii. Time to leave (TTL)
iv. Hop-count to reach destination
vii. Destination IP address.

4 Problem Statement

MANET has many challenges when scaling out on large scale, but it becomes worth a little more concern from the security hotspot as it is vulnerable a plethora of attacks [14, 15]:

i. Session hijacking [16]
ii. Wormhole attack [17]
iii. Blackhole attack [18]
iv. Jamming [19]
v. Eavesdropping [20]
vi. Denial of service [21]
vii. Grayhole attack [22].

Grayhole attack is one of the deadliest attacks against MANETs with regards to:

i. Throughput: The ability of the network to transfer a particular quantity of information per unit time is known as throughput. In other terms, it is the measure of a network's efficiency.
ii. Quality of service (QoS): It is the maximum bandwidth attaining capacity of a network, which affects other parameters such as latency, error rate and uptime [23]. Thus, higher QoS translates to a healthier performance.
iii. Data rate: Also knows as data transfer rate, it is the measure of the number of bits of data transmitted per second over a network. In simpler terms, it is the speed of data transfer over the network, conveyed as bytes per second (Bps or B/s)
iv. Integrity: It enforces that a dataset MUST only be accessed by an authorized and intended user, i.e., if a data is not meant for a particular entity, it must be forbidden for them and it should be private to the legitimate user only [24].

For the sake of analyzing the effects of Grayhole attack on the performance of MANET, we are simulating a dummy network with a number of nodes against a Grayhole attack scenario using AODV routing protocol.

5 Grayhole Attack

Grayhole attack [25, 26] is basically a packet drop attack, which is an extension of Blackhole attack. Here, the routing packets and control are forwarded by the malicious or Grayhole node, but the data packets are completely dropped. This attack uses the method of selective data packet dropping to disguise the compromised node as a legitimate one. This node tries to take part in the data transfer window, and then by advertising a false route, it lures the legitimate nodes to establish the active route through itself. The Grayhole node responds with a route reply after receiving a route request packet and thereby passes a false information that of having the shortest path, which creates an illusion for the source node that the optimum route is through the malicious node and the data packets are redirected toward the malicious node. This series of incidents gives rise to a confusion in the detection and prevention mechanism as packets may as well sometimes drop due to genuine reasons like: congestion, overload, etc. The following are the two ways how Grayhole attacks work:

i. Strictly dropping all the incoming UDP packets.
ii. Randomly/selectively dropping some UDP packets.

Due to its ability to act both as a normal node and switch over to malicious node as needed, a Grayhole node changes its behavior from a legitimate node to a sinkhole, which fools the system to identify whether it is indeed a genuine node or a compromised one. The Grayhole attack takes place in two phases, as below:

i. In this stage, the malicious node exploits the AODV routing protocol table by diverting all the data packets to itself rather than genuine route; thus claiming itself as the shortest route in next hop column.
ii. The attack is launched in this phase where malicious node starts dropping the data packets using a probabilistic method for packet selection. The attacker node changes its behavior rapidly and the malicious node also forward some packets to create an illusion of legitimacy. Hence, this type of attack is pretty difficult to detect.

6 Simulation Environment

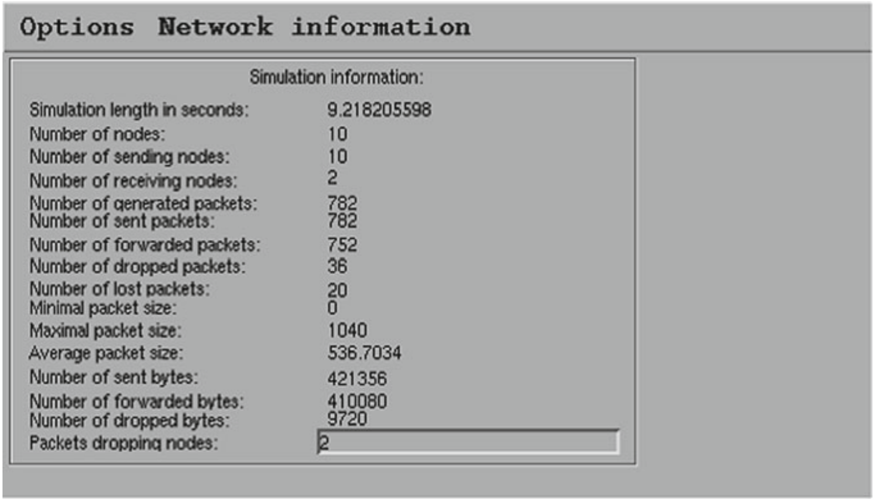
For the purpose of simulation, we are using Network Simulator 2 (NS-2) on a Linux Fedora distribution, which is quite a familiar and popular simulator in MANET research community due to its ease of access and because it supports a variety of network routing protocols. NS-2 is an object-oriented network simulator written using C++ as its backend and object Tcl (OTcl) as its front-end and runs on top of UNIX environment. Below are the details of our attack environment and the parameters at which the system was tuned in to (see Table 1).

Initially, the network is simulated under normal and stable conditions, i.e., without any attack and its throughput is recorded. Later on, we generated an attack of Grayhole

Table 1 Configuration details of the simulation environment

Parameter	Value/type
Number of mobile nodes	10
Link layer type	LL
Antenna type	Omni antenna
Simulation duration	1200 s
Propagation model	Two-way ground
Mobility model	Random waypoint
Interface type	Phy/WirelessPhy
MAC type	Mac/802.11
Interface queue type	Queue/DropTail/PriQueue
Routing protocol	AODV
Channel type	Wireless channel
Simulation area	1000 m × 850 m

nature on the same setup to record and analyze its throughput in order to be able to comprehend the aftermaths of the attack on the network. Here, we noticed that the network throughput drops to zero immediately as soon as the channel in under the attack (Fig. 2).



```
                    Simulation information:
Simulation length in seconds:      9.218205598
Number of nodes:                   10
Number of sending nodes:           10
Number of receiving nodes:         2
Number of generated packets:       782
Number of sent packets:            782
Number of forwarded packets:       752
Number of dropped packets:         36
Number of lost packets:            20
Minimal packet size:               0
Maximal packet size:               1040
Average packet size:               536.7034
Number of sent bytes:              421356
Number of forwarded bytes:         410080
Number of dropped bytes:           9720
Packets dropping nodes:            2
```

Fig. 2 Simulation information of our dummy network

Throughput: The average amount of data transferred between the sender and receiver nodes per unit time within a network is called throughput. It is expressed in terms of kilobytes per second (kbps) and calculated using the following equation (see Eq. 1).

$$\text{Throughput} = \frac{\text{Data transferred (in bytes)* 8}}{\text{Time taken (in seconds)}} \tag{1}$$

7 Attack Simulation and Results

Scenario 1: Without Attack

I. Deploying the mobile of nodes

As shown in Fig. 1, we started deploying nodes to participate in the network. There are no dedicated routers or a similar central administration device present in the network (Fig. 3).

II. A self-arranged ad hoc networks by the deployed nodes

Now as we have deployed a total of 10 nodes in the network (Fig. 4), namely 0 till 9, we observe the nodes interacting with each other as per the AODV protocol.

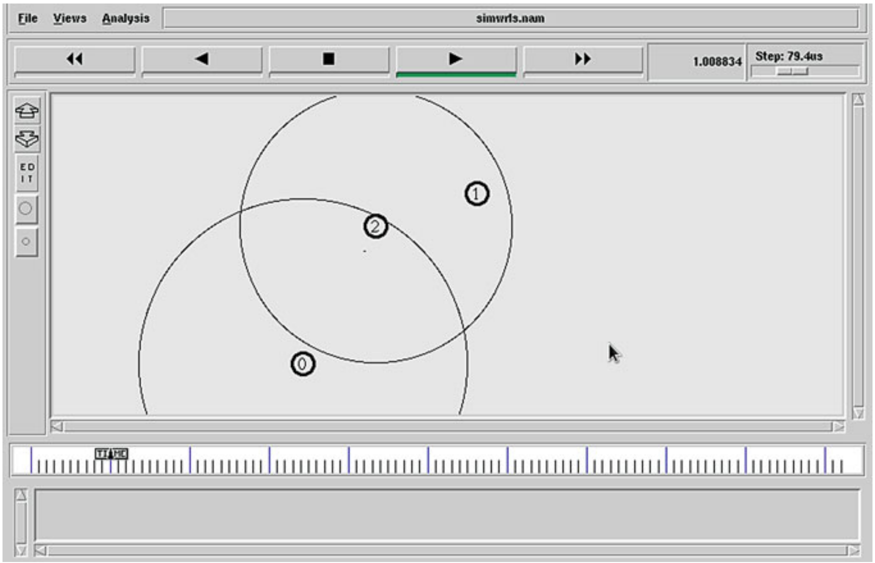


Fig. 3 Deploying the nodes

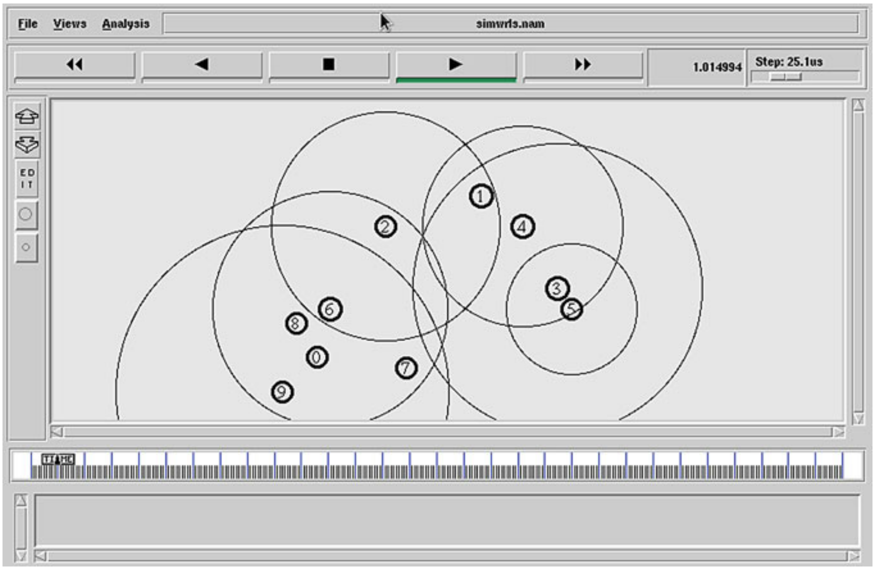


Fig. 4 Individual nodes forming an ad hoc network

III. Identifying the source, destination and forwarder nodes

MANETs are dynamic and self-organizing in nature, i.e., it decides the communication path, thus as seen in Fig. 5, we have a source, a destination and an intermediate node to facilitate the communication as the source and destination are not reachable to each other directly. Whenever there is a scenario like this where the source and destination nodes are unreachable (as their reachable zone is limited), MANETs adapt a multi-hop mechanism to transfer data.

IV. Communication via the established path

In the previous step, we already had our source, node and intermediate nodes identified. In this step, we can actually see the data transfer in action (Fig. 6). There is no loss of data and the communication is happening smoothly. This is an ideal case, without any attack, characterized by a stable throughput and QoS.

Scenario 2: Network under attack

V. Malicious node starts dropping packets

Until now, we were simulating the best case for data transfer with optimum throughput. At this point of time, we launched a Grayhole attack on the network with node ‘2’ dropping the data packets. It is pretty obvious from Fig. 7 that the communication is still happening but the data rate is considerably lower than before, as a lot of data is being drained by the malicious node.

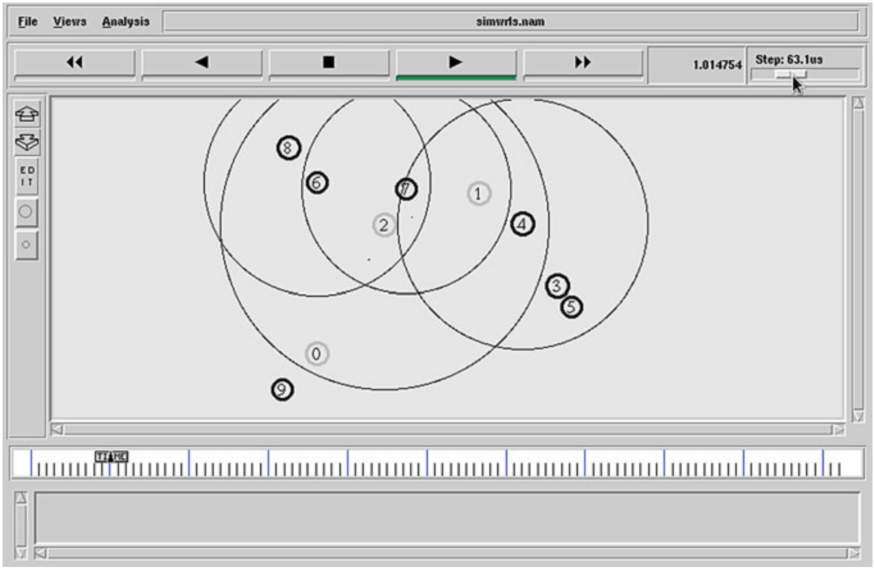


Fig. 5 Identifying sender, receiver and intermediate nodes

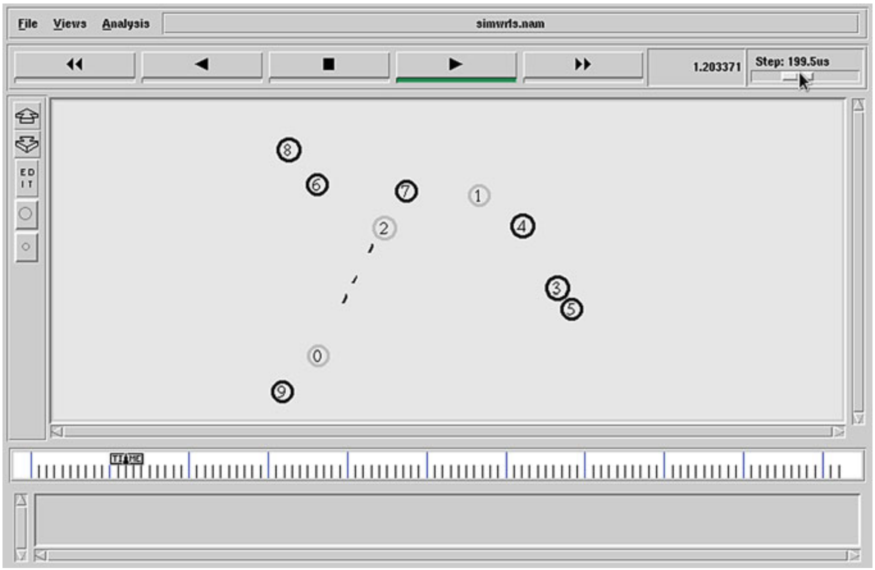


Fig. 6 Data transfer under normal circumstances

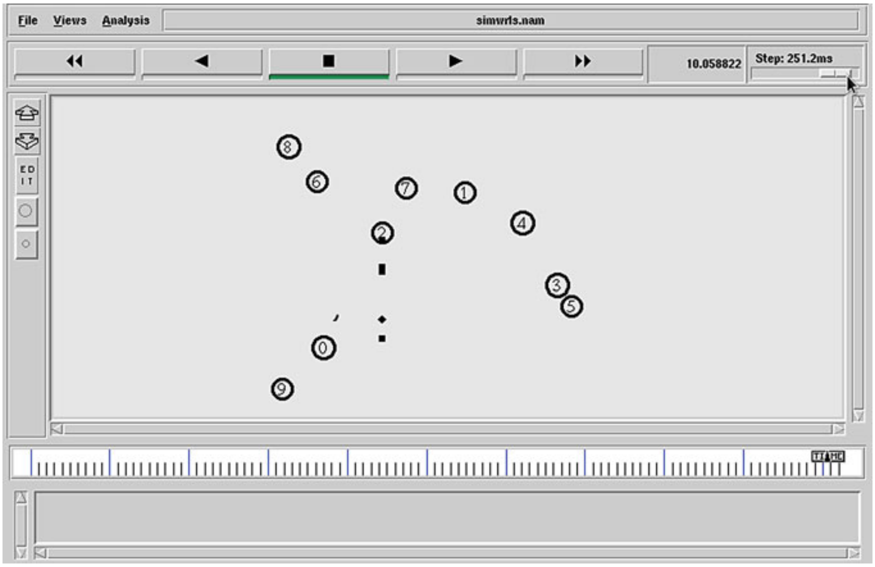


Fig. 7 Malicious node starts dropping data

VI. Loss in throughput

Evident from Fig. 8, we can confirm what we saw in the previous step. The channel's throughput spikes fall miserably, and at the same moment, we launched attack and that continued until the attack persisted.

VII. Trace file of the network scenario at the moment

Figure 9 shows the network trace of the above attack simulation of MANET. From the simulation presented above, it is clear as a mirror that Grayhole attack is indeed a prominent vulnerability to MANETs. These kinds of attacks not only put the data integrity at stake because of the possible leakage in the communication channel, but also pose great threat to the network as a whole, in terms of overall productivity.

8 Conclusion

After a detailed analysis of the performance of MANETs under the effects of a Grayhole attack using AODV protocol via NS-2 simulator, our final inference is that these kinds of ad hoc networks have a strictly linear throughput trend which starts deteriorating dramatically under an attack. Along with throughput, other factors like data rate, QoS, etc., parameters of the network were also affected at an alarming level to be considered as abnormal and concerning. The data transfer within the network kept on falling as long as the attack was kept alive on the network.

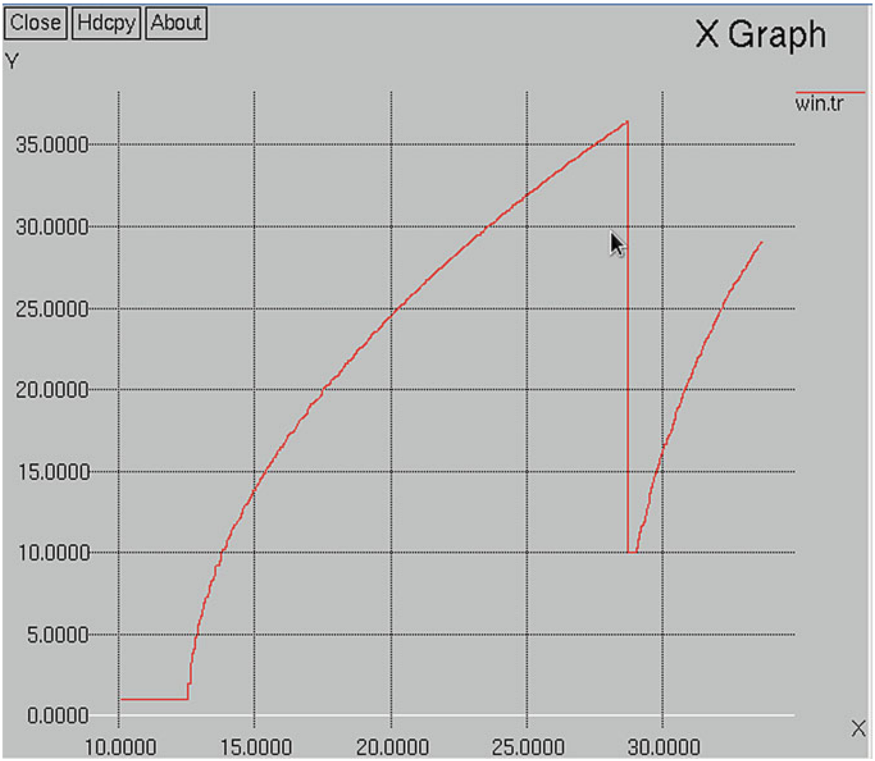
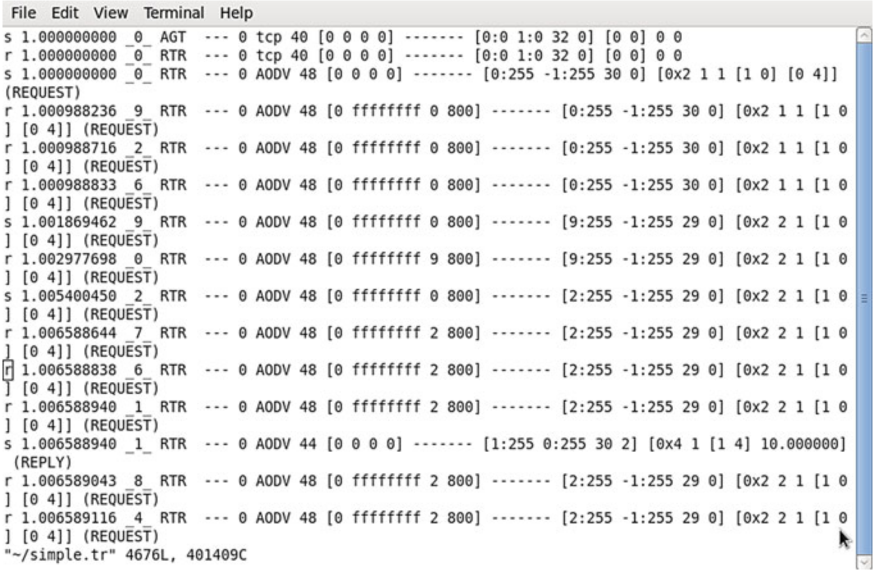


Fig. 8 Network throughput during attack

From the above analysis, it is clearly understandable how a Grayhole attack cannot only hamper the network QoS and throughput, but bully privacy as well. Grayhole attacks are difficult to detect also because the data rate does not drop to zero at once, i.e., the communication keeps on taking place, but it degrades slowly and steadily which also might be misunderstood as a usual network glitch such as channel noise or interference. Until one smells anything fishy, a lot of data might already have been leaked. However, with further advancements in the MANET's immune system and an improved intrusion detection system, it can be guarded against Grayhole attacks.



```
s 1.000000000 _0_ AGT  --- 0 tcp 40 [0 0 0 0] ------- [0:0 1:0 32 0] [0 0] 0 0
r 1.000000000 _0_ RTR  --- 0 tcp 40 [0 0 0 0] ------- [0:0 1:0 32 0] [0 0] 0 0
s 1.000000000 _0_ RTR  --- 0 AODV 48 [0 0 0 0] ------- [0:255 -1:255 30 0] [0x2 1 1 [1 0] [0 4]]
(REQUEST)
r 1.000988236 _9_ RTR  --- 0 AODV 48 [0 ffffffff 0 800] ------- [0:255 -1:255 30 0] [0x2 1 1 [1 0
] [0 4]] (REQUEST)
r 1.000988716 _2_ RTR  --- 0 AODV 48 [0 ffffffff 0 800] ------- [0:255 -1:255 30 0] [0x2 1 1 [1 0
] [0 4]] (REQUEST)
r 1.000988833 _6_ RTR  --- 0 AODV 48 [0 ffffffff 0 800] ------- [0:255 -1:255 30 0] [0x2 1 1 [1 0
] [0 4]] (REQUEST)
s 1.001869462 _9_ RTR  --- 0 AODV 48 [0 ffffffff 0 800] ------- [9:255 -1:255 29 0] [0x2 2 1 [1 0
] [0 4]] (REQUEST)
r 1.002977698 _0_ RTR  --- 0 AODV 48 [0 ffffffff 9 800] ------- [9:255 -1:255 29 0] [0x2 2 1 [1 0
] [0 4]] (REQUEST)
s 1.005400450 _2_ RTR  --- 0 AODV 48 [0 ffffffff 0 800] ------- [2:255 -1:255 29 0] [0x2 2 1 [1 0
] [0 4]] (REQUEST)
r 1.006588644 _7_ RTR  --- 0 AODV 48 [0 ffffffff 2 800] ------- [2:255 -1:255 29 0] [0x2 2 1 [1 0
] [0 4]] (REQUEST)
r 1.006588838 _6_ RTR  --- 0 AODV 48 [0 ffffffff 2 800] ------- [2:255 -1:255 29 0] [0x2 2 1 [1 0
] [0 4]] (REQUEST)
r 1.006588940 _1_ RTR  --- 0 AODV 48 [0 ffffffff 2 800] ------- [2:255 -1:255 29 0] [0x2 2 1 [1 0
] [0 4]] (REQUEST)
s 1.006588940 _1_ RTR  --- 0 AODV 44 [0 0 0 0] ------- [1:255 0:255 30 2] [0x4 1 [1 4] 10.000000]
 (REPLY)
r 1.006589043 _8_ RTR  --- 0 AODV 48 [0 ffffffff 2 800] ------- [2:255 -1:255 29 0] [0x2 2 1 [1 0
] [0 4]] (REQUEST)
r 1.006589116 _4_ RTR  --- 0 AODV 48 [0 ffffffff 2 800] ------- [2:255 -1:255 29 0] [0x2 2 1 [1 0
] [0 4]] (REQUEST)
"~/simple.tr" 4676L, 401409C
```

Fig. 9 Network trace while the attack

References

1. Aarti SST (2013) Study of MANET: characteristics, challenges, application and security attacks. Int J Adv Res Comput Sci Softw Eng 3(5)
2. Ali H, Shahzad W, Khan FA (2012) Energy-efficient clustering in mobile ad-hoc networks using multi-objective particle swarm optimization. Appl Soft Comput 12(7):1913–1928
3. Cho J-H, Chen R, Chan KS (2016) Trust threshold based public key management in mobile ad hoc networks. Ad Hoc Netw 44:58–75
4. Das SK, Tripathi S (2018) Intelligent energy-aware efficient routing for MANET. Wirel Netw 24(4):1139–1159. https://doi.org/10.1007/s11276-016-1388-7
5. Gupta S, Das B (2013) Load based reliable routing in multi-sink sensor networks. Res Inventy Int J Eng Sci 2(12):59–64
6. Jha HN, Gupta S, Maity D (2020) Effect of wormhole attacks on MANET. In: Design frameworks for wireless networks. Springer, Singapore. https://doi.org/10.1007/978-981-13-9574-1_8
7. Sharma R (2016) Int J Comput Sci Inf Technol 7(3):1457–1460 (IJCSIT)
8. Dhaka A, Nandal A, Dhaka RS (2015) Gray and black hole attack identification using control packets in MANETs. Procedia Comput Sci 54:83–91. ISSN 1877–0509
9. Aarti PR (2015) Prevention and elimination of gray hole attack in mobile ad-hoc networks by enhanced multipath approach. Int J Eng Trends Technol (IJETT) 23(5):224–229. ISSN:2231–5381
10. Mittal V (2015) Prevention and elimination of gray hole attack in mobile ad-hoc networks by enhanced multipath approach. Int J Adv Res Comput Eng Technol (IJARCET) 4(5)
11. Yang H, Luo H, Ye F, Lu S, Zhang L (2004) Security in mobile ad hoc networks: challenges and solutions. UCLA Previously Published Works
12. Castellanos WE, Guerri JC, Arce P (2015) A QoS-aware routing protocol with adaptive feedback scheme for video streaming for mobile networks. Comput Commun 77:10–25

13. Perkins CE, Royer EM (1999) Ad-hoc on-demand distance vector routing. In: Proceedings of the 2nd IEEE workshop on mobile computing systems and applications, pp 90–100
14. Sheikh R, Chande MS, Mishra DK (2010) Security issues in MANET: a review. IEEE
15. Goyal P, Parmar V, Rishi R (2011) MANET: vulnerabilities, challenges, attacks, application. IJCEM Int J Comput Eng Manage 11:32–37
16. Lupu TG, Rudas I, Demiralp M, Mastorakis N (2009) Main types of attacks in wireless sensor networks. In: WSEAS international conference. Proceedings. Recent advances in computer engineering
17. Maheshwari R, Gao J, Das SR (2007) Detecting wormhole attacks in wireless networks using connectivity information. In: IEEE INFOCOM 2007–26th IEEE international conference on computer communications, pp 107–115
18. John NP, Thomas A (2012) Prevention and detection of black hole attack in AODV based mobile ad-hoc networks—a review. Int J Sci Res Publ 2(9)
19. Xu W, Trappe W, Zhang Y, Wood T (2005) The feasibility of launching and detecting jamming attacks in wireless networks. In: Proceedings of the 6th ACM international symposium on mobile ad hoc networking and computing, pp 46–57
20. Dai HN, Wang Q, Li D, Wong RC (2013) On Eavesdropping attacks in wireless networks with directional antennas. Int J Distrib Sens Netw 9(8):760834
21. Jhaveri RH, Patel SJ, Jinwala DC (2012) DoS attacks in mobile ad hoc networks: a survey. In: 2nd international conference on advance computing and communication technologies, pp 535–541
22. Sen B, Sharma K, Ghose MK, Sharma A (2015) Gray hole attack in manets. Int J Adv Electron Comput Sci 2(10). ISSN: 2393–2835
23. Basarkod PI, Manvi SS (2015) Mobility and QoS aware routing in mobile ad hoc networks. Comput Electr Eng 48:86–99
24. Ning P, Sun K (2005) How to misuse AODV: a case study of Insider attacks against mobile ad-hoc routing protocols. Ad Hoc Netw 3(6):795–819
25. Nguyen HL, Nguyen UT (2008) A study of different types of attacks on multicast in mobile ad hoc networks. J Ad Hoc Netw 6:32–46
26. Deng H, Li W, Agrawal DP (2002) Routing security in wireless Ad Hoc networks. IEEE Commun Mag 40(10):70–75

Technique to Reduce PAPR Problem in Next-Generation Wireless Communication System

Abhishek Kumar, Vishwas Mishra, Shobhit Tyagi, Priyanka Saini, and Nikhil Saxena

Abstract When the fourth generation for wireless communication networks was developed, it was upgraded to provide both enhanced coverage area and higher data rates to every mobile user with lower latency. However, wireless communication system for the next-generation network will need to challenge new requirement with a greater diversity of application requirements such as ultra-high data rate, ultra-low latency, flexible use of spectrum and spectrum sharing, and battery-powered sensors that needs extremely low energy consumption, and some other control applications that want a very short round trip time (RTT). Due to problems with orthogonal frequency division multiplexing (OFDM) and next-generation demands, OFDM is not used as a promising waveform for next-generation wireless communication network. In these circumstances, alternative multiplexing schemes such as generalized frequency division multiplexing (GFDM), due to flexibility in pulse shape and single cyclic prefix in a multi-path system, GFDM is becoming common every day, making it eligible for 5G wireless technologies. GFDM looks as generalization of OFDM technique. But one of the common drawbacks of every multicarrier system is their high peak to average power ratio (PAPR). The main effect of strong PAPR is instability in the analog to digital converter (ADC) and digital to analog converter (DAC), decreased its performance and raised costs. A PAPR reduction technique such as clipping and filtering that greatly improves the efficiency compared to the initial GFDM signal PAPR. Overall peak re-growth can be reduced by using repeated clip and filter operations. Simulation is performed for this scheme to evaluate this system's PAPR output for different values of roll-off variables.

Keywords Wireless network · Out of band · Generalized frequency division multiplexing · Peak to average power ratio · Signal-to-noise ratio

A. Kumar (✉) · V. Mishra · S. Tyagi · P. Saini
Swami Vivekananda Subharti University, Meerut 250005, India

N. Saxena
University of Cincinnati, Cincinnati, OH 45221, USA

S. K. Das et al. (eds.), *Architectural Wireless Networks Solutions and Security Issues*, Lecture Notes in Networks and Systems 196,
https://doi.org/10.1007/978-981-16-0386-0_16

1 Introduction

Today, wireless communication is one of the fastest growing fields in the world of engineering. Rapid development in the domain of wireless communication system, services and application has drastically altered the way we live, communicate and works. Wireless communication has several variations in term of applications like ad hoc network and wireless sensor network [1–3]. It has several design frameworks and nature-inspired applications and metaheuristic application that solves the several issues of the networks [4–6]. Wireless communication provides a dynamic and broad technological field, which has stimulated incredible excitement and technological progression over last few decades. Initially, communication services offered only voice, but currently along with voice, high data rate transmissions, multimedia service and Internet application are also supported [7, 8].

Currently, 5G (5th cellular networks generation or 5th mobile communication generation) is a term used in some research papers and 5th cellular networks generation or 5th mobile communication generation is a term used in some research papers and initiatives to describe the next major step of wireless communication technology beyond the current 4G standards. Nevertheless, 5G is not an official term used by standardization bodies for any specification or official document such as Wi-MAX Forum, 3GPP or ITU-R or various telecommunication companies around the world. New versions of the existing 3GPP including 4G and LTE Advanced are being introduced; however, they are not considered to be new generations. All of the previous wireless technologies enjoy the convenience of telephone and data sharing, but 5G makes real mobile life. It is expected to improve its offerings and applications [9]. The set of challenges in 5G is given by $C = \{C1, C2, C3, C4, C5, C6, C7\}$.

C1 Serving large amount of user

C2 Effective use of spectrum

C3 Supporting traffic volume 1000 × in ten years

C4 Supporting large number of user

C5 Reducing power consumption

C6 Security cloud

C7 Cost factor

The set of features in 5G is given by $S = \{F1, F2, F3, F4, F5, F6, F7\}$.

F1 Massive MIMO

F2 Reduced latency

F3 Very high data rate

F4 Very high capacity

F5 Improved securities

F6 Device to device communication

F7 Long battery lifetime

We need a dramatic shift in the ideas of cellular architecture to overcome the above difficulties and satisfy the preconditions of 5G network requirements. One of the key ideas for 5G cellular network architecture is to separate the situation outside and indoors. We have been in indoor condition since 80 percent of the time, and there is a huge loss of penetration by construction. So we have got to stay away from that. This is related to the distributed antenna system (DAS) and the massive MIMO network [10].

OFDM is not appropriate for 5G as a result of its low spectral efficiency and synchronization issue. The approach of new applications in 5G wireless systems, for example, Internet of things (IoT), machine-to-machine (M2M) communication and media-rich high transmission capacity applications requires the requirement for new waveforms. Compared to OFDM, signals which present in these waveforms are more localized in frequency and time. In this voyage, we need to search for new waveform contenders for the physical layer (PHY) of 5G correspondence. Some important promising waveform contenders are generalized frequency division multiplexing (GFDM), filter band multicarrier (FBMC), universal filtered multicarrier (UFMC) and bi-orthogonal frequency division multiplexing (BFDM) [10].

With FBMC, the subcarriers have a separate pulse shape to lower the OOB emissions, [11] one of the most examined multicarrier filter systems. From response of transmit filter, length is usually long, as the subcarriers have limited transmission capacity. As a result, only when the large number of symbols is transmitted, then FBMC can achieve good spectral efficiency. Obviously, the above arrangement is not useful for situations of low latency. UFMC is the ongoing proposal on filtering subcarriers groups to reduce OOB emissions. UFMC needs no CP that means time lag is more responsive. It is therefore not suitable for applications where time synchronization is needed to save energy [12]. BFDM [13] uses well-localized pulse shapes on both sides of the transmitter and receiver. In order to achieve entirely localized pulses in both the frequency and time domains, BFDM uses offset QAM (OQAM). Therefore, it is not easy to integrate very high density BFDM with MIMO, one of the key features of 5G technology.

The benefits of the GFDM technique [14] come at the disadvantage of a contrasted enhanced bit error rate (BER) and OFDM procedure whose deterioration is attributable to the way; GFDM is a non-orthogonal waveform. Subsequently, the non-orthogonality of the adjacent subcarriers and time slots results in self-interference. Matched filter (MF), zero forcing (ZF) or minimal mean square error (MMSE) receptor can be extracted to tackle this self-impedance. There are different benefits of GFDM which are given below.

i. Provides the flexibility required by 5G applications
ii. Large spectral and energy efficiency
iii. Block-based transmission using insertion of cyclic prefix
iv. Robust to time synchronization
v. Frequency and time domains multi-user scheduling
vi. To support high mobility and large diversity
vii. Reduce the power consumption
viii. Application like IoT, Tactile Internet, M2M communication, etc.

Yet PAPR is one of the GFDM system's big concerns, which is strong in transmitted signal side that corrupts system output while utilizing a nonlinear high power amplifier (HPA). It is necessary to use the proper PAPR reduction technique on the transmitter side along these lines.

This chapter is explained as following. Section 2 describes the concept of the GFDM device and its operation. The description of the PAPR issue and its results are described in Sect. 3 and the suggested PAPR reduction strategy and the approximation of the outcome of this PAPR reduction for the various roll factor values are described in Sect. 4. The conclusion is ultimately presented in Sect. 5.

2 GFDM System Description

For the 5G PHY layer, GFDM is a promising solution as it can handle a number of prerequisites. To satisfy certain latency conditions [14] in block structure of the MK samples, where subcarriers K convey M sub-symbol, the signal duration must be reduced to meet certain latency requirements [15]. The time–frequency structure can be established to manage the time limits of low latency applications. Thus, this scheme retains all the main benefits of OFDM, but adds additional implementation complexity.

Fettweis et al. proposed GFDM, which is a flexible multicarrier modulation technique [16]. It has also been proposed for 5G networks air interface. GFDM's flexibility makes it possible to cover CP-OFDM and unique carrier frequency domain situations. When $M = 1$, GFDM modifies the OFDM. For $K = 1$ and SC-FDM, SC-FDE is acquired. The important property recognizing the scheme proposed by OFDM and SC-FDE is, however, that similar to SC-FDM it allows the separation in K subcarriers and M sub-symbols of a given time–frequency resource. GFDM relies on the modulation of the independent block. Each block is made up of different subcarriers and sub-symbols [14, 17].

Subcarriers are filtered with a time and frequency domains cyclically shifted prototype filter. This method eliminates OOB emissions and allows the allocation of fragmented and complex spectrum to existing services and other users without significant interference. Subcarrier filtering can produce subcarriers that are non-orthogonal and generate ISI and ICI [13]. This interference can be avoided by an appropriate acceptance process. For example, in different channel models, a matched filter receiver with the iterative interference cancelation [14] can achieve a better SER. From Fig. 1, the OFDM frame has a cyclic prefix for transmitting all K subcarrier symbols and the GFDM frame has a single CP for sub-symbols, all of which we can see are included in K subcarrier. This is possible because a single CP is sufficient to ensure that the linear conversion with the channel is equal to the circular conversion. The MK constellation symbols can be transmitted using only one CP with GFDM, but OFDM requires an M CP interval to transmit MK constellation symbols. In that case, the loss of efficiency due to OFDM and GFDM CP overhead can be described as [18]:

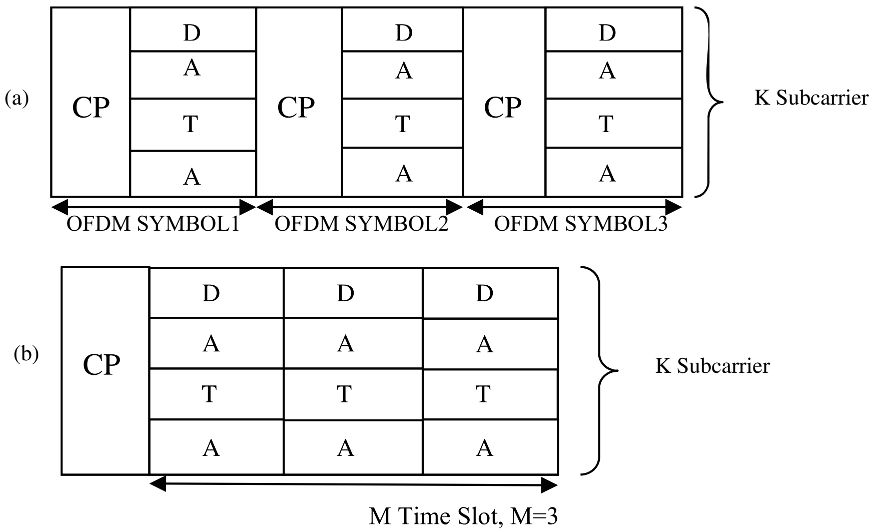


Fig. 1 **a** Three sequential OFDM symbols in one OFDM frame and **b** three sequential GFDM symbols in one GFDM frame

$$\eta_{\text{OFDM}} = \frac{t_{\text{symbol}}}{t_{\text{cp}} + t_{\text{symbol}}} \tag{1}$$

$$\eta_{\text{GFDM}} = \frac{M * t_{\text{symbol}}}{t_{\text{cp}} + M * t_{\text{symbol}}} \tag{2}$$

If t_{symbol} is OFDM symbol length and $1/M$ of GFDM frame duration excluding the CP interval duration is equal to OFDM and GFDM transmission bandwidth.

2.1 GFDM System

The block diagram or signal model are shown in Fig. 2 [19], $\vec{b}$ bits of data given by data source and the data is encoded and mapped. In other words, for this function, M-QAM is the modulation used to symbolize a complicated constellation measured at $2\hat{}\mu$, where μ is the order of modulation. Therefore, vector is obtained, where the block of data consists of elements N. Every factor can be divided into subcarriers K and sub-symbols M that relate as $N = \text{KM}$, leading to $\vec{d} = \left(\vec{d}_0^T, \vec{d}_1^T, \ldots, \vec{d}_{M-1}^T\right)^T$, where each vector is denoted by symbol $\vec{d}_m = \left(\vec{d}_{0,m}^T, \vec{d}_{1,m}^T, \ldots, \vec{d}_{K-1,m}^T\right)^T$. Meaning of T is transpose of matrix. Each and every entity corresponds to the data transmitted on the subcarrier kth and sub-symbol mth in the block, and of these data symbols

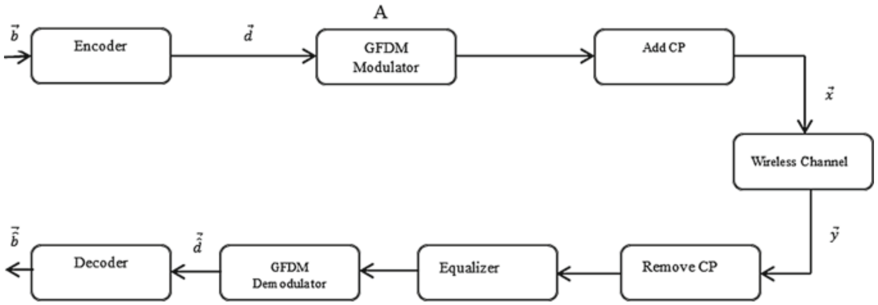


Fig. 2 GFDM transceiver block diagram

are transmitted in domain time and frequency as independently followed with a corresponding pulse shape, where n is defined as the sampling index.

2.2 Modeling of GFDM System

Transmitted signal in GFDM transmitter is given as follows [14]:

$$x[n] = \sum_{m=0}^{M-1} \sum_{k=0}^{K-1} d_{k,m} p\big[(n - MK)_N\big] g_k[n] \tag{3}$$

Various operations inside GFDM modulator as shown in Fig. 3 [19]. If M corresponds to the total number of GFDM symbols within the GFDM frame, K is equal to the total number of GFDM symbols within the GFDM frame [20]. From Eq. (3), $p\big[(n - MK)_N\big]$ represents the circular shift impulse response of pulse shaping filter $p[n]$ by MK with modulo N. $g_n[n]$ is the complex exponential multiplier that shifts the baseband spectrum to the location of the subcarrier kth in Eq. (4) [14].

$$g_n[n] = e^{-j2\pi \frac{kn}{N}} \tag{4}$$

Performance of the GFDM transmitter, $x[n]$ can be written in Eq. (3). Defining $p_m[n] \triangleq p[(n - mk)_N]$ can be given as

$$x[n] = \sum_{m=0}^{M-1} \sum_{k=0}^{K-1} d_{k,m} p_m[n] g_k[n] \tag{5}$$

$$\vec{x} = A\vec{d} \tag{6}$$

where

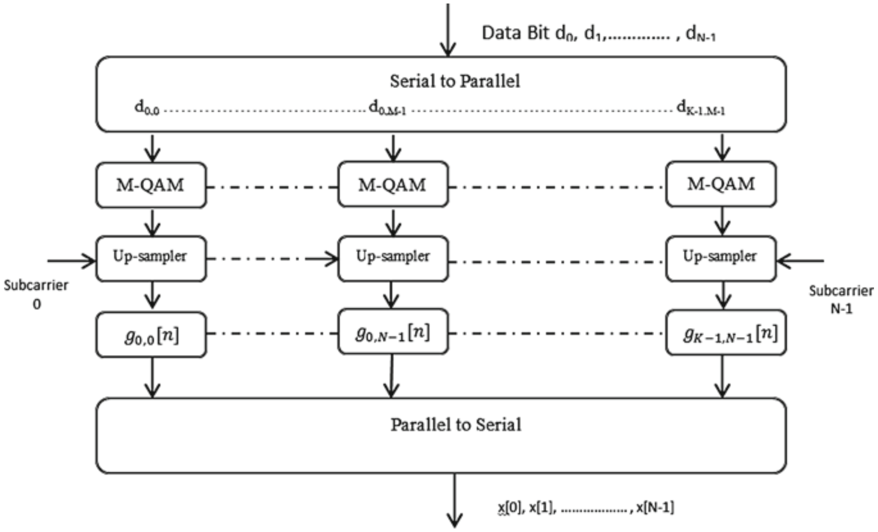


Fig. 3 GFDM modulator

$$\vec{x} = \begin{bmatrix} x(0) \\ x(1) \\ \vdots \\ x(N-1) \end{bmatrix}, \overrightarrow{d} = \begin{bmatrix} d_{0,0} \\ d_{1,0} \\ \vdots \\ d_{K-1,0} \\ d_{0,1} \\ d_{1,1} \\ \vdots \\ d_{K-1,M-1} \end{bmatrix},$$

$$A = \begin{bmatrix} p_{0,0(0)} & p_{1,0(0)} & \cdots\cdots\ p_{K-1,0(0)} \\ p_{0,1(1)} & p_{1,1(1)} & \cdots\cdots\ p_{K-1,1(1)} \\ \vdots & \vdots & \ldots\ldots\vdots \\ \vdots & \vdots & \ldots\ldots\vdots \\ p_{0,M-1(N-1)} & p_{1,M-1(N-1)} & \cdots\cdots\ p_{K-1,M-1(N-1)} \end{bmatrix}$$

where matrix A is defined as transmitter matrix with KM × KM dimension.

2.3 Pulse Shaping Filter

Due to the multi-path propagation in wireless communication, signals spread in time domain and therefore ISI takes place. It significantly increases the error probability at the receiver side. It creates a major obstacle in high data rate communication. The effects of ISI can be minimized, if the symbol duration is allowed to be sufficiently large but it lowers the data rate performance and poor bandwidth efficiency. Techniques are desirable to minimize the effects of ISI without lowering bandwidth efficiency. Pulse shaping is one of such techniques which used before transmission which help to suppress OOB radiation. Before transmission through channel, the signal is passed through a pre-modulated filter which lowers the amount of radiation into the adjacent bands [7].

Rectangular pulses are used in OFDM. Its frequency response theoretically extends to infinity. Practically transmission of rectangular pulses incurs high side lobes which causing significant amount of out-of-band radiations. This results into undesirable ACI. ACI depends on the spectral characteristics of a transmitted signal. A signal with good spectral characteristics should depend on the transmission power within the transmission bandwidth.

Nyquist Criteria for pulse shaping.

Nyquist found out that the effect of ISI can be completely mitigated if the signal is passed through a filter with an appropriate frequency response. Theoretically, it is proved that the minimum bandwidth required to transmit a signal with R symbols per second without ISI is $\frac{R}{2}$ Hz. This becomes possible when the transfer function $S(f)$ assumes a rectangular shape. Then, its single sideband becomes equal to $\frac{1}{2}T$. Its impulse response is represented by $s(t)$ which is a sincfunction. The different pulse shaping filter used in the GFDM are as,

(1) **Raised Cosine**: The most common pulse shaping filter used in mobile communication. The frequency response $S(f)$ is equal to the composite frequency response of transmitter, channel and receiver. Using this filter, ISI can be completely removed if $S(f)$ is designed such that

$$S(f) = \begin{cases} \frac{1}{2B}, & \text{for } 0 \leq |f| \leq f_1 \\ \frac{1}{4B}\left[1 + \cos\left(\frac{\pi}{2B\alpha}(|f| - B(1-\alpha))\right)\right], & \text{for } f_1 \leq |f| \leq 2B - f_1 \\ 0, & \text{for } |f| \geq 2B - f_1 \end{cases} \tag{7}$$

α is the roll-off factor which is defined as $\alpha = 1 - \frac{f_1}{B}$, points to the excess amount of bandwidth 'B' over an ideal shape of a pulse (when $\alpha = 0$). Its indicate the minimum Nyquist bandwidth. As α increases, the required amount of bandwidth also increases. A special case, $\alpha = 1$ indicates that the required amount of excess bandwidth is 100%, this is called full cosine off pulse.

(2) **Root Raised Cosine**: It is variant of RC pulse shaping filter. It is derived by taking the square root of the frequency response of the RC pulse shaping filter. Its waveform occupies a larger dynamic range than a RC filtered waveform. The frequency response $S(f)$ of RRC can be expressed as

$$S(f) = \begin{cases} \frac{1}{\sqrt{2B}}, & \text{for } 0 \le |f| \le f_1 \\ \frac{1}{\sqrt{2B}}\left[1 + \cos\left(\frac{\pi}{4B\alpha}(|f| - B(1-\alpha))\right)\right], & \text{for } f_1 \le |f| \le 2B - f_1 \\ 0, & \text{for } |f| \ge 2B - f_1 \end{cases} \tag{8}$$

3 PAPR Problem

PAPR is characterized as the ratio between the maximum power and the average power for the envelope of a baseband and passband complex signal, $\tilde{s}(t)$, i.e.,

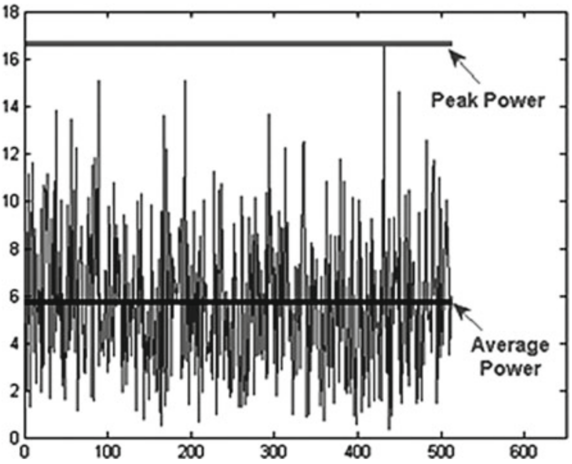
$$\mathrm{PAPR}\{\tilde{s}(t)\} = \frac{\max|\tilde{s}(t)|^2}{E|\tilde{s}(t)|^2} \tag{9}$$

$$\mathrm{PAPR(d}B) = 10\log_{10}(\mathrm{PAPR}) \tag{10}$$

PAPR is also known as crest factor. CCDF is to be calculated as a signal PAPR, because the CCDF is one of the most commonly used PAPR reduction techniques. CCDF means complementary cumulative distribution function [21].

Due to the wide dynamic range of waveforms, PAPR occurs. High PAPR essentially emerges as a result of IFFT pre-processing. Data symbols in all subcarriers here generate high peak signals as shown in Fig. 4.

Fig. 4 Peak signal and average signal



3.1 Complementary Cumulative Distribution Function

The cumulative distribution function (CDF), or the distribution function, is likely to be found in probability theory and statistics, with a value that is not exactly matched, or similar to 'x' represents CDF, to determine the distribution of multivariate random variables. Equation (11) shall be used for each real number 'x' of the CDF of the random variable indicated by 'X'.

$$F_x(x) = P(X \leq x) \tag{11}$$

where on the right side of the equation, the random variable 'X' has a value equal to or below 'x'. The CDF of 'X' is definable as the density function frequency, f as defined in Eq. (12).

$$F(x) = \int\limits_{-\infty}^{x} f(t)\mathrm{d}t \tag{12}$$

where $F(x)$ is cumulative distribution function and$f$(t) is probability density function. The CCDF referred to by $F_\mathrm{c}(x)$ defines the probability of a real random 'X' with a given probability distribution being found to a value greater than 'x'. Equation (13) describes the equation of CCDF of 'X'.

$$F_\mathrm{c}(x) = P(X > x) = 1 - F(X) \tag{13}$$

$F(x)$—Real and imaginary parts of the time domain are given as a mean of zero and a variance of 0.5 in order to carry out, to the degree that a central limit theorem has been performed for a multicarrier signal with a large number of subcarriers (K). In this way, Rayleigh distribution is used for the multicarrier signal amplitude in which a central, two-degree chi-square distribution is used for the distribution of power in the system. At that point, the CDF of the signal sample amplitude is given by [22],

$$F(X) = \left(1 - \mathrm{e}^{-x^2}\right) \tag{14}$$

PAPR's CCDF for data block non-oversampling is given as:-

$$P(\mathrm{PAPR} > x) = 1 - F(X)^K = 1 - \left(1 - \mathrm{e}^{-x^2}\right)^K \tag{15}$$

where it has assumed to be a large number of subcarriers.

3.2 PAPR in GFDM

A discrete GFDM signal [19] is given as

$$x[n] = \sum_{m=0}^{M-1} \sum_{k=0}^{K-1} d_{k,m} p\big[(n - mK)_N\big] g_k[n] \tag{16}$$

$$g_n[n] = \mathrm{e}^{-j2\pi \frac{kn}{N}} \tag{17}$$

where $K \times M = N$ and $p_{k,m}[n] = p[(n - mk) \bmod N] \cdot \mathrm{e}^{-j2\pi \frac{kn}{N}}$ are pulse shape filter $p[n]$ time and frequency conversion equation. $d(k, m)$ is the kth subcarrier data in the block sub-symbol [23]. Therefore, the expression for PAPR of GFDM signal is

$$\mathrm{PAPR}(\mathrm{d}B) = 10 \log_{10} \frac{\max_{0 \le n \le N-1}\big[|x_n|^2\big]}{E\big[|x_n|^2\big]} \tag{18}$$

3.3 Parameters Influencing PAPR

Parameters which influencing the PAPR performance are mentioned in Table 1.

Table 1 Parameters influencing PAPR

Parameters	Influence on PAPR
Modulation schemes	PAPR is linearly dependent on constellation of modulation schemes. It is more for M-QAM than M-PSK
Roll-off factor (α)	It varies from 0 to 1. As 'α' increases, PAPR reduces significantly
The number of subcarriers (K)	PAPR is directly proportional to the K (total number of subcarriers). As it increases, device PAPR also increases due to the greater the number of peaks PAPR can appear [24]
Filter length (L) influence	The filter length L has a high importance for PAPR estimation. The PAPR increases with the filter length
Oversampling factor	The oversampling refers to the no. of samples per symbol period. Oversampling is a technique, which can be used to calculate the more accurate value of PAPR of signal

3.4 Factors for Selecting the PAPR Reduction Strategy

A few considerations are taken into account when selecting the strategy that can efficiently diminish the PAPR just as excellent performance can be sustained. These following factors are to be considered [25] as:

i. Strategies for decreasing PAPR should be permitted to decrease PAPR without adding in-band interference and radiation from OOB.
ii. Low average power: The increased capacity need a high level operational area in HPA and therefore decrease the output of the BER.
iii. No BER performance degradation: The aim of PAPR degradation is to show signs of improvement of system performance.
iv. Power efficiency in the reduction of PAPR should be considered. If more power is needed to operate the technique that reduces the PAPR, then the inherent feature of the BER is destroyed.
v. No spectral spillage.

3.5 PAPR Reduction Method

PAPR reduction strategies can mostly be separated into two domain technologies: a frequency domain technology and a time domain technology. Examples of frequency domain techniques include selective mapping (SLM), partial transmitting sequence (PTS), precoding and more. Examples of time domain technique are clipping and filtering, peak widowing and so on. Contrast between accessible PAPR degradation strategies is appeared in Table 2 [24–28].

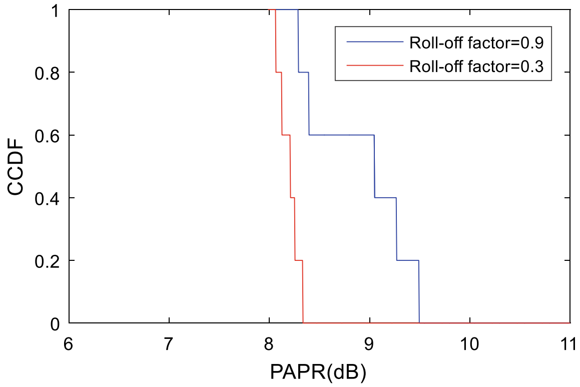
3.6 PAPR in GFDM

Figure 5 is plot between CCDF versus PAPR (likewise as probability vs percentage) in GFDM system with 'α' equal to 0.3 and 0.8, respectively. This plot is determined for the values, where no. of subcarriers = 128, no. of sub-symbols = 5, CR = 4, 16-QAM modulation are utilized. The PAPR of original GFDM signal is not exactly the PAPR of OFDM system as appeared in this figure. In this way, we have to diminish the estimation of PAPR which is examined in the next section. Figure shows the GFDM RC filter system's PAPR with specific 'α'. Increment the PAPR of GFDM with the α extension. The RRC filter is contrasted with the RC filter more with PAPR. The results show that the option of the pulse shaping filter affects GFDM system performance. In this manner, PAPR is impacted by either a similar filter with different 'α' or different filter with a similar 'α'.

Table 2 Comparison between different PAPR reduction techniques

	Clipping and filtering	Coding	PTS	SLM	TR	TI	Precoding
BER performance (degradation)	Yes	No	No	No	No	No	No
Average power (increases)	No	No	No	No	Yes	Yes	No
Computational complexity	Low	Low	High	High	High	High	Low
Expansion of bandwidth	No	Yes	Yes	Yes	Yes	Yes	Yes
Side information	No	No	Yes	Yes	No	No	No
Processing at transmitter sides	Amplitude clipping, filtering	Encoding or table search	IFFTs, complex vector sums	Need more IFFTs than PTS method	Need IFFTs, find value of peak reduction carriers (PRCs)	Need IFFTs, search for maximum point in time	Increase in number of subcarriers
Processing at receiver sides	None	Decoding	Transmission of side information inverse PTS	Transmission of side information inverse SLM	Ignore non-data bearing subcarriers	Modulo-D operation	Expansion of bandwidth

Fig. 5 Plot between CCDF versus PAPR[0] in GFDM with different roll-off factor



4 Problem Solution

The Clipping technique which is used to reduce the PAPR in GFDM signal mostly similar to the OFDM clipping technique [29]. GFDM signal is generated from GFDM modulator where IFFT operation is done with circular shifting of the signal. This signal is then clipped using clipping technique. The clipping is then filtered to the out-of-band strength. The forward FFT converts the signal being clipped back into the distinct frequency domain. In this method, where all band-edge subcarriers are unused and the corresponding components are null and void. The resulting filter is a time-dependent filter that excludes out-of-band frequency but passes components of in-band frequency. It does not cause distortion to the signal within the band. Here, filter works on a symbol-by-symbol basis, and it does not cause any conflict between symbols. The filtering causes re-growth to some peak in transmitter side before transmitting the signal. Proposed model for PAPR reduction using clipping and filtering method is shown in Fig. 6.

The flow diagram of clipping and filtering technique which is applied on the GFDM system is shown in Fig. 7. Finally, clipped GFDM signal is passed over AWGN channel which provide better SNR performance compared to original GFDM signal. Condition of clipping is signal power of GFDM signal is greater than product of CR and mean power of GFDM signal.

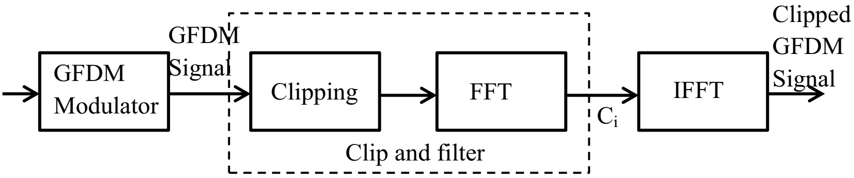


Fig. 6 Block diagram of proposed method for PAPR reduction

Fig. 7 Flowchart of clipping technique

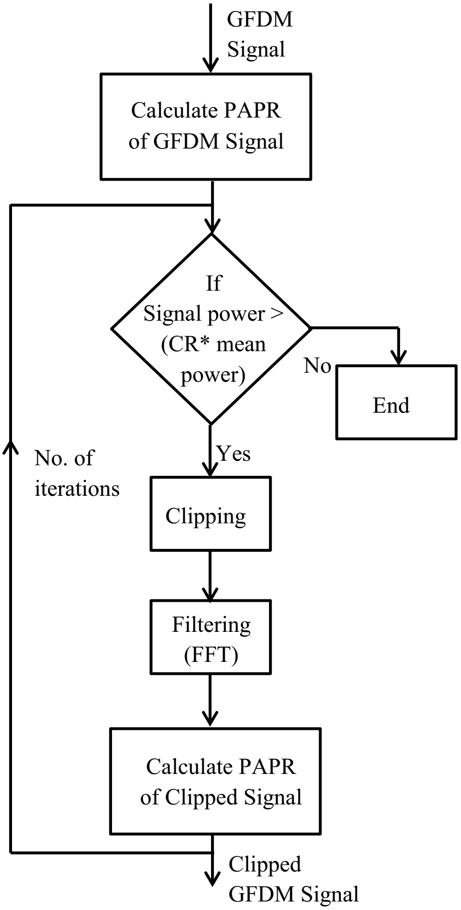
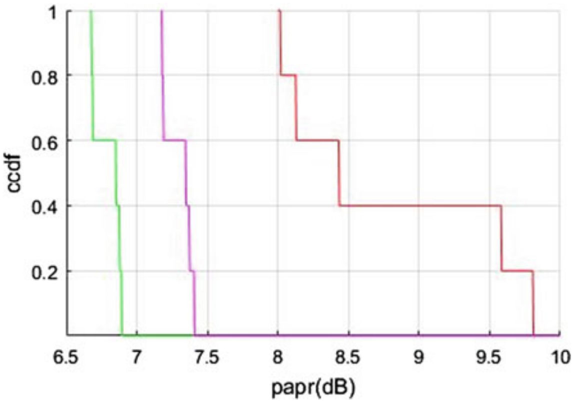

GFDM Signal
Calculate PAPR of GFDM Signal
If Signal power > (CR* mean power)
No
End
Yes
No. of iterations
Clipping
Filtering (FFT)
Calculate PAPR of Clipped Signal
Clipped GFDM Signal

Fig. 8 Clipping on GFDM signal for RRC pulse at $\alpha = 0.3$



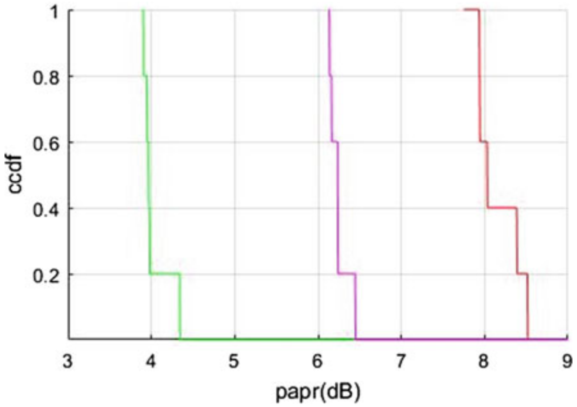


Fig. 9 Clipping on GFDM signal for RC pulse at $\alpha = 0.3$

Table 3 Simulation parameters

Total no. of subcarriers	128
Total no. of sub-symbols	16
Modulation method used	16-QAM
Total no. of iteration	4
Pulse shaping filter used	RC and RRC
Roll-off factor (α)	0.3, 0.8
Clipping ratio	4

Figures 8 and 9 show the repeated clipping and filtering using our proposed method significantly decreases the PAPR for the values of roll-off factor 0.3 for both RRC and RC pulse shaping filter (note the logarithmic scales), and there are two iterative clipping on original PAPR of GFDM system. The parameters of the above simulation are given in Table 3. From the figure, it shows RRC filter has the more PAPR compared to RC filter which is true. The results demonstrate a strong effect on the output of GFDM system, when choosing the pulse shaping filter.

5 Conclusion

Wireless communication network does not stop growing; the evolution from 1 to 4G has developed the quality of human life dramatically. Lots of dreams came to reality by wireless communication technologies and applications. 5G promises to evolve the wireless technology with its speed, reliability and other features. GFDM is the most promising 5G technology waveform competitor that offers high bandwidth and low OOB radiation for the 5G communication network. Multiplexing strategy used in 5G technology still has a problem of PAPR but less than the others. Clipping could be one of the effective methods for the degradation of the PAPR. But this process is nonlinear and could lead to significant bandwidth degradation, which degrading BER performance and OOB noise, reducing spectral performance. Filtering can reduce the spectral splatter after cutting, but it can also cause a peak re-growth. In the wireless communication device with various α-roll factors, pulse shaping filter is a significant factor in the efficiency of GFDM. The performance of PAPR without degrading SNR performance may be improved. All simulation and result are performed using this technique are on MATLAB software.

References

1. De D, Mukherjee A, Das SK, Dey N (2020) Wireless sensor network: applications, challenges, and algorithms. Nature inspired computing for wireless sensor networks. Springer, Singapore, pp 1–18
2. Das SK, Tripathi S (2019) Energy efficient routing formation algorithm for hybrid ad-hoc network: a geometric programming approach. Peer-To-Peer Netw Appl 12(1):102–128
3. Das SK, Tripathi S (2018) Adaptive and intelligent energy efficient routing for transparent heterogeneous ad-hoc network by fusion of game theory and linear programming. Appl Intell 48(7):1825–1845
4. Das SK, Samanta S, Dey N, Kumar R (2020) Design frameworks for wireless networks. Springer
5. De D, Mukherjee A, Das SK, Dey N (2019) Nature inspired computing for wireless sensor networks
6. Binh HTT, Hanh NT, Nghia ND, Dey N (2020) Metaheuristics for maximization of obstacles constrained area coverage in heterogeneous wireless sensor networks. Appl Soft Comput 86:105939
7. Mishra AR (2004) Fundamentals of cellular network planning and optimization, 2G/2.5G/3G…Evolution of 4G. Wiley, New York
8. Pereira V, Sousa T (2004) Evolution of mobile communications: from 1G to 4G. Department of Informatics Engineering of the University of Coimbra, Portugal
9. Chen S, Zhao J (2014) The requirements, challenges, and technologies for 5G of terrestrial mobile telecommunication. IEEE Commun Mag 52(5):36–43
10. Thompson J, Ge X, Wu HC, Irmer R, Jiang H, Fettweis G, Alamouti S (2014) 5G wireless communication systems: prospects and challenges. IEEE Commun Mag 52(2):62–64
11. Farhang-Boroujeny B (2011) OFDM versus filter bank multicarrier. IEEE Signal Process Mag 28(3):92–112
12. Vakilian V, Wild T, Schaich F, ten Brink S, Frigon JF (2013) Universal filtered multi-carrier technique for wireless systems beyond LTE. In: Proceedings of IEEE Globecom workshop, Atlanta, GA, USA, pp 223–228
13. Ayadi R, Siala M, Kammoun I (2007) Transmit/receive pulse-shaping design in gfdm systems over time-frequency dispersive AWGN channel. In: Proceedings of IEEE ICSPC, Dubai, UAE, pp 772–775
14. Datta R, Michailow N, Lentmaier M, Fettweis G (2012) GFDM interference cancellation for flexible cognitive radio PHY design. In: Proceedings of 76th IEEE VTC fall, Québec City, QC, Canada, pp 1–5
15. Kumar A, Gupta M (2015) Key technologies and problem in deployment of 5G mobile communication systems. Commun Appl Electr 1(3):4–7. (Published by Foundation of Computer Science, New York, USA)
16. Fettweis G, Krondorf M, Bittner S (2009) GFDM—generalized frequency division multiplexing. In Proceedings of 69th IEEE VTC spring, Barcelona, Spain, pp 1–4

17. Wunder G et al (2014) 5GNOW: non-orthogonal, asynchronous waveforms for future mobile applications. IEEE Commun Mag 52(2):97–105
18. Sameen M, Khan AA, Khan IU, Azim N, Iqbal N (2016) Comparative analysis of OFDM and GFDM. Int J Adv Comput Tech Appl 4(2):267–274
19. Michailow N, Gaspar I, Krone S, Lentmaier M, Fettweis G (2012) Generalized frequency division multiplexing: analysis of an alternative multi-carrier technique for next generation cellular systems. In: 2012 International symposium on wireless communication systems (ISWCS), pp 171–175
20. Wu J, Ma X, Qi X, Babar Z, Zheng W (2017) Influence of pulse shaping filters on PAPR performance of underwater 5G communication system technique: GFDM. Hindawi Wirel Commun Mob Comput 2017, Article ID 4361589
21. Asseri MI, Tsimenidis CC, Sharif BS, Le Goff SY (2010) Orthogonal-SLM technique for PAPR reduction and recovery without side information in OFDM systems. In: 2010 7th international symposium on communication systems networks and digital signal processing (CSNDSP), pp 162,166
22. Han SH, Lee JH (2005) An overview of peak to average power ratio reductions techniques multicarrier transmission. IEEE Wirel Commun 12(2):56–65
23. Farhang A, Marchetti N, Doyle LE (2016) Low-complexity modem design for GFDM. IEEE Trans Signal Process 64(6)
24. Han SH, Lee JH (2005) An overview of peak to average power ratio reduction techniques for multicarrier transmission. IEEE Wirel Commun 12(2):56–65
25. Michailow N, Fettweis G (2013) Low peak-to-average power ratio for next generation cellular systems with generalized frequency division multiplexing. International symposium on intelligent signal processing and communication systems ISPACS. Naha, Japan, pp 651–655
26. Daumont S, Rihawi B, Lout Y (2008) Root-Raised Cosine filter influences on PAPR distribution of single carrier signals. ISCCSP 2008, Malta, pp 841–845
27. Armstrong J (2001) New OFDM peak-to-average power reduction scheme. Proceedings. IEEE, VTC2001 Spring, Rhodes, Greece
28. Armstrong J (2002) Peak-to-average power reduction for OFDM by repeated clipping and frequency domain filtering. Electron Lett 38(5):246–247
29. SaadHadi Z, Mosa Omran B (2015) Peak-to-average power reduction using repeated frequency domain filtering and clipping in OFDM. Int J Comput Appl 122(11):6–10

Investigation of Memory, Nonlinearity and Chaos in Worldwide Monthly Mobile Data Traffic in Smartphones

Swetadri Samadder and Koushik Ghosh

Abstract The present chapter deals with the time series of worldwide monthly mobile data traffic per smartphone during January, 2014-December, 2019. Firstly, an attempt is made to understand the nature of memory in this time series by taking into account scaling pattern and the issue of persistence or anti-persistence by means of Hurst exponent. Next its self-similarity or self-affinity which is coined as fractal behaviour is analysed using Higuchi's method of fractal dimension estimation. Next the self-organized criticality ('soc') of the present data is analysed with the help of the integrated (cumulative) distribution. To examine the nonlinearity and deterministic/stochastic nature of the governing mechanism, we use delay vector variance (DVV) method. We have taken into account 0–1 test for chaos and recurrence plot (RP) analysis with recurrence quantification analysis (RQA) to test the signature of chaos in the present data. In fine, the proposed chapter employs certain statistical signal processing techniques to realize the memory, self-similarity, self-organized criticality, nonlinearity and chaos in the present time series of worldwide monthly mobile data traffic per smartphone. This study possibly indicates a persistent, self-similar, deterministic, nonlinear and non-chaotic profile with no 'soc' for the present time series.

Keywords Mobile data traffic · Memory · Self-similarity · Self-organized criticality · Nonlinearity · Chaos

S. Samadder
Department of Mathematics, Fakir Chand College, South 24 Parganas, Diamond Harbour 743331, India

K. Ghosh (✉)
Department of Mathematics, University Institute of Technology, The University of Burdwan, Golapbag (North), Burdwan, West Bengal 713104, India

S. K. Das et al. (eds.), *Architectural Wireless Networks Solutions and Security Issues*, Lecture Notes in Networks and Systems 196,
https://doi.org/10.1007/978-981-16-0386-0_17

1 Introduction

The present era is the period of wireless technology, and it has a huge impact in our daily life. In our circadian lives, willingly or unwillingly, we interact with different wireless networks like cell phone networks, wireless local area networks, wireless sensor networks, satellite communication networks, terrestrial microwave networks, etc. After the inception of the Internet, the persistently enhancing volume of network users and their reciprocated exchange of information is turning out to be an imperative conduit of communication. From the perspective of dynamical system, the governing mechanism of a wireless network traffic can be a highly nonlinear occurrence.

Earlier there have been certain attempts to investigate nonlinearity and chaos [1–4] and self-similarity or fractal nature [5, 6] in different forms of wireless network traffic. In the modern era, usage of mobile phones has become a very essential part and parcel in daily life. After the launch of smartphones, mobile Internet data traffic has emerged as an immensely interesting and challenging domain of scientific research.

1.1 Objective and Motivation

In the present correspondence, we have considered the time series of worldwide monthly mobile data traffic per smartphone during January, 2014-December, 2019 [7]. The objective of the present work is to investigate the memory, nonlinearity and chaos in the present time series in order to have a broad understand of the pattern of sharp growth in mobile data traffic worldwide.

1.2 Contribution

The analysis of the present time series starts with the investigation of memory in it by means of scaling analysis, and we have used the concept of Hurst exponent [8–12] for the same. After that we inspect the self-similar (fractal) nature in the present time series by using Higuchi method [13, 14]. Next we apply the integrated (cumulative) distribution analysis [15–17] to identify whether any form of self-organized criticality ('soc') is present in the data or not. Next, we use delay vector variance (DVV) [18, 19] method to inspect the deterministic/non-deterministic profile as well as nonlinearity for the present time series. Then, 0–1 test for chaos [20–24] is implemented on this time series. Finally, recurrenceplot (RP) [25] is used over this data with recurrence quantification analysis (RQA) [26–33] to have a more precise and invasive approach to trace whether any chaos is present in it or not. After demonstrating the working theories and before going to furnish the results with the present time series, we first present the updated and projected scenario of mobile-broadband subscriptions and volume of data usage from the perspective of both global and Indian context with certain graphical profiles to realize the background and importance of this present analysis.

2 Working Theories

In this section, we present the theories we have worked with. First, we present theory of the estimation of Hurst exponent which is followed by theories of estimation of fractal dimension, estimation of self-organized criticality, delay vector variance analysis, 0–1 test for chaos, recurrence plot (RP) analysis and finally recurrence quantification analysis (RQA), respectively.

2.1 Scaling Analysis and Estimation of Hurst Exponent

Hurst exponent plays a very important and convincing role to understand the memory of a time series. This term was coined by H.E. Hurst in a study of hydrology during analysing the hydrological characters of the Nile river in Egypt [8].

For the scaling analysis of a given time series, finite variance scaling method (FVSM) is a very useful technique [8–12]. A well-known version of FVSM is the standard deviation analysis (SDA) [10]. For a given finite time series $\{x(t_n)\}$ (where $n = 1, 2, \ldots., N$ and t represents time), this method gives rise to a sequence of cumulative standard deviations $D\ (t_j)$ associated with the partial time series $\{x(t_n)\}(n = 1, 2, \ldots., j)$ (where $j = 1, 2, \ldots., N$) by the following manner:

$$D(t_j) = [\frac{1}{j}\sum_{n=1}^{j} x^2(t_n) - \{\frac{1}{j}\sum_{n=1}^{j} x(t_n)\}^2]^{\frac{1}{2}}, \quad \text{for } j = 1, 2, 3 \ldots, N \tag{1}$$

For self-similar time series, eventually, we can find that this $D\ (t)$ in Eq. (1) follows a power law by the following way: [8–12]

$$D(t) \propto t^H \tag{2}$$

Here, in Eq. (2), H is the Hurst exponent, and it can be estimated calculating the slope of the line of regression for the log–log plot of $D(t)$ versus t.

Range of value Hurst exponent (H) lies between 0 and 1. $H = 0.5$ implies a Brownian motion and hence a true random walk of the time series. $0 < H < 0.5$ indicates that the time series is of anti-persistent behaviour (negative autocorrelation) and hence ruled by a short memory process. $0.5 < H < 1.0$ occurs for a time series with persistent behaviour (positive autocorrelation) and hence governed by a long memory process. $H = 0$ and $H = 1$ indicate a white noise and a smooth time series, respectively.

2.2 Self-Similarity and Higuchi Method to Determine Fractal Dimension

Higuchi [13, 14] developed a method for calculating the fractal dimension of a given time series. The method is as follows.

We take a finite set of time series $\{x(t_n)\}$ where $n = 1, 2, \ldots., N$ and t represents time taken at a regular interval, i.e. here $t_n = t_1 + (n - 1)\tau$ where τ represents the uniform gap in observation instants.

From $\{x(t_n)\}$, we develop another time series as $\{X(m), X(m + k), X(m + 2\,k), \ldots, X(m + [(N - m)/k].\, k)\}$ where [] refers box function and $k, m \in Z$ with the restriction $m = 1, 2, 3, \ldots, k$. m and k indicate the initial time and the interval time, respectively. In this process, for each k, we obtain k sets of new time series. We fix the length of the curve corresponding to each m for a particular value of k of these new k time series as follows:

$$L_m(k) = \left\{ \sum_{i=1}^{[(N-m)/k]} |X(m + ik) - X(m + (i-1)k|) \frac{N-1}{[(N-m)/k]k} \right\} / k \tag{3}$$

The length of the curve for the time interval k, $\langle L(k) \rangle$ is fixed as the mean value over k sets of $L_m(k)$ as found in Eq. (3). If $\langle L(k) \rangle \propto k^{-D}$, we deduce that the curve is of fractal nature with dimension D. We derive fractal dimension D estimating the slope of the regression line from the plot of $\log\langle L(k)$ against $\log k$.

2.3 Self-Organized Criticality by Means of Integrated (Cumulative) Distribution

Self-organized criticality or ‘soc’ is described as the tendency of large dynamical systems to organize themselves into a critical state, with avalanches or exhibits of ‘punctuations’ of all sizes. In the critical state, events which would otherwise be uncoupled become correlated. ‘soc’ is one of the factors by which complexity arises in a dynamical system. The idea of ‘soc’ was first suggested by Bak et al. [15, 16] to illustrate the mechanism of open and extended driven systems exhibiting avalanche like energy dissipation. The philosophy of ‘soc’ is associated with the idea of scale invariance of the distribution of relaxation events and the simultaneous existence of self-regulatory internal mechanism that can drive the system to a statistical stationary state.

'soc' may be present in a dynamical system provided the system is governed by a power law. The knowledge of the dynamics of the process is decisive for the characterization of the possible 'soc' behaviour. Presence of scaling signature in the signal strengthens the link between the governing dynamics of the signal and 'soc'. Distributions $n(x)$ of events with property x that behave like $n(x)\sim x^{-\tau}$ are most conveniently analysed with the help of the integrated (cumulative) distribution [15–17]:

$$\overline{N}(x) = \int_{x}^{M} n(x)\mathrm{d}x \tag{4}$$

where M is the maximal event encountered in the data set. Eventually, we can find if M is finite

$$\overline{N}(x) \sim x^{-\tau+1}\left[1 - \left(\frac{x}{M}\right)^{\tau-1}\right] \tag{5}$$

Thus, the log–log plot of $\overline{N}(x)$ versus x definitely departs from a straight line as x approaches M.

2.4 Delay Vector Variance (DVV) Analysis

When a time delay is embedded in a time series demonstrated by a temporal signal x_i, the time series can be described in 'phase space' by a set of delay vectors (DVs) $x(k) = \left[x_{k-m\tau}, \ldots, x_{k-\tau}\right]$ (where $k = 1, 2,\ldots, N$), the embedding dimension is given by m, and the embedded time delay lag is denoted by τ. Inside a fixed Euclidian distance τ_d to DV $x(k)$, DVs are assembled which are referred by $\lambda_k(\tau_\mathrm{d})$. Mean target variance σ^{*2} throughout all sets of λ_k: $k = 1, 2, \ldots, N$ is estimated. Optimal embedding dimension m is that embedding dimension for which minimum mean target variance (σ^{*2}) is obtained. The entire ranges of pair wise distances are examined by varying standardized distance [18, 19, 33, 34]. Next, distance axis is standardized by is replacing τ_d by $(\tau_\mathrm{d} - \mu_\mathrm{d})/\sigma_\mathrm{d}$ where μ_d and σ_d are mean and standard deviation, respectively, for the whole range of τ_d, calculated over all pair wise distances between DVs defined by

$$d(i, j) = ||x(i) - x(j)||; i \neq j \tag{6}$$

The DVV plots are produced from the plot of target variance $\sigma^{*2}(\tau_\mathrm{d})$ versus $(\tau_\mathrm{d} - \mu_\mathrm{d})/\sigma_\mathrm{d}$. The estimation of the noise present in the signal is given by minimum value of target variance $\sigma^{*2}_{\mathrm{min}} = \min_{\tau_\mathrm{d}}\left\{\sigma^{*2}(\tau_\mathrm{d})\right\}$. The prevalence of the noise is overriding for the stochastic component of the signal. Hence, stochastic components should deliver larger values of $\sigma^{*2}_{\min}$. On contrary, smaller values of $\sigma^{*2}_{\min}$ indicate that the signal is deterministic. The DVV plots converge to unity at the extreme right as all the DVs are interior to the common universal set, and the variance of the target DVs is identical to the variance of the signal for maximum span.

Iterative amplitude adjusted Fourier transform (IAAFT) [35, 36] can be employed to obtain surrogate signal. Original signal's embedding dimension is used to plot DVV of surrogated signals. A DVV scatter diagram may be composed where target variance $\sigma^{*2}(\tau_\mathrm{d})$ of the original signal corresponds along horizontal axis and mean of target variance $\sigma^{*2}(\tau_\mathrm{d})$ of surrogate signal corresponds along vertical axis. If the DVV plots of the original signal and the surrogate are analogous, then the DVV scatter diagram merges with the bisector line, and the signal is judged to be linear. Conversely, if the two DVV plots are not alike, then the curve deviates from the bisector line, and the signal is judged to be nonlinear. The amount of nonlinearity can be understood by measuring the root mean square error (RMSE) between the $\sigma^{*2}(\tau_\mathrm{d})$ of the original signal and mean of the $\sigma^{*2}(\tau_\mathrm{d})$ of the surrogate signal and is shown below:

$$\mathrm{RMSE} = \sqrt{\mathrm{mean}\left\{\sigma^{*2}(\tau_\mathrm{d}) - \frac{\sum_{k=1}^{N_s}\sigma^{*2}_{s,k}(\tau_\mathrm{d})}{N_\mathrm{s}}\right\}^2} \tag{7}$$

where $\sigma^{*2}_{s,k}(\tau_\mathrm{d})$ is the target variance at span τ_d for the kth surrogate, and the average is considered over all span of τ_d valid in all the surrogate as well as DVV plots [20, 29, 30].

2.5 0–1 Test for Chaos

In 0–1 test [20, 21] for chaos, time series vector is taken as input and '0' or '1' results as output whenever the input time series vector is 'non-chaotic' or 'chaotic', respectively. This test is authentic, easy to apply and robust [22] to determine if deterministic chaos present in a time series.

From the time series $x(k)$ for $k = 1, 2, \ldots, N$, a Fourier transformed series p_n is constructed as follows [23].

$$p_n = \sum_{k=1}^{n} x(k)e^{ikc} \quad \text{where } 1 \le n \le N \tag{8}$$

corresponding to different random values of c. We have generated 100 random values of c in the interval $[\pi/5, 4\pi/5]$ to obtain100 sets of p_n. The smoothed mean square displacement $D_\mathrm{c}(n)$ is obtained as

$$D_c(n) = \frac{1}{N-m} \sum_{k=1}^{N-m} |p_{k+n} - p_k|^2 - \langle x \rangle^2 \frac{1 - \cos nc}{1 - \cos c} \tag{9}$$

where $\langle x \rangle = (1/N) \sum_{k=1}^{N} x(k)$ and n $\leq m \leq N/10 \ll N$. m is not chosen larger then $N/10$ to enhance effectivity of the test. If $x(k)$ is chaotic in nature, then $D_\mathrm{C}\ (n) \propto n$ and plot of Real (p_n) versus Imaginary (p_n) shows a Brownian motion. Otherwise if $x(k)$ is non-chaotic or regular in nature, $D_\mathrm{c}(n)$ follows a bounded function of n, i.e. $D_\mathrm{c}(n)$ does not increase infinitely with n and plot of Real (p_n) versus Imaginary (p_n) represents a bounded motion. As $D_\mathrm{c}(n)$ may be negative due to second term of the right hand side of Eq. (9),

$$D_\mathrm{c}^*(n) = D_\mathrm{c}(n) + \alpha V_\mathrm{damp}(n) \tag{10}$$

has been set to make the test more robust where $V_\mathrm{damp}(n) = \langle x \rangle^2 \sin\left(\sqrt{2n}\right)$.

The amplitude α of the term $V_\mathrm{damp}(n)$ controls the sensitivity of the chaos test to weak noise and weak chaos. The asymptotic growth rate K_c is measured by correlation method to measure the strength of the correlation of $D{*}_\mathrm{c}(n)$ with linear growth n.

$$K_\mathrm{c} = \operatorname{corr}(n, D_\mathrm{c}^*(n)) \tag{11}$$

K_c results binary value 0 or 1.

Finally, median of these values of K_c is computed to obtain

$$K = median(K_\mathrm{c}) \tag{12}$$

If the value of K is close to 0, the time series is interpreted as non-chaotic or regular and conversely, if the value of K is close to 1, the time series is interpreted as chaotic or non-regular [24].

2.6 Recurrence Plot (RP) Analysis

The pictorial approach to identify all those times where a state of a dynamical system recurs is the recurrence plot (RP) [25, 37]. A recurrence is the return of the system in its previously occured state. A recurrence is obsereved when the a system at time i is within the neighbourhood of an earlier point of the system at time j in the phase space. This is picturized within a two-dimensional squared matrix $\boldsymbol{R}$ named as recurrence matrix where both the axes are time axes. For a trajectory of length N and $\overrightarrow{x_i} \in R^m$, $\boldsymbol{R}$ is defined as

$$R_{i,j}(\varepsilon) = \Theta\left(\varepsilon - \left\|\vec{x_i} - \vec{x_j}\right\|\right); i, j = 1, 2, \ldots N \tag{13}$$

where $\Theta(.)$ is the Heaviside step function, ε is the threshold distance and $\|\cdot\|$ is a suitable norm. Hence, elements of ***R*** are either 0 or 1. RP is a visualized representation of ***R*** generated by marking black and white dots for 1 and 0, respectively.

As $R_{i,i}(\varepsilon) = 1 \forall i = 1, 2, \ldots N$, RP contains a long black main diagonal line, line of identity(LOI). But this main diagonal does not contain any useful information about state of recurrence. Other diagonal lines are given by $R_{i+k,j+k} = 1$ for $k =$ 1, 2,.., l where l is the length of the diagonal line. The length of diagonal lines (excluding main diagonal) of RP is a good estimation of predictibility of the system. The RPs with long diagonal lines indicate predictable, and hence, regular signals, short diagonal lines represent system which are sensitively dependent to the initial conditions, i.e., chaotic system and no diagonal line or single point infers that the system has homogeneous distribution of stochastic signals.

2.7 Recurrence Quantification Analysis (RQA)

The recurrence quantification analysis (RQA) [26–33, 37] is an efficient quantitative approach to study nonlinear data. It quantifies RP based on the analyzation of the diagonal lines occurring in its recurrence plot.

We here measure a set of four recurrence variables. First among them is percent recurrence(REC) or recurrence rate(RR) which quantifies the proportion of recurrent points existing within predefined threshold.

REC is expressed as

$$\mathrm{REC}(\varepsilon_i) = \frac{1}{N^2} \sum_{i,j=1}^{N} R_{i,j}(\varepsilon_i) \tag{14}$$

It estimates the probability to find a recurrence point in RP for a fixed threshold and the probability that a particular state will recur by calculating number of black dots in the recurrence plot. REC ranges between 0 and 1. 0 indicates no recurrent point, and 1 indicates all recurrent points in RP. More the REC, more the chance of time series of being regular.

The second recurrence variable is percent determinism (DET) determined by fraction of recurrent points which form diagonal lines of RP.

DET is given by

$$\mathrm{DET} = \frac{\sum_{i,j=1}^{N} D_{i,j}}{\sum_{i,j=1}^{N} R_{i,j}} \tag{15}$$

where

$$
\begin{aligned}
D_{i,j} = {} & 1,\ \text{if } (i,\ j) \text{ and } (i+1,\ j+1) \text{ or } (i-1,\ \ j-1) \text{ are recurrent} \\
& 0,\quad \text{otherwise}
\end{aligned}
\tag{16}
$$

The value of DET is between 0 and 1. As diagonal lines are absent for stochastic time series, long for periodic time series and short for chaotic time series, DET close to 0 and DET close to 1 indicate stochastic and periodic time series, respectively. If the value of DET is neither close to 0 nor close to 1, there is possibility of the time series to be chaotic.

The third recurrence k variable is maximal line length in the diagonal direction(LMAX) which is length of the single longest diagonal line segment in the entire RP excluding the main diagonal.

$$
\text{LMAX} = \text{Max}(l_i), i = 1, 2, \ldots N_l \tag{17}
$$

where l_i is the length of the ith diagonal line and N_l is the number of diagonal lines excluding main diagonal. As diagonal line estimates the range where a segment of the trajectory is close to another segment, LMAX estimates the divergence of the trajectory segments. So, smaller LMAX indicates trajectories tend to be more divergent and hence more chaotic. On the other hand, larger LMAX implies trajectories tend to be more convergent and hence more regular. Based on this fact, it is clear that there is a inversely proportional relationship between LMAX and Lyapunov exponent [25].

The fourth and last recurrence variable is Shannon entropy of the frequency distribution of the diagonal line segments (ENTR).

ENTR is given by

$$
\text{ENTR} = -\sum_{i=1}^{N_l} p(l) \ln p(l) \tag{18}
$$

where $p(l)$ is the probability distribution of lengths of diagonal lines.

Calibrated over integer bins in a histogram., ENTR determines the complexity of deterministic structure of the trajectory. If the value of ENTR is small, low complexity in the dynamics of the trajectory indicates its chaotic nature. On contrary, high entropy indicates more complexity in the trajectory supporting its non-chaotic nature.

3 Mobile-Broadband Subscriptions and Volume of Data Usage: Global and Indian Scenario

Figure 1 demonstrates the active mobile-broadband subscriptions per 100 inhabitants during 2007–2019 [38].

To understand the power law behaviour for the three curves as depicted in Fig. 1, corresponding log–log plots are sketched in Figs. 2a–c, and the corresponding best-fitted straight line with equation (along with the goodness of fit measured in the form of R^2) is provided in each case.

From the estimated equations of the best-fitted straight lines against the log–log plots (base *e*), we can estimate the forecasting power law models for the three patterns as follows:

Number of active mobile-broadband subscriptions per 100 inhabitants in developed countries in a particular year = $16.8728 \times \text{time}^{0.7594}$.

Number of active mobile-broadband subscriptions per 100 inhabitants worldwide in a particular year = $2.7294 \times \text{time}^{1.2505}$.

Number of active mobile-broadband subscriptions per 100 inhabitants in developing countries in a particular year = $0.4785 \times \text{time}^{1.9057}$.

Here in each case time = concerned year-2007 + 1.

Figure 3 represents the total volume of mobile data usage (in million GB) in India by year, starting from 2014 up to the end of 2018 [39]. It shows in 2014 the volume was 828 million GB, while in the years 2015–2018, it creeped up to 1375, 4642, 20,092 and 46,404 million GB, respectively.

Figure 4 gives the increasing profile of the number of wireless Internet subscribers in India during 2007–2019 [40].

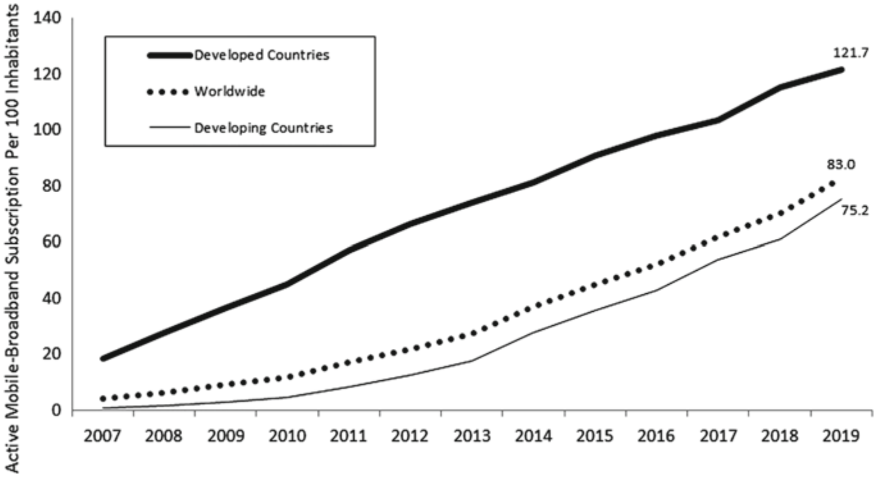


Fig. 1 Active mobile-broadband subscriptions per 100 inhabitants, 2007–2019

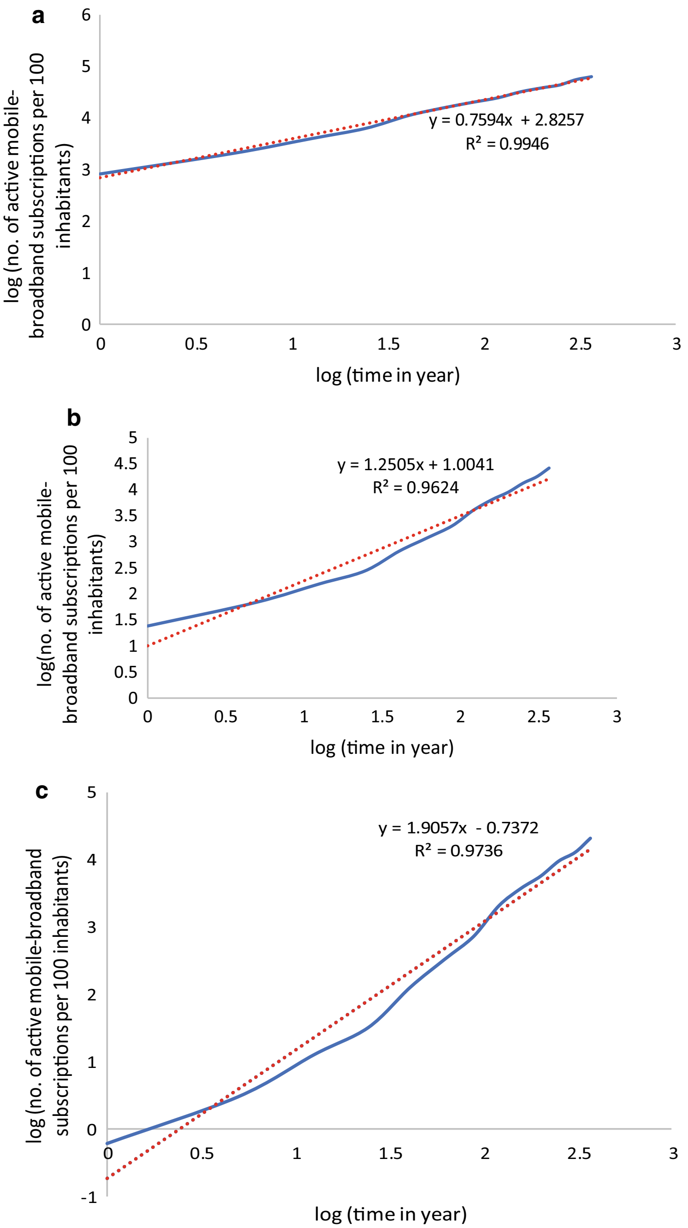


Fig. 2 **a**. Log–log plot of active mobile-broadband subscriptions per 100 inhabitants in developed countries during 2007–2019. **b** Log–log plot of active mobile-broadband subscriptions per 100 inhabitants worldwide during 2007–2019. **c** Log–log plot of active mobile-broadband subscriptions per 100 inhabitants in developing countries during 2007–2019

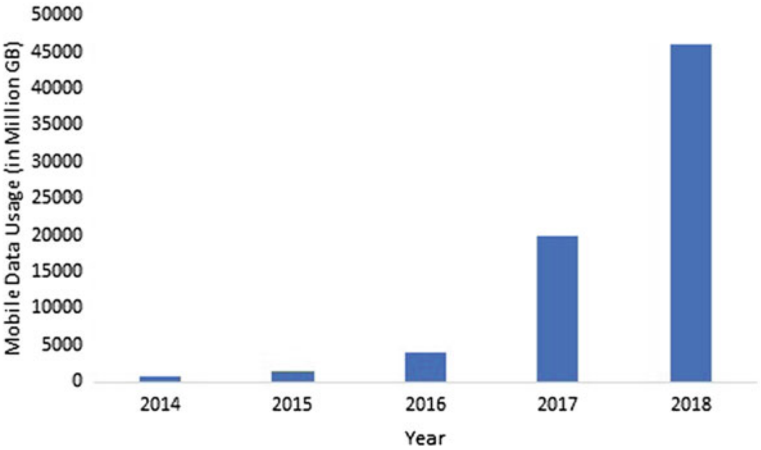


Fig. 3 Volume of mobile data usage in India, by year

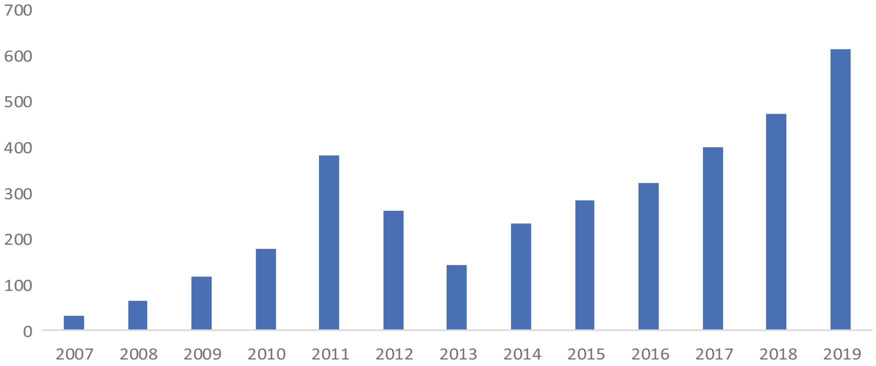


Fig. 4 Wireless Internet subscribers in India (in Millions) (Year: 2007–2019, calculated at the end of March)

To understand the power law behaviour for the profile in Fig. 4 corresponding log–log plots are drawn in Fig. 5, and the corresponding best-fitted straight line with equation (along with the goodness of fit measured in the form of R^2) is provided.

From the estimated equation of the best-fitted straight line against the log–log plot (base e) in Fig. 5, we can estimate the forecasting power law model for the present case as follows:

Number of wireless Internet subscribers in India (in million) in a particular year $= 35.9885 \times \text{time}^{1.0162}$ where time $=$ concerned year-2007 $+$ 1.

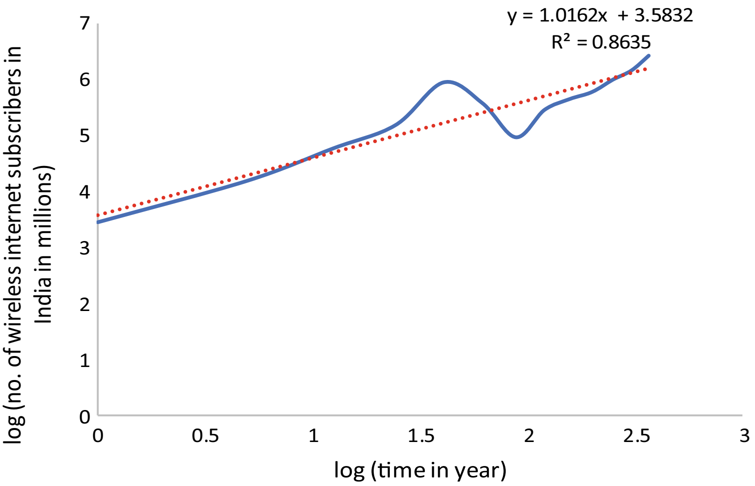


Fig. 5 Log–log plot of wireless Internet subscribers in India (in Millions) (Year: 2007–2019, calculated at the end of March)

4 Results

From Sect. 3, we can get an idea of the rapidly increasing volume of usage of wireless Internet as well as mobile-broadband subscriptions globally as well as in Indian context. Especially, after the inception of smartphones, the usage of Internet has started to show a remarkable inflation day by day. First, we produce Fig. 6 to report the vast and steeply increasing global profile of mobile data traffic in smartphones [7].

To understand the power law behaviour for the profile in Fig. 6, corresponding log–log plots are drawn in Fig. 7, and the corresponding best-fitted straight line with equation (along with the goodness of fit measured in the form of R^2) is provided.

From the estimated equation of the best-fitted straight line against the log–log plot (base 10) in Fig. 7, we can estimate the forecasting power law model for the present case as follows:

Mobile data traffic per smartphone worldwide (in GB/month) in a particular year $= 0.6738 \times \text{time}^{1.3016}$ where time = concerned year-2014 + 1.

Figure 8 demonstrates a finer profile of actual monthly mobile data traffic worldwide per smartphone from January, 2014 to December, 2019 [7].

We use this monthly data for further analysis. First, we go for scaling analysis of this data. Figure 9 demonstrates the profile of this scaling analysis.

From Fig. 9, we get the slope of the best-fitted straight line against the log–log plot of cumulative S.D. versus month as 0.6895 which is the current estimation of Hurst exponent. As the value of the Hurst exponent is here obtained to be greater than 0.5, it suggests that the governing mechanism of the growth in mobile data

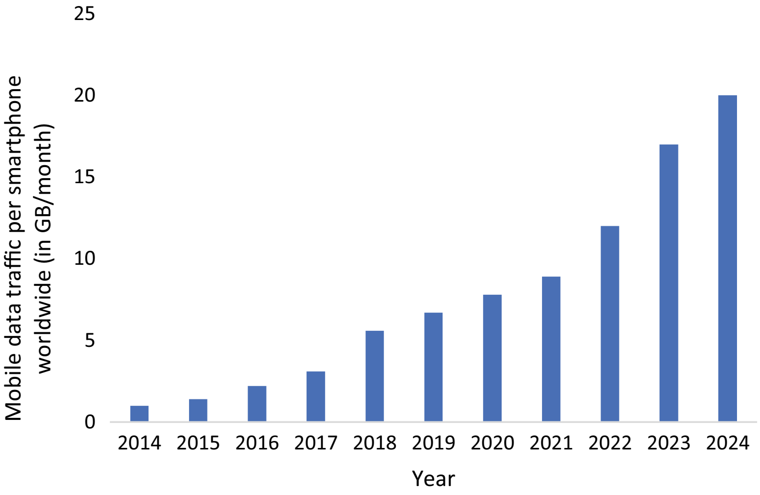


Fig. 6 Mobile data traffic per smartphone worldwide from 2014 to 2024 (in GB/month) [Actual as well as Projected]

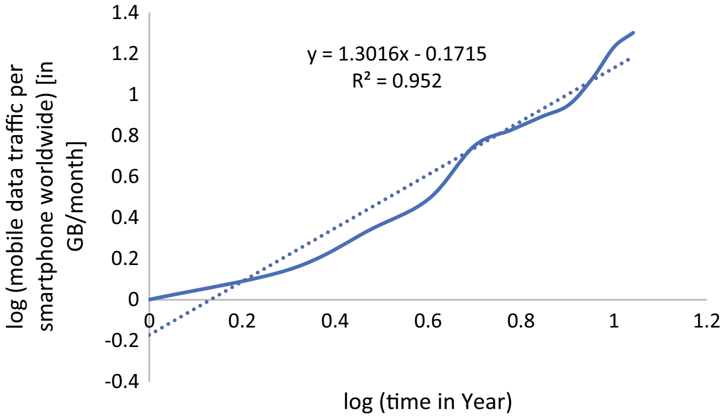


Fig. 7 Log–log plot of mobile data traffic per smartphone worldwide from 2014 to 2024 (in GB/month)

traffic worldwide is a long memory process. Again, as here we get a power law, we anticipate that the present data may have a fractal nature.

So next, we go for fractal analysis of the present time series, and the corresponding result is shown in Fig. 10.

From Fig. 10, we get the slope of the best-fitted straight line against the plot of log $< l(k) >$ versus log (k) as (-1.0658) which gives us the estimate of the corresponding fractal dimension $D = 1.0658$. As here the fractal dimension lies within the interval $1 < D < 2$, we can infer that the present time series is fractal or self-similar in nature.

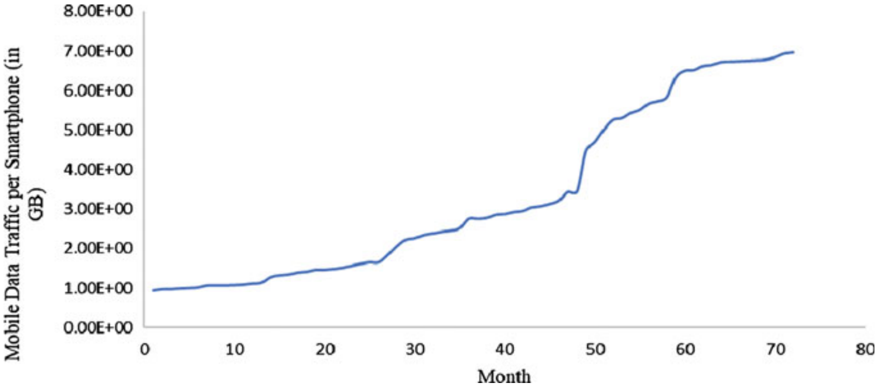


Fig. 8 Monthly mobile data traffic per smartphone worldwide from January, 2014 to December, 2019

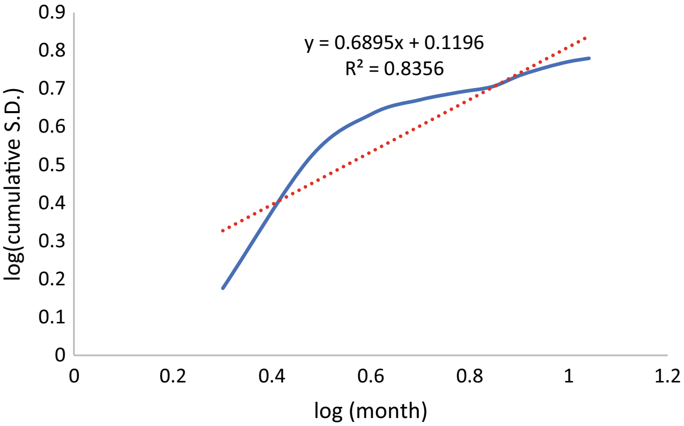


Fig. 9 Scaling analysis of monthly mobile data traffic worldwide per smartphone from January, 2014 to December, 2019 (in GB)

Next, we inspect whether any form of ‘soc’ is present in the data or not. For that, we use the integrated (cumulative) distribution of the events, and the corresponding profile is shown in Fig. 11.

Figure 11 shows the log–log plot of $\overline{N}\,(x)$ versus x to examine ‘soc’ in the data. From the slope of the best-fitted straight line against the plot, we get $(-\zeta + 1) = 1.9012$ which gives $\zeta = -0.9012$. This surely determines that this data does not possess any form of ‘soc’.

In Fig. 12, we present the DVV plot of the present time series, and Fig. 13 demonstrates the DVV scatter plot.

Outcome of the DVV analysis is furnished in Table 1.

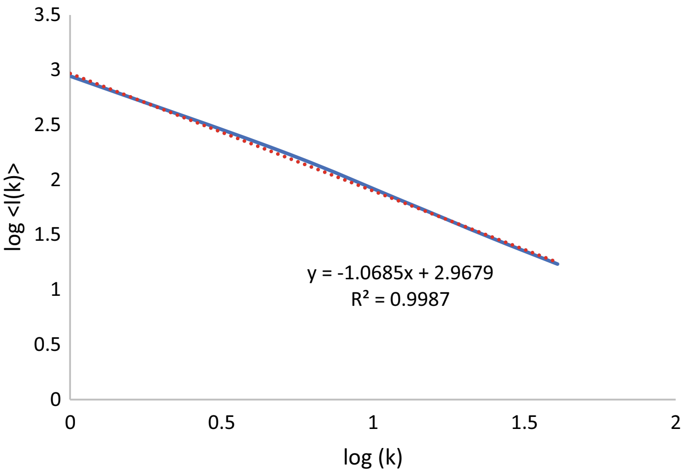


Fig. 10 Fractal analysis of monthly mobile data traffic worldwide per smartphone from January, 2014 to December, 2019 (in GB)

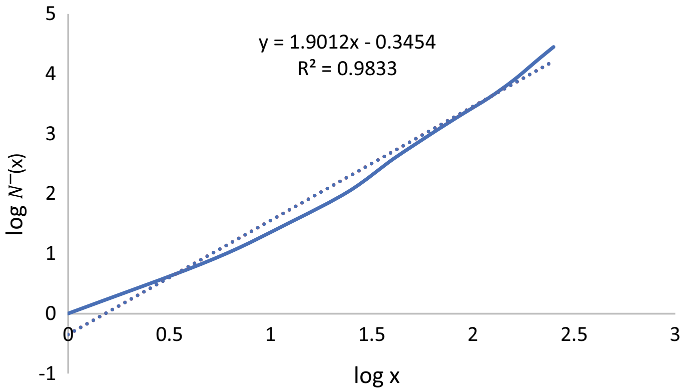


Fig. 11 'soc' analysis of monthly mobile data traffic per smartphone worldwide from January, 2014 to December, 2019 (in GB)

From the DVV Plot, min target variance is very less (0.035). So, the deterministic component of the data overpowers its stochastic component. DVV scatter diagram and moderate value of RMSE clarify nonlinearly of the data. As nonlinearity is observed, we next go for examining whether chaos is present in the time series or not.

Next, we produce Figs. 14, 15, 16 and 17 as the results from 0 to 1 test of chaos.

The profiles of both $D_c(n)$ and $D^*{}_c(n)$ are oscillatory in nature (from Figs. 14 and 16) and hence may be considered as bounded indicating non-chaotic nature. For p_n, a kind of spiral pattern with certain crossovers are observed (Fig. 15) which again

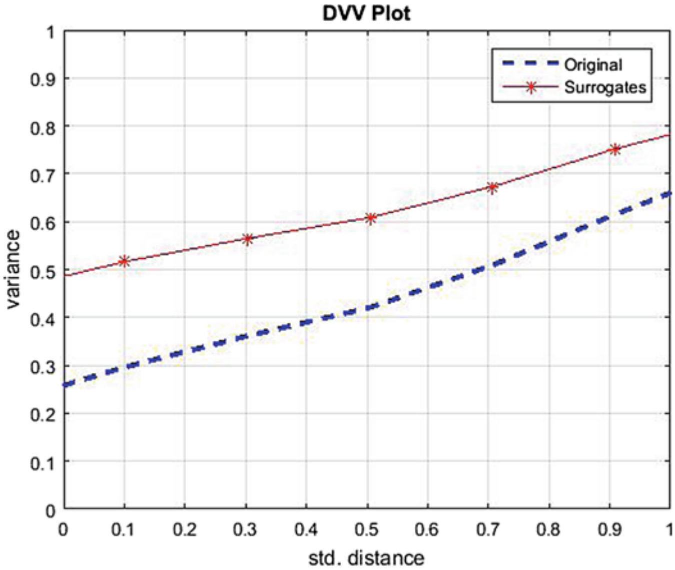


Fig. 12 DVV plot of monthly mobile data traffic per smartphone worldwide from January, 2014 to December, 2019 (in GB)

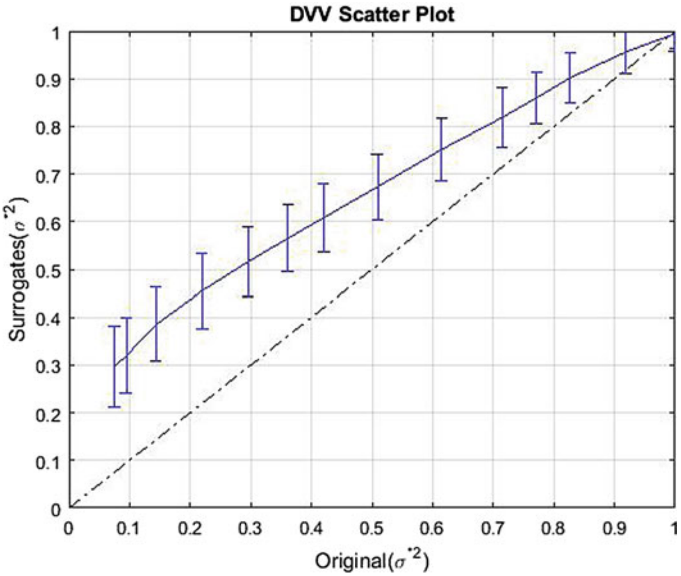


Fig. 13 DVV scatter plot of monthly mobile data traffic per smartphone worldwide from January, 2014 to December, 2019 (in GB)

Table 1 DVV analysis of monthly mobile data traffic per smartphone worldwide from January, 2014 to December, 2019 (in GB)

No. of datapoints	Embedding dimension	Embedding lag	Number of reference DV's	Number of surrogates	Min target variance of original data	RMSE
72	1	1	50	99	0.035	0.087

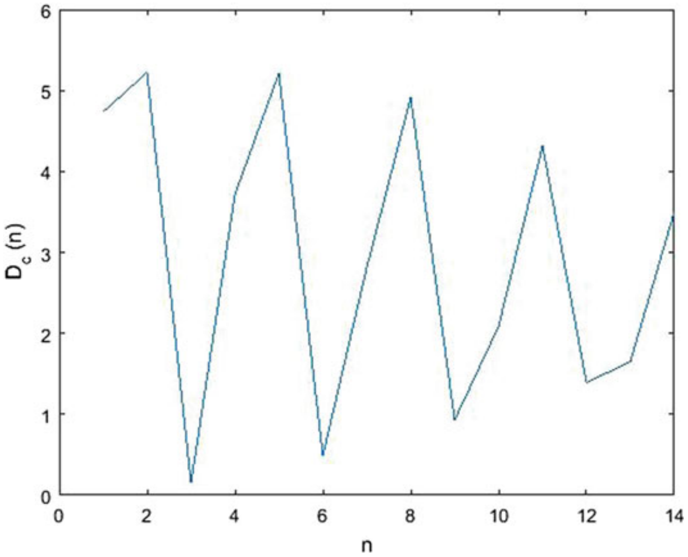


Fig. 14 Plot of D_c (n) versus n

points to the possible absence of chaos. Again, from Fig. 17, we can find K = median (K_c) has the value −0.09 again indicating possible absence of chaos in the data.

Next, we produce the plot of RP analysis for the present time series in Fig. 18.

RP plot (Fig. 18) shows more or less long diagonal lines indicating regular behaviour of the data. Next, we produce Table 2 for the RQA analysis.

The value of REC appearing larger than threshold indicates that the data may be regular. The high value of DET (0.93) possibly shows that the data is deterministic. Again, high values of LMAX and ENTROPY possibly suggest that the present data is non-chaotic in nature.

5 Conclusion

The present chapter produces an in-depth statistical signal processing approach to understand different perspectives of the monthly mobile data traffic per smartphone

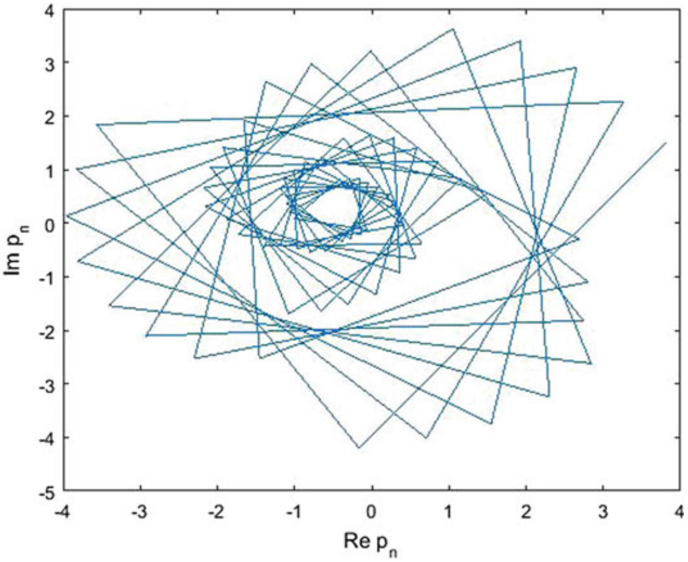


Fig. 15 Plot of p_n in the complex plane

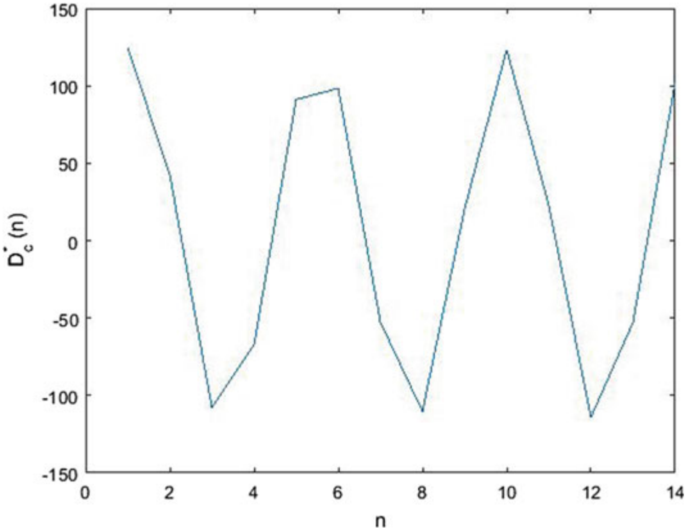


Fig. 16 Plot of $D_{\mathrm{c}}^{*}(n)$ versus n

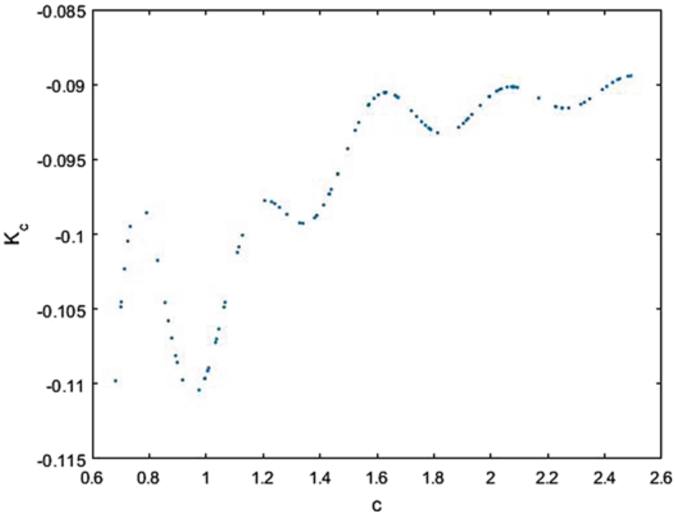


Fig. 17 Plot of K_c versus c

Fig. 18 RP analysis of monthly mobile data traffic per smartphone worldwide from January, 2014 to December, 2019 (in GB)

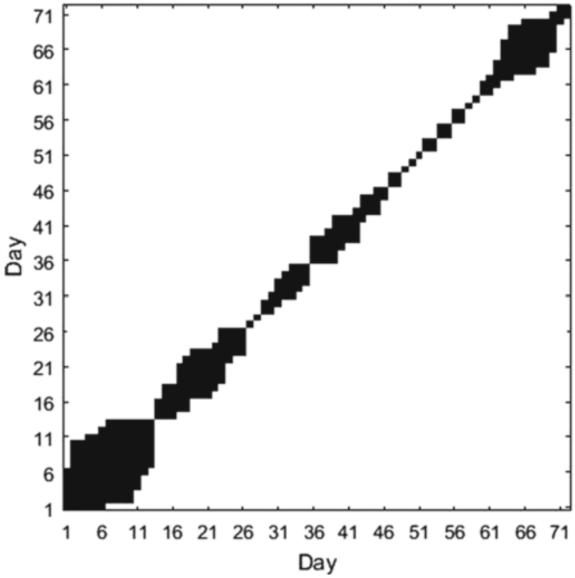


Table 2 RQA analysis of monthly mobile data traffic per smartphone worldwide from January, 2014 to December, 2019 (in GB)

Range after normalizing	Threshold value	REC	DET	LMAX	ENTR
2.83	0.05	0.06	0.93	7.42	1.70

worldwide from January, 2014 to December, 2019 [7]. The study shows a prevailing long memory in the governing mechanism. We can find a self-similarity in the present time series which strengthens and its predictability and perhaps eases the procedure of forecasting. We also can observe that 'soc' is possibly not present in this time series which in turn indicates that the growth as visible in the mobile data traffic in smartphone worldwide is probably an unstable growth and governing mechanism may not be able to survive and self-repair if any substantial damage or perturbation occurs in any of its favourable components at any level of time due to some unavoidable external factors. The study also indicates a possible deterministic and nonlinear profile and absence of chaos. This finding is interestingly different as there are available communications [1–4] which demonstrate chaos for different wireless network traffic. The possible absence of chaos in the present time series should enhance the reliability of predictions both short-term and long-term. The task of forecasting over the present trend can be a very interesting future course of study. The present mode of study can also be extended in designing smart mobile data-based transportation system [41], dynamic mode decision for cellular D2D communications [42], intrusion detection and prevention in wireless sensor networks [43], etc.

References

1. Liu X, Fang X, Qin Z, Ye C, Xie M (2011) A short-term forecasting algorithm for network traffic based on chaos theory and SVM. J Netw Syst Manage 19:427
2. Nikulchev E, Pluzhnik E (2014) Study of chaos in the traffic of computer networks. Int J Adv Comput Sci Appl 5(9):60
3. Liu Y, Zhang J (2016) Predicting traffic flow in local area networks by the largest lyapunov exponent. Entropy 18:32
4. Mukherjee S, Ray R, Samanta R, Khondekar MH, Sanyal G (2017) Nonlinearity and chaos in wireless network traffic, Chaos. Solitons Fractals 96:23
5. Shaikh S (2016) Fractal traffic management in wireless network. Int J Innovations Eng Res Technol 3(1):1
6. Mirzakulova SA, Shuvalov VP, Mekler AA (2017) Studying network traffic using nonlinear dynamics methods. J Theor Appl Inf Technol 95(21):5869
7. Statista: https://www.statista.com/statistics/738977/worldwide-monthly-data-traffic-per-smartphone/#statisticContainer. As retrieved on 20 Apr 2020
8. Hurst HE (1951) Long-term storage capacity of reservoirs. Trans Am Soc Civ Eng 116:770
9. Scafetta N, Grigolini P (2002) Scaling detection in signal: diffusion entropy analysis. Phys Rev E 66:036130
10. Sarkar A, Barat P, Mukherjee P, Bandyopadhyay SK (2005) Scaling analysis of daily sunspot numbers. In: Proceedings of national conference on nonlinear systems and dynamics (held at A.M.U., Aligarh during 24–26 Feb 2005), p 155
11. Saha G, Rakshit K, Ghosh K, Chaudhuri KS (2019) A new proposal on the relation between irregularity index and scaling index in a non-stationary self-affine signal obeying fractional Brownian Motion. Bull Calcutta Math Soc 111(1):79
12. Saha G, Rakshit K, Ghosh K, Chaudhuri KS (2019) A revisit to the relation between irregularity index and scaling index in a stationary self-similar signal obeying fractional Gaussian Noise. J Calcutta Math Soc 15(2):139
13. Higuchi T (1988) Approach to an irregular time series on the basis of the fractal theory. Physica D 31(2):277

14. Higuchi T (1990) Relationship between the fractal dimension and the power law index for a time series: a numerical investigation. Physica D 46(2):254
15. Bak P, Tang C, Wiesenfeld K (1987) Self-organized criticality: an explanation of the 1/f noise. Phys Rev Lett 59:381
16. Bak P, Tang C, Wiesenfeld K (1988) Self-organized criticality. Phys Rev A 38:364
17. Sarkar A, Barat P (2006) Analysis of rainfall records in India: self organized criticality and scaling. Fractals 14(4):289
18. Gautama T, Mandic DP, Van Hulle MM (2004) The delay vector variance method for detecting determinism and nonlinearity in time series. Physica D 190(3–4):167
19. Ahmed I (2010) Detection of nonlinearity and stochastic nature in time series by delay vector variance method. Int J Eng Technol 10(2):22
20. Gottwald GA, Melbourne I (2004) A new test for chaos in deterministic systems. Proc Roy Soc Lond A 460(2042):603
21. Gottwald GA, Melbourne I (2005) Testing for chaos in deterministic systems with noise. Physica D 212(1):100
22. Gottwald GA, Melbourne I (2008) Comment on reliability of the 0–1 test for chaos. Phys Rev E 77(2):028201
23. Gottwald GA, Melbourne I (2009) On the implementation of the 0–1 test for Chaos. SIAM J Appl Dyn Syst 8(1):129
24. Gottwald GA, Melbourne I (2009) On the validity of 0–1 test for chaos. Nonlinearity 22(6):1367
25. Eckmann J-P, Kamphorst SO, Ruelle D (1987) Recurrence plots of dynamical systems. Europhys Lett 4(9):973
26. Webber CL Jr, Zbilut JP (1994) Dynamical assessment of physiological systems and states using recurrence plot strategies. J Appl Physiol 76(2):965
27. Zbilut JP, Webber CL Jr (1992) Embeddings and delays as derived from quantification of recurrence plots. Phys Lett A 171(3):199
28. Atay FM, Altýntas Y (1999) Recovering smooth dynamics from time series with the aid of recurrence plots. Phys Rev E 59(6):6593
29. Marwan N, Wessel N, Meyerfeldt U, Schirdewan A, Kurths J (2002) Recurrence plot based measures of complexity and its application to heart rate variability data. Phys Rev E 66(2):026702.1
30. Marwan N, Romano MC, Thiel M, Kurths J (2007) Recurrence plots for the analysis of complex systems. Phys Rep 438(5):237
31. Marwan N (2011) How to avoid potential pitfalls in recurrence plot based data analysis. Int J Bifurcat Chaos 21(4):1003
32. Hossain KM, Ghosh DN, Ghosh K, Bhattacharjee AK (2015) Complexity in solar irradiance from the earth radiation budget satellite. IEEE Syst J 9(2):487
33. Samadder S, Ghosh K, Basu T (2015) Investigation of nonlinearity and Chaos in prime Indian and American stock exchange indices. Hyperion Int J Econophys New Econ 8(1):65
34. Hossain KM, Ghosh DN, Ghosh K, Bhattacharjee AK (2012) Nonlinearity and chaos in ^{8}B solar neutrino flux signals from Sudbury neutrino observatory. Fractals 20(1):17
35. Kugiumtzis D (1999) Test your surrogate data before you test for nonlinearity. Phys Rev E 60(3):2808
36. Schreiber T, Schmitz A (2000) Surrogate time series. Physica D 142(3):346
37. Webber CL, Marwan N (2015) Recurrence quantification analysis. Springer, pp 3–15
38. ITU World Telecommunication/ICT Indicators database. https://www.itu.int/en/ITU-D/Statistics/Pages/stat/default.aspx. As Retrieved on 20 Apr 2020
39. Mobile Data Usage in India by Year–2014–2018. https://dazeinfo.com/2019/10/08/mobile-data-usage-in-india-by-year-graphfarm/. As Retrieved on 20 Apr 2020
40. Telecom Statistics India (2019) Economics Research Unit, Department of Telecommunications, Government of India
41. Yang W, Wang X, Song X, Yang Y, Patnaik S (2018) Design of intelligent transportation system supported by new generation wireless communication technology. Intelligent systems: concepts, methodologies, tools, and applications, IGI Global, p 715

42. Jayakumar L, Dumka A (2020) Energy aware dynamic mode decision for cellular D2D communications by using integrated multi-criteria decision making model. Int J Ambient Comput Intell 11(3) (to appear)
43. Chandre PR, Mahalle PN, Shinde GR (2020) Deep learning and machine learning techniques for intrusion detection and prevention in wireless sensor networks: comparative study and performance analysis. In: design frameworks for wireless networks. Springer, p 95