

Algorithms for Intelligent Systems

Series Editors: Jagdish Chand Bansal · Kusum Deep · Atulya K. Nagar

Bimal Kumar Roy

Atul Chaturvedi

Boaz Tsaban

Sartaj Ul Hasan *Editors*

Cryptology and Network Security with Machine Learning

Proceedings of ICCNSML 2022

 Springer

Algorithms for Intelligent Systems

Series Editors

Jagdish Chand Bansal, Department of Mathematics, South Asian University,
New Delhi, Delhi, India

Kusum Deep, Department of Mathematics, Indian Institute of Technology Roorkee,
Roorkee, Uttarakhand, India

Atulya K. Nagar, School of Mathematics, Computer Science and Engineering,
Liverpool Hope University, Liverpool, UK

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

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

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

Indexed by zbMATH.

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

Bimal Kumar Roy · Atul Chaturvedi ·
Boaz Tsaban · Sartaj Ul Hasan
Editors

Cryptology and Network Security with Machine Learning

Proceedings of ICCNSML 2022

 Springer

Editors

Bimal Kumar Roy
Applied Statistics Unit
Indian Statistical Institute
Kolkata, West Bengal, India

Atul Chaturvedi
Department of Mathematics
Pranveer Singh Institute of Technology
Kanpur, India

Boaz Tsaban
Department of Mathematics
Bar-Ilan University
Ramat Gan, Israel

Sartaj Ul Hasan
Department of Mathematics
Indian Institute of Technology Jammu
Jammu, Jammu and Kashmir, India

ISSN 2524-7565

ISSN 2524-7573 (electronic)

Algorithms for Intelligent Systems

ISBN 978-981-99-2228-4

ISBN 978-981-99-2229-1 (eBook)

<https://doi.org/10.1007/978-981-99-2229-1>

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2024

This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.

The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Paper in this product is recyclable.

Preface

In 2022, we decided to start the annual series of ICCNSML conferences in order to encourage cryptographers and researchers to propose and work on all the related topics of conference. This is the first volume of papers from our first annual ICCNSML 2022 conference hosted at PSIT, Kanpur 2022, on December 16–18, 2022. We were motivated to launch this series of conferences to attract more researchers to work on related problems on cryptography with Machine Learning (ML). There is a gap in the academics of mathematics and computer science which we hope to bridge with this effort. Mathematicians primarily recognize publications in journals, whereas cryptographers almost always publish their results quickly in proceedings volumes of conferences which are the most prestigious venues for the research area. Many mathematicians are not accustomed to the model of submitting a paper by the conference deadline, presenting the work at the conference, and publishing in the proceedings volume. We wanted to provide a regular annual venue for researchers to contribute to the cryptographic research community at this accelerated pace, and AIS, Springer was an ideal place to recognize all the efforts.

We are at a point in time where it is increasingly important for researchers to be involved in cryptography research, as we set out to determine the next generation of cryptographic systems based on hard mathematical problems which can withstand attacks from a quantum computer once it is built. Recently, in 2017, NIST launched a 5-year international competition to determine Post-Quantum Cryptosystems (PQC). So it is the need of the hour to develop more methods in order to surprise intruders. So we thought that ICCNSML 2022 can play a complimentary role by encouraging researchers to work on and publish interesting results related to current trends of cryptosystems.

The founders of ICCNSML 2022, Prof. Bimal Kumar Roy, Prof. Boaz Tsaban, Prof. Sartaj Ul Hasan along with me, acted as Chairs, and have joined the editorial board for the proceedings. We are excited at the success of the first edition of the conference, which had more than 100 submitted papers and more than 250 registered participants. The 3-day conference included 15 talks/key note session and multiple paper presentation schedules and a poster session, representing 38 accepted papers from various places of India and abroad. The large auditorium was packed with an

audience of more than hundreds of researchers at every given instance. The Technical Program Committee (TPC) worked hard to evaluate the hundred plus submissions and did an outstanding job of selecting many papers worthy of presentation and publication. We hope ICCNSML will continue to be a successful conference and a prestigious venue for presentation of important cryptographic results.

Kolkata, India
Ramat Gan, Israel
Jammu, India
Kanpur, India

Prof. Bimal Kumar Roy
Prof. Boaz Tsaban
Prof. Sartaj Ul Hasan
Prof. Atul Chaturvedi

Contents

Role and Applications of Emerging Technologies in Smart City Architecture	1
Nand Kishore Sharma, Surendra Rahamatkar, and Abhishek Singh Rathore	
A Review on Blockchain-Based Electronic Health Record (EHR) System for Secure Data Storage	15
Vandani Verma and Garima Thakran	
SVM-RF: A Hybrid Machine Learning Model for Detection of Malicious Network Traffic and Files	29
Prashant Mathur, Arjun Choudhary, Chetanya Kunndra, Kapil Pareek, and Gaurav Choudhary	
Key-Insulated Aggregate Proxy Signature	41
P. V. S. S. N. Gopal, T. Gowri, and P. Vasudeva Reddy	
A Comprehensive Study of Cryptocurrency Trend Analysis Based on a Novel Machine Learning Technique	53
Paras Sharma, Adhiraj Gupta, Rakesh Kumar Bajaj, and Prateek Thakral	
Flaws of a Password-Based Three-Party Authenticated Key Agreement Protocol in Post-quantum Environment	63
Sonam Yadav, Vivek Dabra, Pradeep Malik, and Saru Kumari	
Multivariate Aggregate and Multi-signature Scheme	71
Satyam Omar, Sahadeo Padhye, and Dhananjay Dey	
Optical Network Modeling and Performance Using Random Graph Theory	77
Rahul Deo Shukla, Ajay Pratap, and Raghuraj Singh Suryavanshi	
A Survey on Recent Progress in Blockchain Technology	91
Naseem Ahmad Khan, Prateek Singh, Radiant Ambesh, and Md Tarique Jamal Ansari	

Cryptanalysis of Lattice-Based Threshold Changeable Multi-secret Sharing Scheme	99
Ramakant Kumar and Sahadeo Padhye	
Amazon Web Service IOT and Authentication of Edge Devices	109
Meenakshi Srivastava and Arsh	
On Picture Fuzzy Information-Based Hybrid Cryptographic TOPSIS Approach for Best Suitable Cloud Storage and Security Level	125
Himanshu Dhumras, Rakesh Kumar Bajaj, and Varun Shukla	
Network Layer Performance of Hybrid Buffer-Based Optical Router	135
Sumit Chandra, Shahnaz Fatima, and Raghuraj Singh Suryavanshi	
Efficient and Secure Data Aggregation for UAV-to-Ground Station Communication in Smart City Environment	147
Girraj Kumar Verma, Dheerendra Mishra, and Neeraj Kumar	
BBIWMS: A Secure Blockchain-Based Framework for Integrated Water Management System for Smart City	157
B. C Girish Kumar, K. G. Harsha, G. Mahesh, Varun Shukla, and Surendra Talari	
An Intelligent Network Intrusion Detection Framework for Reliable UAV-Based Communication	169
Sujit Beborotta and Sumanta Kumar Singh	
Distributed and Hash-Based Mixers for User Anonymity on Blockchain	179
P. Guna Shekar, Raghwendra Singh, Debanjan Sadhya, and Bodhi Chakraborty	
Implementation and Analysis of Different Visual Cryptographic Schemes	193
Vanashree Gupta and Smita Bedekar	
A New Data Communication Method Using RSA and Steganography	203
Varun Shukla, Manoj Kumar Misra, Shivani Dixit, and Himanshu Dhumras	
Some Computational Attacks on Threshold Secret-Sharing Scheme by Outside Adversaries	213
L. Sreenivasulu Reddy	
Influence of COVID-19 Pandemic on Digital Payment Market Growth	225
Mohammed Kamran Siddiqui and Krishan Kumar Goyal	

Two-Level Security of Color Image Using 9D-Hyperchaotic System and DWT	241
Sonali Singh and Anand B. Joshi	
A Novel Approach For Secure Data Aggregation Scheme in Battlefield Surveillance Using Elliptic Curve Cryptography	265
Abhishek Bajpai and Anita Yadav	
A Recent Survey of Reversible Data Hiding Techniques for 2D and 3D Object Models	279
Amit Verma, Ruchi Agarwal, and Bhogeswar Borah	
Demystifying Facial Expression Recognition Using Residual Networks	295
Pratyush Shukla and Mahesh Kumar	
Block Farm: Blockchain-Based Platform for the Agriculture Supply Chain	309
Udai Bhan Trivedi, Manoj Srivastava, and Manish Kumar	
A Lattice-Based Key Exchange Protocol Over NTRU-NIP	325
Sonika Singh and Sahadeo Padhye	
Blockchain Within the Insurance Industry: A Bibliometric Analysis	335
Lalit Garg, Luca Bugeja, Corinne Marie Formosa, and Varun Shukla	
CNN-LSTM: A Deep Learning Model to Detect Botnet Attacks in Internet of Things	353
Chetanya Kunndra, Arjun Choudhary, Prashant Mathur, Kapil Pareek, and Gaurav Choudhary	
Prediction of Covid-19 Using Artificial Intelligence [AI] Applications	367
R. Kishore Kanna, Mohammed Ishaque, Bhawani Sankar Panigrahi, and Chimaya Ranjan Pattnaik	
A New Authentication Protocol for RFID and Applications in E-Passport	375
Vandani Verma and Garima Jain	
Comprehensive Survey on AQI Prediction Using Machine Learning Algorithms	387
Imran Khan and Rashi Agarwal	
Ransomware 3.0—A Weapon for Next-Generation Information Warfare	397
Mohiuddin Ahmed, A. N. M. Bazlur Rashid, and Al-Sakib Khan Pathan	
Securing Transmission of Medical Images Using Cryptography Steganography and Watermarking Technique	407
Satish Kumar, Pawan Kumar Chaurasia, and Raees Ahmad Khan	

New Commitment-Based Client–Server Key Establishment Protocol 421
Varun Shukla, Surendra Talari, Shishir Kumar, P. Vinooth,
and Harikesh Singh

Role of Virtualization in Secure Network 433
Anju Shukla, Shishir Kumar, and Varun Shukla

Blockchain-Based NFT for Evidence System 441
Aditya Kumar Sharma and Brijesh Kumar Chaurasia

Securing Digital Audio Files Using Rotation and XOR Operations 453
Abdul Gaffar

Author Index 469

About the Editors

Bimal Kumar Roy is a former Director of the Indian Statistical Institute (ISI), Kolkata. He is Founder and General-Secretary, Cryptology Research Society of India (CRSI). He is a cryptologist from the Cryptology Research Group of the Applied Statistics Unit of ISI, Kolkata. He received Ph.D. in Combinatorics and Optimization in 1982 from the University of Waterloo under the joint supervision of Ronald C. Mullin and Paul Jacob Schellenberg. Currently, he is working on Combinatorics, and application of Statistics in Cryptology and Design of Experiments. In 2015, he was awarded Padma Shri, India's fourth-highest civilian honour, recognizing his accomplishments and contribution to education. In 2019, he was appointed as the chairperson of the National Statistical Commission, Ministry of Statistics and Programme Implementation, Government of India. He has been a driving force in advancing the important field of cryptology in India, elevating its visibility to international level. He devoted his career to strengthening India's standing in this timely, fast advancing field.

Atul Chaturvedi is currently working as Professor in Mathematics at PSIT, Kanpur. He has received his M.Sc., M.Phil. and Ph.D. from Dr. B.R.A University, Agra. His research interests include Cryptography and Network Security protocols particularly Lattice based cryptography and Non Commutative Ring based cryptography. He is a life member of Cryptology Research Society of India (CRSI), Indian Society for Technical Education (ISTE) and Indian Science Congress Association (ISCA). He has published various books, research papers in reputed journals and reviewer of many International journals. He has been convener of many national and international conferences/workshops in the area of Cryptology & Network Security under the aegis of various prestigious societies/ organizations like DRDO, ISI Kolkata, AICTE, ISTE, etc. He is copyright owner of many cryptographic algorithms (from Department of Industrial policy and promotion, Ministry of Commerce, Government of India). He has/is guided/guiding many doctoral research fellows in the area of Cryptography and Network Security.

Boaz Tsaban is an Israeli mathematician at the faculty of Bar-Ilan University working as a capacity of Professor. His research interests include selection principles within Set Theory and Nonabelian Cryptology, within Mathematical Cryptology. At the age of 16 he was selected with other high school students to attend the first cycle of a special preparation program in mathematics, at Bar-Ilan University, being admitted to regular mathematics courses at the University a year later. He has completed his B.Sc., M.Sc. and Ph.D. degrees with highest distinctions. Two years as a post-doctoral fellow at Hebrew University were followed by a three-year Koshland Fellowship at the Weizmann Institute of Science before he joined the Department of Mathematics, Bar-Ilan University in 2007. Tsaban's doctoral dissertation, supervised by Hillel Furstenberg, won, with Irit Dinur, the Nessyahu prize for the best Ph.D. in mathematics in Israel in 2003. In 2009 he won the Wolf Foundation Krill Prize for Excellence in Scientific Research.

Sartaj Ul Hasan is working as an Assistant Professor at IIT Jammu. Prior to joining IIT Jammu, he was a Scientist in DRDO, Delhi from February 2007 to January 2018. From September 2011 to December 2012, he was a postdoctoral fellow at Carleton University, Canada, where he worked with Prof. Daniel Panario and Prof. Qiang Wang. He received his Ph.D. in 2010 from the Indian Institute of Technology Mumbai, where he was a student of Prof. Sudhir R. Ghorpade. His main areas of research interests include Finite Fields, Cryptography, and Coding Theory. He is guiding many doctoral candidates in the field of cryptology and editor of many international journals. He has delivered invited talks/guest lecturers in various institutions of international repute.

Role and Applications of Emerging Technologies in Smart City Architecture



Nand Kishore Sharma, Surendra Rahamatkar, and Abhishek Singh Rathore

1 Introduction

In this growing age of technology, many sensors and devices are being used as smart objects to sense phenomena on a single platform. These Smart Objects are implanted with advanced software. They generate and collect a huge amount of multimedia data (audio, images, and videos) for processing. As these objects are distributed and so the motive is to integrate and manage them efficiently for unique identification. And enable the user to access them remotely in real time. The devices or sensors may be considered as Smart Objects if they are having sufficient and necessary computation processing skills. These objects are recognized by their name tag and address which is unique. Wireless-Sensor-Networks and Internet of Things are widely used technologies to connect the sensors via the internet to produce and send data remotely.

1.1 Smart City Architecture with Applications

The smart city is an innovative city and relies on the smart framework as represented in Fig. 1. This framework comprehends the physical infrastructure, networking system, centralized computing center, and data storage system (also responsible to create the

N. K. Sharma (✉) · S. Rahamatkar
Amity School of Engineering and Technology, Raipur, C.G. 493225, India
e-mail: er.nksharma.mtechcs@gmail.com

S. Rahamatkar
e-mail: srahamatkar@rpr.amity.edu

A. S. Rathore
Shri Vaishnav Vidyapeeth Vishwavidyalaya, Indore, M.P. 453111, India
e-mail: abhishekaturjain@gmail.com

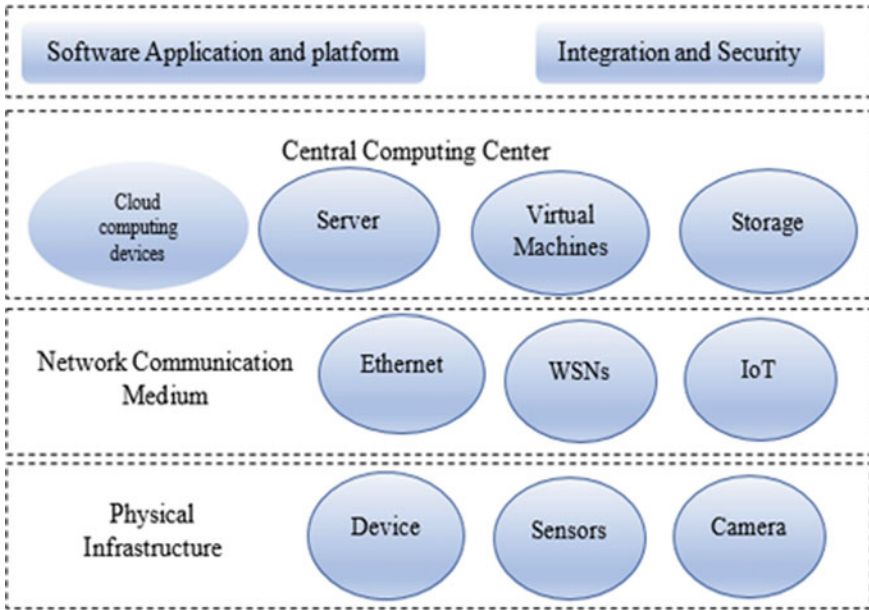


Fig. 1 Smart city framework

replica of data), software applications and platforms, integration and security, and higher level domain use-cases.

The physical infrastructure requires sensors, devices, and cameras to capture and generate data, which is one of the framework's core elements. This information is kept in a central computation center for processing. Several types of network channels or communication mediums are employed to send and receive data from acquisition devices to central computing centers, such as Ethernet, Wireless-Sensor-Networks (WSNs), Internet of Things (IoT), and Fiber optic cables. After passing through several servers, cloud computing devices, and virtual machines in the central computing center, the data is stored in the storage unit. Data is protected by several network standard protocols, data security mechanisms, and APIs. Later, the domain use-cases also indicated as applications may be segmented into—Smart city infrastructure, Smart security solutions, and Smart networks that utilize the data to predict the outcome. A well-structured network and a smart management system are required for all offering applications to make the system intelligent. Although information exchange is essential, only reliable information is required for quick decisions and responses. Table 1 is representing the segment-wise applications.

Every function in the city generates a massive amount of data, which contains some hidden insights about its surroundings. It implies that the data and the technologies used to process it can be imagined as the “umbrella” for the smart city. The technologies used in smart city applications are depicted in Fig. 2.

Table 1 Segment-wise smart city applications

Segments	Related reference	Applications
Smart city infrastructure	[1–3]	<ul style="list-style-type: none"> – Centralized and integrated controlling – Smart mobility services – Smart communication interface – Smart lightening system – Smart waste management system – Smart agriculture – Air pollution monitoring system – Adaptive traffic management – Enhancement in the environmental monitoring system
Smart security solutions	[3–7]	<ul style="list-style-type: none"> – Integrated surveillance in an open and critical area – Road-side traffic management system – Object tracking and monitoring – Monitoring of highway vehicle – Crime detection and monitoring system – Behavioral analysis of people – Abnormal vehicle driving – Over limit speed vehicle tracking – Activity recognition and classification
Smart network	[3, 6–11]	<ul style="list-style-type: none"> – Quick response to emergency services – Road-side assistance – Quick medical and ambulance support for road accidents – Early warning dissemination system – Prevention of unauthorized network access – Remotely monitoring and management of network devices

Every technology is having its framework for every smart city application. In this paper, the author is mainly focusing on deep learning techniques, Machine learning, and IoT enabled with wireless sensors to provide the research directions while selecting the technology. In this research, the author is contributing to the present:

- The technologies involved in smart city applications.
- Various emerging technologies are used in smart city applications.
- Several types of city data.

1.2 Data Analytics

The smart city paradigm entirely depends on data, the fusion of data with upended communication technologies. Enormous challenges come in the data collection, analysis, and distribution [12]. The generated data is heterogeneous and sparse in most

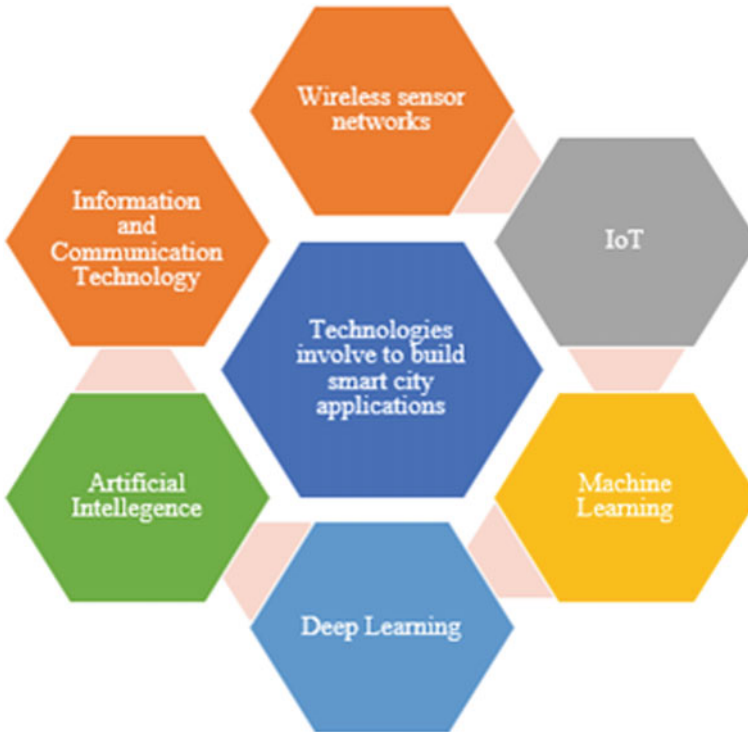


Fig. 2 Trending technologies for smart city applications

cases. The processing of this Big Data in real time is a matter of concern. For data processing, the understanding of data is must, because the smart city application positions on data and its analysis [13]. The data analysis takes place in four phases [14, 15]—Data gaining, data preprocessing and cleaning, data analytics and exploratory data analysis, and outcome of data. In the first phase—the data is captured from the different sensors, which are diverse and heterogeneous. Because only excellent data is required for analytics to produce knowledge, preprocessing techniques are used on data to clean and eliminate impurities. For data analytics, a variety of techniques are used to uncover hidden insights, inferences, patterns, and correlations. In the end, the data was prepared to generate the results and predictions, such as surveillance, estimation of the roadside traffic flow, etc.

The devices involved in all of the aforementioned applications and areas are classified according to their constraints, which include processing power, memory, interoperability, security, data confidentiality, and bandwidth. As a result, numerous challenges are expected to be encountered in the development of an effective system.

2 Background

2.1 Deep Learning

Applications of deep learning are used in smart cities. Using auditory data collection and a deep learning technique, smart cities can anticipate the class labels of activities [16]. Data security is a foremost matter because it is so sensitive. Since this sensitive data can modify by unauthorized access. The deep learning techniques can protect it from cyber-attacks [17]. Table 2 shows the parameters for data cause and attacks with best deep learning practices, and Fig. 3 shows the terminologies and concerns related to the data.

Data is generated in various forms from various types of devices and represented in numerous forms. Table 3 is summarizing the data [13] with data sources, data types, and its representation with deep learning models. Figure 4a depicts the data generated by various sources and Fig. 4b depicts the city data used by deep learning models.

Table 2 Deep learning practices for data cause and attacks

Parameters related to data	Deep learning practices
Data Cause	<ul style="list-style-type: none"> - Convolutional Neural Network - Generative Adversarial Network (GAN) - Long-Short Term Memory (LSTM) - Deep Brief Network - Restricted Boltzmann Machine
Attack detection and Prevention from attacks	<ul style="list-style-type: none"> - Recurrent Neural Network - Convolutional Neural Network - Generative Adversarial Network (GAN) - Long-Short Term Memory (LSTM) - Deep Brief Network - Restricted Boltzmann Machine

Fig. 3 Data cause and attacks

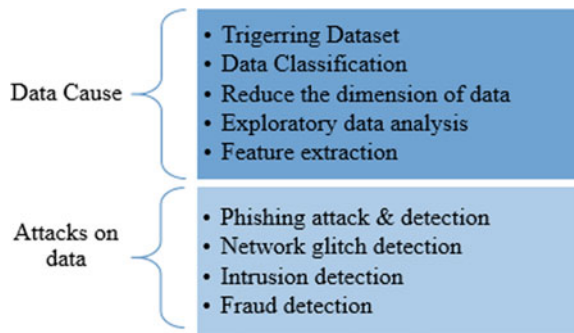


Table 3 Outline of the data used in deep learning techniques

Data source	Type of generated data	Data representation	Data description	Deep Learning model
Camera	Vehicle image data, Vehicle license plate, Face, Parking space images	2-D and 3-D matrix $a_1, 1 \dots a_1, n$ $a_m, 1 \dots a_m, n$	a is the data with dimension $m \times n$, dimension indicates image resolutions	Convolutional Neural Network (CNN) SqueezeNet [18]
Sensors	Road-side vehicle traffic and speed of the vehicle, Human behavior and activity, Human movement	Time categorization data:- { $a_1, a_2, \dots, a_m \dots a_T$ } Vector:- $[\times 1; \times 2; \dots; \times n]$	a is a sequence, a_m sequence at time interval m and T is the length	Recurrent Neural Network (RNN), Stacked Auto Encoder (SAE), Deep Belief Networks (DBN) [19]

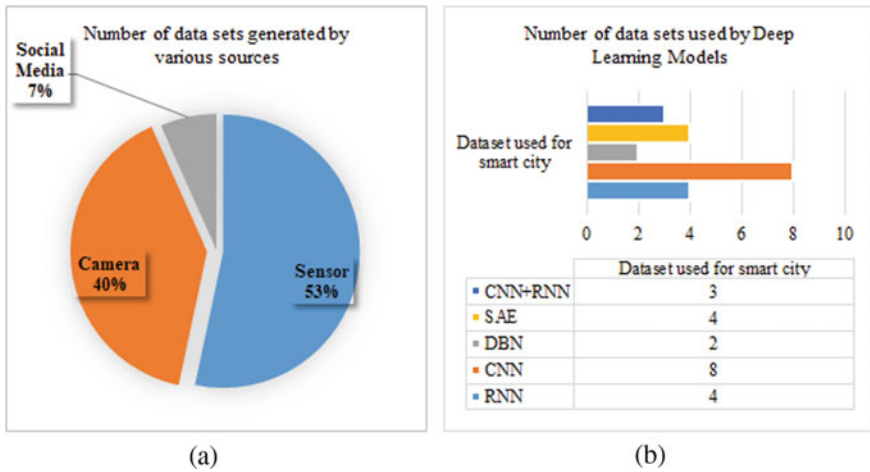


Fig. 4 a Data generated by various sources. b Dataset used by DL models

Deep learning offers several highly potent tools and technologies that can handle data causes and attacks. It is also an exceptionally efficient technique for surveillance applications in adding up to data handling, since manual surveillance takes a lot of time. Multimedia sensors’ involvement enables the acquisition of more accurate and concrete data. For more precise identification in real-time scenarios, modern video, and audio sensors can be used, along with some traditional scalar sensors. Figure 5 displays the effectiveness of various deep learning models for activity recognition when used with sensor data for surveillance.

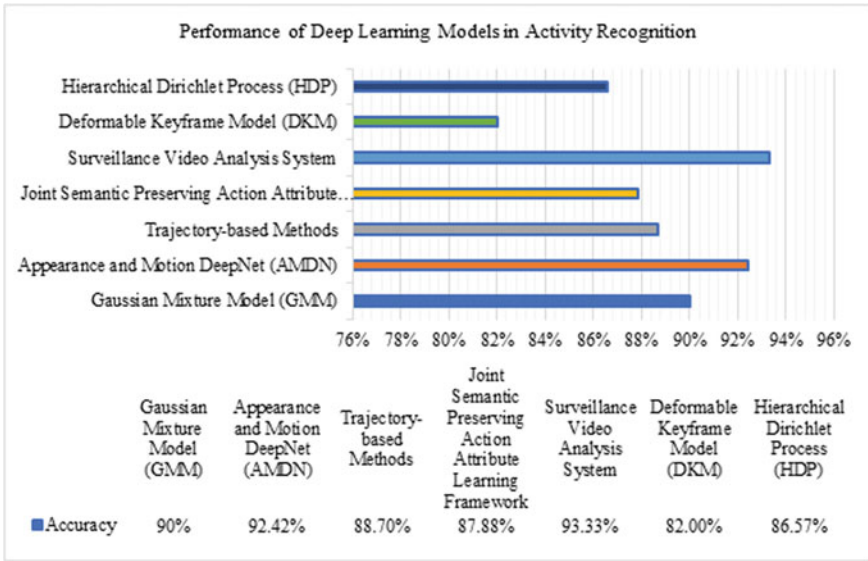


Fig. 5 Performance of deep learning in activity recognition

2.2 Machine Learning

Advanced algorithms that can complete tasks without requiring direct human involvement are referred to as machine learning. Classification, regression, feature extraction, clustering, and anomaly detection are the top issues covered by machine learning algorithms like K-nearest neighbors, K-means, Support-Vector-Regression, Linear regression, Logistic regression, Feed-forward neural network, Random Forest, and decision tree. All mentioned algorithm uses to process the smart city data [20]. Figure 6a is representing the contribution of these algorithms to processing the city data. According to statistics, processing city data mostly involves the use of classification algorithms. A class can be referred to as a target, a label, or a category. Classification is the process of assigning a class to the input data. For instance, if the input vector is c , and k is any class then the discriminant function can be signified as ck . A class can be assigned to input vector x via a discriminant function. if $f(x) > 0$, then assigned class C1 to vector x , otherwise assign class C2 to vector x . The discriminative model is also used for the same work, it acquires the posterior class probability denoted as $p(C_k | x)$ and uses to assign the class. The classification comes under supervised machine learning. The classification is referred to as predictive modeling, where a class label is assigned to input data. Figure 6b represents the types of classification and Fig. 7 is showing the algorithms used for classification with applications.

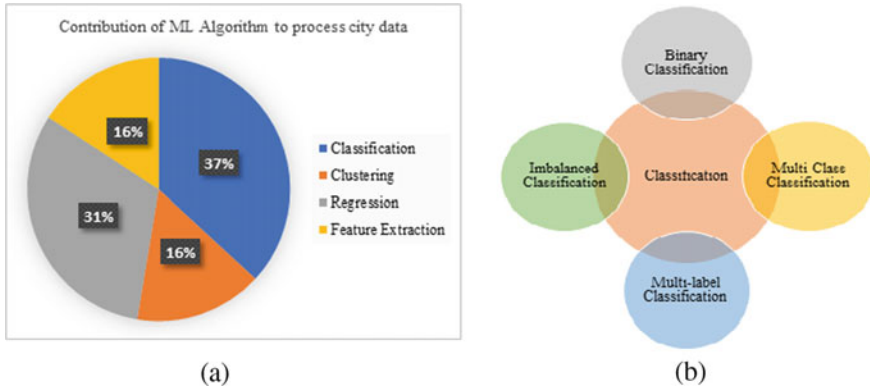


Fig. 6 a Contribution of Machine Learning to process city data, b Types of classification algorithms

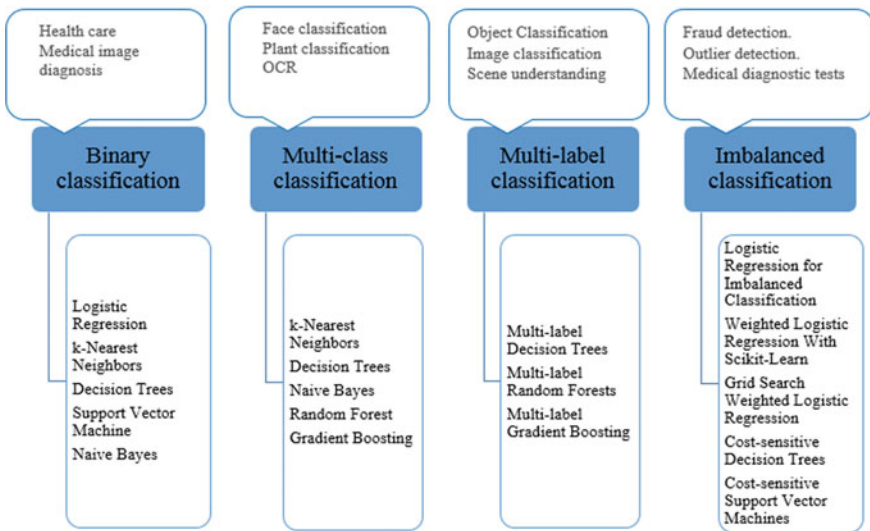


Fig. 7 Classification algorithms with applications

2.3 Wireless-Sensor-Networks and Internet-of-Things

There are numerous wireless-sensor-network technologies available that are utilized to create applications for smart cities, but IPv6 over Low-Power Wireless Personal Area Network is the one that has been proven to be the most promising. It is appropriate to transfer relatively little data using low power, energy-efficient protocols. ZigBee is also well-known for its low-cost computing power. Bluetooth Low Energy (BLE) became the dominant technology since it is mostly utilized in small-range connectivity applications like smart homes and healthcare systems that need very

low power control and monitoring mechanisms. It may be more advantageous to use certain traditional scalar sensors along with new, improved video and audio sensors in this situation. Researchers and users started to investigate ways to receive more exact, accurate, trustworthy, and realistic information in fast-changing environments along with these substantial technological developments. As an outcome, distributed systems with stronger sensor nodes are introduced as Wireless Multimedia Sensor Networks (WMSN). Because the signals can be modulated, conventional Wi-Fi transmissions are no longer useful. RFID, Radar, ultrasonic sensors, and webcams are examples of human-computer interface devices that have advanced in technology and shown potential for future intelligent communication. The wireless multimedia and scalar sensors-based architecture proposed by [21] for a multilayer automatic surveillance system. At the WMS node level, the system consists of two layers. The first layer contains scalar sensors with acoustic, vibratory, and motion detection capabilities. This layer activates the second layer comprising multimedia sensors having audio and video capture capabilities. The challenge is to offer interoperability services within the sensors and smart objects. There are three types of smart objects -Activity, Policy, and Process. There are six main elements to manage all activity [22]. Object-ID and address are used for identification purposes. Electronic Product Codes (EPC) and Ubiquitous Codes (uCode) are few methods available for identification purposes and some addressing methods like IPv6 and IPv4 are employed for addressing. All offerings are characterized into four classes [23] as-Identity-related issues, Aggregation of services, and Collaborative and Ubiquitous Services. Moreover, the security and confidentiality of data become the biggest challenge [24]. Even incomplete and missing data is also a vital problem. Due to the need for heavy and real-time computation energy-efficient systems are required [25]. Although numerous studies have been conducted regarding green communication technologies [26]. More precisely, all emerging technologies for supporting wide-area Machine-to-Machine (M2M) networks are grounded on IoT devices [27], attentive to the standards for data communications, services, and support for Machine-to-Machine [28].

The technologies involved in IoT technology provide many opportunities and facilities for the implementation of its applications [29]. In service-based architecture, middleware is accountable to provide the deployment of devices as a service [30]. The resources used in this technology are limited and expensive [2, 31] even extremely dynamic distributed networks with key management algorithms are not found so appropriate. Consequently, need to save their surrounding environments with limited computing resources such as power and storage. More robust measures must be taken to allow system developers to advance their methods for better security mitigation [32–34] defined four-layer communication architectures that contain— Perception Layer, Communication Layer, Support Layer, and Business Layer. [14, 15, 35] predicted that citywide human mobility is critical. Human mobility modeling is the key building block for creating a smart city. In general, there are two broad categories of human mobility: routine human mobility and irregular human mobility. Routine human mobility takes up a large proportion of an entire historical dataset.

Thus, making an accurate prediction of routine human mobility guarantees that a system should function most of the time.

3 Open Challenges and Issues

Technologies like the Internet of Things and wireless networks are making it possible to develop new applications or to improve existing ones. The collected data may be complete, incomplete, or insufficient to extract information because it was obtained from numerous different sources. However, it requires significant computational power and resources to extract the features from the data; it is not a straightforward task. In this case, it is best to use the associated data to extract the data that can compensate for the weak or missing data. Interoperability and Resource Constraints, Resilience to Physical Attacks and Natural Disasters, Autonomic Control and Scalability, and Information Volume with Privacy Protection are the Challenges that need to be met to meet the Security Needs and to describe the Appropriate Countermeasures. Huge devices with precise real-time processing are needed because of the expansion of smart city applications. Figure 8a and b shows the data privacy in IoT devices involved in cyber-attacks and vulnerabilities sequentially [36].

To deal with the real-time data and produce reliable results, algorithm improvement is also necessary. A significant area of research focused on the gathering of relevant data, data fusion, data processing, and correlation analysis. Because performance is inversely correlated with the data feature set, handling this enormous data is a challenge. Because they do not require human intervention, deep learning, and machine learning techniques are suitable, but they are also subject to resource and computation constraints. Although there are various devices for human-computer interaction and they are frequently used for computations, they still have many restrictions, such as range and natural or artificial phenomena. Combining many information search modalities will produce more accurate results than using just one of them alone.

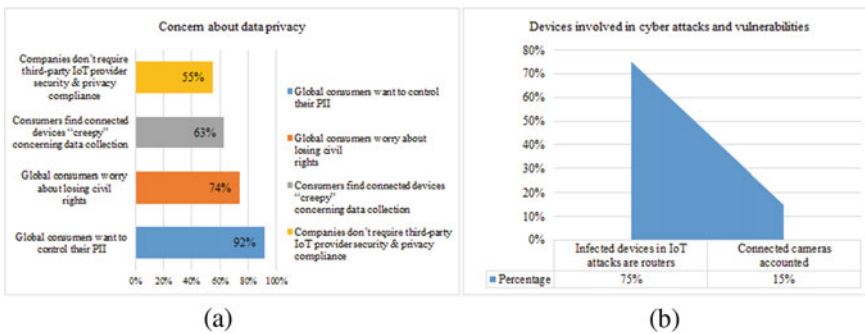


Fig. 8 a Data privacy in the IoT, b Devices involved in cyber-attacks and vulnerabilities

4 Discussion and Conclusion

Analyzing and Monitoring information on networks should be timely. The smart city network architecture is an interconnection of many smart objects to offer innovative services for a smart environment. So, it is enforced to have an effective naming and identification system. By the direct broadcast of their own identity and related quality properties, the extracted data and results will be meaningful for further requirements. Each Smart object involved in the system should be thinking like a human. Wireless Multimedia Sensor Networks (WMSNs) are useful for smart mobility application, but it suffers from energy consumption due to high bandwidth, Quality of Service (QoS), data processing, and data compression at the node level. Hence, in smart mobility, energy-efficient accurate object detection and classification, the fusion of data, scalar data, and advanced sensor nodes is required. New effective models and algorithms are required for data integration and management. There is a need to develop comprehensive modeling, followed by scheming a zero-trust algorithm to moderate either known or unknown categorization of data.

This paper conducted a detailed analysis to compare and discussed various emerging technologies that are involved with smart city applications. There is a ton of research opportunities available in all applications. Every existing technology, including deep learning, machine learning, wireless-sensor-networks, and the Internet of things, should be working to reduce energy usage to provide a better and more efficient solution. All of the currently available technologies that were examined in this study were found to meet some of the essential requirements, but they still have some problems that need to be fixed to offer solutions for specific applications. The author addressed the necessity for data gathering and fusion in this paper's research to improve the outcomes of applications and services for smart cities. The conducted research can help academics and professionals and inspire them to create a new and more effective model, algorithm, and smart devices to close the gaps now present.

Acknowledgements This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors, the research was supported by Technosys Security System Pvt. Ltd. Bhopal, (Madhya Pradesh) India, under research collaboration.

References

1. Impedovo D, Pirlo G (2020) Artificial intelligence applications to smart city and smart enterprise. *Appl Sci* 10:1–5. <https://doi.org/10.3390/APP10082944>
2. Bhattacharya S, Somayaji SRK, Gadekallu TR, Alazab M, Maddikunta PKR (2022) A review on deep learning for future smart cities. *Internet Technol Lett* 5:1–6. <https://doi.org/10.1002/itl2.187>
3. Stübinger J, Schneider L (2020) Understanding smart city—a data-driven literature review. *Sustain* 12:1–23. <https://doi.org/10.3390/su12208460>

4. Tsakanikas V, Dagiuklas T (2018) Video surveillance systems-current status and future trends. *Comput Electr Eng* 70:736–753. <https://doi.org/10.1016/j.compeleceng.2017.11.011>
5. Chaudhary S, Khan MA, Bhatnagar C (2018) Multiple anomalous activity detection in videos. *Proc Comput Sci* 125:336–345. <https://doi.org/10.1016/j.procs.2017.12.045>
6. Ullah Z, Al-Turjman F, Mostarda L, Gagliardi R (2020) Applications of Artificial Intelligence and Machine learning in smart cities. *Comput Commun* 154:313–323. <https://doi.org/10.1016/j.comcom.2020.02.069>
7. Feng C, Arshad S, Zhou S, Cao D, Liu Y (2019) Wi-Multi: a three-phase system for multiple human activity recognition with commercial WiFi devices. *IEEE Internet Things J* 6:7293–7304. <https://doi.org/10.1109/JIOT.2019.2915989>
8. Arshad S, Feng C, Liu Y, Hu Y, Yu R, Zhou S, Li H (2017) Wi-chase: A WiFi based human activity recognition system for sensorless environments. 18th IEEE Int Symp A World Wireless, Mob Multimed Networks, WoWMoM 2017—Conf 2–7. <https://doi.org/10.1109/WoWMoM.2017.7974315>
9. Liu J, Teng G, Hong F (2020) Human activity sensing with wireless signals: a survey. *Sensors (Switzerland)* 20. <https://doi.org/10.3390/s20041210>
10. Zantalis F, Koulouras G, Karabetsos S, Kandris D (2019) A review of machine learning and IoT in smart transportation. *Futur Internet* 11:1–23. <https://doi.org/10.3390/FII11040094>
11. Ha N, Xu K, Ren G, Mitchell A, Ou JZ (2020) Machine learning-enabled smart sensor systems. *Adv Intell Syst* 2:2000063. <https://doi.org/10.1002/aisy.202000063>
12. Javed B, Iqbal MW (2017) Abbas H (2017) Internet of things (IoT) design considerations for developers and manufacturers. *IEEE Int Conf Commun Work ICC Work 2017*:834–839. <https://doi.org/10.1109/ICCW.2017.7962762>
13. Chen Q, Wang W, Wu F, De S, Wang R, Zhang B, Huang X (2019) A survey on an emerging area: deep learning for smart city data. *IEEE Trans Emerg Top Comput Intell* 3:392–410. <https://doi.org/10.1109/TETCI.2019.2907718>
14. Yuan H, Zhu X, Hu Z, Zhang C (2020) Deep multi-view residual attention network for crowd flows prediction. *Neurocomputing* 404:198–212. <https://doi.org/10.1016/j.neucom.2020.04.124>
15. Kang Y, Yang B, Li H, Chen T, Zhang Y (2020) Deep Spatio-temporal modified-inception with dilated convolution networks for citywide crowd flows prediction. *Int J Pattern Recognit Artif Intell* 34. <https://doi.org/10.1142/S0218001420520035>
16. AL Zamil MGH, Samarah S, Rawashdeh M, Karime A, Hossain MS (2019) Multimedia-oriented action recognition in Smart City-based IoT using multilayer perceptron. *Multimed Tools Appl* 78:30315–30329. <https://doi.org/10.1007/s11042-018-6919-z>
17. Chen D, Wawrzynski P, Lv Z (2021) Cyber security in smart cities: a review of deep learning-based applications and case studies. *Sustain Cities Soc* 66. <https://doi.org/10.1016/j.scs.2020.102655>
18. Lee HJ, Ullah I, Wan W, Gao Y, Fang Z (2019) Real-Time vehicle make and model recognition with the residual squeezeNet architecture. *Sensors (Switzerland)* 19. <https://doi.org/10.3390/s19050982>
19. Chang GW, Lu H-J (2020) Integrating gray data preprocessor and deep belief network for say-Ahead PV power output forecast. *IEEE Trans Sustain Energy* 11:185–194. <https://doi.org/10.1109/TSTE.2018.2888548>
20. Mahdavinejad MS, Rezvan M, Barekatin M, Adibi P, Barnaghi P, Sheth AP (2018) Machine learning for internet of things data analysis: a survey. *Digit Commun Networks* 4:161–175. <https://doi.org/10.1016/j.dcan.2017.10.002>
21. Yazici A, Koyuncu M, Sert SA, Yilmaz T (2019) A Fusion-based framework for wireless multimedia sensor networks in surveillance applications. *IEEE Access* 7:88418–88434. <https://doi.org/10.1109/ACCESS.2019.2926206>
22. Gupta S, Kishore Sharma N, Dave M Internet of Thing: a survey on architecture and elements. *Int J Eng Manag Res*
23. Gigli M, Koo S (2011) Internet of Things: services and applications categorization. *Adv Internet Things* 01:27–31. <https://doi.org/10.4236/ait.2011.12004>

24. Bacon L, Ma J, MacKinnon L (2017) IEEE Computer Society, International Association for Computer & Information Science, University of Greenwich, Institute of Electrical and Electronics Engineers. In: Proceedings, 2017 15th IEEE/ACIS International Conference on Software Engineering Research, Management and Applications (SERA) : June 7–9, 2017, the University of Greenwich, London, UK. 395–400
25. Khan R, Khan SU, Zaheer R, Khan S (2012) Future internet: The internet of things architecture, possible applications and key challenges. Proc—10th Int Conf Front Inf Technol FIT 2012 257–260. <https://doi.org/10.1109/FIT.2012.53>
26. Barker P, Hammoudeh M (2017) A survey on low power network protocols for the internet of things and wireless sensor networks. ACM Int Conf Proceeding Ser Part F1305. <https://doi.org/10.1145/3102304.3102348>
27. Dhillon HS, Huang H, Viswanathan H (2017) Wide-area wireless communication challenges for the Internet of Things. IEEE Commun Mag 55:168–174. <https://doi.org/10.1109/MCOM.2017.1500269CM>
28. Gazis V (2017) A survey of standards for machine-to-machine and the Internet of Things. IEEE Commun Surv Tutor 19:482–511. <https://doi.org/10.1109/COMST.2016.2592948>
29. Triantafyllou A, Sarigiannidis P, Lagkas TD (2018) Network protocols, schemes, and mechanisms for internet of things (IoT): Features, open challenges, and trends. Wirel Commun Mob Comput 2018. <https://doi.org/10.1155/2018/5349894>
30. Ngu AH, Gutierrez M, Metsis V, Nepal S, Sheng QZ (2017) IoT middleware: a survey on issues and enabling technologies. IEEE Internet Things J 4:1–20. <https://doi.org/10.1109/IJOT.2016.2615180>
31. Mamdouh M, Elrukhsi MAI (2018) Khattab A (2018) Securing the Internet of Things and Wireless Sensor Networks via Machine Learning: A Survey. Int Conf Comput Appl ICCA 2018:215–218. <https://doi.org/10.1109/COMAPP.2018.8460440>
32. da Costa KAP, Papa JP, Lisboa CO, Munoz R, de Albuquerque VHC (2019) Internet of Things: A survey on machine learning-based intrusion detection approaches. Comput Networks 151:147–157. <https://doi.org/10.1016/j.comnet.2019.01.023>
33. Radoglou Grammatikis PI, Sarigiannidis PG, Moscholios ID (2019) Securing the Internet of Things: Challenges, threats and solutions. Internet of Things (Netherlands) 5:41–70. <https://doi.org/10.1016/j.iot.2018.11.003>
34. Radek Kuchta RN (2014) Smart city concept, applications and services. J Telecommun Syst Manag 03. <https://doi.org/10.4172/2167-0919.1000117>
35. Fan Z, Song X, Xia T, Jiang R, Shibasaki R, Sakuramachi R (2018) Online Deep Ensemble Learning for Predicting Citywide Human Mobility. Proc ACM Interactive, Mobile, Wearable Ubiquitous Technol 2:1–21. <https://doi.org/10.1145/3264915>
36. Crane C No Title. <https://securityboulevard.com/2019/09/20-surprising-iot-statistics-you-dont-already-know/>. Accessed 22 Sep 2022

A Review on Blockchain-Based Electronic Health Record (EHR) System for Secure Data Storage



Vandani Verma and Garima Thakran

1 Introduction

Cryptography is the underlying foundation of the blockchain. The upsurge of blockchain technology as a trustworthy and transparent mechanism for storing and distributing data is opening up new possibilities for addressing serious private information, safety, and data breaches in a variety of fields, including healthcare. Messages were written in codes thousands of years ago to protect them from enemies, and this is where cryptography comes in. Numerous research papers were published in the 1980s and 1990s that theorized the use of cryptography in accordance with secure data chains for the creation of digital currencies. In 1982, David Chaum [1] proposed the digital cash and blind signatures that allow someone to sign a document and prove their ownership while at the same hiding the information in the document. In 1990, David founded DigiCash which created an untraceable digital currency [2] using cryptography, private and public keys, and signatures. Later, DigiCash was declared bankrupt in 1998. In 1997, Adam Back developed hash cash, a proof-of-work algorithm for reducing email spam. Before sending an email, the sender had to prove they had solved a computer puzzle. This consumed computing power and resources, increasing the cost of sending bulk spam emails. He described it more formally in a 2002 paper [3]. In 1998, Nick Szabo [4] proposed a decentralized digital currency called “bit gold.” This included proof-of-work blended with a network of computers that recognized the proof-of-work as legitimate and incorporated it with a time and date into the next puzzle. Bit gold was never a real currency; it existed only in theory. In 1998, Another paper published by Wei Dai [5] illustrated the groundwork for cryptocurrencies, together with Bitcoin, and it is cited in Satoshi Nakamoto’s Bitcoin paper. It was the work during the 1980s to 2000s that laid the groundwork for Bitcoin and the blockchain. In 2008, Satoshi Nakamoto [6] in his paper outlined

V. Verma (✉) · G. Thakran

Department of Mathematics, Amity Institute of Applied Sciences, Amity University, Noida, India
e-mail: vandaniverma@yahoo.com

the creation of Bitcoin and blocks of transactions linked in chains. In 2009, When Satoshi Nakamoto developed the Bitcoin network and the initial blockchain, Bitcoin became a lot more than an idea. This first blockchain was a key aspect of Bitcoin, restricting double spending and serving as a distributed public ledger for all Bitcoin network transactions. Nakamoto also mined the first block on the Bitcoin network known as the “genesis block.”

In a blockchain, all the records (transaction record or a medical record) are stored in blocks. When a block is filled with data, it is added to the chain of previous blocks, and a new block is created for the data entry. Hence, blocks added to the blockchain cannot be altered as they are permanently added to the blockchain and any change in the block is to be notified to each of the previous user. Blockchain technology’s scalability and decentralized implementation make it a valuable tool for enhancing record-keeping processes. Many studies [7–9], proposed changes and developed new approaches to improve and apply a variety of use cases, including smart contracts, supply chain management, and healthcare.

2 Characteristics of Blockchain

- **Immutability:** It’s an essential feature of the blockchain that ensures the integrity of the digital ledger by creating immutable ledgers. Previously, in money transfer, the transaction details can be easily altered, and it also needed a trusted third party to guarantee the integrity of information. In, blockchain, each block relates to the previous block. Thus, minimizing any possibility of block alteration.
- **Decentralization:** The blockchain network is decentralized, i.e., it doesn’t need any trusted third party or governing authority to look after all the transactions. Thus, decentralized distributed ledger solves the issue of single point failure and need of third party to maintain the integrity of transactions.
- **Enhanced Security:** Every information in the blockchain is hash based which is irreversible, so once the can transaction details are hashed and added to the block and published, it becomes impossible to tamper with the transaction details without changing the hash value. So, is someone wants to alter data, he must corrupt every block in the network.
- **Distributed Ledgers:** Every information about a transaction and the participant are shared among all the participants involved in that transaction. Thus, any malicious change in transaction can be easily discovered and makes it transparent/temper proof.

3 Motivations for Blockchain-Based EHR System

EHRs typically include a patient's medical history, personal data like weight and age, results of lab tests, and other things. Verifying the confidentiality and privacy of patient information is crucial. The implementation of healthcare systems in practice is fraught with significant difficulties. The risk of insiders disclosing patient personal information to another organization exists. The key goals for the implementation of secure blockchain-based EHR system are as privacy, security, confidentiality, integrity, availability, auditability, accountability, authenticity, and anonymity. Existing blockchain-based research in the healthcare industry focuses on the primary components listed below to achieve the objectives: Data storage, Data sharing, Data audit, and Identity manager.

3.1 Data Storage

Using blockchain technology is one option to increase security in the EHR system. However, given that blockchain may be public information, there may also be possible privacy concerns for all the encrypted data contained in the public ledger. Personal information is encrypted using public key cryptography within the MediBchain blockchain, which is supported by a healthcare platform. The use of cryptography is not entirely secure. For a select few limited devices, cryptography has a high machine cost. The public ledger's stored encryption text could be cracked by malicious attackers. The loss of a personal key renders data control impossible for the bearer. The EHR systems will use the blockchain to transport medical data. If the data is stored directly in the blockchain, the computational cost and storage load are raised due to the fixed and constrained block size. To solve the storage problem, we use an off-chain store design, where huge volumes of original, encrypted data are stored using reliable third party systems. It can ease the strain on the blockchain's storage system and increase data confidentiality and privacy.

3.2 Data Sharing

The healthcare industry depends on a wide variety of knowledge sources that are documented in various systems including hospitals, clinics, labs, etc. To be used for medical purposes, healthcare information must be stored, retrieved, and altered by various healthcare providers. Interoperability of EHR is the degree to that EHR is known and employed by totally different suppliers as the browse every other's information. It can be classified into three levels: Syntactic interoperability, symmetric interoperability, cross-domain interoperability. One of the main obstacles is the absence of standardized interoperability standards for data sharing between

completely unrelated companies. The risk of a single point attack and data leakage exists with centralized systems. Patients also cannot keep their own personal information in their hands to communicate with a trusted third party. It ought to result in unauthorized use of personal information by some businesses. Additionally, different organizations that lack trust in their collaborations do not appear to be willing to exchange information, which could affect the event of knowledge sharing. Protection of users' data and privacy must be ensured, and ownership of knowledge must be returned to them. Secure access control will promote data sharing. One of the typical strategies to promote data sharing is for secure access control mechanisms that only authorized entities may access. The act of granting authorization to legitimate users so they can access the protected resources is known as share data authorization. This mechanism's access policies are focused on who is doing what action, on what data item, and for what reason.

3.3 Data Audit

Healthcare systems rely on audit log management as a security measure. Since there are some outliers that happened as a result of third parties abusing their power or acting dishonestly. When conflicts emerge, the audit log can be used as evidence to hold users responsible for their transactions with patient records. The blockchain's ledger and smart contracts can offer immutable control for all access requests to establish accountability and traceability. The lack of or manipulation of clinical trials, medical research, and pharmaceutical data severely undermines patient and healthcare provider trust. Blockchain's transparency and accountability can monitor past trial logs and prevent the storage of only the positive results of clinical studies. Audit log offers trustworthy proof of criminal behavior to improve the security of access control models. By obtaining knowledge about interpersonal relationships and hospital work processes, it also helps to improve healthcare services.

3.4 Identity Manager

Membership verification is the initial stage in ensuring the security and privacy of any system before gaining access to any resource. To ensure that certain permissions are granted to data requesters with valid identities, identity authentication is always carried out in the beginning. Users have undergone authentication, biometric authentication, and identity verification before sharing any data. Public key infrastructure, which relies on dependable third parties, is frequently utilized and is based on public key cryptography methods. To manage individual data in the EHR system and to guarantee identification, integrity, and correctly connected individual information, a central master patient index is used. Identity registration is carried out through smart contracts that use public key cryptography to link a legitimate form of

identity information to a specific ETHEREUM address. Member identification with anonymity in a permissioned blockchain was also designed. Finding and relying on a reliable third party that authenticates user identity and completes authentication honestly without running the danger of actual identity leaking is challenging. Most systems use various authentication techniques. Some of them might not be appropriate for an IoT setting. The trend for enhancing the efficiency of blockchain-based EHR systems, particularly in the IoT environment, is the lightweight authentication protocol. Privacy-preserving membership verification through appropriate cryptographic methods and transaction privacy of blockchains without disclosing real identities should be given attention.

4 Blockchain-Based Electronic Health Record (EHR) System

The term “electronic health record” (EHR) refers to the collection of a patient’s health data in the form of digital medical records storing personal health-related information. However, the challenge is to maintain privacy and security of data in such systems. Digitalization is increasing day by day in every field. People are intrigued by digitalizing the healthcare sector furthermore. An electronic health record (EHR) is an electronic version of a patient’s history that is managed over time and will embody all the vital information related to patient’s personal and medical history.

A decentralized management system is required for the health system, which has numerous stakeholders. Blockchain technology has the potential to be that decentralized health management system in which all parties involved would have controlled access to the same health records without any central authority. Since the information cannot be corrupted once it has been saved to the blockchain, the blockchain’s unchangeable nature considerably enhances the security and privacy of the health information contained on it. Every piece of health information on the blockchain is properly ordered and key-encrypted. Additionally, medical records are stored on blockchains utilizing cryptographic keys that help protect patients’ identities. Patients must have access to their information and maintain awareness. Patients would like the assurance that their health information don’t seemed to be misused by other stakeholders and should have a method to detect when such misuse happens. Blockchain helps to fulfill these requirements.

Decentralized blockchain [10] helps in implementing distributed healthcare applications that do not place confidence in a centralized authority. Besides this, the fact that the information in the blockchain is mirrored across all nodes in the network generates an atmosphere of transparency and openness, enabling all stakeholders, and thus patients, to comprehend how their data is used, by whom, when, and how. Furthermore, because information is recorded in the public Ledger and all nodes in the blockchain network have Ledger backups, blockchain-based systems can resolve the restriction of a single point of failure. User can also prevent their real identities

in the sense of pseudo-anonymity. The importance of EHR can also be seen by the latest coronavirus pandemic where distant patient monitoring is increasingly utilized to handle the situation.

There are currently two types of blockchain-based eHealth systems [11]: permissioned blockchain-based eHealth systems and public blockchain-based eHealth systems. To manage EMR storage and sharing, permissioned blockchain-based eHealth systems rely on a modest number of super nodes. Despite its elevated capacity, permissioned blockchain is by no means ideal for secure medical data sharing because it relies on centralized authority (a group of corporations with a common interest that will oversee the entire system). As a result, the data integrity of data in permissioned blockchain is undermined, raising the possibility of a central authority reversing blockchain records. On the other hand, designs based on public blockchain offer greater security and openness but at the cost of scalability. Since the public blockchain is powered by cryptocurrencies, a particular number of coins must be exchanged to include transactions and participate in block mining. Such methods work well for keeping clean organizations like banks, but they offer no incentive for medical facilities. The low efficiency of data retrieval is another barrier to the development of public blockchain-based eHealth solutions. We can perform a direct search on the information kept in the main database. On the blockchain, however, we must first search the block before searching the necessary transactions that are contained in the block. These systems use blockchain technology to trade medical data between medical institutions using smart contracts and scripting language. The medical data is dispersed in blockchain systems because of the data exchange and storage mechanism. It is ineffective in this situation to search through a sea of block data to find patient medical records. Additionally, it takes a long time to access a patient's complete medical records in such systems. Even worse, public blockchain-based eHealth systems struggle to complete transactions quickly. Some Chameleon hash function [10] create new block structures that detail each patient's complete medical history and use the local databases of medical institutions' EMRs to protect them via proxy re-encryption methods. Only accredited healthcare organizations are permitted access to patient EMRs. The proxy re-encryption was presented to ensure the security of data sharing. In these schemes, one party wants a trusted third party to transform the cipher text encrypted with its public key into cipher text encrypted with the other party's public key. Then, other one could decrypt the cipher text with its own private key, i.e., the data is being shared. During the process, the key is not disclosed. Therefore, the data encrypted is private and secure. The specific steps of this process are as follows:

- Party 1 encrypts the text with own public key, i.e., $C_1 = E_1(M)$, where M is the message party 1 wants to share with party 2, and E is an asymmetric encryption algorithm.
- Party 2 sends the request to party 1, and then party 1 generates one conversion key K .
- Party 1 sends C_1 and conversion key K to the agent.

- The agent converts the cipher text C_1 into C_2 using encryption key K . Here, C_2 is the cipher text of M encrypted with party 2's public key.
- The agent sends the cipher text C_2 to party 2.
- Party 2 decrypts C_2 with its own private key to get the plaintext M .

5 Review of Blockchain-Based EHR Systems

In this section, we will discuss the model proposed by Liu et al. [12] that is a medical data sharing and protection scheme based on the private blockchain of the hospital. The two-way proxy re-encryption technology is utilized in the scheme. Also, it has provided a symptoms-matching mechanism for patients with the same disease symptoms. The scheme makes use of two-way proxy re-encryption technology and also established a system for identifying patients with the same disease symptoms. The network's three participants are the system manager (S1), the hospital (H1), and the user (U1). The health management department functions as a trusted third party, creating the master key and system parameters. Hospital (H1) registers with S1 first, and then develops its private and public keys. If a user (U1) sees a doctor in a hospital (H1), he or she must first register with H1 and establish a private key. When the treatment is completed, the doctor will release the outcomes in the blockchain. If they pass the server's verification, the medical results of U1 will be saved in the H1 block chain. If a doctor in any hospital wishes to inquire about the patient U1's medical history, both the doctor and the patient should apply to the S1. SM will compute the conversion key and generate the cipher text of the medical history records, which will have been re-encrypted with the doctor's public key. The encrypted message is then sent to the doctor by S1. Finally, any two patients, US-1 and US-2, could perform mutual authentication and establish a passcode for their subsequent session. This scheme includes six phases: setup, hospital join phase, user join phase, data join blockchain phase, data search and sharing step, and patients' session process (Fig. 1).

This section also explores the studies about the blockchain-based EHR and precisely discusses the contribution of different scientists along with their limitations in the EHR system. We also discuss the various attacks and requirements of blockchain that these schemes can withstand and fulfill in Table 1. We analyze these schemes based on Security (A1), Anonymity (A2), Privacy (A3), Integrity (A4), Authentication (A5), Controllability (A6), Auditability (A7), and Accountability (A8) in Table 2.

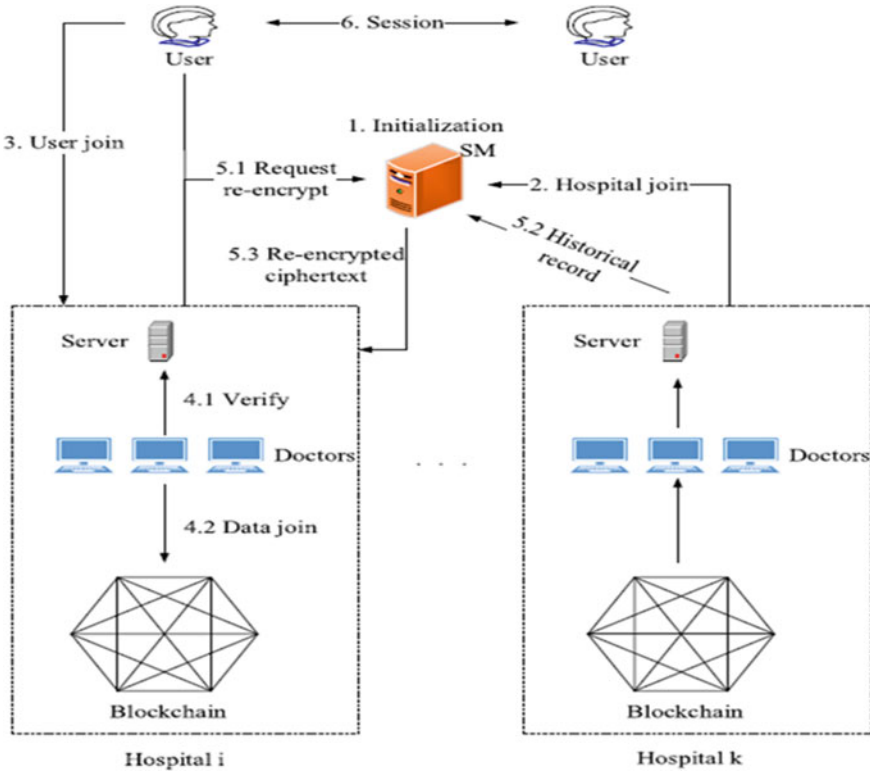


Fig. 1 Medical data sharing and protection scheme architecture Liu et al. [12]

6 Benefits of Blockchain to Healthcare Applications

- **Decentralization:** The stakeholders in the healthcare system are dispersed. They require a decentralized management system. With no centralized authority over the global health information, blockchain will become that redistributed health information management system where all stakeholders will have controlled access to the medical records.
- **Increased data security and privacy:** The blockchain significantly increases the security and privacy of the medical data stored there. Since data saved in a blockchain cannot be edited or corrupted. Every piece of health data on the blockchain is encrypted, time-stamped, and added following a formal account sequence [21–25]. The privacy and identity of patients are protected by utilizing cryptographic keys to store health information on a blockchain.
- **Ownership of health data:** Patients must be required to own their information and must be aware of how it is being used. Patients must be forced to guarantee that other stakeholders are not abusing their health information. Every patient should have a way to know when their information is being used. Through the use of

Table 1 Main contributions and limitations of blockchain-based EHR systems

Ref.#	Contribution	Limitations
[13]	<ul style="list-style-type: none"> - Healthcare Clara is versatile and simple to incorporate utilizing an indicator-centric storage model, and it is protected from confidentiality and integrity attacks by being kept in a private block chain cloud - MPC may be used to compute on encrypted data without leaking any data - It makes it possible for patients to safely control their own data 	<ul style="list-style-type: none"> - High-cost PKE computation - Complexity of key management - Possibility of user's password and data leakage
[14]	<ul style="list-style-type: none"> - Maintain the integrity and accountability of sensitive healthcare data - Data on patients can be protected via cryptographic techniques - Give patients their old control over private data - The patient's true identity can be safeguarded through the use of pseudo-anonymity 	<ul style="list-style-type: none"> - High-cost MPC computation - Data leaking without the owner's consent may result from the replication of data for requestors
[15]	<ul style="list-style-type: none"> - Significantly lessen the amount of encryption keys stored in the blockchain - Use different keys to significantly improve the privacy of the data in the block - Why Without matching symmetric keys, the chances of the attackers successfully decrypting cipher messages are reduced 	<ul style="list-style-type: none"> - Once the corresponding symmetric key is lost all of data will be exposed or corrupted
[16]	<ul style="list-style-type: none"> - The information contained in nail image data can be utilized to identify individuals and aid in future studies of health and disease - For quick and precise biometric authentication, use the SVM and random forest tree method - Use blockchains to safeguard the confidentiality and integrity of sensitive data 	<ul style="list-style-type: none"> - Bottlenecks may appear in the resource-limited IoT devices - The potential for nail image data to be exposed in the public ledger of a blockchain
[17]	<ul style="list-style-type: none"> - Utilizing machine learning approaches, wearable device data quality can be enhanced - Large datasets are stored in an off-chain storage database - Improve data security and privacy - Users have the authority to regulate and share their private health information 	<ul style="list-style-type: none"> - Data leakage caused intentionally or unintentionally by users who decrypted the desired information
[18]	<ul style="list-style-type: none"> - Permit patients to only exchange certain signed medical records - To protect a user's true identity, use multiple public keys for various transactions - Patient transactions that are voluntary and anonymous - Tracking of malicious requestors is possible 	<ul style="list-style-type: none"> - Affect transaction processing directly because it takes time to construct a new block
[12]	<ul style="list-style-type: none"> - Easy to use and low cost - Cloud computing marked the rise of EHR - Attribute-based encryption is used to protect data 	<ul style="list-style-type: none"> - Data leakage risk

(continued)

Table 1 (continued)

Ref.#	Contribution	Limitations
[19]	<ul style="list-style-type: none"> - No express assertion of the patient’s identity for privacy in the signature - Remove the ability to forge the verifier - Block attempts at collusion 	<ul style="list-style-type: none"> - High-cost computation
[20]	<ul style="list-style-type: none"> - Papers with both clinical and technical designs - Medical SCM and drug traceability 	<ul style="list-style-type: none"> - Non-peer-reviewed literature included by the authors as mentioned in [20] - It can affect robustness of the data

Table 2 Security comparison of blockchain-based EHR systems

Ref.#	A1	A2	A3	A4	A5	A6	A7	A8
[13]	Yes	Yes	Yes	Yes	No	Yes	Yes	No
[14]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
[15]	Yes	Yes	Yes	Yes	No	Yes	No	No
[16]	Yes	Yes	Yes	Yes	Yes	No	Yes	No
[17]	Yes	Yes	Yes	Yes	Yes	Yes	No	No
[18]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
[12]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
[19]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
[20]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

smart contracts and robust cryptographic keys, blockchain enables the fulfillment of these criteria [26–30].

- **Availability:** Blockchain records are distributed among several nodes. Being protected from information loss and information corruption, the storage of health data on a blockchain is warranted.
- **Transparency and trust:** The open and transparent nature of the blockchain fosters a culture of trust surrounding distributed healthcare apps. This results in the care stakeholders accepting blockchain apps.
- **Data verifiability:** It is possible to check the accuracy and reliability of these records without having access to the blockchain records themselves. This capability is extremely beneficial in healthcare settings where record verification is required, such as in the management of the pharmaceutical supply chain and the filing of insurance claims.

7 Conclusion

Service providers, patients, and other stakeholders must all have access to unified secure information sharing technologies in order to make informed healthcare decisions. In this study, we share insights on blockchain-based technologies and their potential applications in the healthcare sectors. As previously stated, the digitization of records opens up fresh possibilities for investigating medical trends and assessing quality. There are various benefits of blockchain for support services. The use of blockchain technology improves connection while enhancing security and privacy and lowering costs. Blockchain-based medical records will improve diagnostic accuracy, allow for more informed treatment decisions, and provide a more cost-effective solution. It is critical to bring attention to the obstacles impeding the growth of blockchain implementations in healthcare. Our thorough review of the literature revealed both the benefits and drawbacks of blockchain technology for the healthcare sector.

References

1. Chaum D (1982) Blind Signatures for Untraceable Payments. In: *Advances in Cryptology: Proceedings of CRYPTO '82*, Santa Barbara, California, USA, August 23–25, 1982, pp. 199–203. doi: https://doi.org/10.1007/978-1-4757-0602-4_18
2. Chaum D, Roijakkers S (1990) Unconditionally secure digital signatures. In: *Advances in Cryptology—CRYPTO '90*, 10th Annual International Cryptology Conference, Santa Barbara, California, USA, August 11–15, 1990, Proceedings, vol. 537, pp. 206–214. doi: https://doi.org/10.1007/3-540-38424-3_15
3. Back A (2002) Hashcash-A Denial of Service Counter-Measure. [Online]. Available: <https://www.researchgate.net/publication/2482110>
4. Szabo N (2022) Bit gold: towards trust-independent digital money. Accessed Oct. 28, 2022. [Online]. Available: <https://nakamotoinstitute.org/bit-gold/>
5. Dai W (2022) B-Money, an anonymous, distributed electronic cash system. Accessed: Oct. 28, 2022. [Online]. Available: <http://www.weidai.com/bmoney.txt>
6. Nakamoto S (2008) Bitcoin: A Peer-to-Peer Electronic Cash System. 2008. https://www.usssc.gov/sites/default/files/pdf/training/annual-national-training-seminar/2018/Emerging_Tech_Bitcoin_Crypto.pdf
7. Uddin M, Salah K, Jayaraman R, Pesic S, Ellahham S (2021) Blockchain for drug traceability: Architectures and open challenges. *Health Inform J* 27(2). doi: <https://doi.org/10.1177/14604582211011228>
8. Dwivedi SK, Amin R, Vollala S (2020) Blockchain based secured information sharing protocol in supply chain management system with key distribution mechanism. *J Inform Security Appl* 54. doi: <https://doi.org/10.1016/j.jisa.2020.102554>
9. Khezr S, Moniruzzaman M, Yassine A, Benlamri R (2019) Blockchain technology in healthcare: A comprehensive review and directions for future research. *Appl Sci* 9(9). doi: <https://doi.org/10.3390/app9091736>
10. Zou R, Lv X, Zhao J (2021) SPChain: Blockchain-based medical data sharing and privacy-preserving eHealth system. *Inf Process Manag* 58(4):102604. <https://doi.org/10.1016/j.ipm.2021.102604>
11. Jin H, Luo Y, Li P, Mathew J (2019) A review of secure and privacy-preserving medical data sharing. *IEEE Access* 7:61656–61669. <https://doi.org/10.1109/ACCESS.2019.2916503>

12. Liu X, Wang Z, Jin C, Li F, Li G (2019) A blockchain-based medical data sharing and protection scheme. *IEEE Access* 7:118943–118953. <https://doi.org/10.1109/ACCESS.2019.2937685>
13. Yue XH, Wang D, Jin L, Jiang W (2016) Healthcare data gateways: found healthcare intelligence on Blockchain with novel privacy risk control. *J Med Syst* 40(10). doi: <https://doi.org/10.1007/s10916-016-0574-6>
14. M. S., B. A., K. S., al Omar Abdullah, Rahman (2017) MediBchain: A Blockchain Based Privacy Preserving Platform for Healthcare Data. In: *Security, Privacy, and Anonymity in Computation, Communication, and Storage*, pp. 534–543
15. Liang X, Zhao J, Shetty SS, Liu J, Li D (2017) Integrating blockchain for data sharing and collaboration in mobile healthcare applications. In: *2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, pp. 1–5
16. Lee SH, Yang CS (2018) Fingernail analysis management system using microscopy sensor and blockchain technology. *Spec Collect Artic Int J Distrib Sens Netw* 14(3):2018. <https://doi.org/10.1177/1550147718767044>
17. Zheng X, Mukkamala RR, Vatrappu R, Ordieres-Mere J (2018) Blockchain-based personal health data sharing system using cloud storage. In: *2018 IEEE 20th International Conference on e-Health Networking, Applications and Services (Healthcom)*, pp. 1–6. doi: <https://doi.org/10.1109/HealthCom.2018.8531125>
18. Liu J, Li X, Ye L, Zhang H, Du X, Guizani M (2018) BPDS: a blockchain based privacy-preserving data sharing for electronic medical records. *IEEE Global Communications Conference (GLOBECOM)* 2018:1–6. <https://doi.org/10.1109/GLOCOM.2018.8647713>
19. Shi S, He D, Li L, Kumar N, Khan MK, Choo KKR (2020) Applications of blockchain in ensuring the security and privacy of electronic health record systems: A survey. *Comput Security* 97. Elsevier Ltd, Oct. 01, 2020. doi: <https://doi.org/10.1016/j.cose.2020.101966>.
20. Ng WY et al. (2021) Blockchain applications in health care for COVID-19 and beyond: a systematic review. *The Lancet Digital Health*, vol. 3, no. 12. Elsevier Ltd, pp. e819–e829, Dec. 01, 2021. doi: [https://doi.org/10.1016/S2589-7500\(21\)00210-7](https://doi.org/10.1016/S2589-7500(21)00210-7)
21. Shukla V, Chaturvedi A, Srivastava N (2019) Nanotechnology and cryptographic protocols: issues and possible solutions. *Nanomater Energy* 8(1):1–6. <https://doi.org/10.1680/jnaen.18.00006>
22. Misra MK, Chaturvedi A, Tripathi SP, Shukla V (2019) A unique key sharing protocol among three users using non-commutative group for electronic health record system. *J Discret Math Sci Cryptogr* 22(8):1435–1451. <https://doi.org/10.1080/09720529.2019.1692450>
23. Shukla V, Chaturvedi A, Misra MK (2021) On authentication schemes using polynomials over non commutative rings. *Wireless Pers Commun* 118(1):1–9. <https://doi.org/10.1007/s11277-020-08008-4>
24. Chaturvedi A, Shukla V, Misra MK (2018) Three party key sharing protocol using polynomial rings. *5th IEEE Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON)*, 2018, 1–5. DOI: <https://doi.org/10.1109/UPCON.2018.8596905>
25. Shukla V, Misra MK, Chaturvedi A (2022) Journey of cryptocurrency in India in view of financial budget 2022–23. *Cornell university arxiv*, 2022, 1–6, DOI: <https://doi.org/10.48550/arXiv.2203.12606>
26. Shukla V, Srivastava N, Chaturvedi A (2016) A bit commitment signcryption protocol for wireless transport layer security (wtls). In: *IEEE international conference on electrical, computer and electronics engineering*, pp. 83–86. DOI: <https://doi.org/10.1109/UPCON.2016.7894629>
27. Shukla V, Chaturvedi A, Srivastava N (2019) A secure stop and wait communication protocol for disturbed networks. *Wireless Pers Commun* 110:861–872. <https://doi.org/10.1007/s11277-019-06760-w>
28. Chaturvedi A, Shukla V, Misra MK (2021) A random encoding method for secure data communication: an extension of sequential coding. *J Discret Math Sci Cryptogr* 24(5):1189–1204. <https://doi.org/10.1080/09720529.2021.1932902>

29. Shukla V, Chaturvedi A, Srivastava N (2019) A new one time password mechanism for client-server applications. *J Discret Math Sci Cryptogr* 22:1393–1406. <https://doi.org/10.1080/09720529.2019.1692447>
30. Shukla V, Misra MK, Chaturvedi A (2021) A new authentication procedure for client-server applications using HMAC. *J Discret Math Sci Cryptogr* 24(5):1241–1256. <https://doi.org/10.1080/09720529.2021.1932908>

SVM-RF: A Hybrid Machine Learning Model for Detection of Malicious Network Traffic and Files



Prashant Mathur, Arjun Choudhary, Chetanya Kunndra, Kapil Pareek, and Gaurav Choudhary

1 Introduction

In the past couple of years information technology has seen a massive boom, along with this tremendous growth, cyberspace has also seen its fair share of malware, which are responsible for disrupting regular IT work. The term ‘malware’, originates from the term ‘malicious software’ and can be used to describe any software that is designed to hinder any computer resources in any form. It can be a piece of simplistic software that changes the time of a running computer, without the user’s knowledge, or can be as sophisticated as ‘*Stuxnet*’, which was used to target Iran’s nuclear enrichment program [1].

The number of malware infections has increased from 12.4 million infections in 2009 to a whopping 812.67 million infections in 2018 alone [2]. Even during the COVID-19 pandemic, when the whole world came to a standstill, cybercrime saw exponential growth [3] owing to complete lockdowns, work-from-home mandates, and an exhausted, confused, and untrained population. Cybercriminals took this as an opportunity to further their cause. With The COVID-19 pandemic as their backdrop cybercriminals inculcated the fear of the pandemic into their tactics. Cyberspace saw a surge of COVID-19-themed phishing campaigns, malware being delivered through COVID-19-themed applications [25], and a widespread rampant abuse of the fear instilled within the general public by the pandemic. Organizations that were working on battling the pandemic were also a lucrative target of cyber criminals as seen in

P. Mathur (✉) · A. Choudhary · C. Kunndra · K. Pareek
Sardar Patel University of Police Security and Criminal Justice, Jodhpur, India
e-mail: mtcs20pm@policeuniversity.ac.in

A. Choudhary
e-mail: a.choudhary@policeuniversity.ac.in

G. Choudhary
Technical University of Denmark, Lyngby, Denmark

the case of Dr. Reddy's Laboratories, whose data servers were attacked days after they were approved to conduct trials of Russian made COVID-19 vaccine [26].

As per Kaspersky, a cybersecurity giant, there were a total of 666,809,967 attempts to launch malicious software via online services in 2020 [27] and 687,861,449 attempts in 2021 [28], 2022 is anyone's best guess. Among all the malware classes, the most notorious is the 'Ransomware' category. In this attack vector, the attacker designs the malware in such a way that it encrypts the victim's files and demands a ransom from the victim to decrypt those files, cryptocurrencies such as Bitcoin, Ethereum, etc. are demanded from the victim as the ransom, owing to the use of cryptocurrencies as a ransom it becomes difficult to trace the culprits behind the attack, due to the anonymous natures of all cryptocurrency transactions. As per Cisco, the average ransom paid against a ransomware attack in the year 2020 was nearly \$312,493, with \$10 million being the highest ransom paid, by 2031, ransomware are estimated to cost \$250 billion annually, with the likelihood of a ransomware attack happening every two seconds [4].

Malware not only causes financial loss, but malware attacks can also leave the victim in a state of disarray, corporate victims can face service downtimes, critical data losses, and even loss of reputation, the list of setbacks caused by a malware attack goes on and on. Owing to the advancement in technology, malware authors have become more sophisticated, even with expensive sophisticated defenses in place, they can infect victims with relative ease. In all retrospect, malware have become more of a nuisance than a threat. Thus there arises a dire requirement to create an advanced intelligent system that is capable of identifying and stopping malware attacks before they are executed. This can be done using varying methods, such as scanning network traffic, scanning files, and monitoring user activity on a system. The solution should be capable of such features and should be on active lookout for the same. In this research we have tried to work on these issues and tried to come up with an effective solution that can help with the process of malware identification.

The paper is divided into 7 sections, Sect. 2 discusses the related work carried out in the field of malware detection using machine learning. Section 3 provides a brief overview on malwares, their types and their interaction with the victim. It also provides a brief classification of malwares based on their activities. Section 4 elaborates on our proposed machine learning model that is used to identify malwares. Section 5 explains about the experiments conducted by us, while Sect. 6 discusses the evaluation metrics used to determine the usability of our model, Sect. 7 lays down the results of the experiments performed. Section 7 is followed by the conclusion.

2 Related Work

As technology evolved from its nascent stages, an increased number of issues associated with it also emerged. Malwares over the years has become one such predominant issue. With the increasing sophistication in malwares, traditional detection mechanisms are not able to effectively detect malwares. In order to overcome this hurdle,

machine learning comes into picture. There have been numerous researches in the field using various machine learning and deep learning techniques. This section discusses some of the recent work done in the fields of malware detection using machine learning.

Liu et al. [5] proposes a machine learning model that is composed of three components that perform data processing, making decisions and malware detection. Their first module “data processing” is responsible for extracting features from the inputs. The second layer is used to detect suspicious nature of the malware, finally the third layer uses Shared nearest neighbor (SNN) to categorize input into malware families. Their proposed model gives an accuracy of 86.7% for new malware samples. Their model is trained, validated and tested on their own dataset collected in their home computer lab using Anubis, Kingsoft and ESET NOD32.

Rodrigo et al. [6] proposes BrainSheild, a hybrid machine learning model that employs a three neural network architecture packed with Relu activation function, ADAM optimizer to detect malwares in the android environment. The first neural network is used to for static analysis on the input and has an accuracy of 92.9%, their second neural network is used to carry out dynamic analysis of the input and shows an accuracy of 81.1%, their last neural network running their proposed model gives an accuracy of 91.1%. They use Omnidroid dataset to train, test and validate their proposed model.

Similar to [6], Kim et al. [24] proposes a deep learning based model for detection of malwares in the android environment, they use CNN to extract common features from the API call graph of the application and then use a lightweight classifier, Jaccard similarity algorithm, to classify the application based on similar characteristics. Their proposed model is trained, tested and validated on android applications downloaded from Google Play store and VirusShare and has an accuracy of 91.27%.

Hardy et al. [7] proposes a Stacked AutoEncoder (SAE) based deep learning model. The proposed model has two phases: “unsupervised pre-training” and “supervised backpropagation”. Their model is trained, validated and tested using a dataset obtained from Comodo Cloud Security Center and has an accuracy of 95.64%.

Kan et al. [8] proposes a light-weight deep CNN model that detects malwares based on their grouped instructions. Their model takes raw inputs and groups the input based on instruction sets, the CNN model is used to classify the input as malicious. Their model is trained, validated and tested against a private dataset of 70,000 samples and has an accuracy of 95%.

Table 1 provides a brief overview of the recent research work done on the topic of malware detection using machine learning.

3 Overview

Malware can be considered as any software that is designed to bring harm to the victim by performing malicious actions without the knowledge of the victim. There are various ways to classify malware, here we will classify malware based upon

Table 1 Comparative overview of papers on malware detection using machine learning

Paper	Dataset	Accuracy (%)
Liu et al. [5]	ESET NOD32, VX Heavens	86.7
Hardy et al. [7]	Private dataset	95.64
Kan et al. [8]	Private dataset	95
Rodrigo et al. [6]	Omnidroid dataset	91.1
Kim et al. [24]	Google Play store + VirusShare	91.27
Our model	CTU-13, UNSW-NB15, MMCC	95.92

two most common classification methods, namely, the classification based on the general characteristics and the classification based on the actions a particular malware performs on the victim.

Based on the general characteristics such as propagation type, and general functions malware can be classified into the following categories—

- **Virus**—A virus is a malicious software program that attaches itself to a safe to execute file often by altering the code of the said file when the file is executed the malicious code is also executed, it then replicates to other files often infecting them in the same process [29]. To exist a computer virus must attach itself to a host file.
- **Worm**—A worm is also a replicating malware, but unlike a virus, it doesn't require a host program to propagate itself, it copies throughout the system and some even have the capabilities to propagate themselves over a network [30].
- **Trojan**—A Trojan malware takes its name from the infamous tale of 'The Trojan Horse' that was used by the Greeks during the Trojan War. This form of malware impersonates a legitimate and safe-to-use file tricking victims into executing it [31].
- **Rootkit**—A rootkit is a fairly advanced and stealthy malware, unlike other categories of malware, a rootkit is fairly hard to detect and remove from the system as it is designed to embed itself deep into the operating system, often employing legacy API calls to evade detection from antivirus software [32].
- **Keyloggers**—Keyloggers are predominantly used in Spyware, a keylogger is a piece of software that keeps track and logs all of the victims' keystrokes [33], based upon the intention of the author they can be considered as malicious or benign, a keylogger that is used to collect and exfiltrate all victim's PII is considered a malicious keylogger.
- **Backdoor**—This malware opens up an alternate communication channel between the victim and the attacker in a way that the attacker can bypass all authentication and security mechanisms put into place by the victim [34].
- **Mobile malware**—This umbrella term is used to categorize all malware that is present in the mobile device ecosystem, this category can include mobile ransomware, spyware, backdoors, or Trojans as well. Since the usage of mobile devices has increased exponentially since their conception, they too are a treasure

trove of PII for the attackers and hence a very lucrative target [35]. Owing to such significance, it is also the need of the hour to protect mobile devices from malware and malicious actors. ‘AbstractEmu’ is a fairly recent and extremely dangerous android malware that impersonates 19 types of safe-to-use applications and locks the victim out of their device due to its ability to gain root access to the device [36].

Based on the actions performed by the malware on a victim can be classified into the following categories—

- **Droppers**—This malware is designed to infect the victim with another piece of malware, primarily via covert methods. In other terms, droppers, download another piece of malware and infect the victim using that malware [16].
- **Launchers**—This malware is designed to covertly execute another malware [17].
- **Ransomware**—This malware encrypts user data using strong encryption techniques and then demands a ransom to decrypt, usually the ransom is asked to be paid via cryptocurrency, making the ransom untraceable. One of the prime examples of this category of malware is the infamous WannaCry ransomware [20].
- **Fearware**—This malware is used to instill fear in its victims, they do so by using varying methods, such as displaying threatening messages to victims or damaging their data. Ransomware can also be put in the umbrella category of fearwares. The COVID-19 pandemic saw an unprecedented increase in the use of fearwares to further cybercrimes [21].
- **Bots and Botnets**—This category of malware takes control of a device and executes commands from the attacker, unlike other compromised devices, a bot is part of a large network called the botnet and the said network is under the control of the attacker. An attacker uses a large network of bots to perform nefarious activities such as carrying out Distributed Denial of Service attacks (DDoS) as seen in the case of the infamous Mirai botnet.
- **Spyware**—These malwares are used to exfiltrate victims’ PII(Personal Identifiable Information) and can even use them to steal their identity. Pegasus malware [22] is an iOS spyware that falls under this category.

There can be various other classifications of malware based on various factors such as the API calls a malware makes or whether the malware can mutate itself to avoid detection.

Traditional malware detection techniques rely on signature-based detection methods, such as pattern matching of unique strings specific to malware [18] to identify malware, some may even analyze function calls by performing static analytical operations on a file, but signature-based detection methods are unable to detect new strains of malware [19] and can be easily evaded. Threat actors hide their malicious code or obfuscate it to avoid being discovered by these techniques. The conventional method of detection is hampered by obscurity. Threat actors have come up with sophisticated and unique ways to deliver malware, they most commonly use emails or MS office documents to deliver malware either in the form of malicious links or in

the form of embedded macros. Once installed malware, based on its design can set up a remote connection to the attacker or can establish connections to a command and control server (C2 server), after installation, data can be exfiltrated or the malware can remain dormant until it receives some commands either from the attacker or a C2 server. Malware software updates can also be sent to the victim to install a newer version of the malware. SolarWinds supply chain attack is a prime example of how attackers attack and interact with victims, in the said attack, the supply chain of SolarWinds was compromised and attackers inserted a malicious code with the legitimate software update [23], it remained undetected for quite some time and attackers used the initial compromise to perform nefarious activities. All these interactions can be found in the network traffic. All these entities, namely, files being downloaded, the structure of the file, actions performed by malware, and network traffic among many others are key factors in determining an anomalous entity.

Manually looking out for all these threats is a very tedious and exhilarating task. Artificial intelligence can ease and speed up this process of threat detection, with robust machine learning or deep learning algorithms in place, erroneous detections can be minimized and can protect organizations and individuals to a much greater extent. The **Support Vector Machine** and the **Random Forest Classifier** algorithms are combined in our proposed hybrid model, which is capable of quickly identifying harmful files and malicious network traffic. It is trained, validated, and tested using three datasets, namely: **CTU-13** [9], **UNSW-NB15** [10–14], and **Microsoft Malware Classification Challenge (MMCC)** [15] datasets. The model created using the **MMCC** dataset is used to categorize harmful files, whereas the model created using the **CTU-13** and **UNSW-NB15** datasets are used to detect malicious network traffic (Fig. 1). For ease of understanding, we have classified the problem statements this research intends to solve into the following categories—

- **P1**—Problem of detection of malicious network traffic.
- **P2**—Problem of detection of malicious files.

4 Proposed Model

Support Vector Machine (SVM) is a linear model that can be used to tackle the problems of classification and regression. It can solve both linear and nonlinear problems and due to its adaptability, it can be used to solve a variety of problems. SVM works on a very basic concept: By drawing a line or hyperplane through the dataset, the method separates input into classes. The input data point is then plotted on the hyperplane and in whichever sector the data point lies, it belongs to that class.

Random forest is a supervised learning approach. It is a popular machine-learning algorithm. It can also be used to tackle the problems of regression and classification. It is based on ensemble learning, which is a technique used to solve complex problems by combining several classifiers into a single more refined classifier, this process

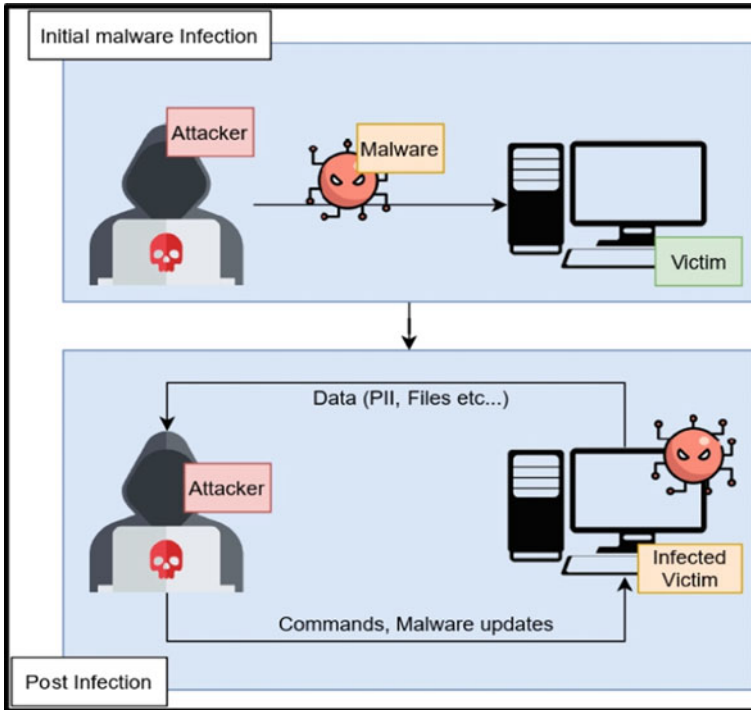


Fig. 1 A typical malware interaction with a victim

increases the overall performance of the combined classifier, making the learning process more efficient.

Our hybrid model uses SVM and RF algorithms and can be used to solve both problems, P1 and P2, SVM takes the raw data as an input and classifies the input based on its features, this step helps us in filtration and pooling our input, the output of the SVM layer is fed into the Random Forest classifier layer. RF enhances the classification done by SVM layers by fine-tuning the output of SVM layers, making the classification more precise and accurate. Figure 2 depicts the logic flow of our proposed model.

5 Experiments

To compute our SVM-RF model's effectiveness we compared its performance to KNN, SVM, RF, CART, CNN, and DF models, using the same datasets and in the same computational environment. All of the tests were performed on a 2.6 GHz Intel i7-8850H CPU with 16 GB RAM, 1 TB of hard disk space, and Ubuntu 20.04 LTS operating system. As a result of our experiments, we found out that our proposed

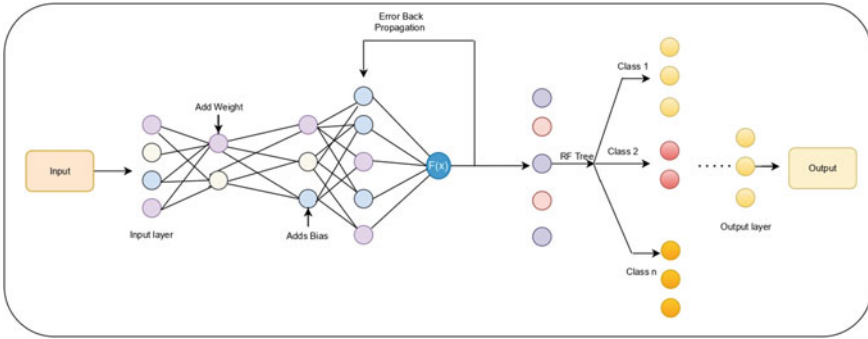


Fig. 2 Proposed SVM + RF model

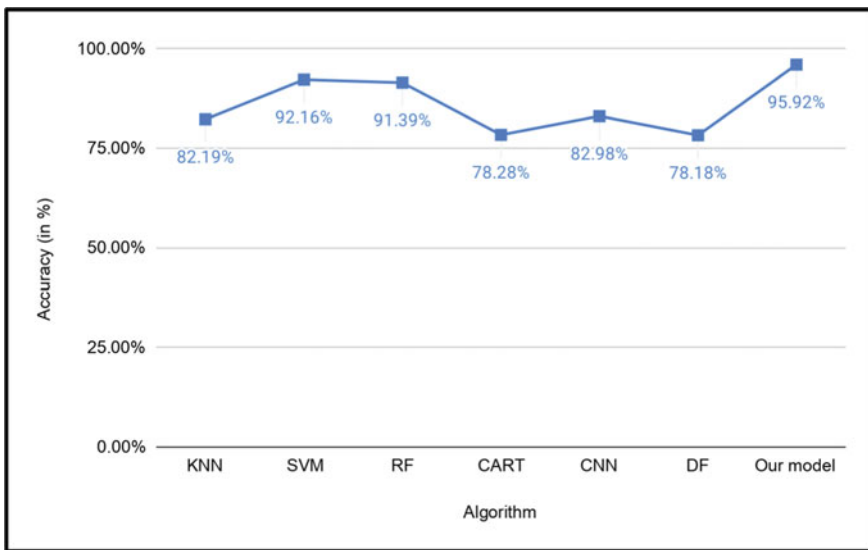


Fig. 3 Accuracy of various algorithms for problem P1

hybrid model showed the best results and accuracy for malicious network traffic and file detection as compared to other models. Figures 3 and 4 show a comparative result of our experimental runs.

6 Evaluation Metrics

Before evaluating the proposed model, we need to understand the following terms—

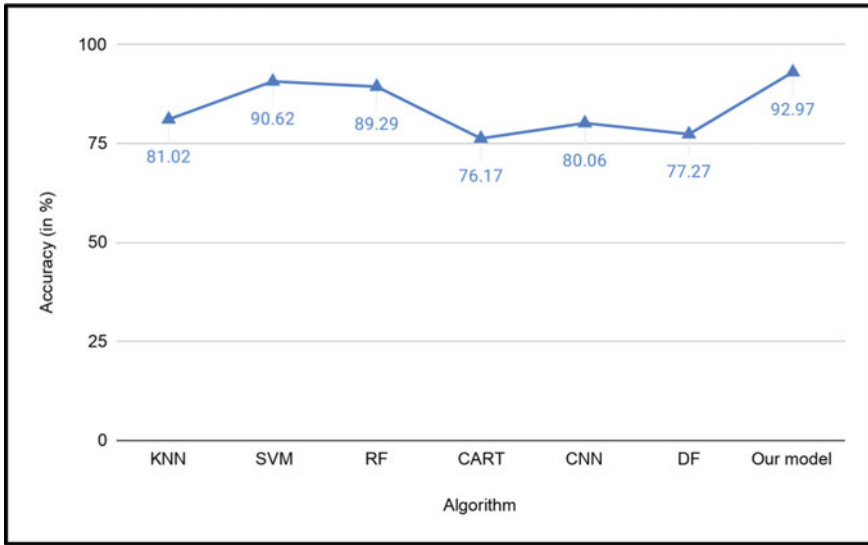


Fig. 4 Accuracy of various algorithms for problem P2

- **True Positives**—True positives (TP) refers to the prediction that is predicted to be true and is also true.
- **True Negatives**—True negatives (TF) refers to the prediction that is predicted to be false and is also false.
- **False Positives**—False positives (FP) refer to the prediction that is predicted to be true but is false.
- **False Negatives**—False negatives (FN) refers to the prediction that is predicted to be false but is true.

Keeping the above terms in mind, we use the following evaluation metrics to test the effectiveness of our proposed algorithm and to compare it with other algorithms—

- **Accuracy**—The ratio of the entire number of accurate forecasts—including both genuine positives and negatives—to the total number of predictions is known as accuracy. The formula provides a definition—

$$Acc = (TP + TN) / (TP + TN + FP + FN) \quad (1)$$

- **Error Rate**—The ratio of all inaccurate predictions—including false positives and negatives—to all forecasts is known as the error rate. The formula provides a definition—

$$Err = (FP + FN) / (TP + TN + FP + FN) \quad (2)$$

7 Results

Tables 2 and 3 depict the accuracy of various models on problems P1 and P2. In both the scenarios our proposed model performs better than other models with an accuracy of 95.92% while detecting malicious network traffic and an accuracy of 92.97% for detecting malicious files, under the current experimental conditions, our model is better suited for detecting malicious network traffic.

Table 2 Experimental results of various algorithms for problem P1

Dataset	Algorithm	Accuracy (%)	Error rate
CTU-13, UNSW-NB15	KNN	82.19	3.97
	SVM	92.16	2.28
	RF	91.39	2.96
	CART	78.28	3.48
	CNN	82.98	3.81
	DF	78.18	4.82
	SVM + RF	95.92	1.62

Table 3 Experimental results of various algorithms for problem P2

Dataset	Algorithm	Accuracy (%)	Error rate
MMCC	KNN	81.02	4.61
	SVM	90.62	2.89
	RF	89.29	3.01
	CART	76.17	4.89
	CNN	80.06	3.92
	DF	77.27	4.88
	SVM + RF	92.97	1.95

8 Conclusion

Detection of malwares is a challenging technological problem. This study presents a hybrid machine learning approach that can be used to detect malicious files and malicious network traffic. This model can help organizations and individuals in making their technology infrastructure more secure and reliable. As per the accuracy of the results we found that model is better suited for detection of malicious network traffic as our experiments show that the model gives an accuracy of 95.92% while detecting malicious network traffic and an accuracy of 92.97% while detecting malicious files. This model can be used to create a hybrid all-in-one security solution that can protect against malicious files and detect malicious network traffic, thus making an organization and even cyberspace a more secure space.

References

1. Michael Holloway (2015) URL: <https://large.stanford.edu/courses/2015/ph241/holloway1/>
2. 2021 Cyber Security Statistics Trends & Data (2021) URL: <https://purplesec.us/resources/cyber-security-statistics/>
3. Harjinder Singh Lallie et al. (2021) "Cyber security in the age of covid-19: A timeline and analysis of cyber-crime and cyber-attacks during the pandemic". *Comput & Secur* 105: 102248
4. Rachel Ackerly (2021) The cost of ransomware attacks: Why and how you should protect your data. URL: <https://umbrella.cisco.com/blog/cost-of-ransomwareattacks>
5. Liu et al. (2017) "Automatic malware classification and new malware detection using machine learning". *Front Inf Technol & Electron Eng* 18(9): 1336–1347
6. Rodrigo, Coarentin et al. (2021) "BrainShield: A hybrid machine learning-based malware detection model for android devices. *Electronics* 10(23): 2948
7. William Hardy et al. (2016) "DL4MD: A deep learning framework for intelligent malware detection". In: *Proceedings of the international conference on data science (ICDATA)*. The Steering Committee of The World Congress in Computer Science, Computer, p 61
8. Kan Z, "Towards light-weight deep learning based malware detection". In, et al (2018) *IEEE 42nd annual computer software and applications conference (COMPSAC)*. vol. 1. *IEEE* 2018:600–609
9. The CTU-13 Dataset. A labeled dataset with botnet, normal and background traffic. URL: <https://www.stratosphereips.org/datasets-ctu13>
10. Nour Moustafa, Gideon Creech, Jill Slay (2017) "Big data analytics for intrusion detection system: Statistical decision-making using finite dirichlet mixture models". *Data analytics and decision support for cybersecurity*. Springer, pp 127–156
11. Nour Moustafa, Jill Slay (2016) "The evaluation of network anomaly detection systems: Statistical analysis of the UNSW-NB15 data set and the comparison with the KDD99 data set". *Inf Secur J: Glob Perspect* 25(1–3): 18–31
12. Nour Moustafa, Jill Slay (2015) "UNSW-NB15: A comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set)". In: *2015 Military communications and information systems conference (MilCIS)*. *IEEE*. pp 1–6
13. Nour Moustafa, Jill Slay, Gideon Creech (2017) "Novel geometric area analysis technique for anomaly detection using trapezoidal area estimation on large-scale networks". *IEEE Trans Big Data* 5(4): 481–494
14. Mohanad Sarhan et al. (2020) "Netflow datasets for machine learning-based network intrusion detection systems". *arXiv preprint arXiv:2011.09144*

15. Microsoft Malware Classification Challenge (2015) URL: <https://www.kaggle.com/c/malware-classification/rules>
16. Kwon, Bum Jun, Jayanta Mondal, Jiyong Jang, Leyla Bilge, Tudor Dumitras (2015) “The dropper effect: Insights into malware distribution with downloader graph analytics.” In: Proceedings of the 22nd ACM SIGSAC conference on computer and communications security, pp 1118–1129
17. Sikorski, Michael, Andrew Honig (2012) Practical malware analysis: The hands-on guide to dissecting malicious software. no starch press
18. Venugopal D, Guoning H (2008) Efficient signature based malware detection on mobile devices. *Mob Inf Syst* 4(1):33–49
19. Bazrafshan, Zahra, Hashem Hashemi, Seyed Mehdi Hazrati Fard, Ali Hamzeh (2013) “A survey on heuristic malware detection techniques.” In: The 5th conference on information and knowledge technology, pp 113–120. IEEE
20. Chen, Qian, Robert A (2017) Bridges. “Automated behavioral analysis of malware: A case study of wannacry ransomware.” In: 2017 16th IEEE international conference on machine learning and applications (ICMLA), pp 454–460. IEEE
21. Tripathi, Rahul (2017) “‘Fearware’ in the times of covid-19 pandemic.” *The Economic Times*. The Economic Times, <https://economictimes.indiatimes.com/tech/internet/fearware-in-the-times-of-covid-19-pandemic/articleshow/75664689.cms?from=mdr>
22. Agrawal, Mayank, Gagan Varshney, Kaushal Pratap Singh Saumya, Manish Verma “Pegasus: Zero-click spyware attack—its countermeasures and challenges”
23. Wolff, Evan D, Growley KM, Gruden MG (2021) “Navigating the solarwinds supply chain attack.” *Procure Lawyer* 56(2)
24. Kim, Jinsung et al. (2022) “MAPAS: a practical deep learning-based android malware detection system.” *Int J Inf Secur*: 1–14
25. Naidoo R (2020) A multi-level influence model of COVID-19 themed cybercrime. *Eur J Inf Syst* 29(3):306–321
26. Bharadwaj, Swati (2020) “Hyderabad: Cyber Hit on Dr Reddy’s labs as covid vaccine work begins: Hyderabad news—Times of India.” *The Times of India*, <https://timesofindia.indiatimes.com/city/hyderabad/cyber-hit-on-dr-reddys-labs-as-covid-vaccine-work-begins/articleshow/78818872.cms>
27. Kaspersky (2020) “Kaspersky security bulletin 2020. Statistics: 26
28. Kaspersky (2021) “Kaspersky security bulletin 2021. Statistics: 26
29. David M Chess, Steve R White (2000) “An undetectable computer virus.” *Proc Virus Bull Conf* 5
30. Kerr PK, Rollins J, Theohary CA (2010) The stuxnet computer worm: Harbinger of an emerging warfare capability. Congressional Research Service, Washington, DC
31. Etaher, Najla, George RS Weir, Mamoun Alazab (2015) “From zeus to zitmo: Trends in banking malware.” 2015 IEEE Trustcom/BigDataSE/ISPA. vol. 1. IEEE
32. Embleton S, Sparks S, Zou CC (2013) SMM rootkit: A new breed of OS independent malware. *Secur Commun Netw* 6(12):1590–1605
33. Ladakis, Evangelos, et al. (2013) “You can type, but you can’t hide: A stealthy GPU-based keylogger.” In: Proceedings of the 6th European workshop on system security (EuroSec)
34. Zhang, Yin, Vern Paxson (2000) “Detecting backdoors.” *USENIX Secur Symp*
35. Felt, Adrienne Porter, et al. (2011) “A survey of mobile malware in the wild.” In: Proceedings of the 1st ACM workshop on security and privacy in smartphones and mobile devices
36. Alerts, Cyware (2021) “ABSTRACTEMU—the rooting malware with a global spread: Cyware Hacker News.” Cyware Labs, Cyware Labs, <https://cyware.com/news/abstractemu-the-rooting-malware-with-a-global-spread-92f9b995>

Key-Insulated Aggregate Proxy Signature



P. V. S. S. N. Gopal, T. Gowri, and P. Vasudeva Reddy

1 Introduction

Ever since, Public Key Cryptosystem (PKC) devised by Diffie and Hellman [1] in 1976, the cryptographic research took a rapid progress. Shamir [2] devised the notion of Identity-based PKC (IPKC), in 1984. In such cryptosystem, signer's public key comprises of binary sequence linked to their identity, like name, mobile number etc. Accordingly, the public key is verified explicitly without accompanying the matching public key certificate. Further, private keys are issued by the trusted party, termed the Key Control Centre (KCC). By the invention of IPKC, many encryption and signature schemes with bilinear pairings of elliptic curves were constructed [3, 4].

Most of the schemes were constructed under opinion that the private keys remain perfectly secure. The whole system's security will no longer be confidential, if suppose the KCC is compromised. To overcome such situation, Dodis et al. [5], devised a cryptosystem via key-insulated mechanism, in 2002.

The basic structure of the system [5] is split life time of master private key as distinct time periods, in which the long term private keys not used for signing directly called helper keys are maintained by a device that is physically-secure, the helper. To perform cryptographic operations, the signers store their interim private keys in a powerful but computationally limited device. Further, this mechanism revives

P. V. S. S. N. Gopal (✉)

Department of BS & H (Mathematics), Kallam Haranadhareddy Institute of Technology (A),
Guntur, Andhra Pradesh 522019, India
e-mail: gopalcrypto786@gmail.com

T. Gowri

Department of EECE, GITAM Deemed to Be University, Visakhapatnam, Andhra
Pradesh 530045, India

P. V. Reddy

Department of Engineering Mathematics, AU College of Engineering, Andhra University,
Visakhapatnam, Andhra Pradesh 530003, India

the momentary private key on distinct time periods through an interaction involving signer and helper; keeping public key unaffected all over. Thus, a compromise of some periods leaves the other unharmed. Hence, this mechanism effectively minimizes the harm caused by revelation of private key until it changes.

Based on scheme [5], the first signature scheme in Identity-based framework using key-insulated mechanism was constructed by Zhou et al. [6], in 2006. Later, many signature schemes and their extensions were constructed [7–9].

Boneh et al. [10], devised an aggregate signature, in 2003, which is single compressed signature attained on combining different n (signatures; signers; messages). Such signature is verified by anyone; convince themselves that the n signer's undeniably signed the n original messages.

Mambo et al. [11] devised a proxy signature in PKI based setting, in 1996. Later, Zhang et al. [12] constructed the first proxy signature scheme in 2003, in ID-based framework. In such a scheme, proxy signer signs on message in support of original signer, attained on receiving a warrant consisting of implicit description of signing rights issued to the former by the latter. Tiwari et al. [13] carried out an analysis on generalization of the proxy signature in 2013.

Wan et al. [14], in 2009, presented a Proxy Signature scheme using Key-insulated mechanism in Identity-based framework (IKPS), that needs 4 pairing computations in proxy signature verification phase proven secure in random oracle paradigm without use of Forking lemma [15].

Lin et al. [16], in 2013, presented an Aggregate Proxy Signature scheme in Identity-based framework (IAPS) on realizing warrant-based delegation. This scheme needs 3 pairing computations in the aggregate signature verification phase and uses Forking lemma [15] in its security reduction.

To handle the issues of key disclosure in proxy signature and maintaining the merits of aggregate signatures, in this article, we construct the first efficient Key-insulated Aggregate Proxy Signature scheme in Identity-based framework (IKAPS) that uses bilinear pairings of elliptic curves. The constructed scheme involves only 3 (constant) pairing calculations in its key-insulated aggregate proxy signature verification phase. Further, we demonstrate that the constructed scheme's security is tightly secure to the hardness of Computational Diffie-Hellman problem [17, 18] in random oracle paradigm without the use of Forking lemma [15].

The rest of paper is categorized as follows: devoted Sect. 2, to some preliminaries including computational hard problems. The constructed IKAPS scheme along with schematic diagram is exhibited under Sect. 3. The constructed scheme's security and its proof of correctness are exhibited in Sect. 4. Efficiency analysis of the constructed scheme is depicted in Sect. 5 and conclusion exhibited finally under Sect. 6.

2 Preliminaries

We summarize the symbolizations and their depiction used in the work; some essential notions; necessary hard problems under this section.

Table 1 Various symbolizations and their depiction used in the constructed scheme

Symbolizations	Depiction
\mathcal{G}_a	Additive cyclic group
\mathcal{G}_m	Multiplicative cyclic group
\in_R	Picked at random from the respective set
$ \mathcal{G} $	Order of group
ID_i	The signer S_i 's identity
d_{ID_i}	The ID_i 's private key
$PSIK_{ID_i,0}$	Proxy signer's initial private key
$HPK_{ID_i,t}$	Proxy helper's private key in t a time period
$PSUK_{ID_i,t}$	Proxy signer's update signing key in time period t
$\{S_i\}_{i=1,2,\dots,n}$	An aggregate collection of proxy signers
$\{M_i\}_{i=1,2,\dots,n}$	An aggregate collection of messages
σ_i	A key-insulated proxy signature on the message M_i by S_i
$\{\sigma_i\}_{i=1,2,\dots,n}$	An aggregate collection of key-insulated proxy signatures
σ	A key-insulated aggregate proxy signature

2.1 Symbolizations and Their Depiction Used in the Constructed Scheme

The symbolizations and their depiction used in the constructed scheme are presented in the following Table 1.

2.2 Bilinear Map

Let $(\mathcal{G}_a, +)$, (\mathcal{G}_m, \cdot) be as mentioned in 2.1, of equal prime order q , and P (say) generates \mathcal{G}_a . A function $e : \mathcal{G}_a \times \mathcal{G}_a \rightarrow \mathcal{G}_m$ is called bilinear map if the below laws are satisfied:

- I. **Bilinear:** $\forall U, V \in \mathcal{G}_a, \forall x, y \in_R \mathbb{Z}_q^*, e(xU, yV) = e(U, V)^{xy}$.
- II. **Non-Degeneracy:** $\exists U \in \mathcal{G}_a, \ni e(U, U) \neq 1$.
- III. **Calculable:** $\forall U, V \in \mathcal{G}_a, e(U, V)$ is calculated by effective algorithm.

On formulating appropriate modifications in Weil/Tate pairing, one works on such using elliptic curves of finite fields.

2.3 Complexity Assumptions

We now exhibit some compulsory hard problems which are used in the constructed scheme's security reduction, in the following.

- **Computational Diffie-Hellman (CDH) Problem:** $\forall c, d \in Z_q^*$, given $P, cP, dP \in \mathcal{G}_a$ evaluate $cdP \in \mathcal{G}_a$. For \mathcal{A} , an adversary in polynomial-time is of advantage (Adv) described as t , the run time in opposition to the CDH problem in \mathcal{G}_a , i.e., $Adv_{CDH}(t) = \Pr[\mathcal{A}(P, cP, dP) = cdP/P, cP, dP \in \mathcal{G}_a]$.
- **Computational Diffie-Hellman (CDH) Assumption:** the (t, ε) -CDH assumption believed to hold in the group \mathcal{G}_a if no \mathcal{A} with Adv at least ε in t -time can break the CDH problem.

3 The Constructed IKAPS Scheme and Its Schematic Diagram

This section refers to the constructed IKAPS scheme, which involves eight algorithms as portrayed below.

1. **Setup:** For $l \in_R Z^+$ security parameter, the KCC run the setup algorithm as portrayed below:
 - Picks two cyclic groups $\mathcal{G}_a, \mathcal{G}_m$ under the binary operations addition, multiplication respectively, of same prime order say $q \geq 2^l$.
 - Picks P a generator of \mathcal{G}_a and $e : \mathcal{G}_a \times \mathcal{G}_a \rightarrow \mathcal{G}_m$ a bilinear map.
 - Picks $s, hpk \in_R Z_q^*$, calculates $P_{pub} = sP, P_{hlp} = hpkP$ as appropriate overall system's, helper's public keys, $g = e(P_{pub}, P)$.
 - Picks the hash functions $\mathcal{H}_1, \mathcal{H}_2 : \{0, 1\}^* \rightarrow \mathcal{G}_a, \mathcal{H}_3 : \{0, 1\}^* \times \mathcal{G}_m \rightarrow Z_q^*, \mathcal{H}_4 : \{0, 1\}^* \times \mathcal{G}_a \times \mathcal{G}_m \rightarrow Z_q^*$.
 - Publishes the system's parameters which are made public as $\mathcal{PP} = \langle l, \mathcal{G}_a, \mathcal{G}_m, q, P, e, P_{pub}, P_{hlp}, g, \mathcal{H}_1, \mathcal{H}_2, \mathcal{H}_3, \mathcal{H}_4 \rangle$, holds $\langle s \rangle, \langle hpk \rangle$ with itself securely.
2. **Key Ext:** The KCC run this algorithm to produce public and private keys of a signer S_i with identity ID_i for $i = 0, 1, 2, \dots, n$. On attaining ID_i of S_i , it calculates $Q_{ID_i} = H_1(ID_i), d_{ID_i} = sQ_{ID_i}$ as appropriate public, private keys of S_i , sends d_{ID_i} to S_i securely.
3. **Initial Proxy Key Gen:** The KCC and the original signer carry out this algorithm. At first, S_0 the original signer prepares a warrant ω with all the necessary information about the allocation rights to the proxy signers $\{S_i\}_{i=1,2,\dots,n}$. The signer S_0 creates a signature $\sigma_0 = (\mathcal{U}_0, \mathcal{V}_0)$ for ω on calculating $\mathcal{U}_0 = g^{r_0}$ where $r_0 \in_R Z_q^*, h_0 = \mathcal{H}_3(ID_0, M, \omega, U_0)$ and $\mathcal{V}_0 = h_0 d_{ID_0} + r_0 P_{pub}$. Finally, S_0 sends $\{ID_0, \omega, \sigma_0\}$ to each proxy signer S_i . Now, S_i can verify the authenticity of σ_0 as below:

$$e(P, \mathcal{V}_0) = e(P, h_0 d_{ID_0} + r_0 P_{pub}) = e(P_{pub}, h_0 \mathcal{H}_1(ID_0)) \mathcal{U}_0.$$

Now, KCC calculates $PSIK_{ID_i,0} = h d_{ID_i} + hpk \mathcal{H}_2(ID_i, 0)$ where $h = \mathcal{H}_4(ID_i, \mathcal{U}_0, \mathcal{V}_0, \omega)$, transmits $PSIK_{ID_i,0}$, $\langle hpk \rangle$ appropriate to proxy signer as their initial proxy signing key and helper as their helper private key securely. Here, '0' of $PSIK_{ID_i,0}$, denote the initial time period.

4. Proxy Key Upd:

- **Helper Key Upd:** At time period t , helper of the proxy signer S_i , calculates a helper key $HPK_{ID_i,t} = hpk[\mathcal{H}_2(ID_i, t) - \mathcal{H}_2(ID_i, t-1)]$, forwards it to S_i .
- **Proxy Signer Key Upd:** Now, S_i updates their private key $PSUK_{ID_i,t} = HPK_{ID_i,t} + PSIK_{ID_i,t-1}$. Finally, the proxy signer wipe away the values $HPK_{ID_i,t}$ and $PSIK_{ID_i,t-1}$.

5. Key-insulated Proxy Sign Gen:

On acquiring message $\mathcal{M} \in \{0, 1\}^*$, in time period t , proxy signer S_i works as below:

- Picks an integer $r_i \in_R Z_q^*$, and calculates

$$\begin{aligned} \mathcal{U}_i &= g^{r_i}, \quad h = \mathcal{H}_4(ID_i, \mathcal{U}_0, \mathcal{V}_0, \omega), \quad h_i = \mathcal{H}_4(ID_i, \mathcal{M}, \omega, \mathcal{U}_0, \mathcal{V}_0, t), \\ \mathcal{V}_i &= h_i PSUK_{ID_i,t} + r_i P_{pub}. \end{aligned}$$

- Outputs $\sigma_i = (\mathcal{U}_i, \mathcal{V}_i)$ the key-insulated proxy signature (IKPS) on \mathcal{M} , signed by S_i in t .

6. Key-insulated Proxy Sign Ver:

Any signer run this algorithm that takes message, identity pairs (\mathcal{M}_i, ID_i) , key-insulated proxy signature (σ_i, t) as input. The verification is done as follows:

- Calculates $h = \mathcal{H}_4(ID_i, \mathcal{U}_0, \mathcal{V}_0, \omega)$, $h_i = \mathcal{H}_4(ID_i, \mathcal{M}, \omega, \mathcal{U}_0, \mathcal{V}_0, t)$.
- Verify $e(P, \mathcal{V}_i) = e(P_{hlp}, h_i \mathcal{H}_2(ID_i, t)) e(P_{pub}, h h_i \mathcal{H}_1(ID_i)) \mathcal{U}_i$ valid or not. It outputs '1', for σ_i valid, else '0'.

7. Key-insulated Agg Proxy Sign Gen:

Each proxy signer $\{S_i\}_{i=1,2,\dots,n}$ presents their key-insulated proxy signature (σ_i, t) in t . Now, any authorized signer calculates $\mathcal{U} = \prod_{i=1}^n \mathcal{U}_i$, $\mathcal{V} = \sum_{i=1}^n \mathcal{V}_i$ and outputs $\sigma = (\mathcal{U}, \mathcal{V})$ as the IKAPS.

8. Key-insulated Agg Proxy Sign Ver:

Any verifier verifies IKAPS (σ, t) for 't' as follows.

- Calculates $h = \mathcal{H}_4(ID_i, \mathcal{U}_0, \mathcal{V}_0, \omega)$, $h_i = \mathcal{H}_4(ID_i, \mathcal{M}, \omega, \mathcal{U}_0, \mathcal{V}_0, t)$
- Verify $e(P, \mathcal{V}) = e(P_{hlp}, h_i \mathcal{H}_2(ID_i, t)) e(P_{pub}, h h_i \mathcal{H}_1(ID_i)) \mathcal{U}$ for validity. It outputs '1', for σ valid, else '0'.

Now, we present the schematic diagram of IKAPS scheme (Fig. 1).

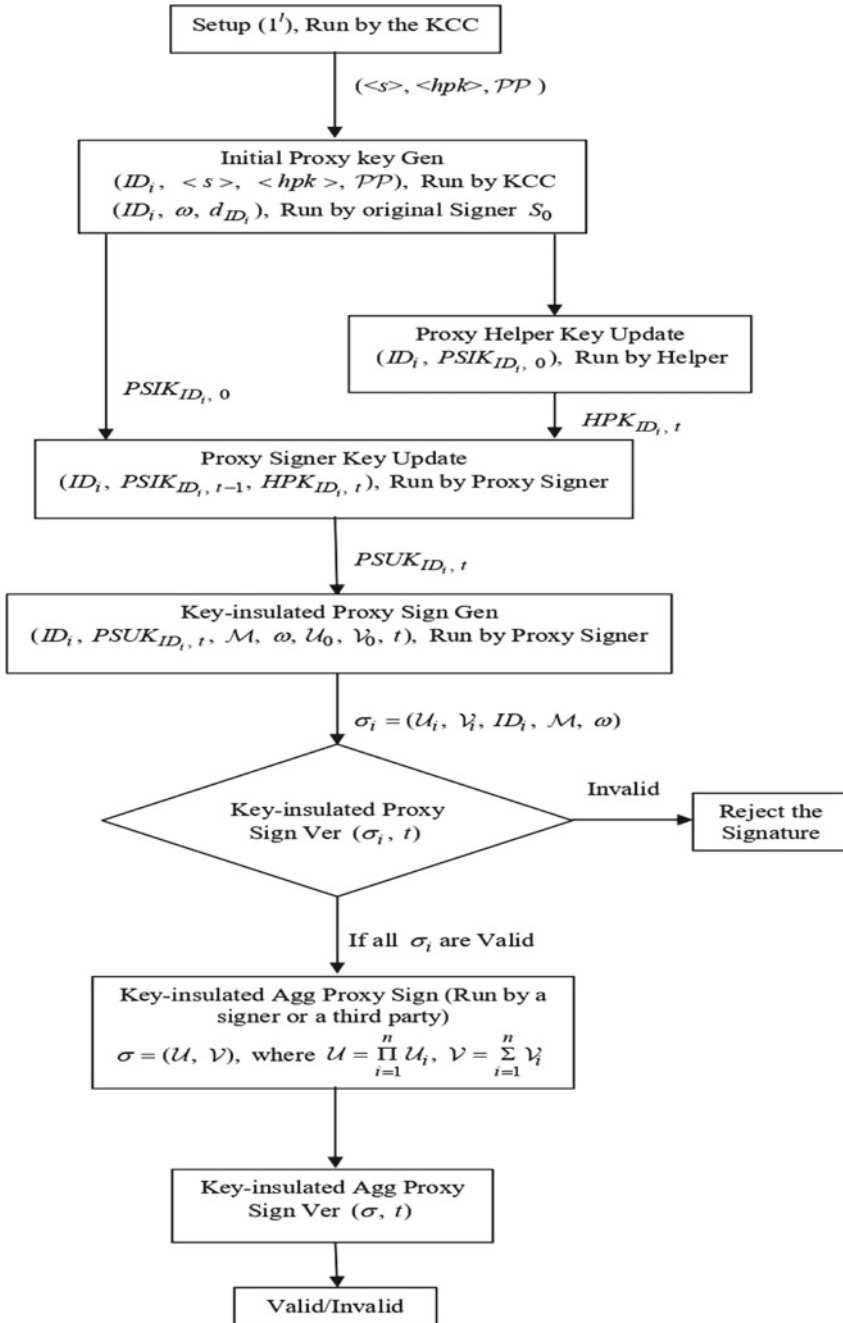


Fig. 1 Schematic diagram of the Constructed IKAPS Scheme

4 Security Analysis

This section briefs proof of correctness as well the security reduction, of constructed IKAPS scheme.

4.1 Proof of Correctness

For IKPS:

$$\begin{aligned} e(P, \mathcal{V}_i) &= e(P, h_i \text{PSUK}_{ID_i, t} + r_i P_{pub}) \\ &= e(P_{hlp}, h_i \mathcal{H}_2(ID_i, t)) e(P_{pub}, h_i \mathcal{H}_1(ID_i)) \mathcal{U}_i. \end{aligned}$$

For IKAPS:

$$\begin{aligned} e(P, \mathcal{V}) &= e(P, \Sigma(h_i \text{PSUK}_{ID_i, t} + r_i P_{pub})) \\ &= e(P_{hlp}, \Sigma h_i \mathcal{H}_2(ID_i, t)) e(P_{pub}, \Sigma h_i \mathcal{H}_1(ID_i)) \mathcal{U}. \end{aligned}$$

4.1.1 Security Reduction

Theorem: Assume \mathcal{A} a forger, in polynomial time can forge the constructed IKAPS scheme with non-insignificant Adv . Next, there is some \mathcal{B} an algorithm, which can output given CDH instance with the same Adv and time.

Proof: Let \mathcal{A} cracks the constructed IKAPS scheme. An algorithm say \mathcal{B} is provided with $xP, yP \in \mathcal{G}_a$ and its objective is to output $xyP \in \mathcal{G}_a$. For this, \mathcal{B} replicates proxy signer to attain valid proxy signature from \mathcal{A} , to solve the CDH problem.

Setup: \mathcal{B} puts $P_{pub} = xP$, forwards the \mathcal{PP} to \mathcal{A} .

Queries: \mathcal{A} queries $\{\mathcal{H}_i\}_{i=1,2,3,4}$ hash functions, proxy key gen and proxy sign ver at their convenience. We presume that before making any initial proxy private key, proxy signing queries on ID ; \mathcal{H}_1 query was made on it earlier. For responding to such, \mathcal{B} evolves as below.

– \mathcal{H}_1 – **Queries:** \mathcal{B} possesses a list \mathcal{L}_1 , empty initially, (ID, c, d, v) of tuples to evolve with the queried hash \mathcal{H}_1 function. On getting a query on \mathcal{H}_1 for $ID \in \{0, 1\}^*$, by \mathcal{A} , \mathcal{B} evolves as below.

1. If \mathcal{L}_1 comprises queried ID , \mathcal{B} evolves with $\mathcal{H}_1(ID) = v$.
2. Else, \mathcal{B} flips arbitrary coin $d \in \{0, 1\}$ with $\Pr[d = 0] = \frac{1}{q_{KE} + q_s + N}$.
3. Now, \mathcal{B} picks $c \in_R \mathbb{Z}_q^*$, for $d = 0$ calculates $v = c(yP)$ and $v = cP$ for $d = 1$.
4. Inserts (ID, c, d, v) to \mathcal{L}_1 , forwards $\mathcal{H}_1(ID) = v$ to \mathcal{A} .

- \mathcal{H}_2 – **Queries:** \mathcal{B} possesses a list \mathcal{L}_2 , empty initially, of tuples (ID_f, t, k, kP) , to evolve with the queried hash \mathcal{H}_2 function by \mathcal{A} . On getting a query on (ID_f, t) by \mathcal{A} , \mathcal{B} evolves as below.
 1. If \mathcal{L}_2 comprises the queried tuple, then \mathcal{B} evolves with $\mathcal{H}_2(ID_f, t)$.
 2. Else, \mathcal{B} picks $k \in_R Z_q^*$, calculates $\mathcal{H}_2(ID_f, t) = kP$, inserts (ID_f, t, k, kP) to \mathcal{L}_2 , forwards kP to \mathcal{A} .
- \mathcal{H}_3 – **Queries:** \mathcal{B} possesses a list \mathcal{L}_3 of tuples (ID_e, ω, U_e, h_3) . On getting a query by \mathcal{A} on H_3 , \mathcal{B} picks $h_3 \in_R Z_q^*$, calculates $\mathcal{H}_3(ID_e, \omega, U_e) = h_3$, inserts to \mathcal{L}_3 , forwards h_3 to \mathcal{A} .
- \mathcal{H}_4 – **Queries:** \mathcal{B} possesses a list \mathcal{L}_4 of tuples $(ID_f, U_e, V_e, \omega, h_4)$, empty initially, to evolve with the queried hash \mathcal{H}_4 function. On getting a query on (ID_f, U_e, V_e, ω) by \mathcal{A} , \mathcal{B} evolves as below.
 1. If \mathcal{L}_4 comprises the queried tuple, then \mathcal{B} evolves with $\mathcal{H}_4(ID_f, U_e, V_e, \omega) = h_4$.
 2. Else, \mathcal{B} picks $h_4 \in_R Z_q^*$, calculates $\mathcal{H}_4(ID_f, U_e, V_e, \omega) = h_4$, inserts to \mathcal{L}_4 forwards to \mathcal{A} .

Also, \mathcal{B} possesses a list \mathcal{L}_5 of tuples $(ID_f, \mathcal{M}, \omega, U_e, V_e, t, v')$ empty initially, to evolve with the queried hash \mathcal{H}_4 function. On getting a query on $(ID_f, \mathcal{M}, \omega, U_e, V_e, t)$ by \mathcal{A} , \mathcal{B} evolves as below.

1. If \mathcal{L}_5 comprises the queried tuple, then \mathcal{B} evolves with $\mathcal{H}_4(ID_f, \mathcal{M}, \omega, U_e, V_e, t) = v'$.
 2. Else, \mathcal{B} picks $v' \in_R Z_q^*$, calculates $\mathcal{H}_4(ID_f, \mathcal{M}, \omega, U_e, V_e, t) = v'$, inserts to \mathcal{L}_5 , forwards v' to \mathcal{A} .
- **Initial Proxy Key Queries:** \mathcal{B} possesses a list \mathcal{L}_6 , empty initially. On getting a query to $\{(ID_e, ID_f), \omega\}$ by \mathcal{A} , \mathcal{B} evolves as below.
 1. \mathcal{B} retrieve the tuples $(ID_e, c_e, d_e, v_e), (ID_f, c_f, d_f, v_f)$ from the list \mathcal{L}_1 . If $d_e = 0$ or $d_f = 0$, \mathcal{B} halts and outputs failure.
 2. Else, it infers that $\mathcal{H}_1(ID_e) = c_eP$ and $\mathcal{H}_1(ID_f) = c_fP$ as determined earlier.
 3. Now, \mathcal{B} retrieve the tuples $(ID_f, c_f, d_f, v_f), (ID_f, t), (ID_f, U_e, V_e, \omega)$ from $\mathcal{L}_1, \mathcal{L}_2, \mathcal{L}_4$ respectively, calculates $d_{ID_f, 0} = c_0P_{pub} + k_0P_{hlp}$, forwards $d_{ID_f, 0}$ to \mathcal{A} .
Now, \mathcal{B} inserts $(ID_f, d_{ID_f, 0})$ to \mathcal{L}_6 .
 - **Helper Key Update Query:** \mathcal{B} possesses a list \mathcal{L}_7 , empty initially. On getting a helper key query on ID_f by \mathcal{A} , in t , \mathcal{B} retrieve (ID_f, t, k, kP) from \mathcal{L}_2 , calculates $HPK_{ID_f, t} = hpk(k_tP - k_{t-1}P)$, forwards $HPK_{ID_f, t}$ to \mathcal{A} .
Now, \mathcal{B} inserts $(ID_f, HPK_{ID_f, t})$ to \mathcal{L}_7 .

– **Proxy Key Update Query:** \mathcal{B} possesses a list \mathcal{L}_8 , empty initially. On getting a update key query of a proxy signer ID_f by \mathcal{A} , in a time period t , \mathcal{B} evolves as below.

1. \mathcal{B} retrieves $(ID_f, d_{ID_f,0}), (ID_f, HPK_{ID_f,t})$ from $\mathcal{L}_6, \mathcal{L}_7$ respectively.
2. Calculates $d_{ID_f,t} = c_f P_{pub} + k_f P_{hlp}$.

– **Proxy Sign Queries:** On getting query to $((ID_e, ID_f), \mathcal{M}, \omega, t)$, i.e., the proxy signature on \mathcal{M} with warrant ω for ID_f by \mathcal{A} , in t , \mathcal{B} evolves as below.

1. Picks $n_e, n_f \in_R Z_q^*$, calculates $\mathcal{U}_e = g^{n_e} \mathcal{U}_f = g^{n_f}$ inserts $(ID_e, \omega, \mathcal{U}_e, h_3), (ID_f, \mathcal{U}_e, \mathcal{V}_e, \omega, h_4)$ to $\mathcal{L}_3, \mathcal{L}_4$ respectively.
2. Examines \mathcal{L}_5 for $(ID_f, \mathcal{M}, \omega, \mathcal{U}_e, \mathcal{V}_e, t, v')$ and retrieve the value determined earlier.
3. Examines \mathcal{L}_8 for $(ID_f, d_{ID_f,t})$ and retrieve the value determined earlier.
4. Fixes $\mathcal{V} = v'(h_4 c_f P_{pub} + k_f P_{hlp}) + n_f P_{pub}$.
5. Forwards to \mathcal{A} , the queried proxy signature $\sigma = (\mathcal{U}, \mathcal{V})$. Answers to the proxy sign queries are all valid and also the output σ as observed below.

$$\begin{aligned} e(P, \mathcal{V}_f) &= e(P, v'(h_4 c_f P_{pub} + k_f P_{hlp}) + n_f P_{pub}) \\ &= e(P_{pub}, v/h_4 \mathcal{H}_1(ID_f)) \hat{e}(P_{hlp}, v/\mathcal{H}_2(ID_f, t)) \mathcal{U}_f. \end{aligned}$$

– **Output:** Ultimately, \mathcal{A} on admitting failure halts, as \mathcal{B} does, or returns a forged aggregate proxy signature σ^* , on \mathcal{M}^* , in t^* . \mathcal{B} retrieves $(ID_e, c_e, d_e, v_e), (ID_f, c_f, d_f, v_f)$ from \mathcal{L}_1 . If $d_e^* = 1$ or then \mathcal{B} output fails. Else, retrieves $(ID_e^*, \omega, \mathcal{U}_e^*, h_3^*), (ID_f^*, \mathcal{U}_e^*, \mathcal{V}_e^*, \omega, h_4^*), (ID_f^*, \mathcal{M}^*, \omega, \mathcal{U}_e^*, \mathcal{V}_e^*, t^*, v')$ from $\mathcal{L}_3, \mathcal{L}_4, \mathcal{L}_5$ respectively.

If $d_e^* = 0$ and $d_f^* = 1$, then $\mathcal{H}_1(ID_e^*) = c_e^* P$ and $\mathcal{H}_1(ID_f^*) = c_f^* (bP)$.

Now, \mathcal{B} calculates and produces the involved:

$$\begin{aligned} e(P, \mathcal{V}_f^*) &= e(P_{pub}, v'^* h_4^* \mathcal{H}_1(ID_f^*)) e(P_{hlp}, v'^* \mathcal{H}_2(ID_f, t^*)) \mathcal{U}_f^* \\ &= e(P, v'^* h_4^* c_f^* (xyP) + v'^* k^* P_{hlp} + n_f^* P_{pub}). \end{aligned}$$

Implies, $\mathcal{V}_f^* = v'^* h_4^* c_f^* (xyP) + v'^* k^* P_{hlp} + n_f^* P_{pub}$ and so

$$xyP = (v'^* h_4^* c_f^*)^{-1} (\mathcal{V}_f^* - v'^* k^* P_{hlp} - n_f^* P_{pub}).$$

This suffices the depiction of Theorem and of \mathcal{B} .

Table 2 Efficiency table

Scheme	Key update phase	Key-insulated Agg Proxy Sign Gen Phase	Key-insulated Agg Proxy Sign Ver Phase
IKAPS Scheme	$1T_m + 2T_a$ = 0.002868ms	$nT_m + (n - 1)T_a$ = (0.00159n - 0.001278)ms	$3T_p + (2n - 2)T_a$ = (35.944833 + 0.002556n)ms

5 Efficiency Analysis

The computational effectiveness of the constructed IKAPS scheme is based on evaluation time of exhausting operations. For this, we take in to account the experimental results carried out by Chen et al. [9], in view of time taken for evaluating different operations as follows: $1T_p \approx 11.982463$ ms (milli seconds), $1T_m \approx 0.000312$ ms, $1T_a \approx 0.001278$ ms. Here a pairing operation symbolized T_p , a scalar multiplication symbolized T_m in \mathcal{G}_a , a point addition symbolized T_a in \mathcal{G}_a . We incorporate these to our constructed scheme as depicted in Table 2.

There are only 3 pairing calculations involved in the key-insulated aggregate proxy signature verification phase of the constructed IKAPS scheme and is a constant irrelevant to the number of proxy signers participate in signing. Also, the communication overhead of the constructed IKAPS scheme is $|\mathcal{G}_a| + |\mathcal{G}_m| = 256$ bytes, i.e., length of the signature is constant.

6 Conclusion

To shield a signature scheme from diverse attacks, one needs to keep securely private keys of the system. To evade harm by key disclosure problem in aggregate proxy signature schemes, we constructed the first efficient IKAPS scheme using pairings in this article. Further, the security of the constructed IKAPS scheme is attained without Forking lemma and so gives tight reductions to the CDH problem.

References

1. Diffie W, Hellman M (1976) New directions in cryptography. IEEE Trans Inf Theory IT 22(6): 644–654
2. Shamir A (1985) Identity-based cryptosystems and signature schemes. In: Blakley GR, Chaum D (eds.) Advances in Cryptology-CRYPTO 1984, LNCS, vol. 196. Springer-Verlag, Berlin Heidelberg, pp 47–53
3. Sahu RA, Padhye S (2011) ID-based signature schemes from bilinear pairing: A Survey. Front Electr Electron Eng China 6(4):487–500
4. Gopal PVSSN, Vasudeva Reddy P, Gowri T (2012) New identity based signature scheme using bilinear pairings over elliptic curves. 3rd IEEE IACC-2013, pp 362–367

5. Dodis Y, Katz J, Xu S, Yung M (2002) Key-insulated public key cryptosystems. In: Knudsen LR (ed.) EUROCRYPT 2002, LNCS, vol. 2332. Springer-Verlag, Berlin Heidelberg, pp 65–82
6. Zhou Y, Cao Z, Chai Z (2006) Identity based key-insulated signature. In: Chen K et al. (eds.) ISPEC 2006, LNCS, vol. 3903. Springer-Verlag, Berlin Heidelberg, pp 226–234
7. Vasudeva Reddy P, Gopal PVSSN (2017) Identity-based key-insulated aggregate signature scheme. *J King Saud Univ Comput Inf Sci* 29: 303–310
8. Gopal PVSSN., Vasudeva Reddy P (2015) Efficient ID-based key-insulated signature scheme with batch verifications using bilinear pairings over elliptic curves. *J Discret Math Sci Cryptogr* 18(4): 385–402
9. Chen Y, Yao T, Ren H, Gan Z (2022) Unidirectional identity-based proxy re-signature with key insulation in EHR sharing system. *Comput Model Eng Sci* 131(3):1497–1513
10. Boneh D, Gentry C, Lynn B, Shacham H (2003) Aggregate and verifiably encrypted signatures from bilinear maps. In: Bihan E (eds.) EUROCRYPT 2003, LNCS, vol. 2656. Springer-Verlag, Berlin Heidelberg, pp 416–432
11. Mambo M, Usuda K, Okamoto E (1996) Proxy signatures: Delegation of the power to sign messages. *IEICE Trans Fundam Electron Commun Comput Sci* E79-A (9): 1338–1354
12. Zhang F, Kim K (2003) Efficient ID-based blind signature and proxy signature from bilinear pairings. In: Proceedings of the 8th Australasia Conference on INFORMATION SECURITY AND PRIVACY (ACISP'03), vol. 2727. pp 312–323
13. Tiwari N, Padhye S (2013) Analysis on the generalization of proxy signature. *Secur Commun Netw* 6(5):549–566
14. Wan Z, Lai X, Weng J, Liu S, Hong X (2009) Identity-based key-insulated proxy signature. *J Electron* 26(6):853–858
15. Pointcheval D, Stern J (2000) Security arguments for digital signatures and blind signatures. *J Cryptol* 13(3):361–396
16. Lin YC, Wu TC, Tsai JL (2013) ID-Based aggregate proxy signature scheme realizing warrant-based delegation. *J Inf Sci Eng* 29:441–457
17. Goh EJ, Jarecki S (2003) A signature scheme as secure as the Diffie-Hellman problem. In: Bihan E (ed) EUROCRYPT 2003, LNCS, vol. 2656. Springer-Verlag, Berlin Heidelberg, pp 401–415
18. Katz J, Wang N (2003) Efficiency improvements for signature schemes with tight security reductions. In: Proceedings of the 10th ACM Conference on COMPUTER AND COMMUNICATIONS SECURITY 2003. ACM Digital Library, pp 155–164

A Comprehensive Study of Cryptocurrency Trend Analysis Based on a Novel Machine Learning Technique



Paras Sharma, Adhiraj Gupta, Rakesh Kumar Bajaj, and Prateek Thakral

1 Introduction

Cryptocurrencies are blockchain-based currencies that have got many applications today. Blockchain [1] is a vital solution to decentralization and cryptocurrencies are decentralized currencies that are not regulated by any authority, rather it is owned and used by people. Cryptocurrencies can also be mined by miners and thus mint new currencies. Cryptocurrencies got into trend less than two decades ago and now some businesses and people also accept cryptocurrencies instead of money. Cryptocurrencies are also considered assets on the basis of their uptrend for so many years, on that account cryptocurrencies are used as a trading entity by many traders worldwide. Due to the high volatility and price fluctuations, cryptocurrencies may be considered a risky entity to trade with. High price fluctuations are common in cryptocurrencies and their volatility makes it difficult to predict the possible direction of the price.

Mainly cryptocurrencies got into light after 2013, though it is totally different from traditional money, many countries have started accepting a few cryptocurrencies in lieu of money. Some major companies have started accepting cryptocurrencies as a method to buy their services and products. Some projects like Cardano are really contributing to society. Cardano's blockchain technology is used by the Ethiopian government to provide education to children by providing them with digital grade verification using Cardano's blockchain technology. Another cryptocurrency, Ethereum, is used by many computer science and blockchain engineers and companies to develop blockchain technologies all over the world. As blockchain technology is fairly transparent and is nearly impossible to tamper with the record of once fed, blockchain technology is the future of the world and it is going to be used widely in many real-world applications starting from educating children to conducting government elections. As of today, there are thousands of cryptocurrencies that are regularly launched under several projects and some of them have the potential to bring a change in the world and serve a purpose. The most famous of all cryptocurrencies and one of the oldest is Bitcoin (BTC), invented in 2009 by an anonymous person named

P. Sharma · A. Gupta · R. Kumar Bajaj (✉) · P. Thakral
Jaypee University of Information Technology, Wakanaghat Solan 173 234, Himachal Pradesh, India
e-mail: rakesh.bajaj@juitsolan.in

Satoshi Nakamoto, bitcoins are the most traded of all the cryptocurrencies. Others include Ethereum (ETH), which is also a decentralized blockchain-based currency, Ethereum's blockchain technology is used widely in projects and real-world applications. Other cryptocurrencies like Cardano (ADA), Litecoin (LTC), and Solana (SOL) are also on the boom and are traded worldwide.

In the early years, due to the unavailability of complete data and fewer visualization tools, it was not feasible to build a model to predict the prices, with the advancements of tools and technologies like machine learning like in [2] and deep learning, it has become a lot easier to predict the values of the assets. Many researchers have worked in this field and made models which can predict the future values of cryptocurrencies using different techniques. A lot of research work has been done in this field; the cryptocurrency industry has been on a boom since 2015; and various new technologies like web3, metaverse, and NFTs are built on the basis of these technologies. Thus cryptocurrency is a matter of interest to many researchers, especially in the field of trading.

The research done in [3] has analyzed the time series of the Bitcoin process by the use of Bayesian Neural Networks (BNNs) which describes volatility in a time series format. To improve the variability, the author has suggested the use of different ML algorithms. The research done in [4] has used the "R2N2" model which includes residuals from vector autoregression in the RNN feature set. They have considered arbitrage opportunities in the market by taking a basket of cryptocurrencies. The work done by researchers in [5] predicts high-frequency exchange rates of cryptocurrencies using a Deep Learning model. They have focused on predicting the one-minute exchange rates of Bitcoin-Ethereum currency. Most of the researchers have used either a single model or a hybrid model to predict the rates of the cryptocurrencies, in the proposed work of this paper, we have applied two machine learning models over a different set of cryptocurrency datasets. The main aim of the proposed model in this paper is to comparatively analyze the working of the models in different datasets of different cryptocurrencies. The researchers in [6] have focused on the emerging phenomenon of cryptocurrencies. The researchers have discussed the trends in academic research related to cryptocurrencies and have tried to highlight the contributed work to the literature. The researchers in [7] have proposed a new clustering-based methodology that provides additional views of the financial behavior of cryptocurrencies. The researchers have applied three different partial clustering algorithms to analyze the trend.

2 Preliminaries

A **blockchain** is a particular sort of Digital Ledger Technology (DLT) that is made up of an expanding collection of data, known as blocks, that are safely connected to one another using encryption. Each block includes transaction information, a timestamp, and a cryptographic hash of the one before it. The timestamp demonstrates that the transaction data was there at the moment the block was produced. Each block links

to the ones before it because each block carries information about the block before it. This causes the blocks to effectively create a chain. Once a transaction is blocked, it cannot be altered without also undoing all subsequent blocks, making blockchain transaction reversal impossible.

“**Machine learning (ML)** [8] is a topic of study focused on comprehending and developing ‘learning’ methods, or methods that use data to enhance performance on a certain set of tasks. It is considered to be a component of artificial intelligence. Without being expressly taught to do so, machine learning algorithms create a model using sample data, also referred to as training data, in order to make predictions or judgments. Machine learning algorithms are utilized in a wide range of applications, including speech recognition, email filtering, computer vision, and medicine, when it is challenging or impractical to create traditional algorithms to carry out the required functions.”

Bitcoin(BTC) [9] is a decentralized digital cryptocurrency which was invented by an anonymous person named Satoshi Nakamoto in 2009. Bitcoin is a blockchain-based cryptocurrency. It is a peer-to-peer network. Bitcoin transactions are verified by network nodes.

Litecoin(LTC) is a cryptocurrency based on a peer-to-peer network. It is a decentralized open-source software project released under the MIT/X11 license. Litecoin was among the earliest alternatives to Bitcoin. Litecoin was initially launched in 2011. Litecoin shares a modified codebase of Bitcoin.

“**Dogecoin(DOGE)** is a cryptocurrency created by software engineers Billy Markus and Jackson Palmer, who decided to create a payment system as a ‘joke’, making fun of the wild speculation in cryptocurrencies at the time. It is considered both the first ‘meme coin’, and, more specifically, the first “dog coin.” Despite its satirical nature, some consider it a legitimate investment prospect. It was introduced on December 6, 2013, and quickly developed its own online community, reaching a market capitalization of over 85 billion on May 5, 2021.”

Software developers Billy Markus and Jackson Palmer came up with the idea for **Dogecoin (DOGE)** as a “joke,” mocking the irrational investment in cryptocurrencies at the time. It is regarded as the first “meme coin” as well as the first “dog coin,” more specifically. It was first released on December 6, 2013, and by May 5, 2021, it had amassed a market valuation of more than 85 billion. The initial intention of making it was not at all to make an asset class, but due to its pump in 2020, it was viewed as a financial asset by many.

Solana was suggested in a white paper written by Anatoly Yakovenko and released in November 2017. Solana’s first block was made on March 16, 2020. A fork in the network brought about by a spike in transactions caused the Solana blockchain to go offline on September 14, 2021, and various validators’ perceptions of the network’s condition varied. The following day, on September 15, 2021, the network was brought back online.

Ethereum is an open-source blockchain that supports smart contracts. Ethereum is a decentralized cryptocurrency. The platform’s native cryptocurrency is Ethereum.

Ether's market valuation is second only to that of bitcoin among cryptocurrencies. 2013 saw the creation of Ethereum by programmer Vitalik Buterin.

“**Linear regression analysis** [10] is used to predict the value of a variable based on the value of another variable. The variable you want to predict is called the dependent variable. The variable you are using to predict the other variable's value is called the independent variable. This form of analysis estimates the coefficients of the linear equation, involving one or more independent variables that best predict the value of the dependent variable. Linear regression fits a straight line or surface that minimizes the discrepancies between predicted and actual output values. There are simple linear regression calculators that use a ‘least squares’ method to discover the best fit line for a set of paired data. You then estimate the value of X (dependent variable) from Y (independent variable).”

“**Bayesian ML** is a paradigm for constructing statistical models based on Bayes' Theorem

$$p(\theta|x) = p(x|\theta)p(\theta)/p(x) \quad (1)$$

Generally speaking, the goal of Bayesian ML is to estimate the posterior distribution $p(\theta|x)$ given the likelihood $p(x|\theta)$ and the prior distribution, $p(\theta)$. The likelihood is something that can be estimated from the training data.”

3 Proposed Methodology

This proposed method as shown in Fig. 1 shows the use of linear regression model and Bayesian model on the same cryptocurrency dataset and then comparatively analyzes the results of both the models and picks the best and optimized model. Different cryptocurrency datasets are used for both of the models for versatility and to avoid overfitting and underfitting. This proposed methodology starts with A. Picking up the dataset followed by B. visualizing the dataset, and printing its original values. Further C. Linear Regression is applied over each dataset followed by D. Bayesian model is also on each dataset of the cryptocurrencies. E. Comparative analysis is performed on the results of both the models and the best and optimal model is chosen and the final results are explored.

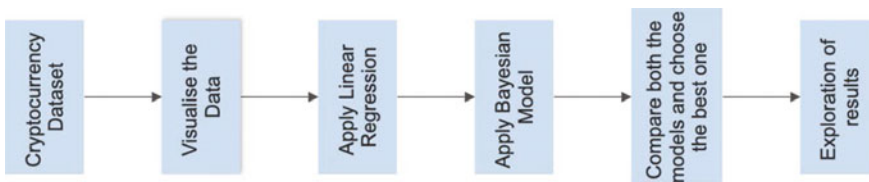


Fig. 1 Proposed methodology

SNo	Name	Symbol	Date	High	Low	Open	Close	Volume	Marketcap	
0	1	Bitcoin	BTC	2013-04-29 23:59:59	147.488007	134.000000	134.444000	144.539993	0.000000e+00	1.603769e+09
1	2	Bitcoin	BTC	2013-04-30 23:59:59	146.929993	134.050003	144.000000	139.000000	0.000000e+00	1.542813e+09
2	3	Bitcoin	BTC	2013-05-01 23:59:59	139.889999	107.720001	139.000000	116.989998	0.000000e+00	1.298955e+09
3	4	Bitcoin	BTC	2013-05-02 23:59:59	125.599998	92.281898	116.379997	105.209999	0.000000e+00	1.168517e+09
4	5	Bitcoin	BTC	2013-05-03 23:59:59	108.127998	79.099998	108.250000	97.750000	0.000000e+00	1.085995e+09
...
2986	2987	Bitcoin	BTC	2021-07-02 23:59:59	33939.588699	32770.680780	33549.600177	33897.048590	3.872897e+10	6.354508e+11
2987	2988	Bitcoin	BTC	2021-07-03 23:59:59	34909.259899	33402.696536	33854.421362	34668.548402	2.438396e+10	6.499397e+11
2988	2989	Bitcoin	BTC	2021-07-04 23:59:59	35937.567147	34396.477458	34665.564866	35287.779786	2.492431e+10	6.615748e+11
2989	2990	Bitcoin	BTC	2021-07-05 23:59:59	35284.344430	33213.661034	35284.344430	33746.002456	2.672155e+10	6.326962e+11
2990	2991	Bitcoin	BTC	2021-07-06 23:59:59	35038.536363	33599.916169	33723.509655	34235.193451	2.650126e+10	6.418992e+11

2991 rows x 10 columns

Fig. 2 BTC data

A. The datasets of five different cryptocurrencies, namely, Bitcoin (BTC), Ethereum (ETH), Litecoin (LTC), Cardano (ADA), Solana (SOL), and Dogecoin (DOGE) are used here which are taken from Kaggle. The time frame for the Dataset is starting from 2013 and ending in 2021. The dataset contains the Opening price, i.e., the initial price of the cryptocurrency on a particular day, Closing Price, i.e., the last price of the cryptocurrency on the particular day at which it traded, High Price, i.e., the highest price of the cryptocurrency at which it traded on a particular day, Low Price, i.e., the lowest price of the cryptocurrency at which it traded on a particular day, for every day for all these years. The dataset also contains the volume, i.e., the number of cryptocurrency coins traded on a particular day for all the days for all the years.

B. First of all the dataset is stored in a data frame using the pandas library in Python as shown in Fig. 2. The figure shows the various features of the dataset as discussed above.

After cleaning the data, the dataset is visualized in a graphical form using the matplotlib library of Python as shown in Fig. 3. The X-axis depicts the date of the recorded data and the Y-axis depicts the Price on the particular date. The blue line in the graph represents the Open price, the red line represents the Closing price on a particular day, the Green line represents the High price on a particular day, the purple line represents the Low price on a particular day, and the orange line represents the projection.

After visualizing the data in graphical form, the dataset is split into training and testing datasets. The training data would be used for training the model and the testing dataset would be used to test the model and check for its accuracy and other results.

C. After visualizing the dataset in graphical form, the Linear Regression model is applied to the dataset. Linear Regression analyzes the data, analyzes all the data points in the dataset, and finds the best fit line for the data points, the mathematical equation for linear regression is

$$y = a_1x_1 + a_2x_2 + \dots + a_nx_n \tag{2}$$

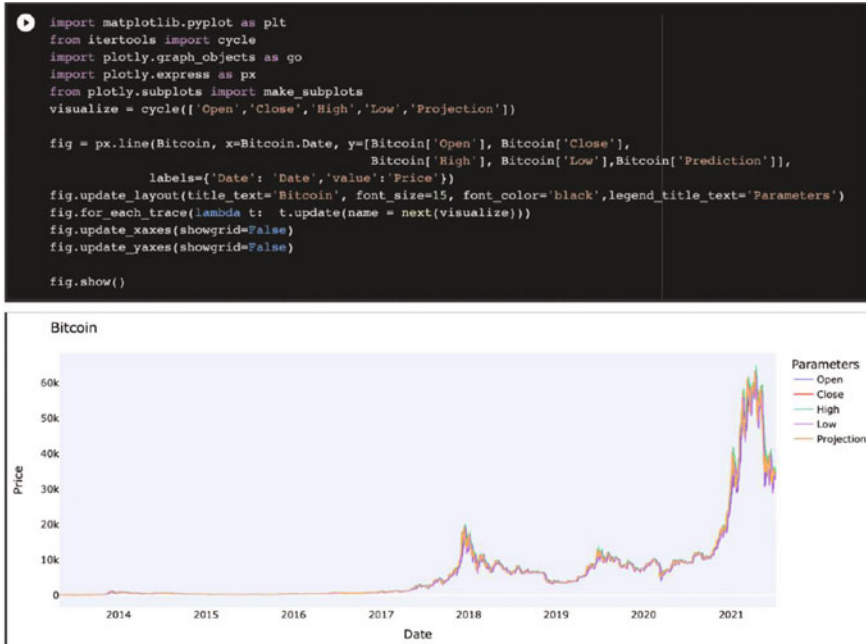


Fig. 3 Visualization of BTC

Here α 's are the linear coefficients, x s are the independent variables, and y is the dependent variable. When plotted graphically, it gives a straight-line fit after considering all the variables. The linear regression model is applied using sklearn library of Python. The results are recorded for the linear regression model as shown in Fig. 4. The predictions based on the linear regression model are duly noted and recorded for comparison in the later stages.

D. After recording the results of the Linear Regression model, we train the Bayesian model using the same training dataset. The Bayesian model is a statistical model where probability is used to represent the uncertainty within the model regarding the parameters of the model. Bayesian model is based on the Bayes theorem, according to which if the probability of an event that has already happened is given and another event is to happen is correlated with the event which has already occurred, then the probability can be calculated according to the mathematical formula:

$$P(A|B) = P(B|A)P(A)/P(B) \quad (3)$$

Here Bayesian model is applied using sklearn.linear model library in Python. After applying the model, its accuracy is calculated and recorded for comparative analysis in a later stage as shown in Fig. 5.

```
[ ] #create & train the model
from sklearn.linear_model import LinearRegression
linReg = LinearRegression()
linReg.fit(x_train_Bitcoin,y_train_Bitcoin)

LinearRegression()

from sklearn.model_selection import cross_val_score
model_lg_acc = cross_val_score(estimator=linReg, X=x_train_Bitcoin, y=y_train_Bitcoin, cv=5, n_jobs=-1)
model_lg_acc

array([0.98802639, 0.98281019, 0.98563612, 0.99013973, 0.98404487])

[ ] print("Accuracy=>",(linReg.score(x_test_Bitcoin, y_test_Bitcoin)))

Accuracy=> 0.986078427671101

R squared coefficient

It is a statistical measure of how well the regression predictions approximate the real data points.

[ ] linReg_confidence = linReg.score(x_test_Bitcoin,y_test_Bitcoin)
print("Linear Regression Confidence: ",linReg_confidence)
print(linReg_confidence*100, "% ")

Linear Regression Confidence: 0.9876677654796077
98.76677654796076 %
```

Fig. 4 Visualization of BTC

```
# Importing modules that are required
from sklearn.datasets import load_boston
from sklearn.model_selection import train_test_split
from sklearn.metrics import r2_score
from sklearn.linear_model import BayesianRidge

model_bayes = BayesianRidge()
model_bayes.fit(x_train_Bitcoin, y_train_Bitcoin)

# Model making a prediction on test data
prediction = model_bayes.predict(x_test_Bitcoin)

# Evaluation of r2 score of the model against the test set
print("r2 Score Of Test Set : (r2_score(y_test_Bitcoin, prediction))")

r2 Score Of Test Set : 0.9876675859388294

[ ] model_bayesian_acc = cross_val_score(estimator=model_bayes, X=x_train_Bitcoin, y=y_train_Bitcoin, cv=5, n_jobs=-1)
model_bayesian_acc

array([0.98405184, 0.98693925, 0.98703882, 0.98576673, 0.986421 ])

[ ] print("Accuracy=>",(model_bayes.score(x_test_Bitcoin, y_test_Bitcoin)))

Accuracy=> 0.9876675859388294
```

Fig. 5 Visualization of BTC

E. After applying both models the accuracy and results are compared of both the models and the best and optimal model is chosen. The results are discussed in Fig. 6 here.

The comparative analysis is given in Table 1 which shows the comparison of the results achieved after applying the different models.

```

=====
                        OLS Regression Results
=====
Dep. Variable:          y          R-squared:                0.986
Model:                  OLS        Adj. R-squared:           0.986
Method:                 Least Squares   F-statistic:             1.811e+05
Date:                   Fri, 02 Sep 2022   Prob (F-statistic):      0.00
Time:                   12:00:23         Log-Likelihood:          -21896.
No. Observations:      2538           AIC:                     4.380e+04
Df Residuals:          2536           BIC:                     4.381e+04
Df Model:               1
Covariance Type:       nonrobust
=====
                    coef    std err          t      P>|t|    [0.025    0.975]
-----
const              80.2252     31.126      2.577    0.010    19.190    141.260
xl                  0.9959      0.002    425.555    0.000     0.991     1.001
=====
Omnibus:              751.214   Durbin-Watson:           2.057
Prob(Omnibus):        0.000   Jarque-Bera (JB):        57958.473
Skew:                 -0.443   Prob(JB):                0.00
Kurtosis:             26.394   Cond. No.                1.54e+04
=====

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[2] The condition number is large, 1.54e+04. This might indicate that there are
strong multicollinearity or other numerical problems.

```

Fig. 6 Visualization of BTC

Table 1 Comparative analysis

Model	Accuracy	R ² Score
Linear regression	0.984051	0.987038
Bayesian	0.986421	0.985766

We achieved an accuracy of 98.64 percent in the Bayesian model and an accuracy of 98.4 percent in the Linear Regression model. The R2 scores of the models were 0.987038 in linear regression and 0.985766 in Bayesian models.

4 Conclusions and Scope for Future Work

The optimized implementation of Linear Regression and the Bayesian model has been performed over five datasets of different cryptocurrencies. A comparative analysis was performed on the results of both models. The execution of the data preprocessing and filtering gave a clear and precise dataset to work on. The usage of linear regression and the Bayesian model was done successfully.

This project can be useful for everyone in the world who is interested in trading and investing in cryptocurrencies. A web application can be made and this project can be listed on different trading research platforms like trading view. Thus the traders can use a future prediction of the cryptocurrencies in which they are interested. Also, more technical indicators can be applied over the prediction part and thus making the predictions more versatile and changing the basis and dimensions of it.

Declarations and Compliance with Ethical Standards

Ethical approval: This article does not contain any studies with human participants or animals performed by any of the authors.

Funding Details: The authors declare that the research carried out in this article has no source of funding.

Conflict of interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Authorship contributions: The authors have equally contributed to the design and implementation of the research, to the analysis of the results, and to the writing of the manuscript.

Acknowledgements We are very much thankful to the anonymous reviewers for suggesting the points/mistakes which have been well implemented/corrected for the necessary improvement of the manuscript. We sincerely acknowledge our deep sense of gratitude to the Editorial office and reviewers for giving their valuable time to the manuscript.

References

1. <https://www.pwc.com/us/en/industries/financial-services/fintech/bitcoin-blockchain-cryptocurrency.html>
2. <https://thecleverprogrammer.com/2021/12/27/cryptocurrency-price-prediction-with-machine-learning>
3. Jang H, Lee J (2018) An empirical study on modelling and prediction of bitcoin prices with bayesian neural networks based on blockchain information. *IEEE Access* 6:5427–5437
4. Samuel P, Andrew S, Ian S (2018) Hybrid autoregressive-recurrent neural networks for algorithmic trading of cryptocurrencies, Stanford, CS230
5. Ibarra IA, Ramos B (2018) High-frequency exchange rate forecasting using deep learning on cryptocurrency markets, Stanford
6. Giudici G, Milne A, Vinogradov D (2020) Cryptocurrencies: market analysis and perspectives. *J Ind Bus Econ* 47:1–18
7. Lorenzo L, Arroyo J (2022) Analysis of the cryptocurrency market using different prototype-based clustering techniques. *Financ Innov* 8–7
8. <https://medium.com/@mukherjeesparsha007/stepping-stones-for-machine-learning-wipython-98f2dcec4cf4>
9. <https://medium.com/analytics-vidhya/predict-bitcoin-price-using-machine-learning-model-288f111eb452>
10. <https://www.ibm.com/in-en/topics/linear-regression>

Flaws of a Password-Based Three-Party Authenticated Key Agreement Protocol in Post-quantum Environment



Sonam Yadav, Vivek Dabra, Pradeep Malik, and Saru Kumari

1 Introduction

Recently, Islam and Basu [3] proposed a password-based three-party authenticated key agreement protocol for mobile devices in a post-quantum environment (PB-3PAKA) protocol. The formal security of the PB-3PAKA protocol is demonstrably secure in Random Oracle Model (ROM). The PB-3PAKA [3] protocol establishes a session key between two mobile users using fresh pair of keys in every session.

Key computation and communication costs are costly, so the key reuse is known to enhance performance during real-world deployments to cut the cost. The resumption mode in TLS v1.2 permits key reuse which drastically decreases online computations. An efficient 0-round-trip time (RTT) resumption mode is suggested in TLS v1.3 draught version 7 [5]. It allows TLS to establish a secure connection without incurring round-trip costs. According to TLS v1.3 version 7, the majority of key exchange computations and communication costs are saved by reusing public and private key pairs. Resumption mode is used to establish the overwhelming majority of TLS connections in the real world. But, this feature-induced security vulnerability in the existing post-quantum key exchange protocol. The key reuse vulnerability has been first identified by Kirkwood et al. [4] in the post-quantum environment. In their work, the reuse of public/private keys is shown to break the security of the protocol. Therefore, authors [4] advise that public-key validation is necessary for the RLWE-based key agreement protocol.

Ding et al. [1] proposed an attack. This attack is known as a signal leakage attack (SLA) and it is against the RLWE-based reconciliation schemes where the public/

S. Yadav (✉) · P. Malik

Department of Mathematics, Faculty of Science, Shree Guru Gobind Singh Tricentenary University, Gurugram 122505, Haryana, India
e-mail: sonamyadav20jan@gmail.com

V. Dabra

Department of Computer Science and Engineering, Panipat Institute of Engineering and Technology, Panipat 132103, Haryana, India

S. Kumari

Department of Mathematics, Chaudhary Charan Singh University, Meerut 250004, Uttar Pradesh, India

private keys are reused. In this approach, the adversary initiates multiple sessions with the honest party to recover the honest party's private key. Using a $2.q$ number of queries with the honest party, the adversary can recover the honest party's secret key.

Continuing the above work, Ding et al. [2] improved signal leakage attack (SLA), and this improved attack is known as i-SLA, in which $2.q$ number of queries is reduced to q number of queries. Now, the secret of the reused public key of the honest party can be recovered with fewer queries.

Influenced by these researchers, we found that Islam and Basu's [3] proposed protocol is vulnerable to dishonest user's attack, signal leakage attack.

2 Review of Islam and Basu's Protocol

In this section, we introduce Islam and Basu's PB-3PAKA protocol [3]. The PB-3PAKA protocol has four phases: initialization phase, user registration phase, authenticated key agreement phase, and password change phase.

Table 1 shows the notations of Islam and Basu's PB-3PAKA protocol. Figures 1 and 2 describe the user registration and authenticated key agreement phase, respectively.

Table 1 Notations of Islam and Basu's [3] protocol

Notation	Meaning
q	Large prime number
a	Random element sampled from R_q
χ_β	Discrete Gaussian Distribution
A/B	Initiator/Responder
S	Server
\mathcal{A}	Adversary
x_i	Public key of i , $i \in \{A, B\}$
r_i	Secret key of i
s	Server's secret key
Cha	Characteristic function
Mod_2	Modular function
ID_i	Identity of U_i/S
PW_i	Password of User i
SK	Session key
\mathbb{Z}	Set of integer numbers
\mathbb{Z}_q	\mathbb{Z} modulo q
\oplus	Bitwise XOR
$H(\cdot)$	Collision resistance function
D	Password dictionary, where $PW_i \in D$

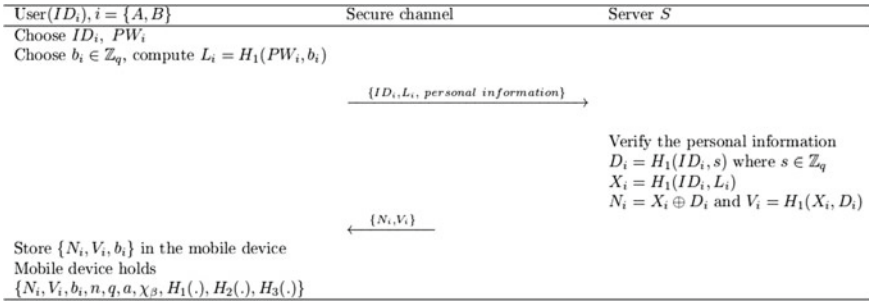


Fig. 1 User registration phase of Islam and Basu’s [3] PB-3PAKA protocol

1. *Initialization phase:*

During initialization phase, the server S selects three one-way hash function, $H_1, H_2, H_3 : \{0, 1\}^* \rightarrow \{0, 1\}^l$ as well as public parameters $\{n, q, a, \chi_\beta\}$. Also, server S selects a secret key s , where $s \in \mathbb{Z}_q$.

2. *User registration phase:*

In user registration phase, the user $\{A, B\}$ chooses a identity ID_i , password PW_i from the dictionary D and $b_i \in \mathbb{Z}_q$ and compute L_i . After that sends $\{ID_i, L_i, \text{personal information}\}$ to server S . On the server’s side, server S verifies the personal information of the user $\{A, B\}$ and computes D_i, X_i, N_i and V_i .

3. *Authenticated key agreement phase:*

In authenticated key agreement phase, the user $\{A, B\}$ computes his public keys x_A and x_B and the parameters Σ_A and Σ_B . $\{ID_A, T_A, x_A, \sigma_A\}$ and $\{ID_B, T_B, x_B, \sigma_B\}$ are sent to the server.

On the server side, the server authenticates the user $\{A, B\}$ and sends its identity and parameters $\Sigma_{S_A}, \Sigma_{S_B}$ to users A and B , respectively.

User $\{A, B\}$ authenticates to the server and computes t_A, t_B as well as signal functions w_A, w_B . Lastly, User A and User B send messages to each other, authenticate each other, and finally generate a session key.

3 Cryptanalysis of Islam and Basu’s Protocol

In this section, we describe the cryptanalysis of Islam and Basu’s [3] protocol based on a password-based three-party authenticated key agreement protocol for mobile devices in post-quantum environment. After examining the protocol, we find that the protocol is vulnerable to dishonest user’s attack and signal leakage attack. These attacks are described below as follows.

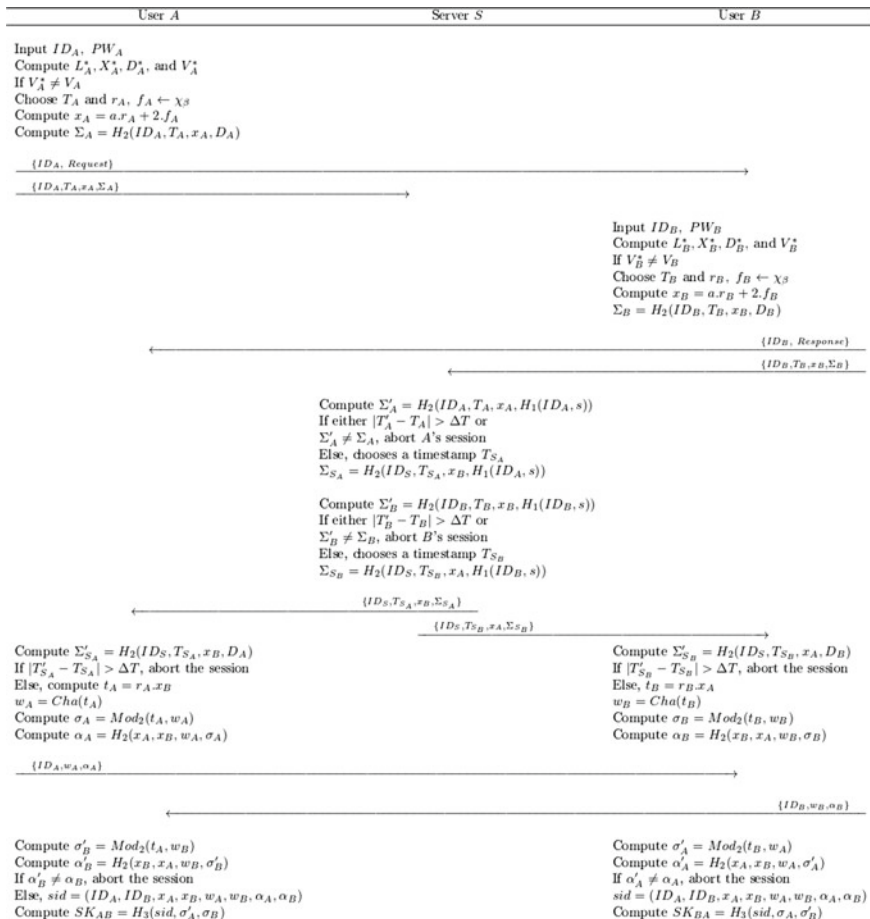


Fig. 2 Authenticated key agreement phase of Islam and Basu's [3] PB-3PAKA protocol

3.1 Dishonest User's Attack

Dishonest user's attack is feasible due to the registration phase of Islam and Basu protocol (see Fig. 1 for a complete description of the registration phase). We show that the adversary correctly recovers the server's secret key s where $s \in \mathbb{Z}_q$. Here, there are two users and one server. The first is user A and the second is user B . We assume that either of these two users is an adversary (Eve). The following steps of dishonest user's attack are as follows:

Step 1: First of all, the adversary \mathcal{A} chooses its ID_A and password PW_A and along with it also chooses a random element $b_A \in \mathbb{Z}_q$. Now, the adversary computes the parameter $L_A = H_1(PW_A, b_A)$ using the hash function on its password

- PW_A and random element b_A (see Fig. 1). After this, \mathcal{A} sends her $\{ID_A, L_A\}$, and personal information to the server through a secure network.
- Step 2: The server receives the $\{ID_A, L_A\}$ and personal information of the adversary and verifies the adversary's personal information. Now, the server computes a parameter D_i using the hash function on his master secret key s where $s \in \mathbb{Z}_q$ and the adversary's ID_A . With this, the server computes a parameter X_i using the hash function on the adversary's ID_A and L_A . Lastly, the server computes the parameters $V_i = H_1(X_i, D_i)$, $N_i = X_i \oplus D_i$ and sends these parameters V_i, N_i to the adversary.
- Step 3: Now, the adversary has the knowledge of V_i, N_i as well as the value of $X_i = H_1(ID_A, L_A)$ because the server has computed the parameter X_i using the hash function on (ID_A, L_A) (see Fig. 1). Further, Adversary \mathcal{A} can easily find the value of D_i by using X_i and N_i parameters.
- In the protocol, the master secret key of server s belongs to \mathbb{Z}_q , and the adversary puts the value of s from 0 to $q - 1$ in $H_1(ID_i, s)$ to match the value of D_i . If adversary guesses the correct value of s , then adversary recovers the server's master secret key s .

3.2 Signal Leakage Attack and Improved Signal Leakage Attack

In TLS v1.3, the key exchange computations and communication costs are saved by reusing public and private key pairs. Resumption mode is used to establish the overwhelming majority of TLS connections in the real world. The security is compromised when keys are reused in TLS, due to this the PB-3PAKA [3] protocol is vulnerable to a signal leakage attack and improved signal leakage attack. Therefore, the adversary can retrieve the user's secret key (see Fig. 2 for a complete description of the signal leakage attack).

Attack overview:

Islam and Basu's PB-3PAKA protocol has two parties, A and B and one server S . As of TLS v1.3, we reuse the secret keys r_A and r_B , respectively, of both parties, A and B in the Islam and Basu's PB-3PAKA protocol. Suppose party A plays the role of an adversary (Eve) and party B is an honest party. Adversary wants to recover the secret key r_B of the honest party B . She generates her malicious public key x_A and sends it to party B . Party B computes $t_B = r_B \cdot x_A$ and signal function w_B using the adversary's malicious public key and sends the signal function w_B to the adversary. So, the adversary retrieves the secret key r_B by observing the signal function w_B sent by the party B . For detailed description, see below attack.

Signal Leakage Attack:

Let the value of the adversary's secret key r_A is 0 and the value of adversary's error term f_A to be 1. By which the public key of the adversary will be $x_A = k$, where the value of $k \in \mathbb{Z}_q$. Now, the adversary sends its ID_A , public key x_A , and other parameters T_A (Timestamp), $\Sigma_A = H_2(ID_A, T_A, x_A, D_A)$ to the server.

Similarly, party B also derives its public key and other parameters and sends them to the server. Also, sends its $\langle ID_B \text{ and response} \rangle$ to the adversary.

Now, the server sends the public key of the adversary to party B and the public key of party B to the adversary.

After receiving the public key of the adversary sent by the server, party B computes t_B where $t_B = r_B \cdot x_A$ (here, r_B is the secret key of party B). In addition, it also computes the signal function

$$w_B = Cha(t_B)$$

modular function

$$\sigma_B = Mod_2(t_B, w_B)$$

and the parameter $\alpha_B = H_2(x_B, x_A, w_B, \sigma_B)$ (see Fig. 2).

Now,

$$\begin{aligned} t_B[i] &= r_B \cdot x_A[i] \\ &= r_B(a \cdot r_A + k \cdot f_A)[i] \\ &= k \cdot r_B[i] \end{aligned}$$

where $k \in \{0, \dots, q-1\}$.

As soon as the adversary will vary the value of k , likewise she will guess the value of $k \cdot r_B[i]$ correctly, because the number of the signal w_B changes for every coefficient of $r_B[i]$. When there is a change in the signal for any i th coefficient of $r_B[i]$, then the number of that change is exactly $2 \cdot r_B[i]$. But the value of $+1$ and -1 only gives signal change of the same number, due to which the adversary can only guess the value up to \pm sign. For the value of $-r_B$, the value of k changes in the reverse direction which is a multiple of r_B .

Therefore, to find out the exact value of the r_B coefficient, the adversary initiates the q number of sessions with party B with its public key (for more details, see [1]).

Improved Signal Leakage Attack:

Attack details:

In the beginning, adversary \mathcal{A} sends its $\langle ID_A \text{ and request} \rangle$ to party B . Moreover, she derives her public key $x_A = a \cdot r_A + k \cdot f_A$, here r_A and f_A are the adversary's secret key and the error term, respectively (see Fig. 2). Now, two cases arise here. In the first case, the adversary takes the value of r_A as 0, and in the second case, the value of r_A is taken as very small depending on error distribution.

In the improved signal leakage attack case, the adversary chooses the value of error term f_A as 1 and selects the value of secret key r_A according to the error distribution. Therefore, the public key of the adversary is $x_A = a.r_A + k$ so that the public key of the adversary cannot be distinguished. Here,

$$\begin{aligned} t_B &= r_B \cdot x_A \\ &= r_B(a.r_A + k.f_A) \\ &= a.r_A.r_B + k.r_B \end{aligned}$$

and signal function

$$\begin{aligned} w_B &= Cha(t_B) \\ &= Cha(a.r_A.r_B + k.r_B) \end{aligned}$$

As adversary \mathcal{A} iterates over k values, $a.r_A.r_B$ remains constant.

Consequently, \mathcal{A} continues to observe the signal changes of $r_B[i]$ while she varies the values of k toward the positive values and starts from $k = 0$. After this, the adversary records the first signal change in w_B at $k = k_1$.

The adversary then varies k toward the negative values and observes the first signal change in w_B and in this direction it records the first signal change at $k = k_2$.

Now, the period of region T or T^c in multiples of $r_B[i]$ is $k_1 - k_2$. The period of the signal change is $k_1 - k_2$, due to which the value of $r_B[i]$ up to the \pm sign is revealed by $\frac{q}{2.(k_1 - k_2)}$. The process of changing the signal continues till the signal becomes stationary after the change, and then adversary can query a small constant number here more than $\frac{q}{2}$ times.

In this way, the adversary can recover $r_B[i]$ up to the sign by doing $\frac{q}{2} + c$ queries. Here c is a small value because as the value of k increases, the value stabilizes and $k.r_B[i]$ moves away from the boundary point. Now, adversary performs $q + c$ queries so that \mathcal{A} can recover the exact value of the secret (for more details, see [2]).

4 Conclusion

We have studied Islam and Basu's [3] proposed protocol based on a password-based three-party authenticated key agreement protocol for mobile devices in post-quantum environment (PB-3PAKA). It has been found that their protocol is vulnerable to dishonest user's attack. The security is compromised when keys are reused in TLS v1.3, and due to this the PB-3PAKA protocol is vulnerable to signal leakage attack. In future, we will propose an improved protocol to overcome the above-identified attacks on Islam and Basu's proposed protocol.

References

1. Ding J, Alsayigh S, Saraswathy R, Fluhrer S, Lin X (2017) Leakage of signal function with reused keys in RLWE key exchange. In: 2017 IEEE international conference on communications (ICC). IEEE, pp 1–6
2. Ding J, Fluhrer S, Rv S (2018) Complete attack on RLWE key exchange with reused keys, without signal leakage. In: Australasian conference on information security and privacy. Springer, pp 467–486
3. Islam SH, Basu S (2021) PB-3PAKA: password-based three-party authenticated key agreement protocol for mobile devices in post-quantum environments. *J Inf Secur Appl* 63:103026
4. Kirkwood D, Lackey BC, McVey J, Motley M, Solinas JA, Tuller D (2015) Failure is not an option: standardization issues for post-quantum key agreement. In: Workshop on cybersecurity in a post-quantum world, p 21
5. Rescorla E (2018) The transport layer security (TLS) protocol version 1.3. Technical report

Multivariate Aggregate and Multi-signature Scheme



Satyam Omar, Sahadeo Padhye, and Dhananjay Dey

1 Introduction

Digital signature is an important cryptographic primitive that is used for authentication, integrity, and non-repudiation, in which one signer signs the message using his private key, and the authenticity of the message can be checked by applying the public key of the signer on the signature. But there are the situations where a group of some persons needs to sign on a message collectively, these type of signatures are called multi-signatures. If different messages are assigned to different group members to sign, then the collective signature of the group is called aggregate signature. The advantage of these aggregate/multi-signature over the collection of all the individual signatures is the constant collective signature size that is not increased by increasing group size.

The concept of multi-signature was first given by Itakura and Nakamura in 1983. Since then, various multi-signature schemes [1, 8, 11] have been proposed over years, but all based on number-theoretic problems which will not be safe against quantum computers [10]. Multivariate public key cryptography (MPKC) is an efficient alternative to produce quantum-resistant digital signatures. There are various efficient and secure generic signature schemes like UOV [7], Rainbow[4], GUI [9], and many more, which can be used to design society-based signature schemes like group signature, ring signature, multi-signature, etc. MPKC produces smallest size signatures among all other post-quantum techniques. Although there exist two multivariate-based aggregate signature schemes [2, 3] already, but their signature sizes increase linearly with group size. To the best of our knowledge, there is no constant size

These authors contributed equally to this work.

S. Omar (✉) · S. Padhye

Department of Mathematics, Motilal Nehru National Institute of Technology, Prayagraj 211004, Uttar Pradesh, India

e-mail: satyamomar840@gmail.com

D. Dey

Department of Mathematics, Indian Institute of Information Technology, Lucknow 226002, Uttar Pradesh, India

multi-signature or aggregate signature in MPKC. We compare our scheme with these schemes in Sect. 5.

Outline: In Sect. 2, preliminary concepts of multi-signature and MPKC have been discussed. In Sect. 3, the specific construction of the proposed scheme is given. Thereafter, the security of the scheme has been discussed in Sect. 4. In Sect. 5, the scheme is comparatively analyzed. Finally, the paper is concluded in Sect. 6.

2 Preliminary

Definition 1 (*MQ Problem*) Given a quadratic system of equations having l variables and v equations over a finite field \mathbb{E} such as

$$D_i(x_1, x_2, \dots, x_l) = y_i, \quad i = 1, 2, \dots, v.$$

Then, to find the solution $(x_1, x_2, \dots, x_l) \in \mathbb{E}^l$ of the above system of equations is called multivariate quadratic problem (MQ problem). This problem is proven to be np-complete [5].

Definition 2 (*MPKC Digital Signature Construction*) It has three algorithms as follows:

Key Generation: The signer chooses two invertible affine transformations μ and ν over \mathbb{E}^v and \mathbb{E}^l , respectively, and an easily invertible quadratic system of equations $D : \mathbb{E}^l \rightarrow \mathbb{E}^v$. Then he computes $\bar{D} = \mu \circ D \circ \nu$. The private key is (μ, D, ν) and the corresponding public key is \bar{D} . This \bar{D} is computationally hard to invert due to inclusion of μ and ν .

Signing: To sign on the message m , the signer computes $\sigma = \bar{D}^{-1}(M) = (\mu \circ D \circ \nu)^{-1}(M) = \nu^{-1} \circ D^{-1} \circ \mu^{-1}(M)$, where $M = H(m)$, and $H : \{0, 1\} \rightarrow \mathbb{E}^v$ be a cryptographically secure hash function.

Verification: The verifier accepts the signature σ , if the equality $\bar{D}(\sigma) = M$ holds; otherwise rejects.

Definition 3 (*Aggregate Signature*) It consists of three algorithms, namely, **Agg KeyGen** in which private and public keys are constructed for a group of signers; **Agg Sign** in which all the group members provide their individual signatures on the respective messages to a combiner who then combines all the individual signatures, and outputs a constant size signature; **Agg Ver** in which the verifier verifies the aggregate signature using the public keys of all the group members.

Multi-signature can be seen as particular case of aggregate signature by taking the same message for all group members to sign.

3 Proposed Scheme

Let \mathbb{E} be a finite field, and l, v, w be the positive integers. $G = (G_1, G_2, \dots, G_w)$ be a group of w members who act as signers while constructing an aggregate signature. $H : \{0, 1\}^* \rightarrow \mathbb{E}^v$ be a collision-resistant hash function.

Agg KeyGen: Each group member G_i contains a private key (μ_i, D_i, v_i) , and the corresponding public key $\bar{D}_i = \mu_i \circ D_i \circ v_i \forall i = 1, 2, \dots, w$, where $\mu_i : \mathbb{E}^v \rightarrow \mathbb{E}^v$, $v_i : \mathbb{E}^l \rightarrow \mathbb{E}^l$ be the invertible affine transformations, and $D_i : \mathbb{E}^l \rightarrow \mathbb{E}^v$ be an easily invertible (solvable) quadratic system of equations. Another easily invertible quadratic system of equations $D^* : \mathbb{E}^{l+2v} \rightarrow \mathbb{E}^v$ is used as combiner function. C plays the role of a combiner who may or may not be the member of the group G . The map D^* is publicly known.

Agg Sign: Let $(m_1, m_2, \dots, m_w) \in \{0, 1\}^*$ be a set messages to sign. The construction of aggregate signature on (m_1, m_2, \dots, m_w) is as follows:

1. Each signer G_i computes hash $M_i = H(m_i \parallel D^*)$, $\forall i = 1, 2, \dots, w$.
2. Each signer G_i computes $\sigma_i = \bar{D}_i^{-1}(M_i) = (\mu_i \circ D_i \circ v_i)^{-1}(M_i) = v_i^{-1} \circ D_i^{-1} \circ \mu_i^{-1}(M_i)$, and sends it to C .
3. C obtains $\sigma_1, \sigma_2, \dots, \sigma_w$ and verifies these individual signatures as $\bar{D}_i(\sigma_i) = M_i$, $\forall i = 1, 2, \dots, w$. Then he computes

$$\begin{aligned}
 Q_1 &= D^*(M_1 \parallel \sigma_1 \parallel M_1) \\
 Q_2 &= D^*(Q_1 \parallel \sigma_2 \parallel M_2) \\
 Q_3 &= D^*(Q_2 \parallel \sigma_3 \parallel M_3) \\
 &\vdots \\
 Q_{w-1} &= D^*(Q_{w-2} \parallel \sigma_{w-1} \parallel M_{w-1}) \\
 Q_w &= D^*(Q_{w-1} \parallel \sigma_w \parallel M_w).
 \end{aligned}$$

4. C outputs $\sigma = Q_w$ as the aggregate signature on (m_1, m_2, \dots, m_w) .

Agg Ver: To verify the aggregate signature $\sigma = Q_w$ corresponding to the set of messages (m_1, m_2, \dots, m_w) and the group $G = (G_1, G_2, \dots, G_w)$, the verifier proceeds as follows:

1. Computes hash of messages $M_i = H(m_i \parallel D^*)$, $\forall i = 1, 2, \dots, w$.
2. Computes $D^{*-1}(Q_w)$ which may or may not have unique solution. If he gets unique solution, parses the solution to obtain Q_{w-1} and σ_w ; otherwise selects that solution which has M_w as its rightmost v components, and then parses the selected solution to obtain Q_{w-1} and σ_w . Similarly, computes $D^{*-1}(Q_{w-1})$, and obtains Q_{w-2} and σ_{w-1} . Repeating this process $w - 1$ times, he obtains up to Q_1 and σ_2 . At last, he obtains σ_1 by parsing the solution of $D^{*-1}(Q_1)$.

3. Now, the verifier has w components $(\sigma_1, \sigma_2, \dots, \sigma_w)$ corresponding to every group member. He verifies all these components as $\bar{D}_i(\sigma_i) = M_i, \forall i = 1, 2, \dots, w$. If the equalities hold for all $i = 1, 2, \dots, w$, accepts the signature; otherwise rejects.

4 Security

In this section, we discuss the security requirements of an aggregate/multi-signature.

Correctness: The correctness of the proposed scheme is straightforward. With given signature $\sigma = Q_w$, the verifier computes all the σ_i 's by repeated application of D^{*-1} on Q_w, Q_{w-1}, \dots, Q_1 . After that, the verification of all the σ_i 's by verifying the equality $\bar{D}_i(\sigma_i) = M_i$ is motivated by $\sigma_i = \bar{D}_i^{-1}(M_i)$ computed in **Agg Sign**.

Theorem 1 (Unforgeability) *The proposed scheme is existentially unforgeable against chosen message attack, if the underlying signature scheme is secure.*

Proof Let \mathcal{A} be a computationally bounded adversary who can forge the proposed aggregate signature, and \mathcal{C} be a challenger who responds to the hash queries and aggregate signing queries of \mathcal{A} . Here, \mathcal{C} would get the solution of MQ problem, if \mathcal{A} give a forgery on a message for which he has not made a signing query. The hash function H is taken as random oracle. Both \mathcal{C} and \mathcal{A} have the access of public keys, and now we see that how the proposed scheme can be proved secure using contradiction. We assume that the set of messages queried for aggregate sign query has already been queried for hash query.

Hash Query: Let \mathcal{A} make a hash query for an arbitrary set of messages (m_1, m_2, \dots, m_w) . \mathcal{C} chooses σ_i uniformly at random from \mathbb{E}^l corresponding to each m_i , and sends $\bar{D}_i(\sigma_i)$ as $H(m_i \parallel D^*)$ for all $i = 1, 2, \dots, w$. So, ultimately \mathcal{C} sends $(\bar{D}_1(\sigma_1), \bar{D}_2(\sigma_2), \dots, \bar{D}_w(\sigma_w))$ as the hash image of (m_1, m_2, \dots, m_w) to \mathcal{A} . If \mathcal{A} asks for the hash query of the target message set $(m'_1, m'_2, \dots, m'_w)$, the challenger \mathcal{C} sends random values (y_1, y_2, \dots, y_w) for which he wants the solutions of MQ problem. \mathcal{C} keeps a record of these sent responses in a hash query response list.

Agg Sign Query: When \mathcal{A} makes an aggregate hash query for that arbitrary set of messages (m_1, m_2, \dots, m_w) , \mathcal{C} checks its hash query response list, and retrieves the corresponding $(\sigma_1, \sigma_2, \dots, \sigma_w)$. Now, \mathcal{C} computes Q_w using D^* as \mathcal{C} computes in **Agg Sign**, and outputs Q_w .

The verifier can easily verify the aggregate query response by retrieving $(\sigma_1, \sigma_2, \dots, \sigma_w)$ from Q_w and then $\bar{D}_i(\sigma_i) = H(m_i \parallel D^*), \forall i = 1, 2, \dots, w$, as hash images of all the m_i 's have been sent as $\bar{D}_i(\sigma_i)$'s.

Forgery: Let \mathcal{A} output a valid aggregate signature $\sigma^t = Q_w^t$ on the target set of messages $(m'_1, m'_2, \dots, m'_w)$. So, \mathcal{C} uses Q_w^t and D^* to compute $(\sigma_1^t, \sigma_2^t, \dots, \sigma_w^t)$, and checks that $\bar{D}_1(\sigma_1^t) = H(m'_1 \parallel D^*) = y_1, \bar{D}_2(\sigma_2^t) = H(m'_2 \parallel D^*) = y_2, \dots, \bar{D}_w(\sigma_w^t) = H(m'_w \parallel D^*) = y_w$. Initially, we have assumed that

the underlying multivariate signature scheme is secure, it means \mathcal{C} gets the solution of some instances of MQ problem, which are $\bar{D}_1(x_1) = y_1, \bar{D}_2(x_2) = y_2, \dots, \bar{D}_w(x_w) = y_w$, where $x_i = \sigma_i^t, \forall i = 1, 2, \dots, w$. This leads to a contradiction.

Impersonation: Let \mathcal{I} be an impersonator who wants to impersonate a signer $G_a (1 \leq a \leq w)$ from the group G . For that \mathcal{I} need to output σ_a satisfying the equation $\bar{D}_a(\sigma_a) = H(m_a \parallel D^*)$ for the given message m_a . But, as we have already discussed above that the underlying signature scheme is secure, so to find such σ_a is equivalent to solve an instance of MQ problem which is computationally hard. So, \mathcal{I} cannot impersonate any of the signer from G .

5 Parameters and Comparison

We compare the signature size of our scheme with the existing multivariate-based aggregate signature schemes [2, 3] for different security levels. Table 1 shows that the signature size in our scheme is significantly smaller than the schemes [2, 3]. We take the signature size of the schemes [2, 3] for 20 users using UOV parameters [7]. In the proposed scheme, the aggregate/multi-signature size is just l field elements. Particularly, using Rainbow signature scheme [4] as underlying signature scheme for the finite field $\mathbb{E} = GF(31)$, the parameters $l = 79, v = 52$ are used for 128-bit security. So, our signature size will be $79 * 5 = 395$ bits = 0.048 kbyte, as one element of $GF(31)$ is 5-bit long.

6 Conclusion and Future Scope

In this work, we have proposed a multivariate aggregate/multi-signature scheme. The signature size of the proposed scheme is constant, i.e., does not depend on the size of the group. Moreover, the signature size is smaller than the existing multivariate aggregate signature schemes. In the proposed scheme as well as existing schemes, the verification cost is not independent of group size. So, we suggest the researchers to design the multivariate aggregate/multi-signature having constant time verification

Table 1 Signature size comparison (Kilo Bytes)

Security (bit)	Scheme [2]	Scheme [3]	Our scheme (UOV)	Our scheme (Rainbow)
80	1.62	1.62	0.113	0.032
100	1.90	1.90	0.140	0.039
128	2.55	2.55	0.187	0.048

as the future scope of this paper. For the time being, the proposed scheme can be proved very handy due to its quantum-resistant MPKC structure.

References

1. Boneh D, Drijvers M, Neven G (2018) Compact multi-signatures for smaller blockchains. In: International conference on the theory and application of cryptology and information security, pp 435–464. Springer
2. Bansarkhani RE, Mohamed MSE, Petzoldt A (2016) MQSAS—a multivariate sequential aggregate signature scheme. In: ISC 2016, LNCS, vol 9866. Springer, pp 426–439
3. Chen J, Ling J, Ning J, Peng Z, Tan Y (2020) MQ aggregate signature schemes with exact security based on UOV signature. In: INSCRYPT-2019, LNCS, vol 12020. Springer, pp 443–451
4. Ding J, Schmidt DS (2005) Rainbow, a new multivariate polynomial signature scheme. In: ACNS 2005, LNCS, vol 3531. Springer, pp 164–175
5. Garey Michael R, Johnson DS (1991) Computers and intractability, a guide to the theory of NP-completeness. W.H. Freeman
6. Itakura K, Nakamura K (1983) A public-key cryptosystem suitable for digital multisignatures. In: NEC research and development, vol 71. NEC, pp 1–8
7. Kipnis A, Patarin L, Goubin L (1999) Unbalanced oil and vinegar schemes. In: EUROCRYPT 1999, LNCS vol 1592. Springer, pp 206–222
8. Le D-P, Yang G, Ghorbani A (2019) A new multisignature scheme with public key aggregation for blockchain. In: 17th international conference on privacy, security and trust (PST). IEEE, pp 1–7
9. Petzoldt A, Chen MS, Yang BY, Tao C, Ding J (2015) Design principles for HFEv-based signature schemes, ASIACRYPT 2015—Part 1, LNCS, vol 9452. Springer, pp 311–334
10. Shor P (1997) Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. SIAM J Comput 26(5):1484–1509
11. Zhang Z, Xiao G (2004) New multisignature scheme for specified group of verifiers. In: Applied mathematics and computation, vol 57. Elsevier, pp 425–431

Optical Network Modeling and Performance Using Random Graph Theory



Rahul Deo Shukla, Ajay Pratap, and Raghuraj Singh Suryavanshi

1 Introduction

The need for Internet traffic is increasing extremely quickly in the telecommunications world of today. The continuous expansion of data-centric applications has increased the requirement of higher bandwidth. Among various techniques, the key method that might be helpful when creating networks to meet the growing need for bandwidth is optical packet/burst switching. This technology is regarded as the future technology due to its high bandwidth utilization, low latency, and high throughput [1, 2]. Switches used in optical network implementation might be entirely optical or electrical in design. Switching's primary goal is to direct the packet to the proper destination port. The main challenge with an optical network is designing a switch and router configuration that can successfully conduct switching operations at high data speeds. Due to the infeasibility of optical processors, the present technology uses a blended method in which control operations are carried out by electronics while data propagates in the optical domain. This mix method is known as photonic packet switching technology. The control operations of big photonic packet switches in the photonic packet switching technology are likely to be handled by electronics, while packet routing and buffering are done optically [3, 4].

R. D. Shukla (✉)

Research Scholar, Amity Institute of Information Technology, Amity University Uttar Pradesh, Lucknow Campus, Lucknow, India
e-mail: rahuld.shukla@gmail.com

A. Pratap

Amity Institute of Information Technology, Amity University Uttar Pradesh, Lucknow Campus, Lucknow, India

R. S. Suryavanshi

Department of Computer Science & Engineering, Pranveer Singh Institute of Technology, Kanpur, India

A number of multinational companies are working on the highly demanding cloud computing. Network connectivity and dispersed computing resources are the two concepts that appear most frequently in descriptions of cloud computing. These two problems have drawn a lot of attention in recent years across a wide range of industries, concurrently with the Internet's enormous popularity and the increasing demands for processing enormous amounts of data, including the big data idea [5]. The concept of "big data" is used to describe the exponential rise in the amount of information that is available and utilized from a variety of sources.

The role of transport network is quite crucial component in the cloud computing concept because it connects disparate computer resources. The rapid growth of cloud computing necessitates a rigorous evaluation of the present network infrastructure from the standpoint of cloud computing demands. The demands of cloud environments cannot be effectively met by the present transport networks as per [6]. Three needs are specifically highlighted by the authors of [6] for transport networks that are cloud-ready.

- It must be flexible enough to provide the necessary capacity as needed.
- Multilayer-focused network administration.
- Cross-strata features that allow for cooperative resource optimization of the cloud-based application's resources and the connectivity-supporting underlying network.

Additionally, current networks are primarily designed to handle unicast (one-to-one) traffic, whereas various cloud computing applications generate novel traffic patterns like anycast (one-to-one-of-many) flows. When processing is subsequently concentrated in a limited number of locations (data centers), we may have a significant increase in the volume of traffic on network links around these locations, necessitating the use of high-capacity network technology.

The optical communication technology has the capacity to deal with cloud computing issues. This paper investigates the networking modeling using the random graph theory.

2 Related Works

In optical networks, nodes are connected via fiber links and data propagates on different wavelengths. At each node wavelengths are separated and passed through the switch and after dropping, buffering they again appear at the output of the switch where they are combined and pushed into the networks (Fig. 1) [7, 8]. Without wavelength conversion, the switching functionality cannot be fully achieved and the optical network's performance goes down as shown in [9, 10]. The light path maintains the same wavelength on all of the fiber links it uses if wavelength conversion is not possible; this leads to underutilized channel capacity. By allowing a link to utilize multiple wavelengths along its course, wavelength conversion enhances

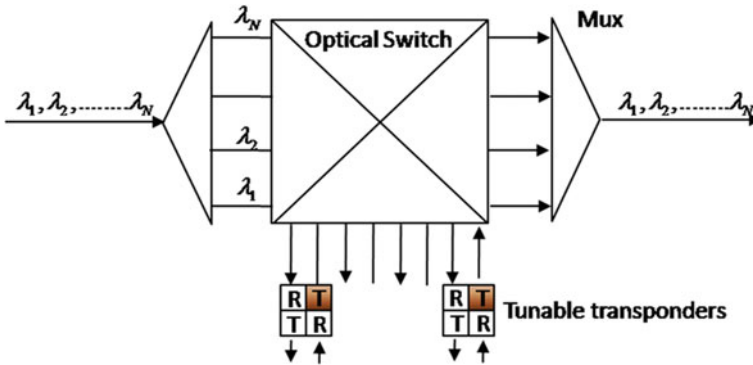


Fig. 1 Schematic representation of node design in optical networks

network blocking performance [6]. It is important to note that the cost of full wavelength conversion at each node is very high, therefore limited wavelength conversion is also considered [6].

Wavelength converters are only present in a small portion of nodes in these networks [11–14]. The optical switches are connected to create optical core networks. Numerous optical switch architectures have been presented in the past. There is a lot of competition among the data packets, so different strategies are utilized to reduce packet drop. A buffer-less switch architecture proposed by Proietti et al. delivers a negative acknowledgement to the sender in the event of contending packets and conducts retransmission for conflicting packets [15]. A large number of packets will be sent again due to the buffer-less design. The network will see more traffic as a result of packet retransmission and negative acknowledgment. The resolution of contention and streamlined traffic flow are both aided by packet buffering.

The concept of fiber delay lines (FDLs) was born as a result of the lack of optical RAM. By including a sequence of set length delays where data are kept for a brief period of time, optical FDL buffering prevents contention [16, 17] (Fig. 2). Storage in FDL is extremely constrained because of noise build-up. Due to FDL's bulky design, a high size buffer is not practical, and packets that cannot be held are ultimately destroyed. As a result, use of electronic data storage as a substitute strategy was proposed. Electronic buffering is a technique for managing congestion in which conflicting packets are buffered into a single shared electronic buffer [18, 19] (Fig. 3). Complex and power-hungry components are required for an electronic loopback buffer, as well as numerous optical and electrical conversions. For the resolution of optical packet contention, several researchers have also proposed a hybrid buffering approach. For short-term packet storage, an optical buffer is used, whereas for long-term packet storage, an electronic buffer is used [18, 19]. During full wavelength conversion, the blocking likelihood is reduced by buffering competing packets. In actual situations, switches are put in networks. Packets must be added or removed in order to increase the optical network system's connectivity.

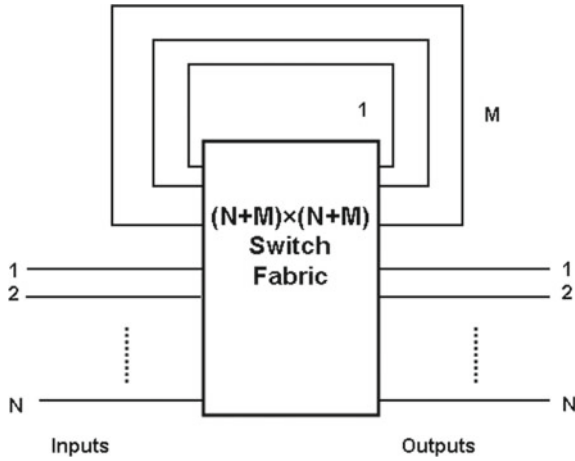


Fig. 2 Schematic representation of optical buffering using FDL

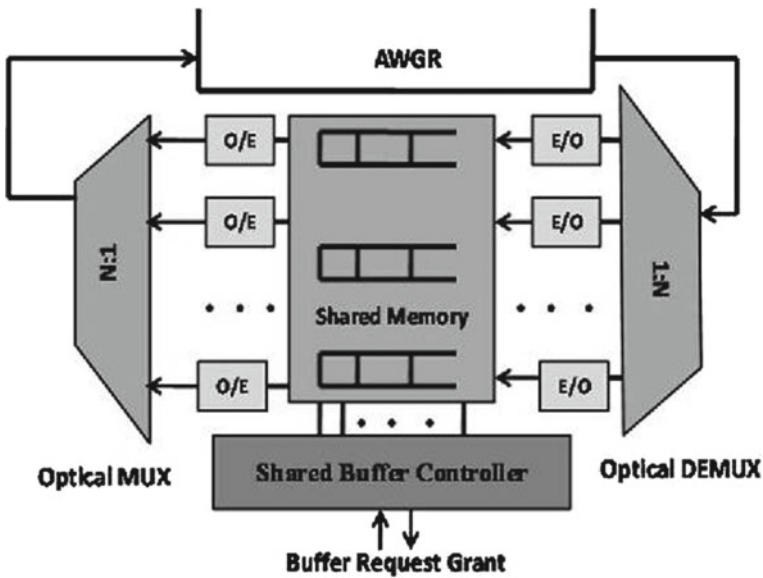


Fig. 3 Schematic representation of electronic buffering using FDL

The PLP of the switch is also examined under various loading and buffering circumstances. In the examination of blocking probability or PLP, load balancing is not taken into account. Using a load balancing approach can lower the likelihood of blocking and improve how network resources are utilized. By distributing traffic evenly over all of the network's links, load balancing prevents traffic jams on those

links. By adopting load balancing, the number of wavelength converters deployed in the network can be decreased.

3 Random Graph Model

The random graph commonly regarded as the basic and oldest approach to modeling network behavior is Erds-Rényi model, [20, 21]. We make random distribution of ‘ n ’ nodes and add each node via an edge with probability ‘ α ’ in this model [22] $d = \frac{\log n}{\log \alpha(n-1)}$ is the normal geodesic distance is typically. For example, the average geodesic distance is 2 when $n = 10,000$ and $\alpha = 0.01$.

The given model proves to be ineffective because it is unrealistic for two nodes to link randomly in a real network. The following model was put out by Watts and Strogatz [23], who dispersed the nodes in a circle. In this model, each node is connected to both its closest and subsequent closest neighbors. Although this model is heavily clustered, it nevertheless deviates from small world features because on average a sizable number of nodes must be visited in order to link to random nodes. The authors suggested rewiring a few more wires linking some randomly picked nodes to shrink this globe. However, this model also misses a crucial aspect of actual networks.

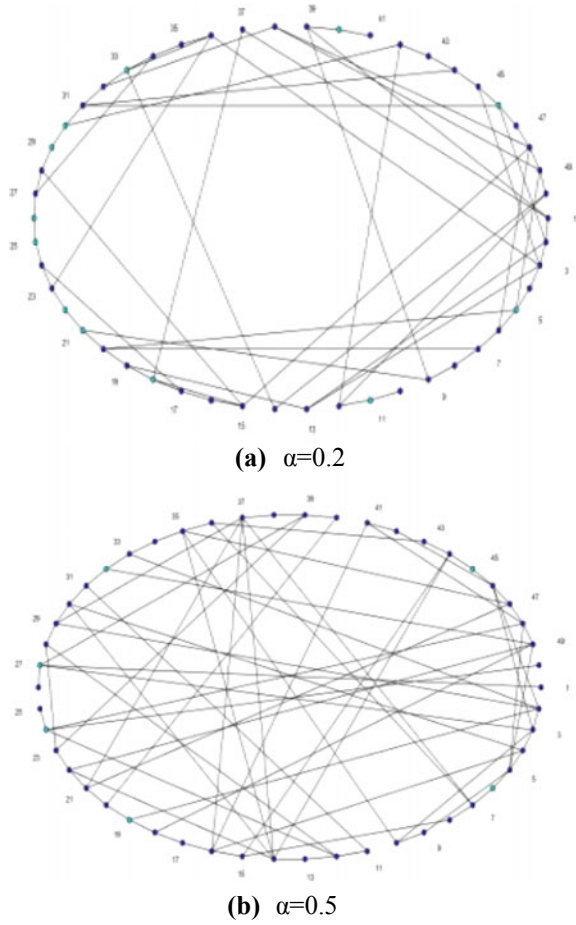
With 50 nodes and a rewiring probability of 0.2 or 0.5, respectively, the Erds-Rényi model is used in Fig. 4(a) and (b) to depict the network and linked nodes. Findings for 300 nodes are similarly displayed in Fig. 5(a) and (b) The greatest eccentricity of each graph vertex is known as the diameter of the graph. In other words, it is the distance along the longest, shortest path connecting any two graph nodes.

Figure 6 shows the relationship between diameter and node count for random graphs with $\alpha = 0$, Erds-Rényi model with $\alpha = 0-1$, and Watts and Strogatz with $\alpha = 1$. The graphic makes it evident that the diameter of a random graph rises linearly while it remains unchanged in the other two circumstances at a diameter value of 5-6.

Figure 7 demonstrates how the number of nodes and the shortest path length interact in random networks, with $\alpha = 0$, Erds-Rényi model with $\alpha = 0-1$, and Watts and Strogatz with $\alpha = 1$. The graph demonstrates that the random graph’s diameter increases linearly but in the other two circumstances, the dimension stays fixed at 4. Due to the fact that the diameter and shortest path between nodes would vary in a real-world network, neither of these models can adequately capture its features.

Barabási and Albert [24] presented an architecture based on two known facts about real networks: networks expand continually when new vertices are added, and these vertices connect specifically to locations that are now widely interconnected. A new node may link to a number of nodes in the current network at each time step in this approach, which assumes a modest number of nodes initially ($t = 0$). A node’s degree determines how likely it is that new links will be added to it, therefore a node with a higher degree is more likely to do so.

Fig. 4 Schematic representation of random graph using Erdős-Rényi model, **a** $\alpha = 0.2$, **b** $\alpha = 0.5$



In a huge proportion of real networks, a normal approximation for the degree power-law exponent. The distribution of outward degree on the Internet is provided by

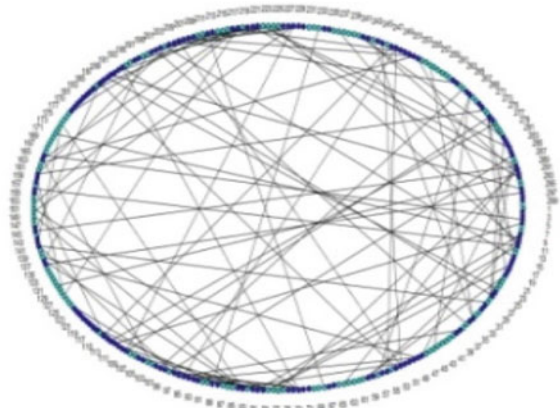
$$d(u) = au^{-\zeta} \quad 2.38 \leq \zeta \leq 2.72 \tag{1}$$

The distribution of inward degree is represented as

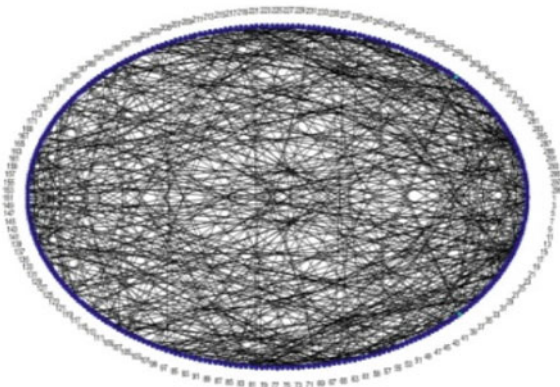
$$d(u) = au^{-\zeta} \quad \zeta = 2.1 \tag{2}$$

The degree vs. fraction of nodes is plotted in Fig. 7. The figure makes it clear that nodes having a high degree are rare. A node having a degree of “1000” is highly improbable with a probability of 10^{-6} . The likelihood of an occurrence with a node degree of “10” is similarly extremely low (0.01). In real-time applications, more

Fig. 5 Schematic representation of random graph using Watts and Strogatz, **a** $\alpha = 0.2$, **b** $\alpha = 1$



(a) $\alpha=0.2$



(b) $\alpha=1$

incoming linkages than outgoing links have been observed. According to various values of, ζ the quantity of outbound links likewise fluctuates. Average degree is given by

$$\langle k \rangle = \sum_{k=1}^{\infty} kp(k) = \sum_{k=1}^{\infty} \frac{1}{k^{\gamma-1}} \tag{3}$$

or

$$\langle k \rangle = 1 + \frac{1}{2^{\gamma-1}} + \frac{1}{3^{\gamma-1}} + \dots + \dots \tag{4}$$

The average node degree can be obtained by Eq. 4, for $\gamma = 2.1$ is 6.61, whereas the average node degrees for $\gamma = 2.38-2.72$ are 3.16 and 2.01, respectively. However,

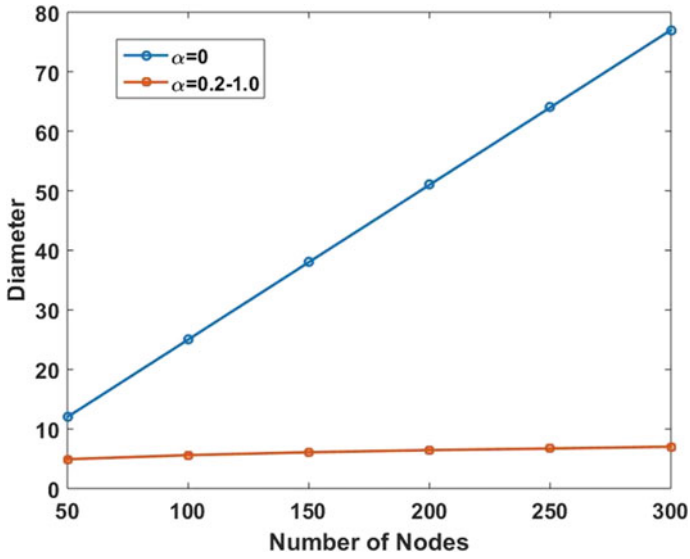


Fig. 6 Network diameter vs. number of deployed nodes for various random graph models

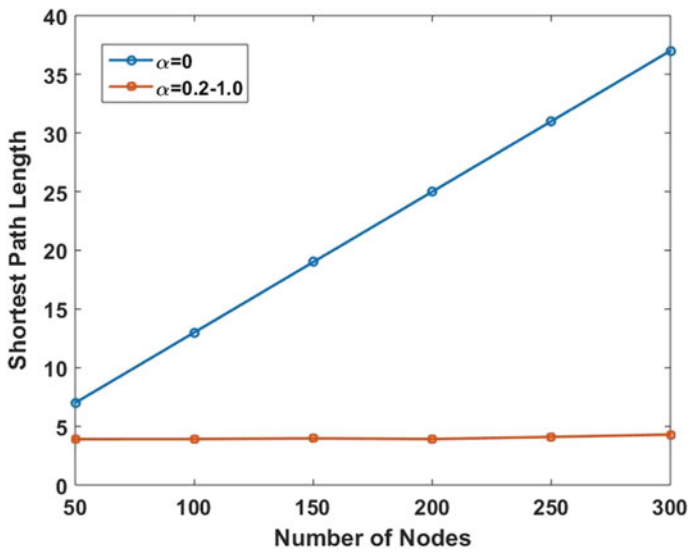


Fig. 7 Shortest path vs. number of deployed nodes for various random graph models

the Internet is generally a small world network (SSN) with an average degree of 4 and $\gamma = 2.2$ (Fig. 8).

The average shortest distance in the Leskovec et al., Web’s experiment, which had 855,802 nodes, was 7, with a diameter of 21 [21].

As detailed above, the inward and outward degrees for networks lie between 2 and 4. Therefore various topologies are designed with minimum of 2 links and maximum of 4 links. In Fig. 9, USA, NSFNET topology is shown with 14 nodes and maximum node degree of 3. In Fig. 10, European Union, EON topology is shown with 28 nodes and maximum node degree of 5. These topologies are designed by considering various factors like geographical distance and wavelength routing criterion, etc.

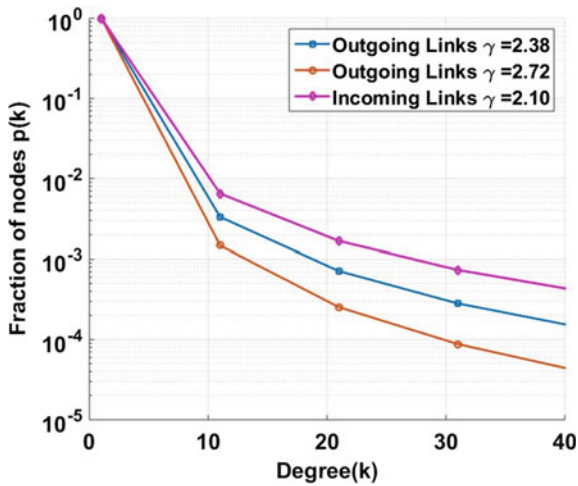


Fig. 8 Fractions of nodes vs. degree

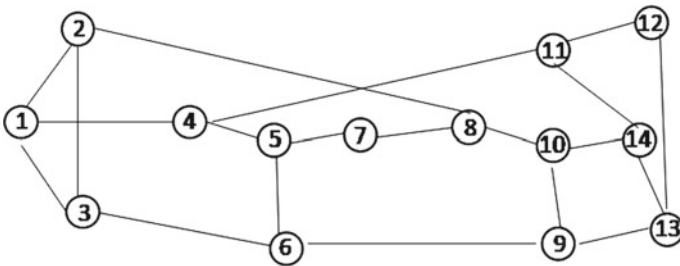


Fig. 9 Schematic of NSFNET topology (USA)

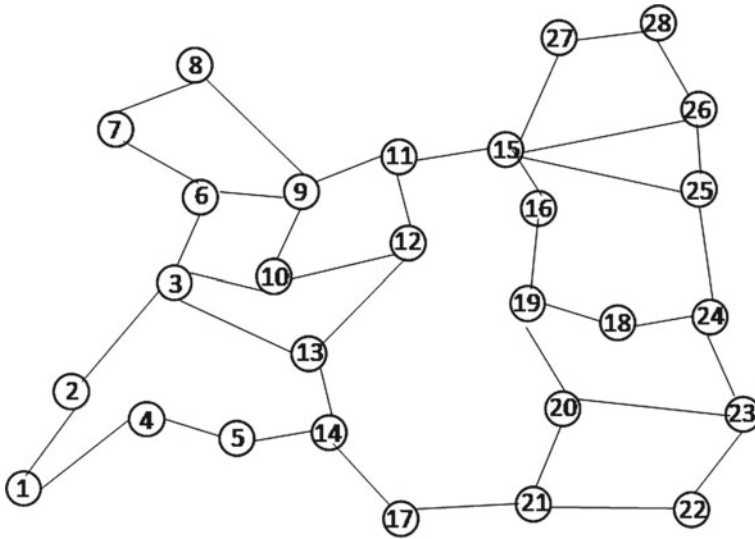


Fig. 10 Schematic of EON topology (European union)

4 Results

As already mentioned, a small-scale network has an average degree of 4, and hence we have to consider 4 input and 4 output links. The traffic model is considered to be random as detailed in [16]. Figure 11 plots PLP against load while varying the buffer from 0 to 16. Here, $B = 0$ makes it very evident that the switch is bufferless, meaning that the likelihood of packet loss is quite high. As a result, it can be concluded that intermediate switches require buffers and that even a little increase in buffer significantly lowers the likelihood of packet loss. Comparing results at the load of “0.6,” the likelihood of a packet being lost for buffer-less switch is 0.21, and for the buffering of 4 packets is 1.4×10^{-3} , and the buffering of 16 packets is less than 10^{-6} . Hence, a buffer upgrade of 16 results in a 200,000 times improvement in packet loss efficiency.

Ultimately, various applications will necessitate varied packet loss rates as well as varying buffer sizes for various numbers of input links. If we require a packet loss probability of 10^{-4} at a load of “0.6,” then maybe the necessary buffer for $N = 4$ is 6.

The maximum number of traversed nodes can increase to 21 as discussed above. As a result, overall packet loss will increase. Let on a node the packet loss is P_L , then the throughput will be $(1 - P_L)$, and if data passes through “ m ” number of switches thus the overall throughput will be $(1 - P_L)^m$. Thus, the final PLP after traversing through m node will be

$$P_L^m = [1 - (1 - P_L)^m] \quad (5)$$

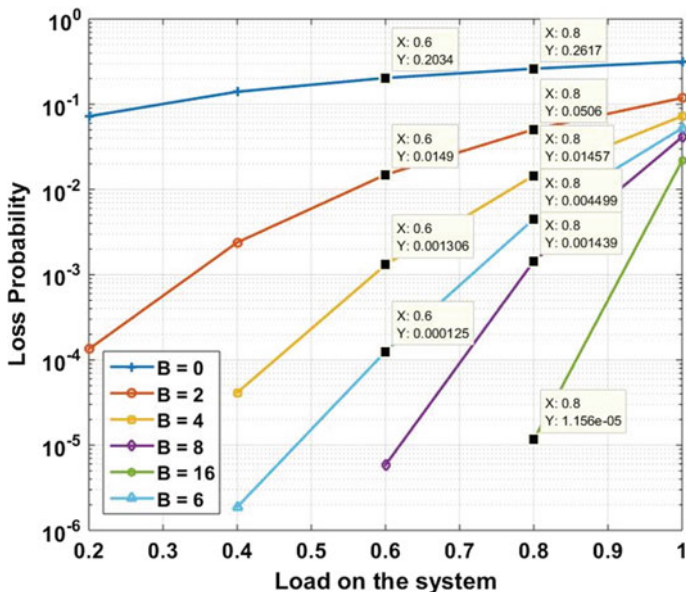


Fig. 11 Loss probability vs. load under varying buffer conditions

In Fig. 12, Effective loss probability vs. number of traversed switches is shown while considering PLP of a node as 10^{-3} and 10^{-4} respectively. It is clearly from the figure that the PLP nearly raises with the number of traversed nodes. For example, after traveling through the 10 nodes the effective PLP becomes 10 times. It can be understood from Eq. 5, as for the smaller value of P_L using the bi-nominal expansion the Eq. 5, can be written as

$$P_L^m = [1 - (1 - P_L)^m] = 1 - (1 - mP_L) = mP_L \tag{6}$$

The previous works deal with the packet loss performance of the individual switches. In this work packet loss performance of the cascaded switches is evaluated and it has been found that in the optical communication system where PLP is very less, the PLP of the cascaded switches rises linearly with the number of switches.

5 Conclusions

In the last several years, Internet usage has exploded. As a result, storage and heating issues are present with servers. Systems are switching from entirely electronic to hybrid technologies to deal with these problems. The design modeling of optical network is proposed by considering random graph theory. Various graph theory

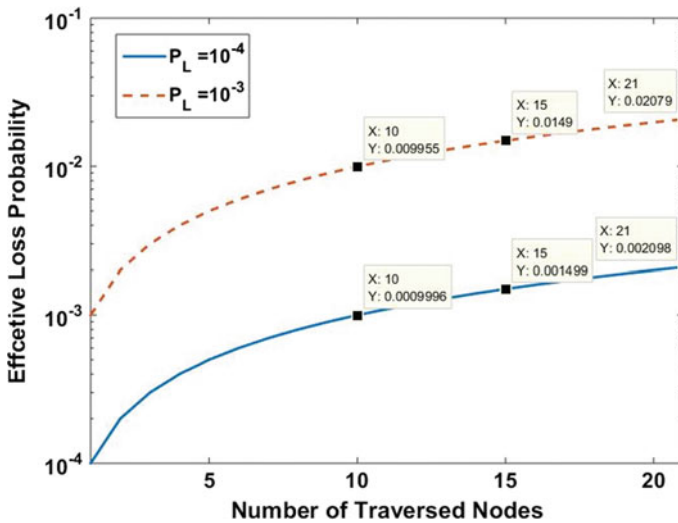


Fig. 12 Effective loss probability vs. number of traversed switches

models are discussed, and it has been found that the inward and outward degree in most of the nodes is 4. The number of cascaded switches is also evaluated. The simulation results are carried out to obtain the packet loss performance for an individual node and as well as for the cascade switches. It has been found that the PLP rises linearly with the number of cascaded switches.

References

1. Singh RK, Srivastava R, Singh YN (2000) Wavelength division multiplexed loop buffer memory based optical packet switch. *Opt Quant Electron* 39(1):15–34
2. Bhattacharya P, Singh A, Kumar A, Tiwari AK, Srivastava R (2017) Comparative study for proposed algorithm for all-optical network with negative acknowledgement (AO-NACK). In: *Proceedings of the 7th international conference on computer and communication technology 2017*, pp 47–51
3. Shukla RD, Pratap A, Suryavanshi RS (2020) Packet blocking performance of cloud computing based optical data centers networks under contention resolution mechanisms. *J Opt Commun* <https://doi.org/10.1515/joc-2019-028>
4. Tucker RS, Zhong WD (1999) Photonic packet switching: an overview. *IEICE Trans Commun E* 82:254–264
5. Kleinberg J, Kumar A (2001) Wavelength conversion in optical networks. *J Algorithms* 38(1):25–50
6. Ramamurthy R, Mukherjee B (2002) Fixed-alternate routing and wavelength conversion in wavelength-routed optical networks. *IEEEACM Trans Netw* 10(3):351–367
7. Yin Y, Wen K, Geisler DJ, Liu R, Yoo SJB (2012) Dynamic on demand defragmentation in flexible bandwidth elastic optical networks. *Opt Exp* 20(2):1798–1804

8. Wang X, Kim I, Zhang Q, Palacharla P, Ikeuchi T (2016) Efficient all-optical wavelength converter placement and wavelength assignment in optical networks. Optical fiber communication conference, W2A–52. Optical Society of America
9. Bonani LH, Forghani-Elahabad M (2016) An improved least cost routing approach for WDM optical network without wavelength converters. *Opt Fiber Technol* 32:30–35
10. Jara N, Vallejos R, Rubino G (2017) Blocking evaluation and wavelength dimensioning of dynamic WDM networks without wavelength conversion. *J Opt Commun Netw* 9(8):625–634
11. Patel AN, Ji PN, Jue JP, Wang T (2012) Routing, wavelength assignment, and spectrum allocation in wavelength convertible flexible optical WDM (WC-FWDM) networks. In: National fiber optic engineers conference, JTh2A–36. Optical Society of America
12. Danielsen SL, Hansen PB, Stubkjaer KE (1998) Wavelength conversion in optical packet switching. *J Lightwave Technol* 16(12):2095–2108
13. Sahu PP (2008) New traffic grooming approaches in optical networks under restricted shared protection. *Photon Netw Commun* 16(3):233–238
14. Chatterjee BC, Sarma N, Sahu PP (2012) Priority based routing and wavelength assignment with traffic grooming for optical networks. *Opt Commun Netw* 4(6):480–489
15. Proietti R, Yin Y, Yu R, Ye X, Nitta C, Akella V, Yoo SJB (1998) All-optical physical layer NACK in AWGR-based optical inter-connects. *IEEE Photon Technol Lett* 24(5):410–412
16. Srivastava R, Singh YN (2010) Feedback fiber delay lines and AWG based optical packet switch architecture. *Opt Switch Netw* 7(2):75–84
17. Pallavi S, Lakshmi M, Srivastava R (2015) Physical layer analysis of AWG based optical packet switch architecture. *J Opt* 44(2):119–127
18. Singh P, Rai JK, Sharma AK (2020) Bit error rate analysis of AWG based add-drop hybrid buffer optical packet switch. In: 2020 2nd IEEE International Conference on Advances in Computing, Communication Control and Networking (ICACCCN). pp 454–458
19. Singh P, Rai JK, Sharma AK (2021) Hybrid buffer and AWG based add-drop optical packet switch. *J Opt Commun*. <https://doi.org/10.1515/joc-2021-0058>
20. Daudin JJ, Picard F, Robin S (2008) A mixture model for random graphs. *Stat Comput* 18(2):173–183
21. Leskovec J, Lang KJ, Dasgupta A, Mahoney MW (2009) Community structure in large networks: Natural cluster sizes and the absence of large well-defined clusters. *Internet Math* 6(1):29–123
22. Newman ME, Watts DJ, Strogatz SH (2002) Random graph models of social networks. *Proc Natl Acad Sci* 99(suppl 1):2566–2572
23. Watts DJ, Strogatz SH (1998) Collective dynamics of ‘small-world’ networks. *Nature* 393(6684):440–442
24. Barabási A-L, Albert R (1999) Emergence of scaling in random networks. *Science* 286(5439):509–512

A Survey on Recent Progress in Blockchain Technology



Naseem Ahmad Khan, Prateek Singh, Radiant Ambesh,
and Md Tarique Jamal Ansari 

1 Introduction

Among the most encouraging emerging innovations of the twenty-first century is blockchain technology. It has numerous advantages, including decentralisation, deception freedom, counterfeiting liberty, as well as auditability. It is ideal for critical anti-counterfeit information storage, data protection, as well as other practical scenarios. Blockchain technology predetermine manipulation as well as data loss security problems in conventional centralised institutions as well as money transfers in finance, medical services, the Internet of Things, real estate rights protection, as well as privacy. The advent of blockchain technology signalled the emergence of research domains as well as innovative decentralised technologies [1–5].

Blockchain is one of the technological innovations that has surfaced in the last decade which has offered a significant deal of assurance. Work is still being done to fully understand the power of blockchain and also where it can be employed successfully. Some believe that blockchain is the key to achieving a decentralised society. Our current surroundings are completely integrated. That is, the couple has the authority to make decisions. Our entire financial framework, for instance, is confined by state-approved banks, as well as decision-making in associations are obviously made by a few individuals on the board of executives. Even the monsters which billions of clients rely on regularly prefer what we see. Decentralised systems are ubiquitous, but power is distributed throughout the system. Bitcoin is a prototype that does not rely on banks or exchanges for financial intermediary; each interaction is distinctive

N. A. Khan · P. Singh · R. Ambesh

Department of Electronics and Communication Engineering, University Institute of Engineering and Technology, Babasaheb Bhimrao Ambedkar University, Lucknow, India

M. T. J. Ansari (✉)

Department of Information Technology, Babasaheb Bhimrao Ambedkar University, Lucknow, India

e-mail: tjtjansari@gmail.com

to each arrangement and the blockchain also stores record, allowing everyone in the system to regulate each interaction. Users can trace it back to its source. This study describes blockchain and discusses some of the applications [6–10].

The Worldwide Blockchain Industry is projected to be worth USD 5.3 billion in 2021 and USD 34 billion by 2026, expanding at a CAGR of 45%. A primary determinant for blockchain technology is the emergence of Blockchain as a Service (BaaS), which enables customers to use cloud-based alternatives for constructing and organising their blockchain apps. Additionally, there has been an increase in cryptocurrency acceptance, which has fueled market growth. The implementation of blockchain technology is widespread, particularly in the financial industrial sector, since it can decrease payment processing time, thereby reducing the intricacies associated with a deal and boosting market expansion. The large starting cost of establishing the method and operational processes, strict regulations requirements in various countries, as well as a scarcity of technical skillsets for incorporating blockchain technology may all be barriers to market expansion [11] (Fig. 1).

Blockchain technology is not just a single technology, it includes cryptography, arithmetic, algorithmic rules and economic models, combining peer-to-peer networks and algorithmic rules to solve old problems. Synchronisation of distributed information. It’s building a unified multi-span infrastructure [12–15].

• **Decentralisation**

In a typical centralised trading system, all trades are sent to a central trusted authority (e.g. central banks) inevitably become value and performance bottlenecks at central

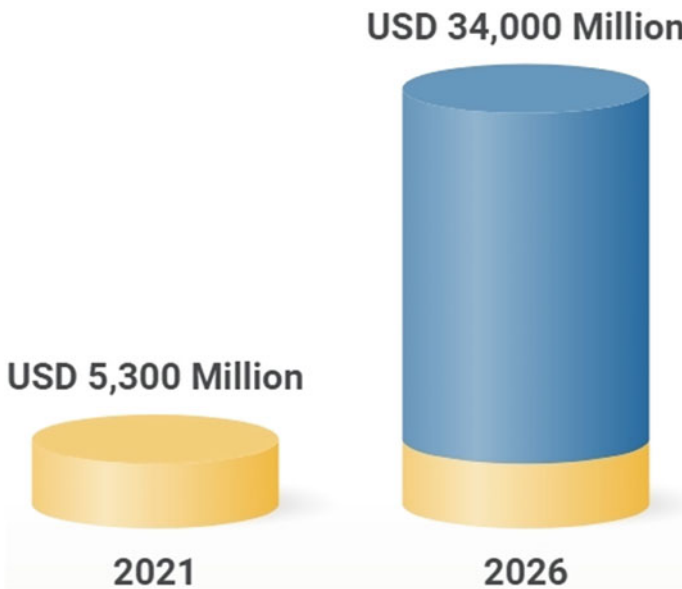


Fig. 1 Global blockchain market 2021–2026 (Source Research and Markets)

servers. Conversely, transactions within the blockchain network are performed between any two peers (P2P), but are not authenticated by a central office. In this way, blockchain significantly reduces server prices (including event value and operational costs), alleviating the performance bottlenecks of central servers.

- **Persistence**

It's almost impractical to change anything, because every interaction that propagates through the system must be confirmed and recorded in the masses that circulate through the system. Also, every space communicated is massively scrutinised by various hubs and exchanges. In this way, we were able to effectively distinguish between each distortion.

- **Anonymity**

Each client trades blockchain coordinates with a generated address. Additionally, customers can make some deliveries to keep a strategic distance from showcasing their personalities. Currently, there is no central repository of customer information close at hand. This component protects the exact level of security of blockchain-locked exchanges. Please note that due to property mandates, blockchains cannot guarantee the best possible protection.

- **Verifiability**

Each transaction on the blockchain is valid and recorded with a time stamp, allowing users to access any node in the decentralised network to easily verify and track previous records. The Bitcoin blockchain allows each transaction to be repeatedly copied onto the previous transaction. Better traceability and more transparency of information storage in the blockchain (Fig. 2).

2 Functioning of Blockchain Technology

Numerous organisations around the globe have been incorporating Blockchain innovation in recent times. But how exactly does Blockchain technology operate? Is this a substantial modification or a minor addition? Blockchain developments are still in their infancy and have the possibilities to be groundbreaking in the coming years. Blockchain is a hybrid of three key technologies:

- Cryptographic keys.
- A peer-to-peer connectivity with a shared ledger.
- A computing method for storing transaction processing as well as records.

Cryptography keys are comprised of two secret keys: private as well as public. Such keys assist in operating effective transactions among two stakeholders. Each person possesses these key pair, that they employ to generate a highly secured authenticity reference. The most significant element of Blockchain technology is its secure individuality. This identity is also known as a 'digital signature' in the

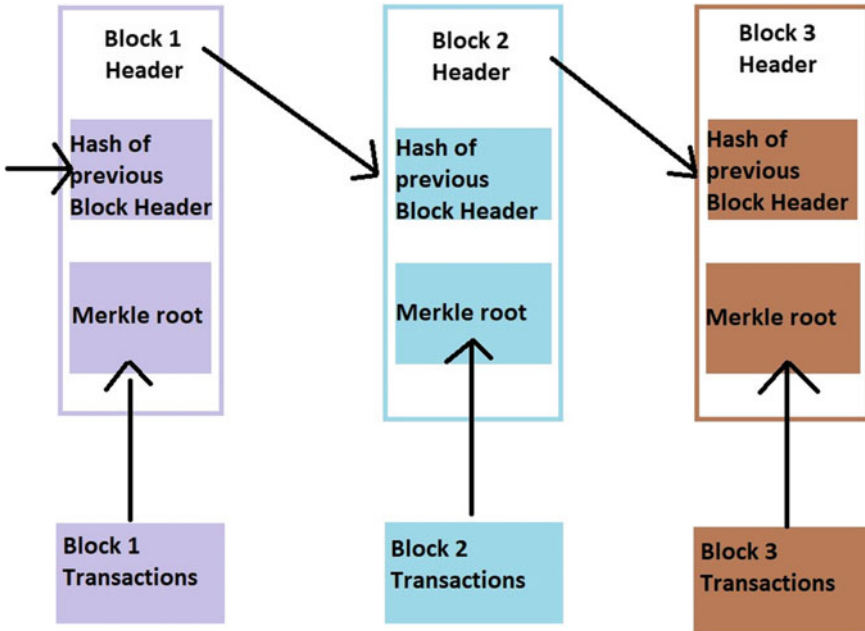


Fig. 2 Blockchain architecture

cryptocurrency space and is employed for authorising and managing transactions. The digital signature is integrated into the peer-to-peer system; a huge number of users acting as authorities utilise the digital certificate to reach agreement on transactions and other concerns. When they approve an operation, it is validated mathematically, resulting in an effective secured transaction among the two network-connected groups. To summarise, Blockchain customers utilise cryptography key code to conduct additional kinds of digital connections across the peer-to-peer system.

3 Types of Blockchain

3.1 Private Blockchain Networks

Private blockchains run on closed networks as well as are best suited to private companies and organisations. Organisations can utilise private blockchains to tailor their responsiveness as well as authorization priorities, network conditions, as well as other critical security features. A private blockchain network is managed by a single authority.

3.2 Public Blockchain Networks

Bitcoin and numerous different cryptocurrencies arose from public blockchains, that also contributed to the widespread adoption of distributed ledger technology (DLT). Additionally, public blockchains aid in the elimination of certain difficulties and issues, including such security vulnerabilities and centralization. DLT distributes data throughout a peer-to-peer network instead of storing it in a single place. A consensus mechanism is employed to confirm the authenticity of details; proof of stake (PoS) as well as proof of work (PoW) are two frequently utilised consensus techniques.

3.3 Permissioned Blockchain Networks

Permissioned blockchain technologies, also referred to as hybrid blockchains, are private blockchains that grant special access to authorised individual persons. Institutions generally set up such kinds of blockchains to receive the best of both worlds, as well as it allows for better configuration when determining who can engage in the network as well as which transactions could be made.

3.4 Consortium Blockchains

Consortium blockchains, like permissioned blockchains, provide both public and private aspects; however, multiple organisations will control a single collaboration blockchain network. Even though these blockchains are more difficult to set up at first, once operational, they can provide greater security. Furthermore, consortium blockchains are ideal for collaboration among multiple organisations.

4 Application of Blockchain

Since the invention of blockchain, much research has been conducted to discover what else this incredible technology can be utilised for. Blockchain applications continue to emerge, a few of which are characterised here [13–17].



Fig. 3 Bitcoin blockchain

4.1 Financial Applications of Blockchain

The first and most important application of blockchain is in financial services. It all began with Bitcoin that used blockchain technology to record banking transactions and eliminate middlemen. As of Bitcoin, different blockchain technologies have given rise to a variety of digital currencies that there are now hundreds of them traded globally. The Bitcoin blockchain is depicted in Fig. 3. When a new transaction is formed, it is transmitted through the network. Miners record such transactions, which are then cryptographically encased into blocks after they have been validated. This block now hashes to the preceding block, which tracks as well as observes users, but all of the data on the web is not ensure security. The Internet of Things (IoT) sector is very interested in blockchain as a decentralised and high maintenance strategy. The amount of IoT nodes is growing by the day, as well as the amount of information gathered. Data security is constantly a concern, and blockchain can assist in protecting and managing this data.

4.2 Smart Contracts

Blockchain with smart contracts, as the name implies, can remove the requirement for lawyers as well as middlemen. All sides can have access to smart contracts, as well as any changes to the contract should be made after achieving Consensus. Smart contracts could be useful in both business and private transactions.

4.3 Blockchain and Internet of Things

The Internet has become such a big component of everyone's lives that we sometimes don't realise how interconnected everything is. Most devices, including smart watches, intelligent fridges, camera systems, as well as mobile phones, are connected to the web. The Internet of Things (IoT) is a network of smart sensors and devices that are linked to the internet and share information to simplify our lives. There's no doubting that the Internet of Things has made our surroundings intelligent for us,

yet it has also created us susceptible. Consider living in a home automation where everything is connected and all the gadgets are monitoring and observing you to assist, however all the information is on the internet that is not secure. Blockchain, as a decentralised and temper proof technology, is very appealing to the Internet of things (IoTs) sector. The quantity of nodes in IoT is growing by the day, as is the amount of data collected. Data security has been a concern; blockchain technology can help us secure and manage this data.

4.4 Blockchain and Voting

Blockchain has become a major topic in conversations about secure voting. Even though e-voting solves the majority of the challenges associated with conventional voting, issues such as voter confidentiality, illegal voting and the high cost of legacy electronic voting platforms continue to stay significant concerns. Blockchain, through smart contracts as well as encryption, could indeed make voting more safe, transparent, as well as private for voters.

4.5 Medical Data

Blockchain technology has truly inspired the therapeutic sector to confirm and implement restorative information collected from patients. Restorative data is critical, as well as any blunder or change can have disastrous consequences. Data can indeed be free and accessible for usages with Blockchain network without fear of modification.

5 Conclusion

With several practical uses for the advanced technologies already in place and being researched, blockchain is ultimately gaining recognition for itself, thanks in large part to bitcoin as well as cryptocurrency. There's no doubt that blockchain is a high priority as of late, regardless of the fact that there are just several concepts we need to see, a few concerns have just been enhanced especially with new system's creation on the application level, making it all the more establish and reliable. The legislature must create comparing legislation for this advancement, and efforts should be made to prepare for comprehend blockchain breakthroughs, preventing it from having a significant impact on the current foundation. At the same time as we acknowledge the benefits that blockchain rapid advancement deliver to us, we must be concerned about the incense as well as security advantages that it may possess.

References

1. Ansari MTJ, Khan NA (2021) Worldwide COVID-19 vaccines sentiment analysis through twitter content. *Electron J Gen Med* 18(6)
2. Ansari TJ, Pandey D (2017) An integration of threat modeling with attack pattern and misuse case for effective security requirement elicitation. *Int J Adv Res Comput Sci* 8(3)
3. Ansari MTJ, Agrawal A, Khan RA (2022) DURASec: durable security blueprints for web-applications empowering digital India initiative. *EAI Endorsed Trans Scalable Inf Syst* e25-e25
4. Ansari MTJ, Al-Zahrani FA, Pandey D, Agrawal A (2020) A fuzzy TOPSIS based analysis toward selection of effective security requirements engineering approach for trustworthy healthcare software development. *BMC Med Inform Decis Mak* 20(1):1–13
5. Ansari MTJ, Baz A, Alhakami H, Alhakami W, Kumar R, Khan RA (2021) P-STORE: extension of STORE methodology to elicit privacy requirements. *Arab J Sci Eng* 46(9):8287–8310
6. Aste T, Tasca P, Di Matteo T (2017) Blockchain technologies: the foreseeable impact on society and industry.
7. Bentov I, Gabizon A, Mizrahi A (2016) Cryptocurrencies without proof of work. In: *International conference on financial cryptography and data security*. Springer, Berlin pp 142–157
8. Courtois NT, Bahack L (2014) On subversive miner strategies and block withholding attack in bitcoin digital currency. [arXiv:1402.1718](https://arxiv.org/abs/1402.1718)
9. Eyal I, Sirer EG (2018) Majority is not enough: bitcoin mining is vulnerable. *Commun ACM* 61(7):95–102
10. Gervais A, Ritzdorf H, Karame GO, Capkun S (2015) Tampering with the delivery of blocks and transactions in bitcoin. In: *Proceedings of the 22nd ACM SIGSAC conference on computer and communications security*, pp 692–705
11. Global Blockchain Market (2021–2026) by component, provider, type, organization size, deployment, application, industry, geography, competitive analysis and the impact of COVID-19 with Ansoff analysis. *Research and Markets - Market Research Reports - Welcome*. <https://www.researchandmarkets.com/reports/5317225/global-blockchain-market-2021-2026-by>. Accessed 4 Sept 2022
12. Kaushik A, Choudhary A, Ektare C, Thomas D, Akram S (2017) Blockchain—literature survey. In: *2017 2nd IEEE international conference on recent trends in electronics, information & communication technology (RTEICT)*, pp 2145–2148. IEEE
13. Khatri S, Alzahrani FA, Ansari MTJ, Agrawal A, Kumar R, Khan RA (2021) A systematic analysis on blockchain integration with healthcare domain: scope and challenges. *IEEE Access* 9:84666–84687
14. King S, Nadal S (2012) Ppcoin: peer-to-peer crypto-currency with proof-of-stake. Self-published paper, August, 19(1)
15. Kshetri N, Voas J (2018) Blockchain in developing countries. *It Prof* 20(2):11–14
16. Tasatanattakool P, Techapanupreeda C (2018) Blockchain: challenges and applications. In: *2018 international conference on information networking (ICOIN)*. IEEE, pp 473–475
17. Zarour M, Ansari MTJ, Alenezi M, Sarkar AK, Faizan M, Agrawal A, Khan RA (2020) Evaluating the impact of blockchain models for secure and trustworthy electronic healthcare records. *IEEE Access* 8:157959–157973

Cryptanalysis of Lattice-Based Threshold Changeable Multi-secret Sharing Scheme



Ramakant Kumar and Sahadeo Padhye

1 Introduction

A secret sharing scheme (SSS) is a way of sharing a secret among some members. These members are known as participants or shareholders. Sharing is in such a way that secrets can be reconstructed if at least a specified number of participants join with their shares. In the (t, n) threshold SSS, the secret can be reconstructed only when at least t out of n participants provide their shares. Applications SSS are in key management [1], cloud computing [2], electronic voting [3], electronic cash [4], and secure multiparty computation [5]. In 1979, Adi Shamir [6] and Blakley [7] independently gave the concept of SSS. Shamir used polynomial interpolation technique, and Blakley used the geometry of hyperplanes, respectively, to propose their SSS. Shamir's (t, n) SSS relies on the fact that a $t - 1$ degree polynomial can be constructed if t points on it are known. Blakley used the idea that the point of intersection of different hyperplanes can be computed if a specific number of these hyperplanes are known. Both of these schemes are unconditionally secure.

The concept of verifiable secret SSS was introduced by Chor [8] to check the honest nature of participants and dealer. Some more verifiable SSS based on number theoretic hard problems are given in [9, 10]. SSS having multi-use and multi-stage nature is also proposed in [11–15]. In a multi-use SSS, the dealer can share a new secret among the participants by making some changes in shares and public values (without using any secure channel) so that a new secret can be computed by using these modified values. In a multi-stage SSS, the dealer can share many secrets among the participants by giving only one share to each participant. The secret reconstructed in any phase does not give any information about other secrets.

In 1994, Shor [16] gave a quantum algorithm that can be used to solve number theoretic problems such as factoring and discrete logarithm problems in polynomial time. So the SSS whose security relies on these hard problems will not be secure against a quantum computer. But, no efficient classical or quantum algorithm exists

These authors contributed equally to this work.

R. Kumar (✉) · S. Padhye

Department of Mathematics, MNNIT Allahabad, Prayagraj 211004, Uttar Pradesh, India
e-mail: ramakantkumar9758@gmail.com

to solve hard problems on lattices, i.e., learning with errors (LWE), shortest vector problem (SVP), closest vector problem (CVP), and short integer solution problem (SIS). Using these hard problems, some lattice-based SSSs have been proposed in [17–22].

In 2011, using LWE, Georgescu [17] proposed an (n, n) SSS. In 2012, Bansarkhani et al. used Ajtai’s one-way function [23] to propose the first lattice-based (n, n) verifiable SSS [18]. This scheme is secure under the hardness of n^c -approximate SVP. Using Babai lattice algorithm [24] in the secret recomputation phase, Khorasgani et al. [19] proposed a threshold SSS in 2014. Using lattice-based short integer solution problem, Dehkordi et al. [20] proposed a lightweight verifiable multi-secret sharing scheme in 2016. In 2017, Pilaram et al. [21] introduced a lattice-based verifiable, multi-stage, multi-use secret sharing scheme. This scheme is proven safe under the worst-case hardness of lattice problems. In 2019, Rajabi et al. used the generalized Compact Knapsack Function to propose a lattice-based verifiable SSS [22]. Gentry et al. [25] introduced a useful non-interactive, publicly verifiable SSS in 2021.

Due to some security reasons, there may be a need to rise the threshold of the scheme before reconstructing the secret. For changing the threshold from t to t' , shares initially shared by the dealer are changed in a fashion that by using new shares, t' participants can construct the secret but lower than t' participants cannot construct the secret. The first threshold changeable SSS was proposed by Martin et al. [26] in 1999. Some threshold changeable SSSs are also given in [27, 28]. Lattice-based threshold changeable SSSs are proposed in [29, 30]. In 2017, Pilaram and Eghlidos gave a technique [31] to change the threshold of their scheme [21]. In this paper, we provide an attack on their threshold changeable technique. We also modify their threshold changeable technique to overcome this attack. We use following notations in this article.

Notations: We use capital bold letters for matrices and small bold letters for vectors. We denote the set of integers by \mathbb{Z} , and \mathbb{Z}_q denotes the set of integers under modulo q . $\mathbb{Z}_q^{n \times m}$ is used to denote the set of matrices of order $n \times m$ with entries from \mathbb{Z}_q , and $\mathbf{0}$ is used to denote a zero matrix of appropriate order. \mathbf{A}^T denotes the transpose of the matrix \mathbf{A} .

2 Pilaram and Eghlidos Secret Sharing Scheme [21]

They proposed a (t, n) threshold multi-stage SSS based on lattices. Any reconstructed secret does not reveal any information about non-constructed secrets. This scheme is secure under the worst-case hardness of lattice problems. The description of this scheme is follows:

Let n and t denote the total number of participants and threshold of the scheme, respectively. Let $S_i \in \mathbb{Z}_q^t$, $i = 1, 2, 3, \dots, m$ be m secrets, and $q = O(n^c)$ for a constant c is a prime number. The dealer randomly chooses and publishes a vector $\mathbf{v} \in \mathbb{Z}_q^t$ such that its last coordinate is 1. For each secret S_i , the dealer computes secret lattice basis $\mathbf{B}_i \in \mathbb{Z}_q^{t \times t}$ s.t.

$$S_i = \mathbf{B}_i \mathbf{v}. \quad (1)$$

This equation has t^2 unknowns and t equations. The dealer chooses $\mathbf{B}'_i \in \mathbb{Z}_q^{t \times (t-1)}$ uniformly at random such that column vectors are LI and computes \mathbf{b}_i as given below:

$$S_i = \mathbf{B}_i \mathbf{v} \implies S_i = [\mathbf{B}'_i, \mathbf{b}_i][\mathbf{v}', 1]^T \implies \mathbf{b}_i = S_i - \mathbf{B}'_i \mathbf{v}', \quad (2)$$

where vector \mathbf{v}' has first $t - 1$ coordinates of the vector \mathbf{v} . Then the dealer selects n public vectors $\eta_j \in \mathbb{Z}_q^t$ for $j = 1, 2, 3, \dots, n$ s.t. any t of these vectors are LI. Let $k \geq \max\{t \log q, n\}$ be an integer. Then the dealer chooses n private vectors $\mathbf{c}_j \in \{0, 1\}^k$ randomly such that first n coordinates of these vectors form LI vectors in \mathbb{Z}_q^n . Then the dealer computes m public matrices $\mathbf{M}_i \in \mathbb{Z}_q^{t \times k}$ for $i = 1, 2, 3, \dots, m$ such that the equations $\mathbf{M}_i \mathbf{c}_j = \mathbf{B}_i \eta_j$ hold for $i = 1, 2, 3, \dots, m$ and $j = 1, 2, 3, \dots, n$. \mathbf{M}_i is computed corresponding to secret S_i for each $i = 1, 2, \dots, m$ as follows:

$$\mathbf{M}_i \mathbf{c}_j = \mathbf{B}_i \eta_j; j = 1, 2, 3, \dots, n.$$

We can write it as

$$\begin{aligned} \mathbf{M}_i [\mathbf{c}_1 \mathbf{c}_2 \dots \mathbf{c}_n] &= \mathbf{B}_i [\eta_1 \eta_2 \dots \eta_n] \implies \mathbf{M}_i \mathbf{C} = \mathbf{B}_i \boldsymbol{\eta} \\ \implies [\mathbf{M}'_i \mathbf{M}''_i][\mathbf{C}' \mathbf{C}'']^T &= \mathbf{B}_i \boldsymbol{\eta} \implies \mathbf{M}'_i = [\mathbf{B}_i \boldsymbol{\eta} - \mathbf{M}''_i \mathbf{C}''] \mathbf{C}'^{-1}, \end{aligned}$$

where \mathbf{C}' is invertible matrix as it has first n rows of \mathbf{C} , and \mathbf{C}'' has last $k - n$ rows of \mathbf{C} . \mathbf{M}'_i has first n column of \mathbf{M}_i and $\mathbf{M}''_i \in \mathbb{Z}_q^{t \times (k-n)}$ is chosen uniformly at random.

For verification purpose, the dealer chooses a matrix $\mathbf{E} \in \mathbb{Z}_q^{t \times k}$ at random and publishes it along with $\mathbf{h}_j = \mathbf{E} \mathbf{c}_j$ for $j = 1, 2, 3, \dots, n$. The dealer also chooses a public hash function H , computes and publishes $H(S_i)$ for $i = 1, 2, 3, \dots, m$.

2.1 Distribution Phase

The dealer sends these share vector \mathbf{c}_j to participant P_j securely, publishes matrices \mathbf{M}_i and vectors η_j for $i = 1, 2, 3, \dots, m$, and $j = 1, 2, 3, \dots, n$. The participant P_j , after receiving his share \mathbf{c}_j , finds $\mathbf{E} \mathbf{c}_j$ and compares it along with \mathbf{h}_j on the bulletin board.

2.2 Combination Phase

Without loss of generality, suppose participants $P_1, P_2, P_3, \dots, P_t$ want to reconstruct the secret S_i . The participant P_j sends pseudoshare $M_i c_j$ to the combiner. After receiving t pseudoshares $M_i c_j, j = 1, 2, 3, \dots, t$, the combiner computes the matrix B_i as

$$B_i = M_i [c_1 c_2 \dots c_t] [\eta_1 \eta_2 \dots \eta_t]^{-1}.$$

After computing B_i , combiner finds the secret S_i as $S_i = B_i v$.

3 Pilaram and Eghlidos Threshold Changeable Technique [31]

Consider the participants want to rise the threshold of the scheme from t to t' such that $t' > t$.

Phase 1: Dimension extension from t to t' : First they rise the size of parameters in the their scheme from t to t' in a way that the original equations still hold

$$S_{i t \times 1} = B_{i t \times t} v_{t \times 1} \implies \begin{bmatrix} S_{i t \times 1} \\ \mathbf{O}_{(t'-t) \times 1} \end{bmatrix}_{t' \times 1} = \begin{bmatrix} B_i & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{bmatrix} \begin{bmatrix} v \\ v'' \end{bmatrix}_{t' \times 1}, \quad (3)$$

$$M_{i t \times k} c_{j k \times 1} = B_{i t \times t} \eta_{j t \times 1} \implies \begin{bmatrix} M_i & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{bmatrix}_{t' \times k'} \begin{bmatrix} c_j \\ c_j'' \end{bmatrix}_{k' \times 1} = \begin{bmatrix} B_i & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{bmatrix}_{t' \times t'} \begin{bmatrix} \eta_j \\ \eta_j'' \end{bmatrix}_{t' \times 1}, \quad (4)$$

where v'' is chosen randomly from $\mathbb{Z}_q^{(t'-t) \times 1}$. For $j = 1, 2, 3, \dots, n$, η_j'' are chosen randomly from $\mathbb{Z}_q^{(t'-t) \times 1}$, and c_j'' are chosen randomly from $\{0, 1\}^{(k'-k) \times 1}$, where $k' \geq \max(t' \log q, n)$.

Phase 2: Sharing zero secret:

$$\mathbf{O}_{t' \times 1} = B''_{t' \times t'} \begin{bmatrix} v \\ v'' \end{bmatrix}_{t' \times 1},$$

$$M''_{t' \times k'} \begin{bmatrix} c_j \\ c_j'' \end{bmatrix}_{k' \times 1} = B''_{t' \times t'} \begin{bmatrix} \eta_j \\ \eta_j'' \end{bmatrix}_{t' \times 1}, \quad j = 1, \dots, n. \quad (5)$$

First part of Eq. 5 is used to find B'' . Then M'' is computed by using this value in 2nd part of Eq. 5.

Phase 3: Parameter Combination: By adding commensurate equations of (3) and (4), we get

$$S'_i = B'_i v', i = 1, 2, 3, \dots, m, \quad (6)$$

$$M'_i c'_j = B'_i \eta'_j, i = 1, 2, 3, \dots, m; j = 1, 2, 3, \dots, n, \quad (7)$$

where

$$S'_{i t' \times 1} = \begin{bmatrix} S_i \\ \mathbf{O}_{(t'-t) \times 1} \end{bmatrix}, i = 1, 2, 3, \dots, m,$$

$$B'_{i t' \times t'} = B'' + \begin{bmatrix} B_i & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{bmatrix}, i = 1, 2, 3, \dots, m,$$

$$v'_{t' \times 1} = \begin{bmatrix} v \\ v'' \end{bmatrix},$$

$$M'_{i t' \times k'} = M'' + \begin{bmatrix} M_i & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{bmatrix}, i = 1, 2, 3, \dots, m,$$

$$c'_{j k' \times 1} = \begin{bmatrix} c_j \\ c''_j \end{bmatrix}, j = 1, 2, 3, \dots, n.$$

$$\eta'_{j k' \times 1} = \begin{bmatrix} \eta_j \\ \eta''_j \end{bmatrix}, j = 1, 2, 3, \dots, n.$$

In this way threshold has changed from t to t' . Now $c'_j, j = 1, 2, 3, \dots, n$ are the modified shares. The vectors v' and $\eta'_j, j = 1, 2, 3, \dots, n$, and matrices $M_i, i = 1, 2, 3, \dots, m$ are the public informations.

4 Proposed Attack on Threshold Changeable Technique

An SSS with threshold t is secure if anyone having lower than t shares cannot get any knowledge about the secret. In the above threshold changeable technique, threshold is changed from t to t' such that $t' > t$. So after changing the threshold, anyone having less than t' shares should not be capable to get any knowledge of the secret. Here, in the modified shares c'_j , the first k entries are from the old shares. So any attacker having any t shares c'_j can get old shares c_j by just keeping the first k entries and discarding the remaining entries of c'_j . By using these $c_j, j = 1, 2, \dots, t$ values

and public values M_i , and η_j , the attacker can compute the private matrix B_i and consequently the secret S_i as follows:

$$B_i = M_i[c_1c_2\dots c_t][\eta_1\eta_2\dots\eta_t]^{-1},$$

$$S_i = B_iv.$$

In this way, any coalition of t'' participants or any attacker having t'' number of modified shares, $t \leq t'' < t'$, can construct the secret.

5 Proposed Modification in the Threshold Changeable Technique

Suppose Pilaram and Eghlidis SSS is used to share a secret $S \in \mathbb{Z}_q^t$ among n participants. Let $c_1, c_2, \dots, c_n \in \{0, 1\}^k$ are the shares of the participants. Let $M \in \mathbb{Z}_q^{t \times k}$, $v \in \mathbb{Z}_q^t$, and $\eta_j, j = 1, 2, 3, \dots, n$ are the public values such that

$$S = Bv$$

$$Mc_j = B\eta_j, j = 1, 2, 3, \dots, n.$$

For changing threshold from t to t' , first we use threshold changeable technique discussed in Sect. 3. Then each participant P_j finds a vector $X^j_{(t'+1) \times 1} \in \mathbb{Z}_q^{t'+1}$ such that

$$M' \begin{bmatrix} X^j_{(t'+1) \times 1} \\ \mathbf{O}_{(k'-t'-1) \times 1} \end{bmatrix} = \mathbf{O}.$$

Breaking matrix M' as

$$\begin{aligned} & \left[M'_{1t' \times (t'+1)} M'_{2t' \times (k'-t'-1)} \right] \begin{bmatrix} X^j_{(t'+1) \times 1} \\ \mathbf{O}_{(k'-t'-1) \times 1} \end{bmatrix} = \mathbf{O} \\ \implies & M'_1 X^j = \mathbf{O}. \end{aligned} \quad (8)$$

There are t' equations and $t' + 1$ variables. Choose one variable randomly from \mathbb{Z}_q and then solve for the remaining unknowns. Participant P_j deletes his earlier share after modifying the share as

$$c'_j = \begin{bmatrix} c_j \\ c''_j \end{bmatrix} + \begin{bmatrix} X^j \\ \mathbf{O} \end{bmatrix}, j = 1, 2, 3, \dots, n. \quad (9)$$

In this way, the threshold of the scheme has changed from t to t' successfully. This threshold changeable technique can be used to rise the threshold of the SSS if needed.

Now we give a toy example to understand the modified threshold changeable technique.

Example: Let there are $n = 3$ participants and threshold of the scheme be $t = 2$. Let $q = 7$ (i.e., all the operations are done under modulo 7) and $k = 6$. Let $S = (2, 5)^T$ be the secret which is shared among the participants. $B = \begin{bmatrix} 3 & 2 \\ 0 & 4 \end{bmatrix}$

be the dealer's private matrix. $c_1 = (1, 0, 1, 0, 0, 1)^T$, $c_2 = (0, 1, 1, 0, 1, 1)^T$, and $c_3 = (0, 0, 1, 1, 0, 1)^T$ are the shares of the participants. $v = (1, 3)^T$, $\eta_1 = (1, 3)^T$, $\eta_2 = (2, 4)^T$, $\eta_3 = (3, 5)^T$, and $M = \begin{bmatrix} 1 & 4 & 5 & 4 & 2 & 3 \\ 2 & 1 & 2 & 3 & 5 & 1 \end{bmatrix}$ are the public informations.

Suppose due to some security reasons, participants want to rise the threshold of the SSS from 2 to 3. $v'' = 1$, $\eta_1'' = 2$, $\eta_2'' = 6$, $\eta_3'' = 1$ are chosen uniformly at random from \mathbb{Z}_7 , modified value of k is $k' = 9$. $c_1'' = (0, 1, 0)^T$, $c_2'' = (1, 0, 1)^T$ and $c_3'' = (1, 1, 0)^T$ are chosen uniformly at random from \mathbb{Z}_2^3 . Then the dealer solves

first part of Eq. 5 and get private matrix $B'' = \begin{bmatrix} 5 & 0 & 2 \\ 1 & 3 & 4 \\ 2 & 1 & 2 \end{bmatrix}$. Then the dealer uses

this value of B'' in second part of Eq. 5 and get $M'' = \begin{bmatrix} 1 & 0 & 0 & 1 & 4 & 2 & 1 & 6 & 1 \\ 0 & 6 & 5 & 2 & 3 & 1 & 2 & 5 & 0 \\ 0 & 0 & 6 & 0 & 1 & 0 & 4 & 3 & 2 \end{bmatrix}$. So

$$M' = M'' + \begin{bmatrix} M_i & O \\ O & O \end{bmatrix} = \begin{bmatrix} 2 & 4 & 5 & 5 & 6 & 5 & 1 & 6 & 1 \\ 2 & 0 & 0 & 5 & 1 & 2 & 2 & 5 & 0 \\ 0 & 0 & 6 & 0 & 1 & 0 & 4 & 3 & 2 \end{bmatrix}.$$

Then each participant P_j finds a vector $X^j_{4 \times 1} \in \mathbb{Z}_q^4$, $j = 1, 2, 3$, such that Eq. 8 holds. Without loss of generality suppose each participant get $X^j = (2, 0, 0, 2)^T$. Finally, the modified shares of the participants are $c'_1 = (3, 0, 1, 2, 0, 1, 0, 1, 0)^T$, $c'_2 = (2, 1, 1, 2, 1, 1, 1, 0, 1)^T$, $c'_3 = (2, 0, 1, 3, 0, 1, 1, 1, 0)^T$. These modified shares correspond to threshold value 3.

6 Conclusion

In this paper, we found the weakness of the threshold changeable technique proposed by Pilaram and Eghlidos. We showed that any coalition of t'' participants or any attacker having t'' number of modified shares, $t \leq t'' < t'$ can construct the secret, i.e., threshold is still t . We also modified the threshold changing technique to overcome this attack. The modified technique can be used to rise the threshold of the scheme whenever required.

References

1. Chunying W, Shundong L, Yiyang Z (2013) Key management scheme based on secret sharing for wireless sensor network. In: Fourth international conference on emerging intelligent data and web technologies (EIDWT), pp 574–578
2. Attasena V, Harbi N, Darmont J et al (2013) Sharing-based privacy and availability of cloud data warehouse, 9themes journées francophones sur les Entrepôts de Données et l'Analyse en ligne (EDA 2013), pp 17–32
3. Schoenmakers B (1999) A simple publicly verifiable secret sharing scheme and its application to electronic voting. In: Advances in Cryptology CRYPTO99. Springer, pp 148–164
4. Stadler M (1996) Publicly verifiable secret sharing. In: Maurer U (ed) Advances in cryptology EUROCRYPT 96. Lecture notes in computer science, vol 1070. Springer, Berlin, pp 190–199
5. Cramer R, Damgård I, Maurer U (2000) General secure multi-party computation from any linear secret-sharing scheme. In: Preneel B (ed) EUROCRYPT 2000. LNCS, vol 1807, pp 316–334
6. Shamir A (1979) How to share a secret. Commun ACM 22(11):612–613
7. Blakley GR (1979) Safeguarding cryptographic keys. In: Proceedings of the 1979 AFIPS national computer conference, vol 48, pp 313–317
8. Chor B, Goldwasser S, Micali S, Awerbuch B (1985) Verifiable secret sharing and achieving simultaneity in the presence of faults. In: SFS '85': proceeding of 26th annual symposium on foundations of computer science, pp 383–395
9. Benaloh Jc (1987) Secret sharing homomorphisms: keeping shares of a secret secret. In: Proceedings on advances in cryptology- CRYPTO86, pp 251–260
10. Feldman P (1987) A practical scheme for non-interactive verifiable secret sharing. In: SFCS '87: proceedings of the 28th annual symposium on foundations of computer science, pp 427–438
11. Blundo C, De Santis A, Di Crescenzo G, Gaggia AG, Vaccaro U (1994) Multi-secret sharing schemes. In: Advances in cryptology CRYPTO94. Springer, pp 150–163
12. He J, Dawson E (1994) Multistage secret sharing based on one-way function. Electron Lett 30(19):1591–1592
13. Chang TY, Hwang MS, Yang WP (2005) A new multi-stage secret sharing scheme using one-way function. SIGOPS Oper Syst Rev 39(1):48–55
14. Das A, Adhikari A (2010) An efficient multi-use multi-secret sharing scheme based on hash function. Appl Math Lett 23(9):993–996
15. Chang TY, Hwang MS, Yang WP (2011) An improved multi-stage secret sharing scheme based on the factorization problem. Inf Technol Control 40(3):246–251
16. Shor PW (1994) Algorithms for quantum computation: discrete logarithms and factoring. In: Proceedings of the 35th annual symposium on foundations of computer science. Series. SFCS 94. Washington, DC, USA: IEEE Computer Society, pp 124–134
17. Georgescu A (2011) A LWE-based secret sharing scheme. IJCA Spec Issue Netw Secur Cryptogr NSC(3):27–29
18. El Bansarkhani R, Meziani M (2012) An efficient lattice-based secret sharing construction. In: Askoxylakis I, Phls H, Posegga J (eds) Information security theory and practice. Security, privacy and trust in computing systems and ambient intelligent ecosystems. Series. Lecture notes in computer science, vol 7322. Springer, Berlin, pp 160–168
19. Khorasgani HA, Asaad S, Eghlidis T, Aref M (2014) A lattice-based threshold secret sharing scheme. In: 2014 11th international ISC conference on information security and cryptology, pp 173–179. <https://doi.org/10.1109/ISCISC.2014.6994043>
20. Dehkordi MH, Ghasemi R (2016) A lightweight public verifiable multi secret sharing scheme using short integer solution. Wirel Pers Commun 91(3):1459–1469
21. Piliaram H, Eghlidis T (2017) An efficient lattice based multi-stage secret sharing scheme. IEEE Trans Dependable Secur Comput 14(1):2–8. <https://doi.org/10.1109/TDSC.2015.2432800>
22. Rajabi B, Eslami Z (2019) A verifiable threshold secret sharing scheme based on lattices. Inf Sci 501:655–661

23. Ajtai M (1996) Generating hard instances of lattice problems (extended abstract). In: Proceedings of the twenty-eighth annual ACM symposium on theory of computing. Series. STOC 96. ACM, New York, NY, USA, pp 99–108
24. Babai L (1986) On Lovasz lattice reduction and the nearest lattice point problem. *Combinatorica* 6(1):1–13
25. Gentry C, Halevi S, Lyubashevsky V Practical non-interactive publicly verifiable secret sharing with thousands of parties, *Cryptology ePrint Archive: Report 2021/1397*
26. Martin KM, Safavi-Naini R, Wang H (1999) Bounds and techniques for efficient redistribution of secret shares to new access structures. *Comput J* 42(8):638–649
27. Zhang Z, Chee YM, Ling S, Liu M, Wang H (2012) Threshold changeable secret sharing schemes revisited. *Theor Comput Sci* 418:106–115
28. Lou T, Tartary C (2008) Analysis and design of multiple threshold changeable secret sharing schemes. In: Franklin M, Hui L, Wong D (eds), *Cryptology and network security, LNCS*, vol 5339, pp 196–213
29. Steinfeld R, Pieprzyk J, Wang H (2007) Latticebased threshold changeability for standard Shamir secret-sharing schemes. *IEEE Trans Inf Theory* 53(7):2542–2559
30. Steinfeld R, Pieprzyk J, Wang H (2006) Lattice based threshold-changeability for standard CRT secret-sharing schemes. *Finite Fields Their Appl* 12(4):653–680
31. Pilaram H, Eghlidos T (2017) A lattice-based changeable threshold multi-secret sharing scheme and its application to threshold cryptography. *Sci Iran* 24:1448–1457

Amazon Web Service IOT and Authentication of Edge Devices



Meenakshi Srivastava and Arsh

1 Introduction

IOT is delivering a compelling technology innovation. Connecting variety of devices with each other and allowing them to participate in data sharing results in new products and services. IOT devices store this collected data in the cloud where various AI applications are used to perform analytics. AWS helps in storing a vast amount of data collected by the IOT devices continuously in real-time. But the optimum potential of these smart devices using AI-enabled applications to perform analysis on this data and gain considerable insights such that data provide significant information, but this can solely be attained by using effective security mechanisms. AWS uses a manifold layers to provide deterrent security methods like access control to data and encrypting the data. Amazon web services provide efficient security services to keep the collected data secure.

2 Authentication

Authentication is one of the crucial security parts of the IOT devices. AWS IoT services provide tools for creating secure key authentication. Various credentials like private keys and digital certificates are involved in authenticating edge devices with the AWS platform. Users should not be allowed to set up credentials of services to authenticate their devices with AWS. Also, the credentials should not be revealed to manufacturer of the devices and any entity part of the development process like personnel and delivery pipeline. These devices would become vulnerable if the credentials are disclosed [1].

M. Srivastava (✉) · Arsh
Amity Institute of Information Technology, Amity University Uttar Pradesh, Noida, India
e-mail: msrivastava@lko.amity.edu

It would lead to data leak and modification without authorization. This data could include personal information that is collected from offices, homes. Improper authentication would lead to counterfeiting and impersonation [1].

The major issue involved is how the device will securely validate the authenticity of the AWS endpoint it connects to.

The certificate chain also needs to be created securely and the supplying of private key to the edge device. The device's private key is main or crucial part of the security paradigm. As it will be used to conduct verification during connecting to AWS IoT. In spite of being highly secure, AWS still cannot maintain integrity unless the device is not protected appropriately. Also it is important to ensure that all the keys, i.e. public, private, and processes are kept separate from device operations [1].

AWS uses microchip technology's safe component ATECC608a to provide device-level security and effective authentication mechanism. The component provides various hardware-level security methods and techniques like storing all the keys and certificates in a safe environment and removing vulnerabilities. This component is used to generate a private key that acts as a root for certificate used for securing networked devices. Basically private key is a unique identifier. While authenticating with the AWS IoT platform, the edge devices use it for communicating data to the platform [4, 6].

The microchip technology's component that is embedded in the device is built in a way that encapsulates the private key such that it becomes impossible to manipulate the private key [4].

The certificates, private key, and the public key infrastructure (PKI) are available through various means. It depends on the user's choice and business requirement as to which options should be considerably better as per the requirement. Either the complete infrastructure of the component-producing firm can be incorporated or AWS native security infrastructure for authentication can be used.

Also, third party and custom CA's can be used as per the context in which it is to be used. And even the authentication model can be fully customized to cope with specific security risk [4].

AWS IoT security model includes authentication mechanisms to authenticate devices with the AWS IoT platform and authorizing them to perform required actions.

Consider using edge device for booking cab online rather than using a cab-booking app on smartphone and providing various details in order to confirm a ride. An edge device like Amazon IoT button can be used to rapidly book rides on just a click of a button. These buttons are programmable and can be configured in the cloud. They can be used for various purposes like alerting or calling someone, ordering services online, order a cab, open gates remotely and check-in & check-out in hotels, etc [5].

When the button is clicked the information about who clicked it and at what location is necessary to book a ride, as it is done in smartphone applications using mobile numbers, usernames to uniquely identify a customer. Similarly, AWS IoT button needs a unique identifier to confirm authenticity. Devices using MQTT network protocol will use X.509 certificate [5].

AWS IoT uses public key cryptography or asymmetric cryptography for creating unique identity and signing documents. The digital signature to a message is used

to verify whether the original message has been received and no modifications or tampering took place. It will also be used to demonstrate possession of the private key. The X.509 certificate exhibits possession of a public key. Each edge device will have a unique identity, which will be done using separate X.509 certificate for each device. Three major options available for generating a certificate devices are, first a certificate generated by the AWS IoT can be used, which includes a public and private key and certificate signed by the AWS IoT Certificate Authority. Second option available for generating certificate is to use one's own Certificate signing authority, in this scenario, the AWS will not be aware about the private key. Last option is to use certificate of Certificate authority, which you find reliable and trustworthy. Root certificate, which is used by the AWS IOT, is also needed to set up a secure and authentic connection with the AWS IoT platform. All the keys and certificates generated earlier need to be stored on the cab booking button or edge device to be used during authentication procedure.

3 Device Security

The major issue involved is how the device will securely validate the authenticity of the AWS endpoint it connects to.

The certificate chain also needs to be created securely and the supplying of private key to the edge device. The device's private key is main or crucial part of the security paradigm. As it will be used to conduct verification during connecting to AWS IoT. In spite of being highly secure, AWS still cannot maintain integrity unless the device is not protected appropriately. Also, it is important to ensure that all the keys, i.e. public, private, and processes are kept separate from device operations [1].

4 Microchip's Trust Platform

AWS uses microchip technology's safe component ATECC608a to provide device-level security and effective authentication mechanism. The component provides various hardware-level security methods and techniques like storing all the keys and certificates in a safe environment and removing vulnerabilities. This component is used to generate a private key that acts as a root for certificate used for securing networked devices. Basically private key is a unique identifier. While authenticating with the AWS IoT platform, the edge devices use it for communicating data to the platform [4, 6].

The microchip technology's component that is embedded in the device is built in a way that encapsulates the private key such that it becomes impossible to manipulate the private key [4].

The certificates, private key, and the public key infrastructure (PKI) are available through various means. It depends on the user's choice and business requirement as

to which options should be considerably better as per the requirement. Either the complete infrastructure of the component producing firm can be incorporated, or AWS native security infrastructure for authentication can be used.

Also third party and custom CA's can be used as per the context in which it is to be used. And even the authentication model can be fully customized to cope with specific security risk [4].

5 AWS IoT Security Model

AWS IoT security model includes authentication mechanisms to authenticate devices with the AWS IoT platform and authorizing them to perform required actions.

Consider using edge device for booking cab online rather than using a cab-booking app on smartphone and providing various details in order to confirm a ride. An edge device like Amazon IoT button can be used to rapidly book rides on just a click of a button. These buttons are programmable and can be configured in the cloud. They can be used for various purposes like alerting or calling someone, ordering services online, order a cab, open gates remotely and check-in & check-out in hotels, etc. [5].

When the button is clicked the information about who clicked it and at what location is necessary to book a ride, as it is done in smartphone applications using mobile numbers, usernames to uniquely identify a customer. Similarly AWS IoT button needs a unique identifier to confirm authenticity. Devices using MQTT network protocol will use X.509 certificate [5].

AWS IoT uses public key cryptography or asymmetric cryptography for creating unique identity and signing documents. The digital signature to a message is used to verify whether the original message has been received and no modifications or tampering took place. It will also be used to demonstrate possession of the private key. The X.509 certificate exhibits possession of a public key. Each edge device will have a unique identity, which will be done using separate X.509 certificate for each device. There are three major options available for generating a certificate devices. First one is a certificate generated by the AWS IoT which includes a public and private key and certificate signed by the AWS IoT Certificate Authority. Second option available for generating certificate is to use one's own Certificate signing authority, in this scenario, the AWS will not be aware about the private key. Last option is to use certificate of Certificate authority, which you find reliable and trustworthy. Root certificate, which is used by the AWS IOT, is also needed to set up a secure and authentic connection with the AWS IoT platform. All the keys and certificates generated earlier need to be stored on the cab booking button or edge device to be used during authentication procedure.

6 Authentication to AWS IoT

When all the keys and certificates are installed on the edge device, i.e. AWS IoT button then it is ready to create a secure connection to AWS IoT platform and communicate with it. The Transport Layer Security protocol (TLS) that is used is to create secure connections while using banking services and other payment gateways are used here to set up a connection between AWS IoT platform and the AWS IoT button and the digital certificate stored at the edge device is used to prove its unique identity and demonstrate its authenticity [5].

The AWS IoT button starts by sending a **HELLO** message to the component of AWS IoT responsible for verifying the identity of edge device [3] (Fig. 1).

The TLS handshake is used to specify the format of the messages and the order in which the messages will be transmitted. Both the parties will comply on a shared secret that will be used for encryption of the transmitted messages. Hello message contains information about, which algorithms from the cipher suite the client, i.e. AWS IoT device will be able to use. The server then uses the message received to decide which cryptographic technique will be used to set up a transmission channel and returns the server certificate along with the cryptographic info [3, 8] (Fig. 2).

The server certificate will be used to authenticate the AWS IoT platform by comparing the public key of the AWS IoT root certificate stored on the edge device with the digital signature of the certificate received from server [3] (Fig. 3).

After verifying that the AWS IoT platform the edge device needs to verify its identity with the AWS IoT and further needs to agree on a shared secret. The edge

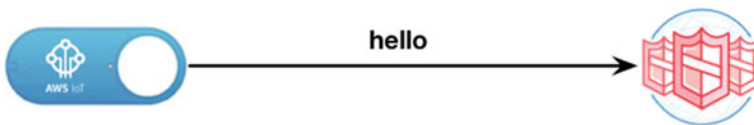


Fig. 1 Sending HELLO to AWS IoT platform [3]

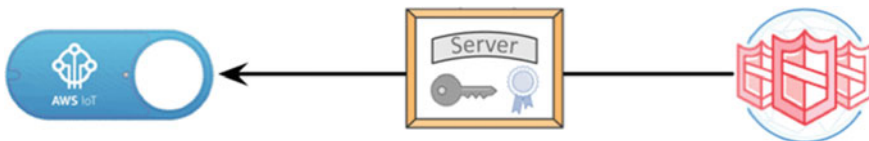


Fig. 2 Server certificate received [3]

Fig. 3 Server certificate and root certificate verification [3]

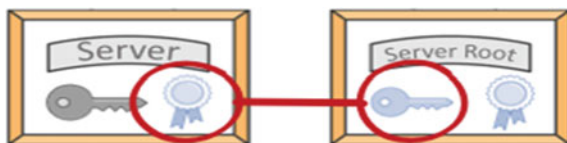




Fig. 4 Device certificate sent to AWS IoT server [3]

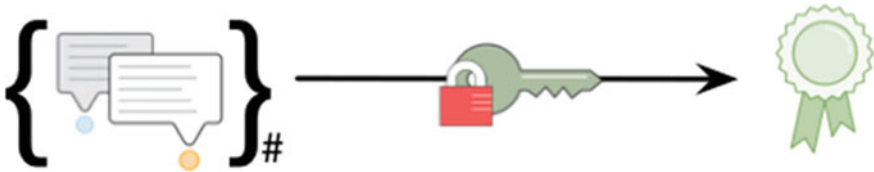


Fig. 5 Computing hash over messages [3]



Fig. 6 Digital signature sent to AWS IoT server [3]

device sends a device certificate to the AWS IoT server for authentication [1, 3] (Fig. 4).

The button then computes a hash upon all the messages transmitted currently to the AWS IoT platform. Then the device using its private key computes a digital signature for this hash [3] (Fig. 5).

The computed digital signature by the AWS IoT button is transmitted to the AWS IoT platform [3] (Fig. 6).

The AWS IoT platform now has all the required parameters to check and authenticate the edge device, AWS IoT button in this case. It now has the digital signature and public key of the edge device, which it received in the device certificate, the messages have also been logged at the server side and a hash for the same is also, calculated which should be same as the hash computed on the edge device. The authenticity of the digital certificate is verified against the device’s public key [3] (Fig. 7).

After all the successful authentication procedures, the AWS IoT platform is now certain of the edge device it is communicating with, cab booking device of a particular passenger in this case. The unique id of the edge device tells the AWS IoT as to which passenger is trying to set a connection. The shared secret to be used is dependent on the cipher suites algorithm that the AWS IoT and the edge device uses, i.e. key exchange algorithm decided during Transport Layer Security (TLS) handshake. The AWS IoT button uses the server’s public key to encrypt the message, which was obtained from the server’s certificate. The server uses its private key to decrypt the

Fig. 7 Digital signature verified against device's public key [3]



message received from the AWS IoT button. After this procedure, both the parties are now able to set up a shared secret and further communication will be secured using this agreed upon shared secret [8].

7 Permission to Book

After all the authentication has been performed, i.e. the AWS IoT platform has been verified and AWS IoT button has authenticated its unique identity using X.509 certificate to the server and secure messages are being transferred, shared secret has been set up. The button is ready for booking a cab. The MQTT message, which the button or edge device sends, looks like “iotdashbutton/G01XXXXXXXXXX”. The second half is the serial number corresponding to the edge device. The security of the AWS IoT button should be in a manner such that minimal authorization is provided to the passengers. Otherwise, there can be a scenario where an atrocious customer can modify or alter the program and impersonate to manipulate the system and book a cab inappropriately. A policy needs to be associated with the device certificate, which will authorize that specific identity. The serial number was contained in the certificate, which was transmitted while authenticating the device [3].

8 AWS IoT Security

AWS IoT includes MQTT that uses less network bandwidth, a less heavy communication protocol and can easily work with irregular connections, uses less code resulting in efficient performance. Also supports web sockets and HTTPS protocols. The mechanism discussed above ensures that no data transmission between AWS IoT button and AWS IoT platform happens without authentication [1]. The encryption and authentication techniques ensure that no unauthorized or fake access attempt can be performed. Also, various levels of permissions are available to be used in

our devices as it may vary depending upon the environment where it is used and the purpose for which it is being used. And various policies are put into place such that social engineering attacks can be ceased. In addition, the authentication techniques can be modified or updated from the AWS IoT platform. Device shadows are used to extract the current state of the devices to ensure the security even when the device is not connected. Intel hardware security ensures that in contrast to software security hardware security is also well maintained and uses secure boot and trusted execution.

After the successful authentication and authorization the channel is secured with mutual authentication and encryption. The message from edge device is transferred to the device gateway that is responsible for communicating with devices via MQTT, HTTP and web sockets, and AWS IoT platform [1].

AWS provides a range of IoT services designed to help customers enhance the security of their devices, networks, and data. These services empower users to implement comprehensive security measures, spanning from safeguarding devices to securing data both during transit and at rest. Additionally, AWS offers a suite of security features that facilitate the implementation of necessary security protocols to meet specific security requirements.

9 Security Components

Components that together help security and authenticating edge devices include Amazon FreeRTOS, AWS IoT Greengrass, AWS IoT Core, AWS IoT device management, AWS IoT Device Defender [2].

AWS IoT core gives secure transmission channels using the most effective Transport Layer Security.

AWS IoT Device Defender administration constantly reviews IoT arrangements to guarantee that setups aren't going amiss from security best practices to keep up and uphold IoT designs—for example, guaranteeing gadget character, confirming and approving gadgets, and scrambling gadget information. The administration Securing IoT with AWS 8 can send an alarm if there are any holes in a client's IoT setup that may make a security hazard, for example, character authentications being shared over different gadgets or a gadget with a denied personality declaration attempting to associate with AWS IoT Core [2].

Amazon FreeRTOS is responsible for encrypting data and management of keys. It uses its libraries to perform these tasks, which leads to secure connections and promotes data security. Secure connection is established using the Transport Layer Security (TLS), which is included in the Amazon FreeRTOS. This OS also has the capability to ensure code at the edge device is not altered and also includes features to include to update the devices remotely and providing security fixes [2].

AWS IoT GreenGrass is responsible for authenticating the edge device and encrypting data to be transmitted based on the decided shared secret. It does not allow data sharing prior to the procedure of identity validation. Mutual device authentication and authorization is performed through this component.

The AWS IoT device requires a device certificate, policy to connect to the Greengrass service. It also gives equipment foundation of trust private key storing for the AWS IoT devices. For IoT Greengrass in AWS, all AWS IoT edge devices should enable full disk encryption like, AES 256-bit keys based on NIST FIPS 142 verified algorithms and pursue main management leading methods [2].

10 Security Challenges

Security dangers and vulnerabilities can possibly bargain the safety and protection of client information in an IoT app. Combined with the developing number of gadgets, and the information produced, the capability of mischief brings up issues that how can we cope-up with security dangers presented by IOT gadgets and gadget correspondence to the cloud and from it itself.

Regular client concerns with respect to dangers focus on the safety and encryption performed on information while it is traveling to and from the cloud, or it is traveling from the edge administrations to the gadget and from it, alongside fixing of gadgets, gadget furthermore, client verification, and access control. Making sure about IoT gadgets is fundamental, not exclusively to keep up information uprightness, in any case, to likewise ensure against assaults that can affect the unwavering quality of gadgets. As gadgets can send enormous sums of delicate information through the networks and clients at the other end are enabled to straightforwardly manage a gadget [3].

To stay aware of the passage of gadgets into the commercial center just as the dangers coming on the web, it is ideal to execute administrations that address every piece of IOT environment and cover in its ability to make sure about and ensure, review and also remediate, or oversee armada organizations of IOT gadgets [6].

11 AWS IoT Greengrass

AWS IoT's Greengrass program allows clients to run neighborhood process, informing, information reserving, match up, and ML derivation capacities for associated gadgets. It validates and encodes gadget information for nearby and cloud interchanges, and information is not traded among gadgets and the cloud in the absence of demonstrated character. The administration utilizes security and accesses the board like the clients know about in AWS IoT's Core, with common gadget confirmation and approval, and secure network with the cloud [7].

AWS IOT approaches, and AWS Identity's and also Access Management approaches guarantee that AWS IoT's Greengrass use is very safe. AWS IoT gadgets need an AWS IOT object, a gadget declaration, and an AWS IoT arrangement to interface with the AWS IoT's Greengrass administration. This permits AWS IoT's

Greengrass center gadgets to safely interface with the AWS IoT cloud administration. It additionally permits the AWS IoT's Greengrass cloud administration to send design data, AWS Lambda works, and oversee memberships to AWS IoT's Greengrass center gadgets. What's more, AWS IoT's Greengrass gives equipment foundation of trust private key stockpiling for edge gadgets [8].

12 AWS IoT Device Defender

This fully managed service from AWS helps customers assess the security measures established for their fleet of IoT devices. The service conducts ongoing evaluations of IoT configurations to ensure that settings align with the highest security standards for maintaining and enforcing IoT architectures. This includes tasks such as verifying device identities, authenticating and validating devices, and encrypting device data.

The administration is able to send an alarm if there are any holes in a client's IoT design that may make a safety hazard, for example, character endorsements being shared over numerous gadgets or a gadget with a renounced personality authentication attempting to interface with AWS IoT Core[4].

With the administration's checking and inspecting abilities, clients can set cautions that make a move to rectify any deflection that is found in gadgets. Like instance, spikes in rush hour gridlock may demonstrate that a gadget is taking an interest in a conveyed disavowal of administration (DDoS) assault. AWS IoT's Greengrass and Amazon's RTOS additionally naturally incorporate with AWS Device Defender to give safety measurements of the gadgets for assessment. AWS Device Defender can send alarms to Amazon CloudWatch, AWS IoT's, and Amazon's Notification Administration, which makes distributing aware of Amazon CloudWatch measurements. In the event that a client chooses to address a caution, AWS IoT's Devices Management may be utilized to undertake alleviating activities, for example, making safety fixes [5].

AWS IoT's Device Defender reviews IOT setups related to client gadgets against a lot of characterized IoT effective safety measures so clients will watch where the safety holes are and control reviews on persistent and on the other hand specially appointed premise. There are likewise security rehearses inside AWS IoT's Device Defender that may be chosen and run as a major aspect about the review. The administration likewise coordinates by different AWS administrations—, for example, Amazon Cloud Watch and Amazon's SNS—for sending security cautions to AWS's IoT whenever a review comes up short or when conduct peculiarities are identified so clients can explore and decide the main driver. For instance, AWS IoT's Device Defender can caution clients when gadget characters are getting to delicate APIs. AWS IoT's Device Defender also can likewise suggest activities that limit the effect of safety issues, for example, renouncing authorizations, rebooting a gadget, resetting manufacturing plant default, or pushing bugs in security fixes to any client's associated gadgets [9].

Clients may likewise be worried about terrible on-screen characters; human or fundamental mistakes and approved clients with vindictive goals can present designs along negative safety impacts. AWS IoT's Core gives the safety building hinders for clients to safely associate gadgets to cloud and also different gadgets. The structure squares permit upholding security controls, for example, validation, approval, review logging, and start to finish encryption. At that point, AWS IoT's Device Defender comes in and serves to consistently review safety setups for consistence with eminent security practices and also clients' own authoritative security-related strategies.

13 AWS IoT Device Management

AWS IoT's Device Management helps clients install, arrange, screen, and distant oversee IoT gadgets at scale. AWS IoT's Device Management coordinates with AWS IoT's Core to effortlessly associate gadgets with the cloud and different gadgets such that clients can distantly deal with their armadas about gadgets. AWS IoT's Gadget Management helps clients locally available modern gadgets by utilizing AWS IoT thing inside the AWS's Management Support or like API to transfer formats that are populated with data such as gadget maker and sequential no. X509 character endorsements, or security arrangements. After this, clients would then be able to arrange the whole armada of gadgets with this data with a couple of snaps in AWS's IoT inside the AWS's Management Console [10].

Through this functionality, customers can organize their device fleet into a hierarchical structure based on factors such as function, security needs, or similar categories. They have the flexibility to group a single device within a room, multiple devices on the same floor, or all devices functioning within a building. These groupings can then be leveraged to control access policies, monitor operational metrics, or execute actions across the entire cluster. Furthermore, an innovative feature referred to as "Dynamic Things" automatically adds devices that adhere to user-defined criteria and removes devices that don't meet these requirements. This process ensures a secure and streamlined approach while maintaining operational integrity. The Dynamic Things feature also simplifies the process of locating device reports based on any combination of device attributes and allows users to perform bulk updates effortlessly.

The clients can likewise push programming and firmware for gadgets in field as to fix security weaknesses and improve gadget usefulness; implement mass updates, also control organization speed; set disappointment edges; and characterize persistent occupations to refresh gadget programming naturally with the goal that they are continually running the most recent variant of programming. Clients can remotely send activities, like gadget restarts or plant retunes, as to fix programming problems in the gadget or reestablish the gadget to its unique settings. Clients can likewise carefully sign documents that are sent to their gadgets, assisting with guaranteeing the gadgets are not bargained.

The capacity to push programming refreshes isn't constrained to the cloud administrations. Indeed, OTA updates occupation in Amazon's FreeRTOS permit clients to utilize AWS IoT's Device Management to plan programming refreshes. Also, clients can likewise make an AWS IoT's Greengrass center update work for at least one AWS IoT's Greengrass center gadgets utilizing AWS IoT's Device Management so as to convey security refreshes, error fixes, and new AWS IoT's Greengrass highlights to associated gadgets.

14 Enhance IOT Using Provable Security

Latest security administrations and advances are also being worked on AWS to assist ventures with making sure about their IoT and edge gadgets. Specifically, AWS has as of late propelled checks inside AWS IoT's Device Defender, fueled by an Artificial intelligence innovation identified as mechanized thinking, which use numerical evidences to check programming is composed effectively and decide whether there is unintentional access to the gadgets. The AWS IoT's Device Defender is a model of way clients can straightforwardly utilize computerized thinking to make sure about their own gadgets. Inside, AWS has utilized robotized thinking to check the memory respectability of program running on Amazon's FreeRTOS and to secure against malwares. Interest in robotized thinking to give versatile confirmation of safe programming, alluded to as proved security, permits clients to work touchy outstanding tasks at hand on AWS [6].

15 Security in Future Systems

All-encompassing security capacities covering the entire lifecycle of an IoT framework, and its segments are required for future IoT frameworks. Improvement of new danger examination and hazard the executives also as self-mending capacities to distinguish and vanquish potential assaults are required. Gathering, coordinating, and handling heterogeneous information from various sensors, gadgets, and frameworks will need new united personality and access the executive's arrangements. Future IoT frameworks ought to be ready to rapidly and suitably react to dangers and assaults, fuse and gain from new danger data, and create and sanction string relief plans. The ability to agreeably analyze issues and execute security plans for different subsystems in the framework, which might be claimed by various elements, is additionally required [7].

Future IoT frameworks ought to likewise have the option to guarantee controllable information possession across big business limits. To protect the security of clients as well as endeavors while handling an enormous measure of information, new information investigation calculations and new cryptographic strategies, for

example, homomorphic or accessible encryption, are required. Sharing risk knowledge data by various frameworks empowers helpful safety efforts that are equipped for acknowledging increasingly firm information on the present and future assaults.

16 AWS IoT Platform

Devices are required to possess credentials in order to access the message broker, and all data transmission must be safeguarded through Transport Layer Security (TLS). The platform offers support for identity principals, such as X.509 certificates commonly employed by AWS IoT devices, as well as Identity Access Management (IAM) users, groups, and roles. Federated identities, used by web and desktop applications, are also encompassed, along with Amazon Cognito identities, frequently used by mobile applications, which enable the integration of other identity providers.

Tls

X.509 Certificate

Aws iam

Federated Identities

Amazon Cognito

This service within the AWS IoT platform facilitates the organization, monitoring, and control of IoT devices. It involves the core architecture and integration of AWS IoT. This service enables devices to enroll in large numbers and categorize them into groups, linking them with access policies. AWS IoT offers a repository for managing items, stored as JSON data. Interaction with this repository is achievable through either the AWS IoT console or the AWS Command Line Interface.

The stage utilizes leads so as to communicate with different AWS administrations. Rules are formed by a trigger created in a SQL like punctuation and at least one activity enacted.

Correspondence to and from AWS IoT's Core is permitted by a distribute/buy-in message representative help. The message intermediary underpins MQTT to distribute and buy-in, and HTTPS just to distribute, both through IPv6 and IPv4. The usage of the message agent depends on MQTT v.3.1, yet, it doesn't bolster QoS2, and it doesn't permit the association of at least two customers with the equivalent customer ID all the while. All themes that start with \$ are held themes, used for gadget shadow activities. The intermediary underpins associations with the HTTP convention by the utilization of RESTful API.

So as to make and cooperate with gadgets AWS IOT gives a Command Line Interface (CLI), AWS IoT's API to fabricate application utilizing HTTPS or HTTP solicitations and Device SDKs. AWS offers administrations for the assortment and the preparing of information records: Amazon's Kinesis Data Stream to constant procedure of gushing information, AWS Lambda to execute server less code, Amazon

Notification Service to transmit or get warnings, and Amazon's Simple Queue Administration to store information in a line.

17 Microsoft AZURE For IoT

Microsoft IoT administrations help endeavors to pick up bits of knowledge from associated gadgets and transform these bits of knowledge vigorously. The organization has items and programming improvement packs set up to address the issue of each person, engineer, and endeavor.

It is an assistance that empowers just-intime provisioning of gadgets to an IoT center point, without requiring human mediation. Gadgets contact the provisioning administration endpoint passing their distinguishing data.

There are three essential zones to be considered in regards to security: gadget, association, and cloud security. The Azure Hub Identity Registry gives secure capacity of gadget characters and every security key; all associations must be started by the gadget to the Center point, not the other way around, and use TLS validation with X.509 endorsement; Azure Active Directory is utilized for client validation and approval for cloud get to.

Azure Hub identity registry

Tls

X.509 certificates

Azure Active directory

18 Google Cloud IoT Core

Google is among the top IoT stage suppliers around the globe, making it simpler for engineers to assemble associated gadgets. The internet searcher mammoth gives Cloud IoT Core as its leader IoT answer for making secure and imaginative arrangements.

The IoT Core offers per-gadget open or private key validation utilizing JSON web tokens and bolsters for Elliptic Curve or RSA calculations to check marks. Concerning correspondence security, the TLS 1.2 convention, utilizing root authentication specialists, is required for MQTT associations. Google Cloud Identity and Access Management (IAM) permits to control, validate, and approve the Cloud IoT Core API get to.

Jwt

Tls

X.509 certificates

Google Cloud IAM

It incorporates enrollment, verification, and approval forms. With the gadget chief, it is conceivable to make and design libraries and gadgets inside them. The gadget vault is arranged with at least one Cloud sub/pub themes to which telemetry occasions are distributed for all gadgets in that vault. A gadget is characterized with metadata, it sends telemetry messages and gets setups, client-characterized mass of information sent from the Cloud.

The stage bolsters MQTT and HTTP for overseeing gadgets and interchanges. By the utilization of MQTT, gadgets send distribute solicitations to a particular subject, while utilizing HTTP, gadgets don't keep up an association with the stage.

19 IBM Watson IoT Platform

It is an overseen administration facilitated on the cloud, empowering secure association, the executives, and preparing of IoT information. Alongside the intensity of IoT, the IBM Watson IoT Platform uses innovations like man-made brainpower (AI) and blockchain to permit endeavors to catch information from gadgets, hardware, and machines. They can additionally utilize this information to pick up experiences and settle on better business choices.

Based on IBM Cloud, the Watson IoT Platform is a versatile IoT administration that can adjust in the blink of an eye when a business needs to develop. It utilizes AI for information investigation with the goal that the information from IoT gadgets can be handled right away, and undertakings can increase important experiences from it. It utilizes blockchain to empower sharing of secure data over the environment. The usage of blockchain innovation builds trust and straightforwardness by approving provenance and occasions in an unchanging record.

20 Conclusion

While AWS IoT implements robust security measures for authenticating edge devices, there remains a potential vulnerability where these devices and communication pathways could be compromised or exploited. This underscores the ongoing need for refining existing techniques. One way to enhance security is by incorporating security measures during the design phase, following established security best practices. This involves the utilization of widely accepted and reputable security frameworks and algorithms for authentication. The selection of specific security mechanisms should be tailored to the unique customer IoT environment, as this approach assists in pinpointing areas of highest risk and prioritizing security as a paramount consideration.

References

1. Li J, Kuang X, Lin S, Ma X, Tang Y (2020) Privacy preservation for machine learning training and classification based on homomorphic encryption schemes. *Inf Sci* 526:166–179
2. Kurniawan A (2018) Learning AWS IoT
3. https://d1.awsstatic.com/whitepapers/Security/Securing_IoT_with_AWS.pdf
4. <https://aws.amazon.com/blogs/iot/understanding-the-aws-iot-security-model/>
5. <https://aws.amazon.com/blogs/apn/implementing-secure-authentication-with-aws-iot-and-microchips-trust-platform/>
6. Li J, Yu Q, Zhang Y (2019) ‘Hierarchical attribute-based encryption with continuous leakage-resilience.’ *Inf Sci* 484:113–134
7. Du L, Li K, Liu Q, Wu Z, Zhang S (2020) Dynamic multi-client searchable symmetric encryption with support for Boolean queries. *Inf Sci* 506:234–257
8. Deng H, Qin Z, Wu Q, Guan Z, Zhou Y (2020) Flexible attributebased proxy re-encryption for efficient data sharing. *Inf Sci* 511:94–113
9. Takabi H, Hesamifard E, Ghasemi M (2016) Privacy preserving multiparty machine learning with homomorphic encryption. In: *Proc NIPS*. Barcelona, Spain, pp 1–5
10. Yin H, Qin Z, Zhang J, Ou L, Li F, Li K (2019) Secure conjunctive multi-keyword ranked search over encrypted cloud data for multiple data owners. *Future Gener Comput. Syst* 100:689–700

On Picture Fuzzy Information-Based Hybrid Cryptographic TOPSIS Approach for Best Suitable Cloud Storage and Security Level



Himanshu Dhumras, Rakesh Kumar Bajaj, and Varun Shukla

1 Introduction

Even while cloud computing has gained a lot of popularity, the lack of adequate security measures prevents the majority of businesses or customers from implementing it. Because of flawed plans, programming or configurations made by designing agencies and service provider firms at various architectural layers, such as infrastructural ground and applications that can compromise the evaluation of the contracted quality of service (QoS), the cloud may suffer from a number of vulnerabilities. Additionally, attackers choose the cloud as a favorite target since it allows them to engage in offensive behavior. Therefore, in order to protect the large cloud market, suitable and more protective cloud security is essential. Security issues are the most important problem that cloud computing is currently experiencing, according to research from the International Data Corporation (IDC) [1]. For securing extremely sensitive data/information and limiting unauthorized/unauthenticated accesses in the cloud or elsewhere, data owners may need to encrypt data before outsourcing it to the commercial public cloud [2]. The conventional plain text keyword search-based data usage services may be rendered useless as a result. The enormous cost of data transfer/exchange capacity/capability in cloud-scale frameworks would make it even more unfeasible to download all the data and decode it locally. Finding a good search engine and protecting the privacy of cloud data that has been encrypted is therefore of utmost importance. Due to various unavoidable security/protection boundaries with various stringent requirements like the “index privacy”, “data protection”, “keyword

H. Dhumras · R. K. Bajaj (✉)
Jaypee University of Information Technology, Wagnaghat, Solan PIN 173 234, Himachal Pradesh, India
e-mail: rakesh.bajaj@juitsolan.in

V. Shukla
PSIT, Kanpur, Uttar Pradesh, India

privacy”, and many more [3, 4], the “encrypted cloud data search framework” is still a challenging task for recent distributed/parallel computing systems.

Over the years, a lot of researchers have put forth several “cloud storage and security frameworks”. However, due to the steadily growing strengths of “cloud users and security concerns”, extensive research for improvement is still ongoing. The first study on cryptography-based cloud storage was published by Kamara et al. [5]. By utilizing the advantages of non-standard/non-traditional cryptographic approaches, such as “attribute-based encryption and searchable encryption”, they created secure cloud storage systems for both consumer and enterprise applications. Their earlier efforts introduced other features like integrity, searchability, and verifiability, which corrected their earlier works’ overemphasis on achieving confidentiality [6]. In a distributed storage system called the cloud, simulation of the data security issue in cloud data storage is proposed by Wang et al. [7], and their research suggested a useful plan with excellent data support and other features like block update, remove, and append. Here, the file distribution system is set up using the erasure-correcting code approach with the goal of ensuring data dependability. Data error localization and storage correctness insurance are integrated into the process. Additionally, their plan is strong, incredibly effective, and resistant to many failures and attacks like “Byzantine failure”. The researchers also offer an improved “public auditing system” with a protocol that strongly supervises all the operations involving dynamic data [8].

As a result, by impersonating the fundamental Markle hash tree, the provable data possession (PDP) or proof of retrievability (PoR) scheme’s existing soundness was enhanced. With the aid of an effective bilinear aggregate signature mechanism and third-party auditing (TPA), this approach was capable of handling numerous auditing tasks. The Boneh–Lynn–Shacham (BLS) algorithm’s high computational cost, however, was a significant problem in this case. Additionally, their model was not capable to support correctness for both dynamic data and public verification. A group of storage servers that are often installed with the aid of hundreds to thousands of servers offers processing power in the cloud computing environment [9]. The authors of this article modeled a typical four-layered cloud-based dataset. Massive physical resources (such as storage and application servers) made up the lowest tier and assisted in strengthening the storage servers. These servers directly handled the next-level virtualization tools and services that allowed sharing of capabilities among the server virtual instances. The virtual occurrences, however, were isolated from one another, creating a fault-tolerant behavior and an isolated security context [10]. For maintaining the unique feature of the “encryption and decryption” keys is crucial since cloud computing relies heavily on replication. This problem has recently become a hurdle for Amazon’s cloud computing platforms. However, the lack of foresight in cryptography might lead to regrettable outcomes [11]. This application places a lot of emphasis on key distribution and encryption. Takabi et al. [12] established a thorough security strategy for cloud computing systems. Their model suggested a few methods for addressing security issues. The model was made up of many security-related elements. Issues including identity management, access control, policy integration across many clouds, and trust management across several

clouds as well as between a cloud and its users were covered in the modules. Guleria et al. [13] presented a parameterized information measure for the Pythagorean fuzzy set with monotonicity and maximality feature along with an algorithm for solving a decision-making problem. The cryptographic assessment and enhancement is certainly a crucial component in the process of cloud computing for security reasons. Maintaining the uniqueness of the encryption and decryption keys is crucial since cloud computing relies heavily on replication. This problem has recently become a hurdle for Amazon's cloud computing platforms. By routinely verifying the hash estimation of the files kept in the huge data storage, Venkatesan et al. [14] suggested an effective multi-agent-based static and dynamic data integrity protection. The multi-agent system was necessary for their proposed model (MAS). The agent in this situation was capable of self-rule, cunning, social aptitude, and other things. Three entities (the client, service provider, and data owner) are included in the suggested architecture, and several agents are used to screen and maintain the data integrity.

The present paper has been structured as follows: Sect. 2 briefly presents very important preliminary definitions and fundamental notions which are available in the literature. Section 3 describes the security classification using the picture fuzzy information with the incorporation of various security parameters. The decision-making algorithms by making use of the TOPSIS technique for storage selection of servers have been done in Sect. 3. The necessary conclusions and advantages have been given in Sect. 4.

2 Preliminaries

In this section, we are presenting the basic notions and definitions of various other fundamental sets which are available in the literature. These preliminaries would help to understand the proposed notions of picture fuzzy hypersoft set and increase the readability for the researchers.

Definition 1 Intuitionistic Fuzzy Set(IFS) [15]. "An intuitionistic fuzzy Set R in V is given by $R = \{v, \rho_R(v), \omega_R(v) | v \in V\}$; where $\rho_R : V \rightarrow [0, 1]$ is the degree of membership of v in R and $\omega_R : V \rightarrow [0, 1]$ is the degree of non-membership of v in R and ρ_R, ω_R satisfies the constraint $\rho_R(v) + \omega_R(v) \leq 1$ ($\forall v \in V$); and $\pi_R(v) = (1 - (\rho_R(v) + \omega_R(v)))$ is called the degree of indeterminacy v in V . We denote the set of all intuitionistic fuzzy sets over V by $IFS(V)$ ".

Definition 2 Picture Fuzzy Set(PFS) [16]. "A picture fuzzy Set R in V is given by

$$R = \{v, \rho_R(v), \tau_R(v), \omega_R(v) | v \in V\};$$

where $\rho_R : V \rightarrow [0, 1]$ is the degree of positive membership of v in R , $\tau_R : V \rightarrow [0, 1]$ is the degree of neutral membership of v in R and $\omega_R : V \rightarrow [0, 1]$ is the degree of negative membership of v in R and ρ_R, τ_R, ω_R satisfies the constraint

$$\rho_R(v) + \tau_R(v) + \omega_R(v) \leq 1 \quad (\forall v \in V);$$

and, $\neg_R(v) = (1 - (\rho_R(v) + \tau_R(v) + \omega_R(v)))$ is called the degree of refusal membership of v in V . We denote the set of all the picture fuzzy sets over V by $PFSS(V)$.

As per the findings by Prasad et al. [17], it is understandable that “A proficient multi-agent-based static and dynamic data integrity protection by periodically confirming the hash estimation of the files stored in the massive data storage. Their proposed model depended on the multi-agent system (MAS). The agent here had a capacity for self-ruling, ingenuity, social ability, and so on. The proposed architecture incorporates three entities (i.e. client, service provider, and data owner) and has different agents to screen and keep up the data integrity.”

Also, Sood et al. [18] stated that “The concept of data security sections, which is followed in our paper in a different manner. Confidentiality, availability, and integrity parameters for cryptography in addition to the message authentication code (MAC) for checking the data integrity are utilized as a part of this procedure. The strategy provides classification, uprightness, authorization, verification, and non-repudiation and anticipates data spillage. The security degree that they provide in ascending order is MAC, classification of data, and execution of index and encryption system.”

3 Security Classification Using Picture Fuzzy Information

This stage follows the encryption stage on the part of the data owner. The data owner would process the data in accordance with his selected security requirements after successfully logging into the CSP. The required P , Q , and R security parameters— P for “proactive threat detection and management”, Q for “quality data backup”, and R for “maximum uptime and lowest downtime” are listed and sent collectively to the CSP in order to be stored. Along with those three parameters are the encrypted message M' , encrypted index IM' containing the user’s most frequently searched phrases, and the secret key K_1 discussed in the previous section.

Here, a fuzzy-based method for storing data with various access kinds on various cloud storage servers depending on the three aforementioned crucial security characteristics is described (i.e., P , Q , and R). With a shorter execution time and without the additional load of dataset training, the proposed fuzzy-based approach has been used to categorize the access kinds where ambiguity or fuzziness will be handled utilizing membership functions. The CSP gives the user an option of various fuzzy variables for each security parameter. These variables will be selected by the user. These choices will be converted into a security factor (S_f) using the suggested method illustrated in the sections that follow.

The user might not be familiar with the process for assigning values to the aforementioned security parameters because they are qualitative in nature. The user will be able to list their needs on a more detailed level with the aid of the fuzzy linguistic variables as shown in Table 1.

Table 1 Linguistic variables for computing the security parameters

Qualitative term	PFNs
“Absolutely bad (AB)”	(0.83, 0.04, 0.11)
“Very very bad (VVB)”	(0.75, 0.05, 0.15)
“Very bad (VB)”	(0.62, 0.1, 0.2)
“Bad (B)”	(0.55, 0.11, 0.25)
“Medium bad (MB)”	(0.50, 0.15, 0.30)
“Medium (M)”	(0.45, 0.20, 0.35)
“Medium high (MH)”	(0.40, 0.22, 0.37)
“High (L)”	(0.35, 0.25, 0.40)
“Very high (VH)”	(0.25, 0.30, 0.43)
“Very very high (VVH)”	(0.15, 0.35, 0.48)

The range for security factor S_f is between $0 \leq S_f \leq 0.399$ for public access type and for private access type it is ranging between $0 \leq S_f \leq 0.799$. Also, for the limited access owner, the range is between $0.8 \leq S_f \leq 1$.

Selection of server storage and data storage

After data encryption on the owner’s end, the data will be sent to the cloud for storage. Instead of being stored at a single server, the suggested approach will divide the data over a number of separate, geographically dispersed storage servers. During this stage, the CSP’s registered storage servers are chosen from a pool of available servers for data storage. Some cloud service providers split apart the data that they get from the data owner. Next, each data component is saved on a distinct storage server with a different storage type, level of security, etc. at a different geographic location. To ensure an effective, secure, and quick data storage process, the task is divided across several access level sites (based on S_f) on various storage servers.

Computation of the weights of the criterion

The area that concerns the most is data security. As a result, the storage server’s level of security is the most crucial one, because they reduce execution time and network bandwidth, processing speed, and time delay. These factors are regarded as the most important factors when storing data concurrently on several storage servers, which aids in determining the additional communication cost in a cloud environment.

Analytic Hierarchy Process (AHP) is used in this case to generate the weights of the criteria through pairwise comparisons for each of the selected criteria (AHP). AHP converts empirically based comparisons into numerical numbers for further analysis and comparison. The most popular scale is the relative importance scale between two criteria, as proposed by Saaty [19]. The scale, which has values ranging from 1 to 9, is shown in Table 2 and it is used to compare one criterion’s importance to another criterion. The consistency ratio (CR) gauges how consistently respondents answer questions on the AHP forms. Using Saaty’s significant scale [19] provided in

Table 2 Scale of significance for criterions

Qualitative term	PFNs
“Extremely important (EI)”	(0.83, 0.04, 0.11)
“Very important (VI)”	(0.60, 0.05, 0.21)
“Important (I)”	(0.53, 0.12, 0.25)
“Less important (LI)”	(0.45, 0.15, 0.30)
“Very less important (VLI)”	(0.30, 0.25, 0.35)

Table 3 Computations of criterion weights

Criteria	Weights
“No. of CPU”	0.0403
“Avg. processing speed”	0.1276
“Security level”	0.3894
“Avg. transmission speed”	0.2611
“Avg. time delay”	0.0998
“Avg. memory utilization”	0.0268

Table 2 to build the pairwise comparison matrix. The final weight for each criterion determined using the AHP technique is shown in Table 3. These weights will then be processed further in the TOPSIS method for the selection of storage servers.

Multiple Storage Servers Selection

For handling the uncertainty found in the process of selection of data storage centers, the role of fuzzy set theories and soft computing decision-making techniques become quite important and can be extensively utilized for better results. In the selection process, a number of quantitative and qualitative criteria must be taken into account. In order to solve the storage server selection problem, a combination of these two approaches is used here. The suggested model uses a picture fuzzy TOPSIS technique, in which the ratings of various storage servers under various criteria are appraised in linguistic words represented by picture fuzzy numbers, to choose the best storage servers under real-time conditions. The implementation of fuzziness for three qualitative criteria termed as “degree of security”, “memory use”, and “time delay”, for which the linguistic scaling/fuzzification are presented in Table 4. Some significant decision factors for storage servers are listed in Table 5.

Each storage server registered with the CSP has a direct relationship with its relevant properties. Some of them are static, while others were captured at the moment. In Table 6, the values of qualitative variables like “degree of security”, “average memory use”, and “average time delay” are taken into account. While the qualitative qualities are expressed in fuzzy linguistic terms, the quantitative attribute values are simply determined from the various storage servers’ current behavior. By using

Table 4 Security level

Qualitative term	PFNs
“Very low (VL)”	(0.23, 0.04, 0.11)
“Low (L)”	(0.51, 0.05, 0.21)
“Moderate low (ML)”	(0.53, 0.12, 0.17)
“Fair (F)”	(0.38, 0.15, 0.30)
“Moderate high (MH)”	(0.29, 0.25, 0.23)
“High (H)”	(0.10, 0.25, 0.35)
“Very high (VH)”	(0.05, 0.25, 0.22)

Table 5 Utilization of memory

Qualitative term	PFNs
“Low (L)”	(0.33, 0.02, 0.12)
“Medium (M)”	(0.54, 0.03, 0.21)
“High (H)”	(0.33, 0.04, 0.17)

Table 6 Time delay

Qualitative term	PFNs
“Close (C)”	(0.73, 0.04, 0.11)
“Adequate (A)”	(0.52, 0.05, 0.18)
“Fair (F)”	(0.37, 0.12, 0.23)

the TOPSIS algorithm, it is feasible to order the storage servers according to their priority in the selection process.

Picture Fuzzy TOPSIS Algorithm for the selection of storage server

Step 1: In the first step, there is the conversion of linguistic variables into picture fuzzy numbers.

Step 2: In this step, there is the conversion of the picture fuzzy numbers to crisp numbers.

Step 3: In this step, the different attribute values which are not in the range of 0 to 1 need to be normalized so that the computations in decision-making can be done. Normalization can be done by $x_{ij} = \frac{y_{ij} - \min y_{ij}}{\max y_{ij} - \min y_{ij}}$, where, y_{ij} is the value of the j th criteria.

Step 4: Now, calculation of the normalized weighted decision matrix: $v_{ij} = x_{ij} \times w_{ij}$, where v_{ij} is the weighted normalized data and w_{ij} denotes the weight of the j th criteria.

Step 5: Next, calculation of positive and negative ideal solution:

$$V_j^+ = \{v_1^+, v_2^+, \dots, v_m^+\}$$

$$V_j^- = \{v_1^-, v_2^-, \dots, v_m^-\}.$$

Step 6: In this step, the computation of the distance of every attribute value v_{ij} from a positive ideal solution (V_j^+).

$$D_i^+ = \sqrt{\sum_{j=1}^n \frac{1}{3}(v_{ij} - v_j^+)^2}; i = 1, 2, \dots, n; j = 1, 2, \dots, m.$$

Step 7: Similarly, the computation of the distance of every attribute value v_{ij} from a negative ideal solution (V_j^-).

$$D_i^- = \sqrt{\sum_{j=1}^n \frac{1}{3}(v_{ij} - v_j^-)^2}; i = 1, 2, \dots, n; j = 1, 2, \dots, m.$$

Step 8: In this step, computation of the coefficient of relative closeness can be done by making use of the following formula:

$$CRC_i = \frac{S_i^-}{S_i^- + S_i^+}; \text{ where, } 0 \leq CRC_i \leq 1; i = 1, 2, \dots, n; j = 1, 2, \dots, m.$$

Step 9: In the final step, ranking of the alternatives can be done on the basis of the values of CRC_i in decreasing order (Table 7).

Table 7 Computation of the ranking orders of the storage servers

Storage server	D_i^+	D_i^-	CRC_i	Order
“(SS ₁)”	0.073	0.017	0.315	7
“(SS ₂)”	0.042	0.054	0.623	3
“(SS ₃)”	0.058	0.047	0.346	6
“(SS ₄)”	0.018	0.074	0.812	1
“(SS ₅)”	0.039	0.052	0.587	4
“(SS ₆)”	0.030	0.058	0.628	2
“(SS ₇)”	0.053	0.061	0.558	5

4 Conclusions and Scope for Future Work

In order to secure the privacy of distributed cloud storage systems, this study proposes a cloud storage framework/technique that uses a 128-bit encryption key generated by synchronizing a deoxyribonucleic acid (DNA) cryptographic approach with the Hill Cipher algorithm. With the aid of a picture fuzzy information-based classification methodology, the data in this document have been categorized in accordance with several security parameters. Additionally, this architecture would let you pick the best storage server from a selection. Further, a picture fuzzy information-based technique for order of preference by similarity to ideal solution (TOPSIS) decision-making algorithm has been implemented to determine the best storage server where the data can be saved, reducing the execution time in the process. Also, the extension of this work can be done by executing various other methods like AHP, WASPAS, VIKOR in different techniques.

Declarations and Compliance with Ethical Standards

Ethical approval: This article does not contain any studies with human participants or animals performed by any of the authors.

Funding Details: The authors declare that the research carried out in this article has no source of funding.

Conflict of interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Authorship contributions: The authors have equally contributed to the design and implementation of the research, to the analysis of the results, and the writing of the manuscript.

Acknowledgements We are very much thankful to the anonymous reviewers for suggesting the points/mistakes which have been well implemented/corrected for the necessary improvement of the manuscript. We sincerely acknowledge our deep sense of gratitude to the Editorial office and reviewers for giving their valuable time to the manuscript.

References

1. Gartner N (2012) Consumers will store more than a third of their digital content in the cloud by 2016. Press Release
2. Velte AT, Velte TJ, Elsenpeter RC, Elsenpeter RC (2010) Cloud computing: a practical approach. McGraw-Hill, New York, pp 44

3. Kamara S, Lauter K (2010) Cryptographic cloud storage. In: Proceedings of the international conference on financial cryptography and data security. Springer, Berlin, Heidelberg, pp 136–149
4. Cao N, Wang C, Li M, Ren K, Lou W (2014) Privacy-preserving multi-keyword ranked search over encrypted cloud data. *IEEE Trans Parallel Distrib Syst* 25(1):222–233
5. Chen R, Mu Y, Yang G, Guo F, Wang X (2016) Dual-server public-key encryption with a keyword search for secure cloud storage. *IEEE Trans Inf Forensics Secur* 11(4):789–798
6. Chase M, Kamara S (2010) Structured encryption and controlled disclosure. In: International conference on the theory and application of cryptology and information security. Springer, Berlin, Heidelberg, pp 577–594
7. Wang C, Wang Q, Ren K, Lou W (2009) Ensuring data storage security in cloud computing. In: Proceeding of the IWQoS, pp 1–9
8. Wang Q, Wang C, Ren K, Lou W, Li J (2011) Enabling public auditability and data dynamics for storage security in cloud computing. *IEEE Trans Parallel Distrib Syst* 22(5):847–859
9. Buyya R, Murshed M (2002) Gridsim: a toolkit for the modeling and simulation of distributed resource management and scheduling for grid computing. *Concurr Comput: Pract Exp* 14(13–15):1175–1220
10. Smith JE, Nair R (2005) The architecture of virtual machines. *Computer* 38(5):32–38
11. Balding C (2008) Is your amazon machine image vulnerable to SSH spoofing attacks?. *Cloud Secur*
12. Takabi H, Joshi JB, Ahn GJ (2010) Securecloud: towards a comprehensive security framework for cloud computing environments. In: Proceedings of the 34th annual computer software and applications conference workshops (COMPSACW). IEEE, pp 393–398
13. Guleria A, Bajaj RK (2018) Pythagorean fuzzy R -norm information measure for multicriteria decision-making problem. *Adv Fuzzy Syst*. Article ID 802301
14. Venkatesan S, Vaish A (2011) Multi-agent based dynamic data integrity protection in cloud computing. In: Proceedings of the international conference on advances in communication, network, and computing. Springer, Berlin, Heidelberg, pp 76–82
15. Atanassov KT (1986) Intuitionistic fuzzy sets. *Fuzzy Sets Syst* 20(1):87–96
16. Coung B (2014) Picture fuzzy sets. *J Comput Sci Cybern* 30(4):409–420
17. Prasad P, Ojha B, Shahi RR, Lal R, Vaish A, Goel U (2011) 3 dimensional security in cloud computing. In: 2011 3rd international conference on proceedings of the computer research and development (ICCRD). IEEE, pp 198–201
18. Sood SK, Sarje AK, Singh K (2011) A secure dynamic identity based authentication protocol for multi-server architecture. *J Netw Comput Appl* 34(2):609–618
19. Saaty TL (2005) The analytic hierarchy and analytic network processes for the measurement of intangible criteria and for decision-making. In: Multiple criteria decision analysis: state of the art surveys, vol 78. Springer, New York, NY, pp 345–405

Network Layer Performance of Hybrid Buffer-Based Optical Router



Sumit Chandra, Shahnaz Fatima, and Raghuraj Singh Suryavanshi

1 Introduction

Due to the data-centric nature of current applications like TV on demand, the internet, etc., there has been a sharp rise in the demand for more bandwidth. There is a bottleneck in the provision of very high-speed data communication due to present technological technologies. As a result, in the near future, a substitute technology will be needed to accommodate such high data rates (160 Gbps). The optical fiber cable has the capacity to carry a lot of data (Tbps). As a result, optical communication is regarded as the newest technology for data transfer.

Figure 1 displays the optical communication technologies' future plan. Digital Cross-Connect (DXC) was first launched in 1990. An Optical add-drop Multiplexer (OADM) was created 5 years later. Then, Optical Cross-Connect (OXC), passive optical networks (PON) were proposed. The concepts of optical label switching (OLS) where labels are used, Optical Burst Switching (OBS) where packets are transmitted in the form of bursts of packets, and OPS were introduced. Some of these technologies have since been evaluated and implemented in various parts of the world as discussed by Kachris [1]. The biggest barrier to OPS deployment is technology. The steadily increasing speed of electrical circuitry makes gigabit rates unsuitable. Thus, optical technology is a solution that has promise but is still in its

S. Chandra (✉)

Research Scholar, Amity Institute of Information Technology, Amity University Uttar Pradesh, Lucknow Campus, Lucknow, India
e-mail: sumit8842@gmail.com

S. Fatima

Amity Institute of Information Technology, Amity University Uttar Pradesh, Lucknow Campus, Lucknow, India

R. S. Suryavanshi

Department of Computer Science & Engineering, Pranveer Singh Institute of Technology, Kanpur, India

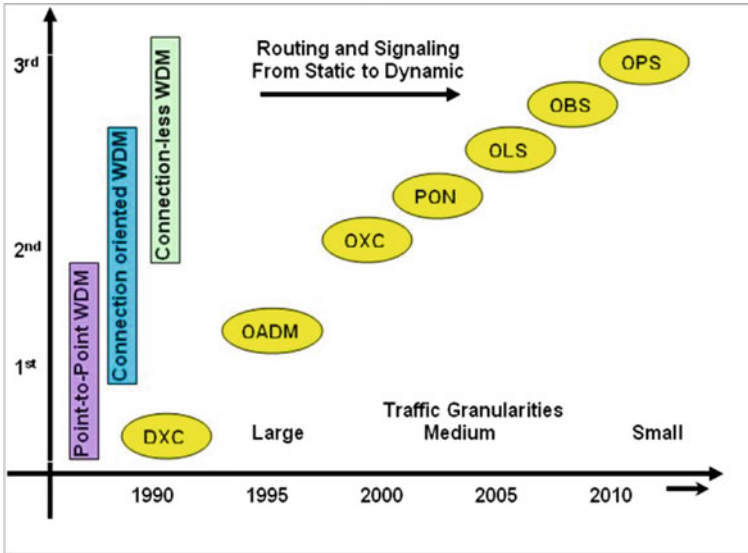


Fig. 1 Roadmap of optical communication

infancy. The three main technological concerns that are crucial to OPS are fast optical packet header processing, routing, and buffering as detailed by Kachris and Tomkos [2].

The speed at which routing may be completed affects OPS deployment. Due to the complex nature of the problem, it is difficult to determine the bare minimum requirements for OPS in routers. The following are the key factors that are essential for OPS systems as discussed by Kachris and Tomkos [3].

Switching Time

Components should be able to be configured for fast optical packet switching in less than 10 ns to maintain minimal overhead.

Throughput

Throughput is the amount of traffic that can move through a network at once without being blocked, and it should be high.

Signal Degradation

It comprises cross-talk, noise deterioration, and optical loss. The transmission distance is constrained by signal degradation. Regeneration of the signal is required to protect it from degradation, but optical 3-Regeneration signals don't exist. Due to data-centric applications, internet traffic has dramatically expanded. All optical communication is still impractical as discussed by Kachris et al. [4]. Currently, only the point-to-point interface of optical communication is used, and routing is still carried out in the electrical domain, wasting a significant amount of energy in the

process of switching from electrical to optical and from optical to electrical as detailed by Hemenway et al. [5].

Due to the complexity of the optical system and the lack of optical counterparts for electrical components, the shift from electrical to optical technology is not seamless as discussed by Proietti et al. [6]. Currently, wavelength division multiplexing is used to build point-to-point networks. It is anticipated that optical routers will take the place of the present electronic routers in the following phase. Due to the lack of optical processors, all-optical router implementation is still not conceivable. As a result, hybrid optical router designs incorporating electrical processors to perform data control functions are anticipated in the upcoming phase.

Packets arrive at the inputs of the switch randomly for any output, so they may select a common output for the exit of the switch. This phenomenon is known as contention. Contention resolution procedures are employed to prevent this collision. The deflection route is currently not chosen because of the high delay. The other two processes, wavelength conversion and contesting packets’ buffering, are thought to be the better options. However, it’s also crucial to remember that for efficient conflict resolution, wavelength conversion and buffering should be combined.

In a similar context, many optical packet router designs have evolved with time and depend on buffering requirements as summarized in Table 1. The optical router design study was proposed by Srivastava et al. [7] and since then, several other router designs have undergone a similar analysis (Fig. 2).

Table 1 Comparison between different switch designs

References	Novelty
Bari et al. [8]	Realization of optical DCs
Xu et al. [9]	Passive optical DCs
Segawa et al. [10]	AWG-based switch design
Sato et al. [11]	Wavelength routing in DC
Srivastava et al. [12]	AWG-based switch design
Srivastava et al. [13]	Optical switch design analysis
Srivastava et al. [14]	Re-circulating type switch design
Srivastava et al. [15]	Concept of optical memory
Singh et al. [16]	Concept of hybrid memory
Srivastava et al. [17]	Concept of dual buffers
Shukla et al. [18]	Re-circulating type switch design
Singh et al. [19]	Hybrid buffer with add-drop
Singh et al. [20]	Optical buffer with add-drop
Chandra et al. [21]	Hybrid switch

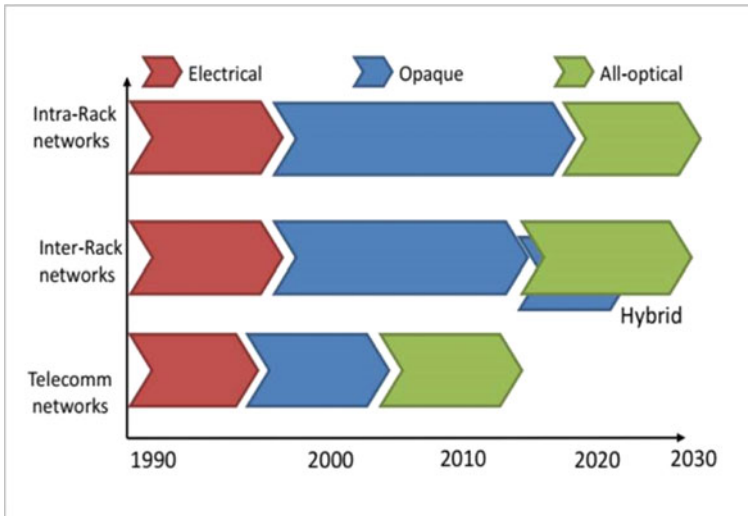


Fig. 2 Timelines for various technologies

2 Literature Survey

A variety of switch designs and buffering methods have been proposed, with the majority of modern work relying on OFC technology. Some of the notable switch designs are highlighted in Table 1. In total, 14 papers are detailed written by nine authors. An optical data center was suggested and created by Bari et al. [8]. A passive optical data center design was put up by Xu et al. in [9]. A Data Center design based on an arrayed waveguide grating was proposed by Segawa et al. [10]. A wavelength-routed DC architecture that is easily adaptable to WDM was proposed by Sato et al. [11]. For the storing of competing packets, Srivastava et al. [12] presented an optical core switch system based on AWG and fiber delay lines. An optical switch that is a broadcast and select type switch was introduced by Srivastava et al. [13], and a thorough mathematical analysis of the optical switch is presented to determine the operation window. A re-circulating type buffer was the foundation for the numerous optical switch designs that Srivastava et al. [14] presented and compared. According to Srivastava et al.'s [15] proposal, optical memory based on FDL has both benefits and drawbacks. A hybrid buffer with optical and electrical memories was suggested by Singh et al. [16]. Mathematical analysis has been done for the purpose of optimizing memories. The idea of a dual buffer was put forward by Srivastava et al. [17] to improve buffering storage for contention resolution. An AWG-based re-circulating type optical buffer design was put out by Srivastava et al. By incorporating WDM inside the buffer, Shukla et al. [18] expanded on Srivastava et al. [17]'s approach. An add-drop-based optical switch with hybrid buffering was introduced and its advantages over more contemporary designs by Singh et al. [19]. A network router-capable add-drop optical switch with optical buffering was presented by Singh

et al. Additionally noted are the problems with switch placement in the network [20]. Chandra et al. presented a hybrid optical switch, which is considered in this paper [21]. This switch design consists of electronic and optical buffer along with the inclusion of negative acknowledgment scheme to avoiding excess dropping of packets. The presented optical switch designs have their own advantages and dis-advantages.

3 Description of the Optical Packet Switch

Figure 3 depicts an optical packet switch based on FDL and an electronic buffer with negative acknowledgment. It is important to note that the buffering time in FDL is of the order of μs while in electronic buffers the buffering time is of the order of ms. Thus, in the case of a longer stay in the buffer ($\sim\text{ms}$), the electronic buffer will be used. The main problem associated with the buffering of electronic buffering is the first conversion of data from optical-to-electrical (O/E) and for the retrieval from electronic buffer electrical to optical conversion. However, the All-Optical Negative Acknowledgement (AO-NACK) scheme prevents this O/E and E/O conversion, but AO-NACK is sent back to the sender, thus increasing downlink traffic (Fig. 4a). To decrease the AO-NACK packets sent back to the sender, an input buffer was proposed by Chandra et al. [21], which AO-NACK packets are temporarily stored in the input buffer for a single slot only, using the buffer structure shown in Fig. 4b. The complete operation of the switch is detailed in [21]. The switching optical buffer will be used to store competing packets first, followed by the input buffer if the optical buffer is full, and the electronic buffer if the switching optical buffer and the input buffer are both full. Due to the sluggish read and write speeds of electronic random access memory, electronic buffers should be avoided. However, since they are very inexpensive compared to the cost of switches overall, they can be added to switches. It is important to keep in mind that the input buffer is only functional when there are no packets waiting at the input line. The input buffering strategy can only be advantageous under low and moderate loading circumstances. In the part after, simulation results are shown to demonstrate the utility of the input buffer.

4 Simulation Results

A computer simulation is used to demonstrate the suitability of the suggested design. The quantity of inputs, the size of the buffer, and the quantity of outputs are crucial variables from the perspective of simulation. A discrete event simulator is developed for computer simulation, and random numbers are applied to the traffic creation. As the simulation develops over time, evaluations of the number of created packets and the number of packets that successfully traversed the switch are made. A Monte Carlo simulation is run [19] to estimate average performance.

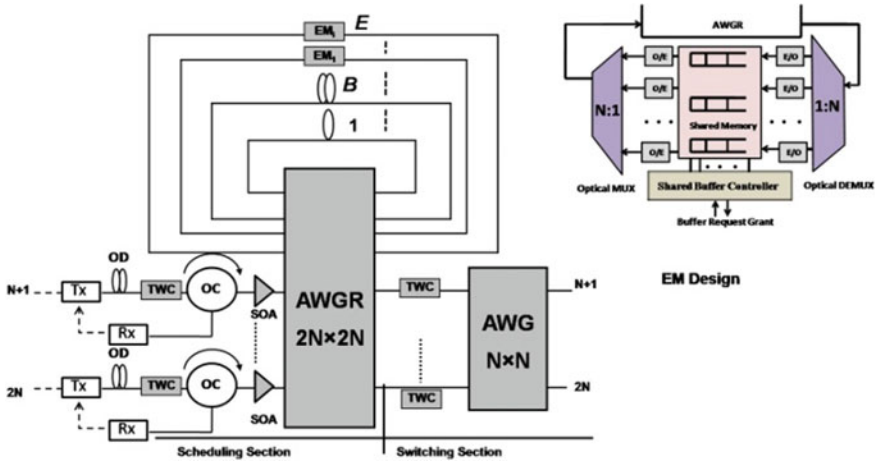


Fig. 3 Illustrative diagram of proposed switch design

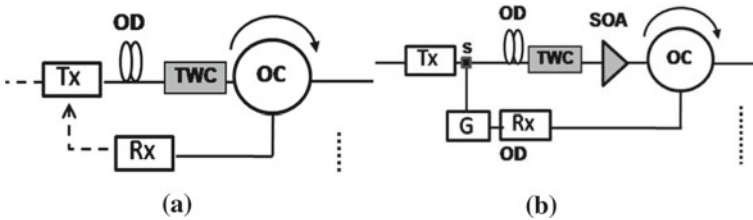


Fig. 4 a AO-NACK scheme b Input buffer scheme

The simulation steps are detailed in Algorithm 1. In Figs. 5 and 6, results for step 1 are shown while considering two loads value of ‘0.2’ and ‘0.8’, respectively. The load refers to the fraction of nodes generating data in a time slot. The load can be defined as the traffic that is arriving on the switch inputs, i.e., ‘0’ load means no traffic while load ‘1’ indicate continuous traffic on each load. For the illustration point of view, only 1000 slots are shown while considering the switch size of 4.

Algorithm 1

Input: Size of switch (N), buffering capacity (B),
 MS = maximum slot for simulation
do $j=1$: MS
For load = ρ ,
Step 1: Packet generation
 With probability ρ
 If packet is generated
 $P=P+1$ (Packet count (P))
 else
 $P=P$
Step 2: Destination assignment
 With probability ρ/N
Step 3: Buffer allotment
 First fill optical buffer
 If optical buffer is full
 Fill input buffer
 If input buffer is full or input line is busy
 Fill electronic buffer
 If electronic buffer is full
 Count lost packet (L)
Step 4: Packet Loss Probability (L/P)
end
end

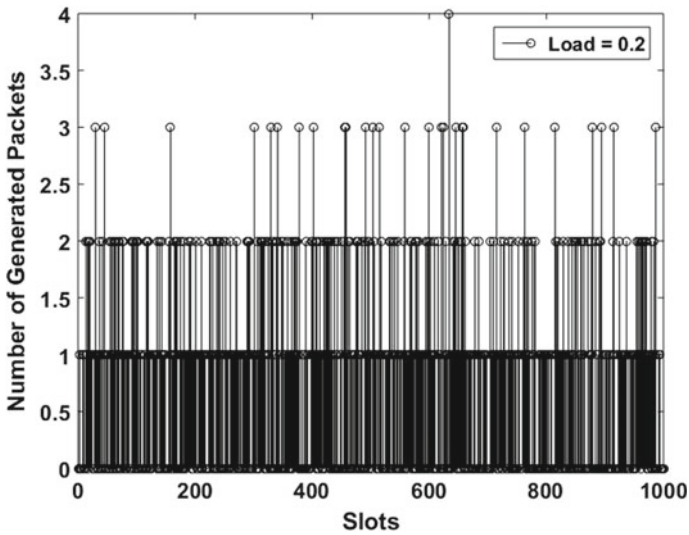


Fig. 5 Number of generated packets versus slots (load = 0.2)

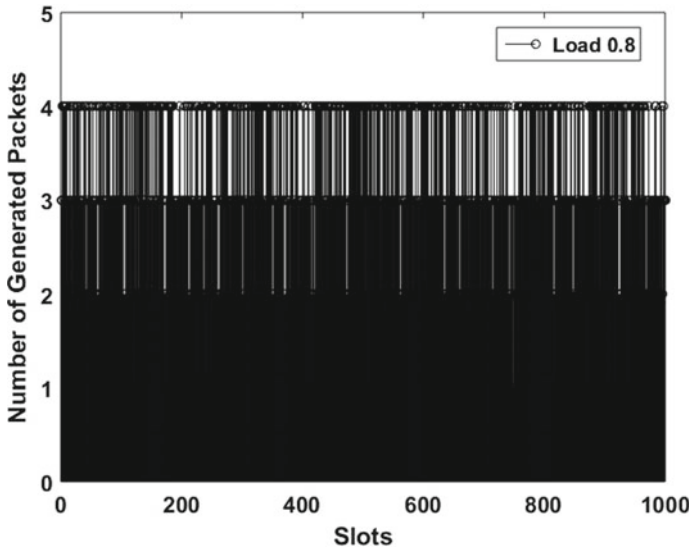


Fig. 6 Number of generated packets versus slots (load = 0.8)

Thus, the maximum number of packets that can be generated is 4. It is clear from Fig. 5, that the out of 1000 slots only once four packets are generated some slots are vacant and most of the times only one packet is generated.

In Fig. 6, the number of generated packets versus slots is shown at a load of 0.8. Here, a minimum of two packets are generated and a significant number of three and four packets are also generated. This is expected as the load increases the number of generated packets will also increase.

In Fig. 7, packet loss performance of electronic and optical buffers is shown at a load of 0.8. The packet loss probability is higher in electronic buffers as compared to optical buffers, and to maintain the same packet loss performance in electronic buffers as in optical buffers, comparatively buffer space will be required.

In Fig. 8, the number of generated packets, lost packets without buffer, lost packets with optical buffer, and lost packets in the presence of both input and optical buffer are shown. The index 1–1000 represents the performance at a load of 0.2 and, similarly, for other loads, an index of 1000 is chosen. At a load of 0.2, the number of generated packets is 792, the number of dropped packets without a buffer is 50, and with an optical buffer, the number of lost packets is zero. It is also clear that as the load increases, the number of generated and lost packets both increases. At the very high load of 1, the number of dropped packets is not zero.

In Fig. 9, packet loss probability (PLP) versus load is shown. Here the packet loss probability is very high in the case of no buffer, which is around 7.1% at the load of 0.2 and nearly 32% at the load of 1. With the use of an optical buffer, PLP is zero till the load of 0.8. While at a load of 0.8, the PLP is 1.1×10^{-4} . Finally, in the case

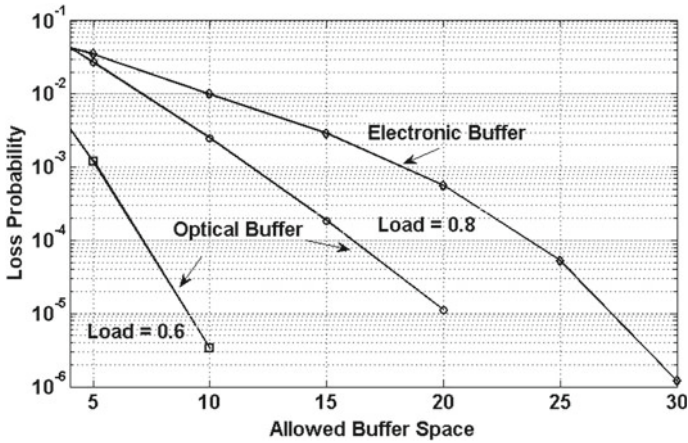


Fig. 7 PLP versus Allowed buffer space

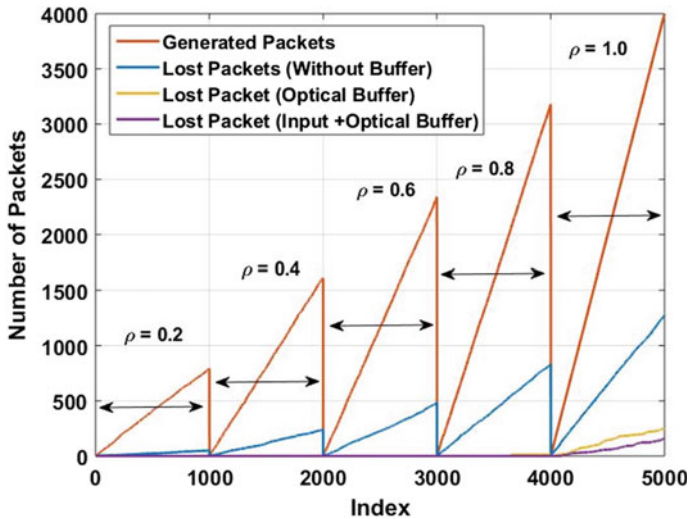


Fig. 8 Number of packets versus index

of input and optical buffer, the PLP is as low as 2.12×10^{-5} . Thus, using the input buffer, the PLP can be further reduced to $1/5$.

In Fig. 10, average delay (AD) versus load is shown. Here, the AD till load of 0.8 is nearly two slots. Thereafter, it rises very sharply. In the case of the optical buffer, it is around six slots, while in the case of the input and optical buffer, the average delay is around five slots. The difference of one slot is evident due to the single slot delay at the input buffer.

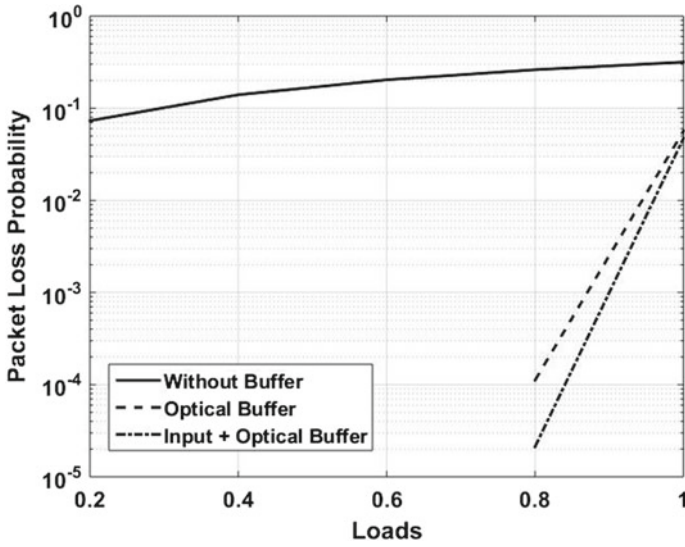


Fig. 9 Packet loss probability versus loads

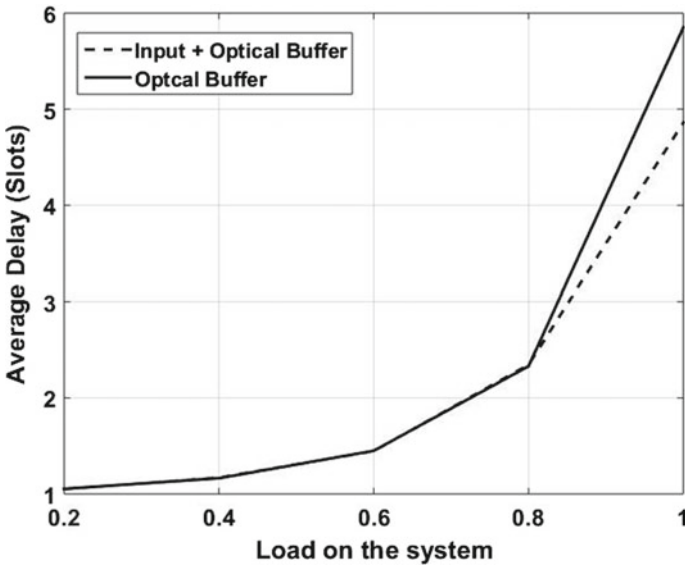


Fig. 10 Average delay versus loads

Table 2 Comparison with state-of-the-art switch designs

Reference	PLP
Singh et. al. [19]	1×10^{-4}
Singh et. al. [20]	1.28×10^{-4}
Proposed (2022)	8×10^{-5}

In Table 2, a comparison with the state-of-the-art method is shown; in the comparison, two papers are considered. In Table 2, switch size $N = 4$ is considered. The PLP, for Singh et al.'s [19] work is 1×10^{-4} . In the considered design, PLP is as low as 8×10^{-5} . As a result, the considered switch design outperforms current state-of-the-art technologies.

5 Conclusions

Optical routers are the main component in the designing of the OPS system, this paper discusses an optical router design, and to lessen the need for repeated transmission of the packets, inclusion of optical buffer of unit slot is introduced at the input of the switch. However, packet can only be stored in the input buffer when no other packets are present at the input line. By using a Monte Carlo computer simulation, the findings are achieved. It has been discovered that the electrical buffer performs worse than the optical buffer in terms of packet loss. The switch optical buffer has a significant effect on the PLP as a whole. The presence of an optical buffer at the switch's input can further reduce PLP, and the average delay of the switch is very less of 6 slots. In the future work, packet loss performance can be further evaluated using bursty traffic modeling.

References

1. Kachris C, Kananaskis, Tomkos, I (2013) Optical interconnection networks in data centers: recent trends and future challenges. *IEEE Commun Mag* 51(9): 39–45
2. Kachris C, Tomkos I (2012) A survey on optical interconnects for data centers. *IEEE Commun Surv & Tutor* 14(4):1021–1036
3. Kachris C, Tomkos I (2013) Power consumption evaluation of all-optical data center networks. *Clust Comput* 16(3):611–623
4. Kachris C, Bergman K, Tomkos I eds (2012) *Optical interconnects for future data center networks*. Springer Science & Business Media
5. Hemenway R, Grzybowski R, Minkenberg C, Luijten R (2004) Optical-packet-switched interconnect for supercomputer applications. *OSA J Opt Netw* 3:900–913
6. Proietti R, Yin CJNY, Yu R, Yoo SJB, Akella V (2012) Scalable and distributed contention resolution in AWGR-based data center switches using RSOA-based optical mutual exclusion. *IEEE J Sel Top Quantum Electron* 19(2), art. No. 3600111
7. Srivastava R, Singh RK, Singh YN (2010) Feedback fiber delay lines and AWG based optical packet switch architecture. *Opt Switch Netw* 7:75–84

8. Bari MF, Boutaba R, Esteves R, Granville LZ, Podlesny M, Rabbani MG, Zhang Q, Mohamed Zhani F (2012) Data center network virtualization: A survey. *IEEE Commun. Surv. & Tutor* 15(2): 909–928
9. Xu M, Liu C, Subramaniam, S (2016) PODCA: A passive optical data center architecture. In: 2016 IEEE International conference on communications (ICC), pp 1–6. IEEE
10. Segawa T, Matsuo S, Kakitsuka T, Shibata Y, Sato T, Kawaguchi Y, Kondo Y, Takahashi R (2010) All-optical wavelength-routing switch with monolithically integrated filter-free tunable wavelength converters and an AWG. *Opt Express* 18(5):4340–4345
11. Sato K, Hasegawa H, Niwa T, Watanabe T (2013) A large-scale wavelength routing optical switch for data center networks. *IEEE Commun Mag* 51(9):46–52
12. Srivastava R, Singh RK, Singh YN (2009) Large capacity optical router based on arrayed waveguide gratings and optical loop buffer. *Opt Quant Electron* 41(6):463–480
13. Srivastava R, Singh RK, Singh YN (2009) Design analysis of optical loop memory. *J Lightwave Technol* 27(21):4821–4831
14. Srivastava R, Singh RK, Singh YN (2008) WDM-based optical packet switch architectures. *J Opt Netw* 7(1):94–105
15. Srivastava R, Mangal V, Singh RK, Singh YN (2006) A modified photonic switch architecture based on fiber loop memory. In: 2006 Annual IEEE India conference, pp. 1–5. IEEE
16. Singh A, Tiwari AK, Srivastava R (2018) Design and analysis of hybrid optical and electronic buffer based optical packet switch. *Sādhanā* 43(2):19
17. Srivastava R, Bhattacharya P, Tiwari AK (2019) Dual buffers optical based packet switch incorporating arrayed waveguide gratings. *J Eng Res* 7(1)
18. Shukla MP, Srivastava R (2018) Arrayed waveguide grating and re-circulating buffer based optical packet switch. *J Opt Commun* 1(ahead-of-print)
19. Singh P, Rai PK, Sharma AK (2021) Hybrid buffer and AWG based add-drop optical packet switch. *J Opt Commun*
20. Singh P, Rai JK, Sharma AK (2022) An AWG based optical packet switch with add-drop of data. *Int J Inf Technol* 14:1603–1612. <https://doi.org/10.1007/s41870-022-00886-0>
21. Chandra S, Fatima S, Suryavanshi RS (2020) Hybrid buffer-based optical packet switch with negative acknowledgment for multilevel data centers. *J Opt Commun*

Efficient and Secure Data Aggregation for UAV-to-Ground Station Communication in Smart City Environment



Girraj Kumar Verma, Dheerendra Mishra, and Neeraj Kumar

1 Introduction

Currently, the development of smart cities is the priority of most countries. The foundation of such development is based on the utilization of Internet of things (IoT) technologies [1]. In this context, the deployment of IoT technologies helps to ensure the smart living style/standards of the citizens [2]. However, the increasing number of vehicles in smart cities has increased road accidents. According to the reports, road accidents will be the fifth leading cause of casualties by 2030 [3, 4]. Besides, the road conditions and weather conditions also increase the cost of transportation and cause the high cost of consumer items. Therefore, to ensure the smooth functioning of road traffic in smart cities, an intelligent traffic system (ITS) has been evolved [5].

In ITS, unmanned aerial vehicles (UAVs) (also called drones) are deployed to record traffic-related information. These drones can communicate with each other and also can communicate with the traffic control office (TCO). The communication between drones is called UAV-2-UAV communication and between drones to TCO is called UAV-to-Ground Station (UAV-2-GS) communication. The communication system for UAV-2-UAV and UAV-2-GS communication is utilizing the latest technologies and is called the Internet of drone (IoD) [6]. Based on the information

G. K. Verma (✉)

Amity University Madhya Pradesh, Gwalior 474005, India
e-mail: girrajv@gmail.com; gkverma@gwa.amity.edu

D. Mishra

Department of Mathematics, Maulana Ajad National Institute of Technology, Bhopal 462003, India
e-mail: dheerendra@manit.ac.in

N. Kumar

Department of Computer Science and Engineering, Thapar University, Patiala, India
e-mail: neeraj.kumar@thapar.edu

received from UAVs, TCO can update the functioning of the traffic management system. Thus, real-time feedback from UAVs causes a real-time improvement in ITS [7].

In this IoD environment, UAVs work like moving nodes in ad hoc networks. However, due to limited storage and power capacity, UAVs are resource-constrained devices. To utilize their resources in optimized manner, UAVs collect the observations from their current locations and send to TCO using a local roadside unit (RSU). This RSU is considered to have a larger space and more energy resourceful device than UAVs. Thus, some part of computation and storage is done by RSU like an edge device. In this IoD-based communication, UAVs share the information to RSU using wireless channels. Further, RSU sends the aggregated information to TCO using wired or wireless links [8].

As most of the communication links in IoD are open channels. Therefore, an unauthorized attacker can easily target the information shared. It can capture, alter, or destroy the sensitive information between UAV-2-GS communication. Sometimes, this attack on shared information can cause a serious threat. For example, attacker can modify the road condition information and send to TCO. This modified information can misguide TCO and result may be a traffic congestion [9]. As we know that traffic congestion results in a high transportation cost. Thus, the manufacture, transport company, or consumer will be in loss. Therefore, the shared information/observations should be secured from such attacks. The security in this context can be achieved by authentication and confidentiality of the data.

1.1 Signcryption and Aggregated Signcryption

To achieve authentication and confidentiality simultaneously, the paradigm of signcryption has been devised in [10]. This pioneering work by Yulian Zhang reduces the cost of encryption and then signature approach by fusing these two operations. Thus, it is suitable to deploy resource-constrained IoD-based UAV-2-GS communication. In UAV-2-GS communication, several UAVs lying in a certain region share the information to a specified TCO. Therefore, the data received from various UAVs should be processed in an efficient manner. The meaning is that the verification and decryption of received data should be performed in a single step like batch verify. To achieve the batch verification in signcryption, Selvi et al. [11] proposed the first identity-based aggregated version of signcryption. However, their scheme is utilizing costly pairing operations. Thus, it can be improved further by removing the use of pairing. In [12], Wang et al. devised a new aggregated signcryption scheme using the paradigm of multilinear maps. This scheme was the first secure in standard model. However, it has no discussion about efficiency. Further, to improve efficiency, Swapna and Reddy [13] proposed an efficient aggregated signcryption. However, still the devised construction was based on pairing. Thus further improvement can be made possible. The first pairing less identity-based aggregated signcryption has been devised in [14] by Abouelkheir and El-sherbiny. As authors have removed the use of pairing,

the scheme is more efficient than previous literature. Later, some more aggregated signcryptions with more features have been devised in the literature [15–18].

1.2 Motivation and Contribution

According to the discussion, in smart city environment, utilization of ITS is the imperative need for smooth transportation. As the functioning of ITS is associated with the data received from various UAVs. Therefore, the security of communication links between UAV-2-UAV and UAV-2-GS is highly important. To secure the links, various key agreement and authentication protocols like [6–9, 19, 20] have been designed in the literature. However, in the case of UAV-2-GS link, several UAVs share information to a single TCO. Thus, to save resources at the receiving end (i.e., TCO), the verification/recovery of the received information should be done by using batch verify. The batch verify facility cannot be availed using key agreement. Therefore, key agreement schemes are insufficient to secure UAV-2-GS links. To secure these links, an efficient and secure data aggregation scheme is required. Therefore, in this paper, an efficient identity-based secure data aggregation scheme for UAV-2-GS communication has been devised. According to our sources (i.e., Internet or literature), the proposed scheme is the first scheme to secure UAV-2-GS communication in smart city ITS scenario.

The outline of the paper is as follows: next Sect. 2 presents the base definitions of foundation and related points. Section 3 introduces the proposed scheme and Sect. 4 discusses the security and efficiency analysis in brief. Section 5 concludes the paper along with future directions.

2 Preliminaries

This section introduces the basic concepts on mathematics and data aggregation in brief.

2.1 Mathematical Background

Let p and q be two primes selected randomly such that $p|(q - 1)$. Suppose E be the elliptic curve defined over the finite field F_p^* and P be the generator of E . Then the following problems are defined as the base of the construction.

- *Computational Diffie–Hellman Problem (CDHP)*: Given an instance (P, aP, bP) of three elliptic points for random unknown $a, b \in F_p^*$ and it is computationally

hard to find abP . The advantage of an algorithm \mathcal{A} to solve CDHP is the probability $Pr[abP \leftarrow \mathcal{A}(P, aP, bP)]$.

- *Discrete Log Problem (DLP)*: Given an instance (P, aP) of two elliptic points for random unknown $a \in F_p^*$ and it is computationally hard to find a . The advantage of an algorithm \mathcal{A} to solve DLP is the probability $Pr[a \leftarrow \mathcal{A}(P, aP)]$.

The security of the proposed data aggregation relies on these computationally hard problems.

2.2 Security Attributes of the Proposed Scheme

For the proposed data aggregation between UAV-2-GS communication, the following security attributes should be considered:

- *Authentication and Integrity*: In the UAV-2-GS communication, authentication of the sender UAV (i.e., source) and data integrity is important.
- *Confidentiality*: Confidentiality of the information shared is another important attribute.
- *Man-in-the-middle (MITM) Attack*: The consideration of MITM attack is also important.

A detailed discussion regarding definition and achieving the goals will be considered in Sect. 4.2.

2.3 Threat Model

To achieve the security attributes, the semantic security along with the unforgettability of the base signcryption should be considered [21–23]. In the current settings, two types of attackers have been defined. The Type-I attacker is an honest but curious KGC. This attacker has access to master secret, however not able to replace key of a drone (user). Another attacker is Type-II, who is malicious drone (user). It has no access to master secret. The proposed scheme is said to be secure against these attackers, if no attacker wins the attack games defined in [14] corresponding to provable security. These games are played between the attacker and the challenger. In the attack games, the two types of attackers has been permitted to put requests to *Key-Gen*, *Signcryption*, and *Designcryption* oracles. The challenger can access the oracles to respond the requests. At last, the challenger can design an algorithm to solve the CDHP or DLP (for a challenged instance). For a detailed description on provable security and various attack games, please refer [14].

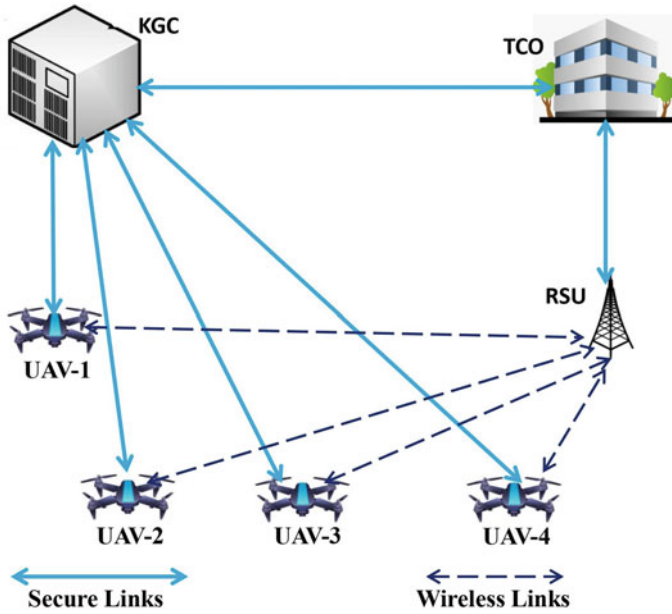


Fig. 1 System model

2.4 System Model

In the devised scheme, four entities, Key Generation Center (KGC), Traffic Control Office (TCO), UAVs, and Road Side Unit (RSU), are involved (Fig. 1). KGC generates the system parameters and keys of all the users. Generally, KGC is a UAV manufacturer company who stores the data in UAV before installation. KGC and TCO can communicate with each other using secure links. TCO is responsible for smooth functioning of ITS. For this purpose, TCO receives the data from all UAVs via RSU and use it for ITS improvement. The links between RSU-to-TCO are wired (Internet-based) links. The work of RSU is to aggregate the data received from various UAVs lying in its range. The links between UAV-2-RSU are the wireless open channels. Therefore, the communication done by using these links is the most insecure. Thus, the purpose of the proposed data aggregation is to secure this communication and to perform an efficient verification at TCO end.

3 Proposed Data Aggregation for UAV-2-GS Communication

The proposal is a modified version of the signcryption devised in [14]. The detailed steps of the scheme are follows:

– **Initialization:** KGC runs it by input a security parameter λ and obtains the outputs as

1. Two random primes p and q such that $p|(q-1)$.
2. An order p subgroup \mathcal{G} of elliptic curve defined by $y = x^3 + ax + b$ (where $4a^3 + 27b^2 \neq 0 \pmod{p}$) over \mathbb{Z}_p^* . P be a generator of \mathcal{G} .
3. Five hash functions $H_0 : \{0, 1\}^* \rightarrow \mathbb{Z}_q^*$, $H_1 : \{0, 1\}^* \times G \rightarrow \mathbb{Z}_q^*$, $H_2 : G \rightarrow \mathbb{Z}_q^*$, $H_3 : \{0, 1\}^* \times \{0, 1\}^* \times G \times G \times \{0, 1\}^* \times G \rightarrow \mathbb{Z}_q^*$, $H_4 : \{0, 1\}^* \times \{0, 1\}^* \times G \times G \times \{0, 1\}^* \times G \times \mathbb{Z}_q^* \rightarrow \mathbb{Z}_q^*$.
4. A random $s \in \mathbb{Z}_q^*$ as master secret key and $P_{\text{pub}} = sP$ as master public key.

Final output is $params = (p, q, G, P, P_{\text{pub}}, H_0, H_1, H_2, H_3, H_4)$.

– **Key-Gen:** Suppose, ID_i be the identity of UAV_i . KGC runs it by input $params$, ID_i and s . The steps are follows:

1. For $x_i \in_R \mathbb{Z}_q^*$, computes $X_i = x_i P$, $S_{ID_i} = sH_0(ID_i) \pmod{q}$, $q_i = H_1(ID_i, X_i)$ and $d_i = (x_i + sq_i) \pmod{q}$.
2. Secret key of UAV_i is (S_{ID_i}, d_i) and public key is X_i .

KGC sends secret keys to UAV_i via secure link.

– **Data-Aggregate:** It is done in two steps.

• **Step 1:** For the message $m_i \in \{0, 1\}^*$, the following steps are done by UAV_i :

1. Selects $r_i \in_R \mathbb{Z}_q^*$ and computes $R_i = r_i P$ and $W_i = r_i H_0(ID_{\text{tco}}) P_{\text{pub}}$.
2. Computes $h_{2i} = H_2(W_i)$, $h_{3i} = H_3(m_i, ID_i, X_i, W_i, ID_{\text{tco}}, X_{\text{tco}})$ and $h_{4i} = H_4(m_i, ID_i, X_i, W_i, ID_{\text{tco}}, X_{\text{tco}}, h_{3i})$.
3. Computes $v_i = (r_i h_{3i} + d_i h_{4i}) \pmod{q}$ and $V_i = v_i P$.
4. Computes $C_i = (m_i \| v_i) \oplus h_{2i}$.

After this, UAV_i sends $\sigma_i = (C_i, R_i, V_i)$ to RSU.

- **Step 2:** RSU receives data from n UAVs and computes $V = \sum_{i=1}^n V_i$.
- RSU forwards $\sigma_{\text{agg}} = ((C_1, C_2, \dots, C_n), (R_1, R_2, \dots, R_n), V)$ to TCO as aggregated data.

– **Verify-Decryption:** The following steps are done by TCO:

1. For $1 \leq i \leq n$, compute $W_i = S_{ID_{\text{tco}}} R_i$ and recover $m_i \| v_i = C_i \oplus H_1(W_i)$.
2. Checks $V = \sum_{i=1}^n h_{3i} R_i + \sum_{i=1}^n h_{4i} X_i + (\sum_{i=1}^n h_{4i} q_i) P_{\text{pub}}$.

If equation holds, accept the ciphertexts as valid.

4 Security and Efficiency Discussion

4.1 Correctness

From the construction of the scheme, $V = \sum_{i=1}^n V_i$, where $V_i = (r_i h_{3i} + d_i h_{4i})P$, $h_{3i} = H_3(m_i, ID_i, X_i, W_i, ID_{\text{tco}}, X_{\text{tco}})$ and $h_{4i} = H_4(m_i, ID_i, X_i, W_i, ID_{\text{tco}}, X_{\text{tco}}, h_{3i})$. Therefore, $V = \sum_{i=1}^n (r_i h_{3i} + d_i h_{4i})P$, i.e., $V = \sum_{i=1}^n h_{3i} R_i + \sum_{i=1}^n h_{4i} X_i + (\sum_{i=1}^n h_{4i} q_i)P_{\text{pub}}$ as $d_i = (x_i + s q_i) \bmod q$. Thus, the Verify-Decryption runs correctly.

4.2 Security Attributes Analysis

As per the discussion in Sect. 2.2, the proposed scheme satisfies the following attributes:

- *Authentication and Integrity*: In the designing of the protocols, during *Data-Aggregate* phase, each UAV_i computes ciphertext $C_i = (m_i || v_i) \oplus h_{2i}$, where $v_i = (r_i h_{3i} + d_i h_{4i}) \bmod q$. This computation is possible with secret key d_i of the UAV_i . At the receiver end, i.e., TCO, verification needs to check $V = \sum_{i=1}^n h_{3i} R_i + \sum_{i=1}^n h_{4i} X_i + (\sum_{i=1}^n h_{4i} q_i)P_{\text{pub}}$. This step is possible only with public key X_i of UAV_i and secret key $S_{ID_{\text{tco}}}$ of TCO. Thus, the generation of ciphertext can be done by legitimate UAV_i only. From the discussion in [14], the base scheme is unforgeable, and therefore authentication is satisfied. For verify purpose, secret key of TCO is needed, so alteration of message is not possible, i.e., integrity is also satisfied.
- *Confidentiality*: As per the scheme [14], the encryption is semantically secure. Therefore, an adversary is unsuccessful to get any observation from the ciphertext. Thus, confidentiality is also satisfied.
- *MITM Attack*: As the base scheme [14] is unforgeable against an adaptively chosen message attack. Besides, the possible alteration to ciphertext will result in rejection during verification/decryption process. Thus, no adversary can impersonate the signer or cannot modify the content. Thus, the scheme is secure against MITM attack.

4.3 Efficiency Analysis

The computational costs of various cryptographic operations have been referred from [24] (shown in Table 1). In the literature, a limited resource device single 798 MHz CPU has been utilized with 256 MB RAM support. Thus, it can be a good choice to emulate an UAV capacity. Based on the discussion, UAVs are resource-

Table 1 Computation costs of various cryptography operations [24]

Operation	OBU/RSU (ms)
Bilinear pairing	67.32
Modular exponentiation	7.87
Modular multiplication	21.63
Hash	0.025
Pairing multiplication	21.63
Scalar multiplication	14.83
Map to point hash	5.23
Elliptic point addition	4.61

constrained devices. UAVs are having less storage capacity, less computation capacity, and limited power backup. Thus, the computation done by UAVs is analyzed. During the ciphertext generation phase, 3 scalar multiplications + 3 modular multiplications + 3 hash functions are computed. The total cost of these operations is $3 \times 14.83 + 3 \times 21.63 + 3 \times 0.025 = 109.455$ ms. Thus, it is not a very big computation time for a resource-constrained device like UAV. If the computation overhead of RSU is considered, it is $\approx (n - 1)4.61$ ms. As RSU is stronger than UAV, it is adjustable computation. Similarly, the cost incurred by TCO is $\approx (66.195n - 6.8)$ ms. It is also not a large consumption time for TCO as it has infinite resources. Therefore, the proposed scheme is practically suitable for UAV-2-GS communication with respect to computational efficiency.

5 Conclusion

Based on the various reports, it is observed that road accidents and traffic congestion are big losses to the world economy. In smart city environment, UAVs are utilized to get real-time traffic data. This data enhances the functioning of ITS by inserting the feedback analysis. However, the links between UAV-to-TCO are wireless. Thus, a secure data aggregation based on a signcryption scheme has been proposed in this paper. The security and efficiency analysis presents the suitability of the proposal for UAV-2-GS communication.

As a future scope, data aggregation for multiple applications using UAV-2-GS communication should be devised.

References

1. Misra MK, Chaturvedi A, Tripathi SP, Shukla V (2019) A unique key sharing protocol among three users using non-commutative group for electronic health record system. *J Discret Math Sci Cryptogr* 22(8):1435–1451
2. Shukla V, Chaturvedi A, Misra MK (2021) On authentication schemes using polynomials over non commutative rings. *Wirel Person Commun* 118(1):1–9
3. WHO, Global status report on road safety. <https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries#:~:text=Every%20year%20the%20lives%20of,a%20result%20of%20their%20injury>. Accessed 08 May 2022
4. European Commission (2021) Road safety thematic report—Fatigue. European Road Safety Observatory. Brussels, European Commission, Directorate General for Transport. https://ec.europa.eu/transport/road_safety/system/files/2021-07/asr2020.pdf. Accessed 08 May 2022
5. Raya M, Hubaux J-P (2005) The security of vehicular ad hoc networks. In: Proceedings of the 3rd ACM workshop on security of ad hoc and sensor networks, pp 11–21
6. Gope P, Sikdar B (2020) An efficient privacy-preserving authenticated key agreement scheme for edge-assisted internet of drones. *IEEE Trans Veh Technol* 69(11):13621–30
7. Alladi T, Chamola V, Kumar N (2020) PARTH: a two-stage lightweight mutual authentication protocol for UAV surveillance networks. *Comput Commun* 160:81–90
8. Alladi T, Bansal G, Chamola V, Guizani M (2020) SecAuthUAV: a novel authentication scheme for UAV-ground station and UAV-UAV communication. *IEEE Trans Veh Technol* 69(12):15068–77
9. Srinivas J, Das AK, Kumar N, Rodrigues JJ (2019) TCALAS: temporal credential-based anonymous lightweight authentication scheme for Internet of drones environment. *IEEE Trans Veh Technol* 68(7):6903–16
10. Zheng Y, Digital signcryption or how to achieve cost (signature & encryption) \ll cost (signature)+ cost (encryption). In: Annual international cryptology conference. Springer, Berlin, Heidelberg, pp 165–179
11. Selvi SSD, Vivek SS, Shriram J et al (2009) Identity based aggregate signcryption schemes. In: Progress in Cryptology—INDOCRYPT, Lecture Notes in Computer Science, Poland, pp 378–397
12. Wang H, Liu Z, Liu Z, Wong DS (2016) Identity-based aggregate signcryption in the standard model from multilinear maps. *Front Comput Sci* 10(4):741–54
13. Swapna G, Reddy PV. Efficient identity based aggregate signcryption scheme using bilinear pairings over elliptic curves. *J Phys: Conf Ser* 1344(1):012010 (IOP Publishing)
14. Abouelkheir E, El-sherbiny S (2020) Pairing free identity based aggregate signcryption scheme. *IET Inf Secur* 14(6):625–632
15. Shukla V, Srivastava N, Chaturvedi A (2016) A bit commitment signcryption protocol for wireless transport layer security (wtls). In: IEEE international conference on electrical, computer and electronics engineering, pp 83–86
16. Yang X, Zhou H, Ren N, Tian T. Homomorphic proxy re-signcryption scheme and its application in edge computing-enhanced IoT. In: 2021 2nd international conference on electronics, communications and information technology (CECIT). IEEE, pp. 644–649
17. Yu H, Ren R (2021) Certificateless elliptic curve aggregate signcryption scheme. *IEEE Syst J* 16(2):2347–54
18. Yang Y, He D, Vijayakumar P, Gupta BB, Xie Q (2022) An efficient identity-based aggregate signcryption scheme with blockchain for IoT-enabled maritime transportation system. *IEEE Trans Green Commun Netw*
19. Zhang Y, He D, Li L, Chen B (2020) A lightweight authentication and key agreement scheme for internet of drones. *Comput Commun* 15(154):455–64
20. Wazid M, Das AK, Kumar N, Vasilakos AV, Rodrigues JJ (2018) Design and analysis of secure lightweight remote user authentication and key agreement scheme in internet of drones deployment. *IEEE Internet Things J* 6(2):3572–84

21. Chaturvedi A, Shukla V, Misra MK (2021) A random encoding method for secure data communication: an extension of sequential coding. *J Discrete Math Sci Cryptogr* 24(5):1189–1204
22. Shukla V, Chaturvedi A, Srivastava N (2019) A new one time password mechanism for client-server applications. *J Discrete Math Sci Cryptogr* 22:1393–1406
23. Shukla V, Misra MK, Chaturvedi A (2021) A new authentication procedure for client-server applications using HMAC. *J Discrete Math Sci Cryptogr* 24(5):1241–1256
24. Verma GK, Gope P, Kumar N (2022) PF-DA: pairing free and secure data aggregation for energy internet-based smart meter-to-grid communication. In: *IEEE Trans Smart Grid* 13(3):2294–2304

BBIWMS: A Secure Blockchain-Based Framework for Integrated Water Management System for Smart City



B. C Girish Kumar, K. G. Harsha , G. Mahesh , Varun Shukla ,
and Surendra Talari 

1 Introduction

15.2 million hectares [1] (mha) of canal-irrigated land were part of the 28.2 million hectares (mha) of undivided India [2] that was used for irrigation before it became independent. India lost the irrigation [3] land that belonged to Pakistan after the country was divided into two halves known as India and Pakistan (Bharadwaj 1990). In 1949–1950, there were 2 million tons of irrigation in total, or around 62 million tons. India needed to become self-sufficient in the production of agriculture [4] and not rely on other nations for the production of irrigation, which developed 5-year plans (Bharadwaj 1990; Vohra 1995). The many small, medium, and large irrigation schemes at the period led to increased crop production ranging from 21 mha to 46.2 mha between 1951 and 1991 (Vohra 1995), which sparked a 2.42% yearly growth in food production [5] with the goal of reaching 180 million tons by 1995. The food-grain irrigation system grew at a rate of 3.3% from 1964–1965 to 1970–1971; as a result, the irrigation area increased.

We've discovered recently that India's rapidly expanding population is to blame for the country's rising water consumption. Due to the large number of people

B. C. G. Kumar (✉)
SJB Institute of Technology, Bengaluru, India
e-mail: girishshekar.89@gmail.com

K. G. Harsha · G. Mahesh
JSS Academy of Technical Education, Noida, UP, India
e-mail: harshakg@jssaten.ac.in

G. Mahesh
e-mail: mahesh@jssaten.ac.in

V. Shukla
Pranveer Singh Institute of Technology, Kanpur, UP, India

S. Talari
GITAM Deemed to Be University, Vishakhapatnam, AP, India

moving to cities from villages for a variety of reasons, including education, business, and occupation, there is a greater demand for drinking water as well as water for home use and food production [6]. Water was needed as industry expanded across India in order to produce completed items and get rid of the pollution they were absorbing. In many arid and semi-arid regions, where groundwater is heavily used and results in significant drops in water levels as well as the deterioration of groundwater quality and further reduction of good quality of groundwater [4], the water resources available from surface and groundwater are insufficient for water supplies. Surface water is used extensively in several basins practically everywhere. Industrial water disposal practices represent a risk of groundwater and surface water resource pollution. Water pollution is a side effect of how municipal garbage from industrial locations is disposed of. India had access to around 6008 m³ of high-quality water per person annually at the time of its independence [6].

India's water resources faced two challenges: a lack of freshwater due to diminishing natural supplies, and increased demand for water for irrigation, drinking water, industrial output, and ecosystem management (Ballabh et al. 1999). The second issue is data sharing disagreements, which affect how data is needed for demand and supply. In order to develop a water management strategy, we need a scientific database of data on water supply in the social range. The world's population is under threat due to a lack of available water and pressure from water-related issues. As a result, an overview ensures that the board can withstand the weight of water. [4] As a result, well-programmed frameworks are required here; there is a need for programmed, strong, empowered frameworks that can be used for water measurement, resulting in increased cultivation and plants [5]. The expected security engineering for water management with strong frameworks employs existing and configured system security [4]. We require optimized water management technology with minimal data duplication. India now has a centralized system for managing water resources; however, there are drawbacks like high costs, poor performance, and unsafe data storage. To overcome these water management challenges, some of the researcher's work also comes with a solution technology called IOT (Internet of Things) but with some limitations.

2 IoT (Internet of Things)

It involved IP-based networking technologies with wireless communication between the sensor actuator linked devices, used in domains such as cloud computing, data analytics, and Moore's law [6], contains people, machines, and information, and has a significant impact on society in a number of ways. Given the variety of views and deployment scales, it is challenging to provide a comprehensive description of IoT, as stated in [7].

The first technique is small-scaled IoT deployment; in this method, devices/things can be uniquely identified and have sensing, actuating, and programming capabilities. The data are widely gathered and can be changed from anywhere in the world.

The second technique is large-scale IoT deployment, in which items or devices may adapt to complex networks by employing standard communication protocols. Things can be identified in the actual or virtual worlds and be capable of sensing and actuating with programmable information of status, location, or other business information without the use of any human interfaces. It is accessed via clever interfaces and may be accessed from anywhere and at any time. Using IoT in water management, leaks can be quickly recognized, water quality can be detected, water monitoring security can be maintained, transparency in water distribution can be easily achieved, and water infrastructure can be predicatively maintained. Water management can be accomplished more efficiently with the help of a smart metering system.

In addition to IoT, there are certain difficulties with security, regulatory, and legal issues.

Standards Security and Privacy Issues: The data generated and communicated digitally necessitate the highest level of security, and data should not be divulged; instead, it should be sent in a private format. Cyber attackers can target data, such as DDOS attacks, data hacking, data theft via the internet, and remotely data hijacking.

Interoperability and Standardization: To build an IoT ecosystem, a variety of combining technologies including sensor, wireless communication, embedded, cloud, and virtualization technologies are used.

Regulatory, Legal, and Right Issues: This concern relates to the continual generation of data by IoT devices, which, if stored and used across international jurisdictions, could lead to information misuse, violations of data user rights, and legal liability. **Emerging Economy and Developmental Issues:** We can expect improvements in economic growth, political power, and social prestige as a result of IoT for national development. To overcome the above challenges, we have come up with a solution technology called Blockchain.

3 Water Quality Detection Using Machine Learning and IoT

The study offers a cheap method for cleaning up water pollution in household overhead tanks. IoT devices are utilized to evaluate the quality of the water, and machine learning algorithms are employed to foretell possible issues with water contamination. The suggested system consists of various sensors that are interfaced with NodeMCU to acquire water characteristics. Before the water becomes tainted, the user is informed. The adopted method is both cost-effective and effective at preventing

water pollution. The project's future objectives include identifying illnesses brought on by a variety of reasons and developing the best strategy for tank cleaning [8].

4 Blockchain

The Blockchain is the most recent technology [9], as described in Nakamoto's article "Bitcoin: A P2P Electronic Cash System (2008)." The Blockchain is decentralized, which means that the data cannot be limited to one system and can be distributed across all systems. It is also immutable, which means that once the data are entered, it cannot be changed by anyone, so the ledger created cannot be disclosed, but we can add new contents to that ledger. According to the time sequence, [10] Blockchain contains distributed and linked chains. The cryptography technique used in Blockchain ensures that data are transferred within a time sequence. The Blockchain is used to store and transmit data securely, consistently, and impenetrably. The block is made up of a head and a body. The block header [9] contains address-based hashing information for the block body as well as metadata about the current block. The header also includes the hash [11] of the previous block, which allows for the collection of previous transaction information with timestamps using the Merkle root hash. The Blockchain's body contains information about the data recorded in the block. The following are the four characteristics of Blockchain (Table 1).

5 Related Work

5.1 *Blockchain Technology with Conflicts*

In water management [14], there are many disputes occurring in the area of water preservation and management. Apart from these disputes, shortage of water is also one more top challenge among others, so it is required to manage the water we have distribution effectively [15]. In India, agriculture is the main backbone for the country, which purely depends on water availability. The [4] irrigation methods such as farming, cultivation, heavy irrigation that can harm and misuse the agricultural land, which tend to make it more drought-prone. If sufficient water is not available for irrigation then Droughts result in less crop production, which results in human life purely, depends on agricultural foods. The people cannot sustain without food and drinking water. So, it is required to maintain a proper water management system [16] to utilize the water efficiently.

Water is a basic commodity required by everybody; by using this resource certain sectional people [17] make it tradable for earning profit and beignets in the society. Governments and municipal corporations, which are centralized institutions, control

Table 1 Blockchain Characteristics [9]

Serial number	Name	Description
1	Data decentralization [12]	A large number of nodes are connected in a peer-to-peer network. Centralized devices are not present, so data belonging to chaining cannot be stored in the same place at the same time
2	No data falsification[13]	Data are distributed across all nodes in Blockchain technology, which prevents others from intervening in data. As a result, attackers find it difficult to attack the data in Blockchain because the data are encrypted
3	Traceability[14]	In a Blockchain technology, the number of nodes is linked cryptographically, resulting in time-stamped blocks with data storage that can be easily traceable. Here, transaction information is recorded fully and can be searched from any node, as well as create queries for any block information for accessing from the Blockchain, resulting in an interaction that can be seen without hiding. The data are open and transparent, making it simple to check the transaction records entered at each node. This kind of feature can facilitate communication and trust between nodes
4	Programmability	The Blockchain also supports on-chain scripting for the creation of application-layer services. Here, Blockchain users can develop smart contracts that offer security for automated activities and transactions. Smart contracts are a type of optimized and embedded algorithm that can be identified by both parties working in a Blockchain. As a result, Blockchain is widely used in smart contract applications. Any transaction that occurs between a buyer and a seller can be entered as an online transaction on nodes

the production and distribution of water. The municipal corporation will distribute the water in their jurisdiction; here customers are forced to pay the water bill based on their quantity of water consumed. The customer did not meet their demand even though they paid the specified bill. The water quality [18] is not meeting up to the mark. Less than one percent of the earth’s water is unsalted and available for human consumption. The desalination process is required to get the purified water. The water may contain lead micropollutants, organic matter available naturally. The waste released by factories leads to contamination of the water resources. If such water is not treated properly so that causes polluted and contaminated water supply and distribution throughout. All these types of issues can be managed by the Blockchain efficiently and eliminate the centralized water trades and also avoid collecting excess amounts from customers and more benefits can be obtained from the Blockchain.

5.2 Blockchain Technology Used to Overcome the Conflicts

Blockchain technology can be used to create a decentralized, open system that distributes water equally throughout several geographic [19] regions. In fact, it may be possible to totally do away with any centralized organizations or intermediary distribution companies that are only in charge of supplying water. An approach would be to create a Blockchain-based platform for uncomplicated engagement between parties and peer-to-peer communication [14] to talk about water allocation, preservation, and distribution. On this platform, participants from businesses, governments, and consumers would all be present. The Blockchain eliminates the need for unreliable third parties and allows for the establishment of a more transparent and equitable system for determining water pricing [15, 20]. Transactions involving the amount of water given drunk, fees paid, and extra water can be made using the Blockchain platform.

6 Intelligent Water Management System Using Blockchain

Framework for a Water Resource Management System: The challenges of the current traditional architecture employed in water management systems will be overcome by Blockchain-Based Integrated Water Management Systems (BBIWMS), as depicted in Fig. 1 [14, 22]. These four components make up the confluence layer [17, 23], application layer, Blockchain network layer, and data accumulation and forwarding layer in the Blockchain framework for managing water resources. For distributed data storage and data sharing between nodes to prevent data tampering, Blockchain technology [21] is employed in the Blockchain network layer and the application layer. The smart contract [14] code is used to automatically identify issues and provide people incentives or punishments. Layer for Information Gathering [8, 22]. This layer, which depends on the monitoring system, offers a range of services for gathering and transmitting real-time data about water resources and integrating it with existing databases.

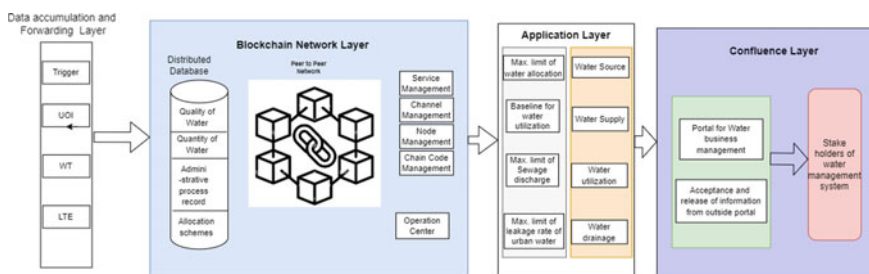


Fig. 1 Blockchain-based integrated water management systems (BBIWMS)

Blockchain Network [23, 24]r Each node examines the outcomes on the block, including water suppliers, waterworks, water consumers, and sewage treatment plants [25]. Each node also has data on water resource information. Using time stamps, a succession of Blockchain are built, and a P2P network is established between the nodes. The block can also be used to document business operations [14] associated with managing water resources. Planning frameworks and index criteria are kept in the block as fundamental information. Data uploads data transfers between nodes, data queries, and the creation of smart contract rules are all handled by the Blockchain network layer. To implement the Blockchain service functions and simplify the management of the water sector, the operation and maintenance of the Blockchain are carried out in cooperation with internet service providers.

To manage water sources, water supply, water utilization, and water discharge in accordance with the requirements for information service, business management, responsibility, traceability, scheduling, configuration, emergency management, and other functions [14], four application modules are established in this layer of application (layer of application [25]).

This tactic supports the three red-line management systems (the upper limit of water resource allocation [2], the baseline of the utilization efficiency of water resources, and the upper limit of sewage discharge).

6.1 Security method

- This layer contains numerous tools, such as triggers and others, that permit forwarding data to subsequent layers, such as the Blockchain layer, where the hashing function can be employed to preserve the uniqueness of accumulated data. This layer stores data on accumulating or storing water. Such information about water quality can be stored in nodes.
- The gathered data in Blockchain technology [26] are transmitted through a number of stages for processing; in this situation, the hashing technique can be used to secure the data by encrypting it with cryptography. Only authorized people who require the data can access it after it has been decoded at the receiver's other end.
- Data accumulation and forwarding layers make up the top layer. This layer contains numerous tools, such as triggers and others, that permit forwarding data to subsequent layers, such as the Blockchain layer, where the hashing function can be employed to preserve the uniqueness of accumulated data. This layer stores data on accumulating or storing water. Such information about water quality can be stored in nodes.
- In Blockchain technology [27, 28], data are amassed and sent through various layers for data processing. Here, hashing technique can be utilized to secure data by encrypting it with cryptography. Only those with permission to access the data may disclose it, and at the other end of the receiver, it is decrypted using the hashing function. Therefore, information cannot be provided to unauthorized or outside parties. The data can be hashed as a single unique value in the levels that

are provided. In this case, everyone can see the public key for the shared data, but only the authenticated users can unlock it to continue editing or changing the data. Data that is being sent as can be encrypted using symmetric encryption. Because the time necessary for a transaction may be recorded, the data entered in the form of a node can be easily traced, resulting in interaction transparency. The data sharing ensures the trust and flexible. The Blockchain technology supports smart contracts, which results in embedding an algorithm optimization.

- With this technology, the application layer can keep track of information regarding the quantity, quality, and size of water needed. If there are any leaks, they should be simple to find and stop to save water waste.
- The confluence layer, which is the final layer in the given architecture, gives the essential information for the stakeholders who supplied funds for company. In this technology, the application layer can maintain the. It also provides information on water availability, water leaks or wastage, water released, and acceptable facts, which are all stored here on a daily basis, for the interested parties.

Multimedia Layer following system implementation, this layer offers a unified integration platform and information gateway. The Water Management Application Platform, a Blockchain-based architecture for managing water resources, consists of four functional modules: Water Source Management, Water Supply Management, Water Utilization Management, and Water Discharge Management. Each management unit may select the suitable function module based on the set permissions. Because distributed data storage enables each node to share the global data, one-sided judgments are avoided. Through the consensus process, information consistency is assured. Thanks to the hashing process, information is exceedingly difficult to alter, and asymmetric encryption [14, 29] provides excellent security. As a result, confidence between the departments and nodes grows. The autonomous water management function of each node, enabled by the flexible smart contract script, ensures high system efficiency. Each node can connect to the Blockchain network and create an account after verification. The related link's management module becomes active when staff logs into the account. Each function module in the smart contract is written as a piece of code to facilitate data requests, information retrieval, analysis, and reasoned decision-making. The application tasks for the four modules are listed (Fig. 2).

Module for Managing Water Sources [3] this module's objective is to guarantee sufficient water supply in both quantity and quality. Reservoirs serve as sources of drinking water; hence their water volume affects the cities' water supply. The sensors supply the block of the water supply node with time-stamped real-time data on reservoir water levels. Information on reservoir water level [15, 30] is compared with control indices, such as the drought limit water level, using an embedded data analysis tool, and the results are uploaded to the Blockchain for storage. Making decisions is another component of the module for managing water sources. Using optimization algorithms, knowledge reasoning, and other technologies, the water supply plan is improved in accordance with the evaluation results and additionally, the block of the node receives for storage the data on the water quality of the water sources

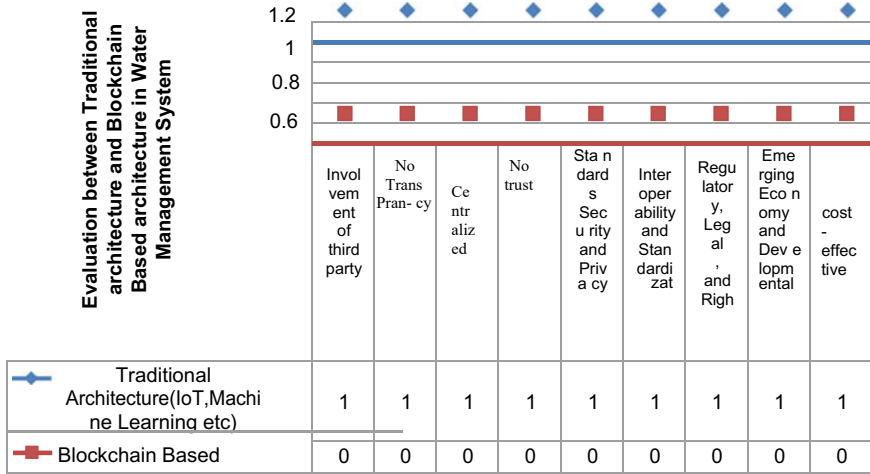


Fig. 2 Evaluation between traditional architecture and Blockchain-based architecture in water management system

collected from sensors. Smart contracts are used to compare the measured data to the water quality goals before sending the evaluation findings to the Blockchain. Making decisions is another task performed by this module [16, 31]. The quality of the water is monitored, and the results are communicated in real-time, in order to promptly spot cases of water contamination. The linkage of real-time monitoring data for reservoir water quantity and quality allows for quick and scientific changes to water delivery plans. In order to ensure the consistency and dependability of water abstraction permits, supply storage and verification services, and enhance the security of electronic licenses, Blockchain technology [17, 32] is being used to administer electronic licenses. In order to prevent manipulation, Blockchain technology also makes sure that electronic licenses and data are preserved permanently on the chain. It is practical and efficient to verify certificates using this method. The Blockchain also keeps a thorough record of each license authorization, ensuring that the data can be tracked and preventing theft. The capacity to track every action has virtually eliminated the public’s security concerns.

7 Conclusion

In light of problems with trust-building, inadequate management efficacy, and high data storage risk in conventional centralized water management, this study investigated the use of Blockchain technology for enhancing water management. It also proposed a framework called a Blockchain-based integrated water management system (BBIWMS). High managerial effectiveness, high trust, and secure data storage are all features of this suggested system. To control water sources, water

supply, water use, and water discharge, it has four functional modules. The advantages of this suggested system are high information transmission efficiency, reliable and secure storage of data about water resources, and excellent traceability of issues with water quality. By solving the problems listed below, we will implement the suggested framework model (BBIWMS) in the upcoming study effort. Data exchange across departments, legal issues with the Blockchain, and the absence of a standardized system are some of the difficulties that come with integrating Blockchain technology into water management systems. It is crucial to challenge conventional wisdom and modify corporate operations to the Blockchain platform. In order to make the necessary adjustments, it is also essential to connect with current information systems and promote the use of Blockchain technology.

Acknowledgements Our goal is to show his holiness “**Jagadguru Sri Shivarathri Deshikendra Mahaswamiji.**” Pontiff of JSS Suttur Math. We convey our heartfelt thanks for their continual exceptional encouragement and support from our Department Head, Prof. (Dr). Mayank Singh and family members for completing this research work.

Funding No funding.

Conflict of Interest The authors affirm that our interests are aligned.

Ethical Approval We, the writers, are not narrowing our investigation in this publication to a certain subject or animal.

References

1. Dogo EM, Salami AF, Nwulu NI, Aigbavboa CO (2019) Blockchain and internet of things-based technologies for intelligent water management system, pp 129–150. https://doi.org/10.1007/978-3-030-04110-6_7
2. Paramathma MK, Pravin AC, Rajarajan R, Velmurugan SP (2019) Development and implementation of efficient water and energy management system for indian villages. In: IEEE Int Conf Intell Tech Control Optim Signal Process. INCOS 2019, pp 1–4. <https://doi.org/10.1109/INCOS45849.2019.8951362>
3. Bordel B, Di. Martin, Alcarria R, Robles T (2019) A blockchain-based water control system for the automatic management of irrigation communities. In: 2019 IEEE Int Conf Consum Electron. ICCE 2019, no. 1, pp 1–2. <https://doi.org/10.1109/ICCE.2019.8661940>
4. Saad A, Benyamina AEH, Gamatie A (2020) Water management in agriculture: a survey on current challenges and technological solutions. IEEE Access 8:38082–38097. <https://doi.org/10.1109/ACCESS.2020.2974977>
5. Narendran S, Pradeep P, Ramesh MV (2017) An internet of things (IoT) based sustainable water management. In: GHTC 2017—IEEE Glob Humanit Technol Conf Proc., vol 2017-Janua, pp 1–6. <https://doi.org/10.1109/GHTC.2017.8239320>
6. Fatima M, Jain S, Chikara A, Luthra M (2019) Review on implementing smart water grid for smart cities in india: challenges and solutions, 2019 5th Int. Conf. Adv Comput Commun Syst ICACCS 2019:216–219. <https://doi.org/10.1109/ICACCS.2019.8728485>

7. Anjana S, Sahana MN, Ankith S, Natarajan K, Shobha KR, Paventhan A (2016) An IoT based 6LoWPAN enabled experiment for water management. *Int Symp Adv Networks Telecommun Syst ANTS 2016-Febru*: 1–6. <https://doi.org/10.1109/ANTS.2015.7413654>.
8. Kaur P, Parashar A (2022) A systematic literature review of blockchain technology for smart villages. *Arch Comput Methods Engineering* 29(4). Springer Netherlands. <https://doi.org/10.1007/s11831-021-09659-7>
9. Treiblmaier H (2020) Blockchain and tourism. *Handb. e-Tourism*: 1–21. https://doi.org/10.1007/978-3-030-05324-6_28-2
10. Tasatanattakool P, Techapanupreeda C (2018) Blockchain: Challenges and applications. *Int Conf Inf Netw 2018-Janua*: 473–475. <https://doi.org/10.1109/ICOIN.2018.8343163>
11. Feng Q, He D, Zeadally S, Khan MK, Kumar N (2018) A survey on privacy protection in blockchain system. *J Netw Comput Appl* 126(May 2018): 45–58. <https://doi.org/10.1016/j.jnca.2018.10.020>
12. Velliangiri S, Karthikeyan Karunya P (2020) Blockchain technology: Challenges and security issues in consensus algorithm. In: 2020 Int Conf Comput Commun Informatics, ICCCI 2020. <https://doi.org/10.1109/ICCCI48352.2020.9104132>
13. Ozdemir AI, Ar IM, Erol I (2020) Assessment of blockchain applications in travel and tourism industry. *Qual Quant* 54(5–6): 1549–1563. <https://doi.org/10.1007/s11135-019-00901-w>
14. Xia W, Chen X, Song C (2022) A framework of blockchain technology in intelligent water management. *Front Environ Sci* 10(June):1–12. <https://doi.org/10.3389/fenvs.2022.909606>
15. Kassou M, Bourekadi S, Khouliji S, Slimani K, Chikri H, Kerkeb ML (2021) Blockchain-based medicaland water waste management conception E3S Web Conf 234: 1–6. <https://doi.org/10.1051/e3sconf/202123400070>
16. Gurudas Vernekar A (2020) Blockchain based water management system. *Int Res J Eng Technol May*: 7505–7507, [Online]. Available www.irjet.net
17. Mahmoud HHM, Wu W, Wang Y (2019) Secure data aggregation mechanism for water distribution system using blockchain. In: ICAC 2019—2019 25th IEEE Int Conf Autom Comput., no. September, pp. 1–6. <https://doi.org/10.23919/ICoAC.2019.8895146>.
18. Tiwari S, Gautam J, Gupta V, Malsa N (2020) Smart contract for decentralized water management system using blockchain technology. *Int J Innov Technol Explor Eng* 9(5):2046–2050. <https://doi.org/10.35940/ijitee.e3202.039520>
19. Abu-Amara F et al (2022) A blockchain solution for water and electricity management. *Mater Today Proc* 63:731–736. <https://doi.org/10.1016/j.matpr.2022.05.106>
20. Pincheira M, Vecchio M, Giaffreda R, Kanhere SS (2020) Exploiting constrained IoT devices in a trustless blockchain-based water management system. In: IEEE Int Conf Blockchain Cryptocurrency, ICBC 2020. <https://doi.org/10.1109/ICBC48266.2020.9169404>
21. Shah J (2017) An internet of things based model for smart water distribution with quality monitoring. *Int J Innov Res Sci.* 6(3):3446–3451. <https://doi.org/10.15680/IJRSET.2017.0603074>
22. Kakkar M, Gupta V, Garg J, Dhiman S (2021) A detection of water quality using machine learning and IoT. *IJERT* 10(11). ISSN 2278–0181. <https://doi.org/10.17577/IJERTV10IS110022>
23. Girish Kumar BC, Nand P, Bali V (2022) Opportunities and challenges of blockchain technology for tourism industry in future smart society. In: 2022 Fifth international conference on computational intelligence and communication technologies (CCICT), pp 318–323. <https://doi.org/10.1109/CCICT56684.2022.00065>.
24. Girish Kumar BC, Nand P, Bali V (2022) Review on opportunities and challenges of blockchain technology for tourism industry in future smart society. In: Bali V, Bhatnagar V, Lu J, Banerjee K (eds) *Decision analytics for sustainable development in smart society 5.0*. Asset Analytics. Springer, Singapore. https://doi.org/10.1007/978-981-19-1689-2_16
25. Hakak S, Khan WZ, Gilkar G.A, Haider N, Imran M, Alkathairi MS (2020) Industrial wastewater management using blockchain technology : architecture, requirements and future directions. (June): 38–43

26. GKBC, Singh A, Patel U, Yadav A, Kumar A (2021) Tourist and hospitality management using blockchain technology. 11(08): 28670–28672
27. GKBC, Mahesha AM, Harsha KG (2021) A review on data storage retrieval using blockchain technology. 11(08):28630–28637
28. GKBC, Yadav S, Singh R, Gupta S, Singh S (2022) Review on driver drowsiness and fatigue detection system. 12(1): 29193–29196
29. Ahamed N, GKBC (2016) Information retrieval using CP-ABE. 6(7):1596–1600
30. Singh AN, Arya A, Sharma A, Girish Kumar BC (2022) Contact tracing for communicable diseases using blockchain. Int J Emerg Technol Innov Res 9(5):j528-j532. (www.jetir.orgUGC and issn Approved), ISSN:2349–5162. Two Papers published in UGC journal 2022
31. Girish Kumar BC, Gupta AK, Verma A, Chaudhary D, Pandey R (2022) An encrypted automatic multiple-choice question generator for self-assessment using natural language processing. Int J Emerg Technol Innov Res 9(5):j337-j343. (www.jetir.org). ISSN:2349–5162
32. Girish Kumar BC, Nand P, Bali V (2023) BBACTFM (Blockchain based accurate contact tracing framework model) for tourism industry. In: Shaw RN, Paprzycki M, Ghosh A (eds) Advanced communication and intelligent systems. ICACIS 2022. Communications in computer and information science, vol 1749. Springer, Cham. https://doi.org/10.1007/978-3-031-25088-0_46
33. Satiya N, Varu V, Gadagkar A, Shaha D (2017) Optimization of water consumption using dynamic quota based smart water management system. In: TENSYP 2017—IEEE Int Symp Technol Smart Cities. <https://doi.org/10.1109/TENCONSpring.2017.8070075>
34. GKBC, Nand P, Bali V BLOBDBM : Blockchain based framework for decentralized business model in tourism industry I. Introduction
35. Ismail S, Dawoud DW, Ismail N, Marsh R, Alshami AS (2022) IoT-based water management systems: survey and future research direction. IEEE Access 10:35942–35952. <https://doi.org/10.1109/ACCESS.2022.3163742>
36. Kumar BCG, Garg M, Saini J, Chauhan K, Malsa N, Malsa K (2023) Accurate rating system using blockchain. In: Shaw RN, Paprzycki M, Ghosh A (eds) Advanced communication and intelligent systems. ICACIS 2022. Communications in computer and information science, vol 1749. Springer, Cham. https://doi.org/10.1007/978-3-031-25088-0_44

An Intelligent Network Intrusion Detection Framework for Reliable UAV-Based Communication



Sujit Bebortta and Sumanta Kumar Singh

1 Introduction

Unmanned aerial vehicles (UAVs) have recently gained popularity, which has paved the way for their deployment in a variety of industries, including applications for border security, crowd surveillance, and vegetation index study. These UAVs have been extensively integrated with the Internet and IoT devices as a result of advancements in the Internet of Things (IoT), making remote data collection and dissemination possible [1, 2]. The IoT-enabled UAVs have served a variety of purposes during the COVID-19 outbreak, ranging from the delivery of medications in remote locations to the monitoring of COVID-19 impacted regions. Due to the sensitivity of the information carried by UAVs, it is crucial to communicate this information securely inside the UAV network [2].

Data from Ground Control Stations (GCSs) or the sensors installed inside the UAVs are mostly used by them to operate. Such data are sent or received wirelessly, which increases the risk of data compromise by hackers during transmission and increases the risk to the entire network. As UAV intrusions increase in frequency, defensive measures against attacks on such systems are urgently required. Intrusion detection systems (IDS) [3–6] may include the solution to the current problem. The threat landscape for UAVs is continually shifting as a result of technological advancements, therefore traditional signature-based IDS detection will be unable to adequately protect the UAV. Depending on the objective, many UAV platforms can be

S. Bebortta (✉)

School of Information and Computer Sciences, Department of Computer Science, Ravenshaw University, Odisha 753003, India
e-mail: sujitbebortta1@gmail.com

S. K. Singh

Department of Computer Science and Engineering, Gandhi Institute of Education and Technology, Baniatangi 752060, Odisha, India
e-mail: sksingh@giet.edu.in

used. Various properties, including payloads, essential sensors, and control systems, to mention a few, could change as a result [7].

Network intrusion has been viewed as a significant security risk that could adversely affect many IoT-based UAV applications in light of the aforementioned problems. Due to the integration of the several heterogeneous technologies, these UAV-based systems are now being extensively exposed to many types of cyberattacks [7]. IoT devices physically connected to the system could be harmed by any system vulnerability given the significance and complexity of establishing IDS. These damages may cause either long- or short-term failures, depending on the size and form of the cyberattack. In order to safeguard user privacy and guarantee that UAVs function as intended, this paper discusses a few security issues relating to UAV networks. An intelligent intrusion detection framework for UAV-based communication networks has been proposed to secure the UAV network against threats. To achieve improved prediction performance, the proposed framework integrates an ensemble of Random Forest (RF) and Artificial Neural Network (ANN) models. A real-world UAV dataset was analysed, which provides fresh research directions for understanding the efficacy of the proposed intelligent architecture [9]. The key performance parameters of the proposed framework, such as attack prediction accuracy, precision, recall, and CPU time, are contrasted with those of other popular frameworks. The study also offers a comprehensive analysis of a few articles of qualitative literature, enabling the development of some future specifications for the current UAV network. Next, the findings from the current study are provided, along with a few prospective directions for future investigation into privacy-preserving UAV networks.

The remaining portions of the article are structured as follows: The related studies reviewed in this research are covered in Sect. 2, the framework for constructing IDS for UAV networks is covered in Sect. 3, and the experimental results and comparison with other cutting-edge studies are covered in Sect. 4. The concluding remarks and potential paths for extending the work are offered in Sect. 5.

2 Related Studies

UAVs have grown in popularity recently in organisational planning to broaden inclusion and achieve execution goals. Utilising IDS frameworks has become essential in modern firms to ensure framework dependability for the intended presentation of such organisations. The authors in [10] create their suggested Intrusion Detection and Prevention System (IDPS) module using the Deep Q-learning Network (DQN) model so that UAVs can intuitively recognise suspicious movement and take immediate steps as necessary to maintain network security. Additionally, the reward function for their model was modified to make it easier to train on the dataset they investigated and to correctly capture the minor classes. An ultra-dense remote UAV network based on Femto Access Points (FAPs) was researched in [11]. It was proposed that the Convolution Neural Network may be recommended by a multi-objective optimization problem (CNN). Additionally, a Q-network was specifically created to allow clients

in the connected UAV-based femtocaching network to deal with requests being intercepted.

With the emergence of the IoT, UAV technology has found extensive use in a variety of industries, including defence, crowd monitoring, vegetation monitoring, and wildlife protection. Concerns about the security of the data being communicated have been highlighted by the coordination of UAV with 5G technology. The authors of [12] examined the security design concerns of UAVs using the CSE-CIC IDS-2018 dataset in light of 5G network and satellite innovation. The intrusion detection module also used a number of machine learning models to evaluate harmful packets in the UAV network. Finally, the effectiveness of their model was shown in terms of F1-score, precision, recall, and prediction accuracy. A blockchain-based decentralised system was suggested for securing data in UAV networks in order to facilitate a continuous self-configurable system [13]. To evaluate the effectiveness of the suggested blockchain-based machine learning approach, the case study of cooperative intrusion detection in UAVs was taken into consideration.

In [14], the quantitative properties of UAVs, such as signal-to-noise ratio (SNR) and energy threshold, were employed to create a feature-based classification model by utilising traditional machine learning algorithms. Furthermore, the gathered spectrogram images were run through the CNN model. Quantitative analysis was used to assess the prediction effectiveness of both methodologies for determining false alarm rates and achieving high classification accuracy for jamming. A multi-agent driven deep learning technique was suggested by the authors in [15] in order to obtain good predictive performance in a UAV context. Their suggested strategy used a deep reinforcement learning model that was approximated across a deep neural network in order to reduce the operational cost of energy delivery to UAV networks.

The development of IDS for UAV network security has, however, had a huge gap that hasn't received much attention. Therefore, the current study suggests an ensemble of RF and ANN models for facilitating on-time detection of attacks in the network in order to accurately identify the cyberattacks imposed over UAVs. To conduct the experimental investigation, the model uses data on actual network traffic that is produced by a UAV network. The results show that the suggested method converges quickly and performs better than other algorithms in terms of accuracy, precision, recall, f-measure, and CPU time.

3 Proposed Framework

This section deals with the proposed framework with an emphasis on the machine learning models using in this study for developing the IDS for UAVs. In Fig. 1, an overview of the proposed intelligent intrusion detection framework is presented. Here, the environment comprises the UAV network operating over some GCS and embedded with several sensor devices for capturing data. The data from the UAV network is first captured by the RF-ANN algorithm to build the training phase for the proposed model. Further, the attacks identified in the UAV network are classified

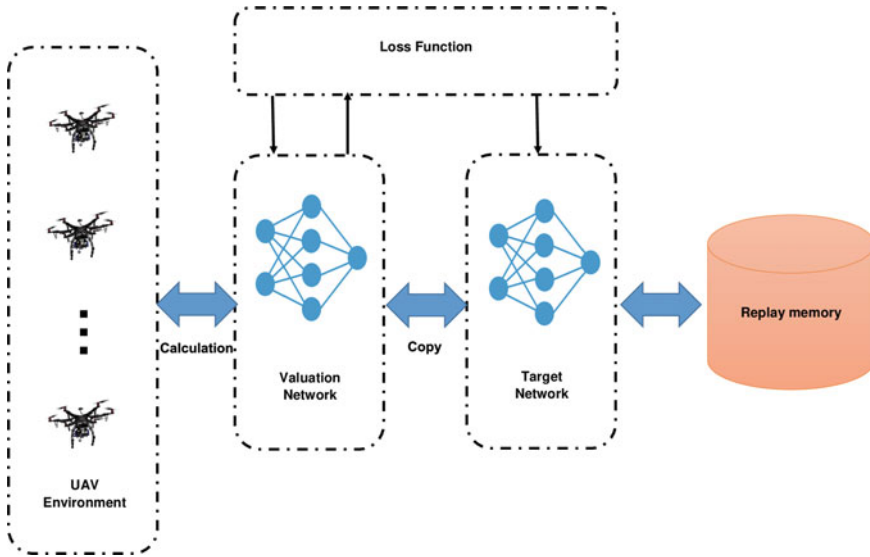


Fig. 1 Proposed framework for UAV-based IDS System

accordingly. The replay memory module acts as an intermediate storage unit for storing the transition sample to train the RF-ANN model.

3.1 Machine Learning Models

In this section, we present the machine learning models, viz., RF and ANN, which we used to construct the proposed framework for identifying the attacks in the UAV network. The models are further tested against different performance metrics to illustrate the efficacy of the proposed model.

3.1.1 Random Forest (RF) Algorithm

The RF algorithm is popularly used as a supervised ensemble learning algorithm to achieve higher predictive performance. The RF algorithms work by combining many decision trees which are generated randomly across the input vector and assist in handling complex datasets by simplifying classification. By utilising the majority voting technique, the forecast outcome from the decision trees is taken into consideration. By choosing the most popular class, the underlying decision trees are capable of anticipating future events. As a result, these features increase the RF classifier's resilience to resolving real-world issues and also highlight its effectiveness for multi-class datasets. The Gini Index, which measures the impurity of classes

for the provided dataset, is the foundation upon which the RF algorithm creates the prediction model. As a result, the Gini Index for a certain dataset DS when choosing random features in class X can be defined as

$$G = \sum_{i=1}^n \frac{f(X_i, DS)}{|DS|} \times \sum_{j=1}^m \frac{f(X_j, DS)}{|DS|} \quad (1)$$

where $f(X_i, DS)$ represents that the attributes selected belong to class X_i .

3.1.2 Artificial Neural Network (ANN) Algorithm

To handle the non-linear UAV network intrusion data considered in this study, we use the artificial neural network (ANN) model. The ANN is a popular method for dealing with complicated non-linear data with a high degree of effectiveness [20, 22]. Several publications have also noted that a time series analysis using this model's effectiveness [22, 23]. This model takes into account three layers: the input layer, the output layer, and the hidden layer. Consequently, if we take O_t we can model it as follows to reflect the output for the ANN,

$$O_t = \sum_{j=1}^l (p(\ln(j)) \times \omega_{jk}) + \rho_k \quad (2)$$

From the Eq. (2) if the bias for the output layer is represented by rho k, assuming that $j = 1, 2, \dots, l$ and $k = 1, 2, \dots, s$ represent the nodes in the hidden and output layers, respectively, then ω_{jk} signifies the weight between the nodes of the hidden layer and output layer. In this instance, the activation function is represented by $p(\ln(j))$ and can be represented as

$$\ln(j) = \sum_{p=1}^s (\theta_p \omega_{pj}) + \rho_j \quad (3)$$

where θ_p stands for the number that corresponds to the p^{th} node in the input layer, with the values $p = 1, 2, \dots, 4, s$. As a result, we use the sigmoid function as the activation function to process the result of Eq. (3), which is denoted by the following representation:

$$P(D|X_i) = \prod_{j=1}^n P(d_j|X_i) \quad (4)$$

Substituting the values of Eqs. (3) and (4) in Eq. (2), we obtain,

$$O_t = \sum_{j=1}^l \left(p \left(\sum_{p=1}^s (\theta_p \omega_{pj}) + \rho_j \right) \times \omega_{jk} \right) + \rho_k \quad (5)$$

4 Results and Discussions

This section presents the experimental outcomes obtained over the UAV data by exploiting the proposed framework. The data comprises cyberattacks over UAVs induced using simulations over PX4 autopilot and Gazebo simulator [9]. The data comprises logged sensor data values captured pervasively across autopilots. The dataset comprises attacks deployed over various UAV platforms; however, this study exploits the Plane UAV platform having a standard plane model constituting of 23198 benign and 1055 malicious sensor values with the simulation type as software-in-the-loop to facilitate interoperability of the proposed framework over different UAVs.

Figure 2 presents the comparative study of different machine learning models like RF and ANN in comparison with the proposed RF-ANN framework. It is observed that the proposed model outperforms the RF and ANN model in terms of recall and F-measure; however, the RF model outperforms all other models in terms of precision. In Fig. 3, the accuracy for different models is compared. It is observed that the proposed framework provides the highest prediction accuracy of 99.372%. Finally, the CPU time for the proposed model is observed to be 76.117 seconds which is faster in comparison to the CPU time for RF and ANN models. Table 1 gives a detailed comparative overview of the different performance metrics considered in this study for the proposed framework and baseline algorithms (Fig. 4).

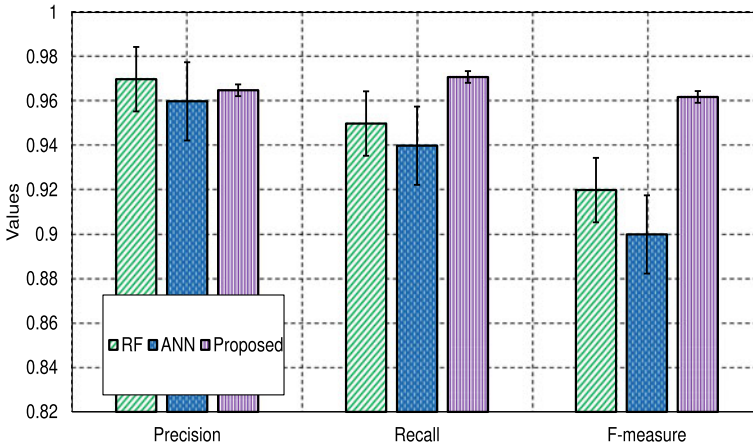


Fig. 2 Comparison of performance for the proposed framework with different machine learning models

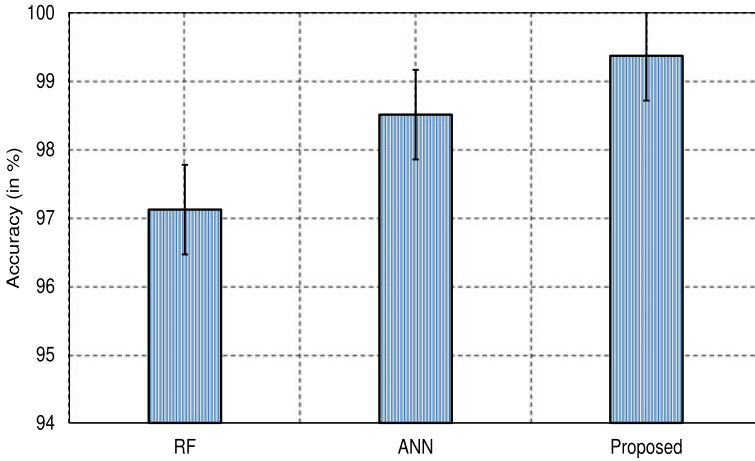


Fig. 3 Comparison of prediction accuracy for the proposed model with different machine learning models

Table 1 Performance metrics for proposed algorithm and baseline algorithms

Models	Precision	Recall	F-measure	Accuracy (in %)	CPU time
RF	0.970	0.952	0.922	97.121	86.665
ANN	0.961	0.0.940	0.901	98.511	88.911
Proposed	0.965	0.971	0.962	99.372	76.115

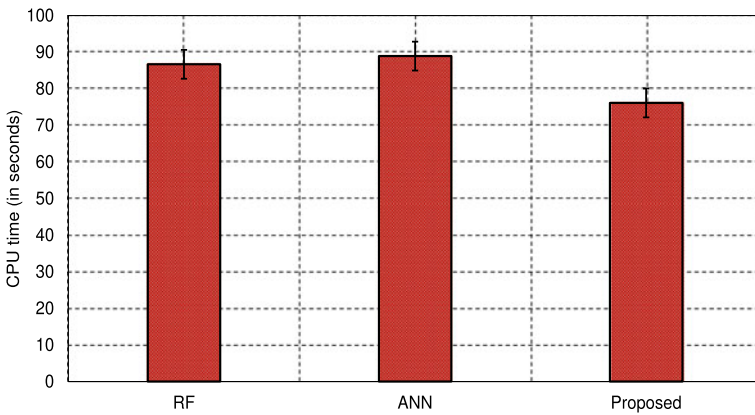


Fig. 4 CPU time for the proposed framework with RF and ANN

5 Conclusions and Future Work

With the growing advancements in UAV technology, the dependence on these technologies has increased drastically in recent times. However, considering the possibility of cyberattacks in heterogeneously connected UAVs, the need for developing IDS to preserve the sensitivity of critical tasks becomes essential for UAVs. This work proposed a machine learning-based hybrid framework by combining RF and ANN algorithms through ensembling technique for achieving higher predictive performance towards the detection of cyberattacks in UAVs. The model was studied in convergence with baseline machine learning algorithms to assess the performance for different parameters like precision, recall, F-measure, accuracy, and CPU time. It was observed that the proposed framework provided the highest predictive performance of 99.372.

In the future, we would like to extend our work by incorporating more precise learning models like the reinforcement learning model over different UAV platforms to achieve higher performance. Further, there are several research gaps that can address other major issues, such as those imposed by attackers which affect the hardware components associated with the compromised UAVs. This paves a new path for future research towards addressing the limitations of IDS for UAV networks.

References

1. Alsamhi SH, Afghah F, Sahal R, Hawbani A, Al-qaness MA, Lee B, Guizani M (2021) Green internet of things using UAVs in B5G networks: a review of applications and strategies. *Ad Hoc Netw* 1(117):102505
2. Sharma R, Arya R (2022) UAV based long range environment monitoring system with Industry 5.0 perspectives for smart city infrastructure. *Comput Ind Eng* 168:108066
3. Bebortta S, Singh SK (2021) An adaptive machine learning-based threat detection framework for industrial communication networks. In: 2021 10th IEEE international conference on communication systems and network technologies (CSNT). IEEE, pp 527–532
4. Bebortta S, Singh SK (2022) An intelligent framework towards managing big data in internet of healthcare things. In: International conference on computational intelligence in pattern recognition. Springer, Singapore, pp 520–530
5. Bebortta S, Singh SK (2022) An opportunistic ensemble learning framework for network traffic classification in IoT environments. In: Proceedings of the seventh international conference on mathematics and computing 2022. Springer, Singapore, pp 473–484
6. Singh SK, Mishra AK. Rain fall prediction using bigdata analytics. *Int J Innov Eng Technol (IJJET)* 151. <https://doi.org/10.21172/ijjet>
7. Sun M, Xu X, Qin X, Zhang P (2021) 10 Aol-energy-aware UAV-assisted data collection for IoT networks: a deep reinforcement learning method. *IEEE Internet Things J* 8(24):17275–17289
8. Bebortta S, Panda M, Panda S (2020) Classification of pathological disorders in children using random forest algorithm. In: 2020 international conference on emerging trends in information technology and engineering (ic-ETITE) 2020. IEEE, pp 1–6
9. Whelan J, Sangarapillai T, Minawi O, Almeahadi A, El-Khatib K (2020) Novelty-based intrusion detection of sensor attacks on unmanned aerial vehicles. In: Proceedings of the 16th ACM symposium on QoS and security for wireless and mobile networks, pp 23–28

10. Bouhamed O, Bouachir O, Aloqaily M, Al Ridhawi I (2020) Lightweight ids for UAV networks: a periodic deep reinforcement learning-based approach. In: 2021 IFIP/IEEE international symposium on integrated network management (IM). IEEE, pp 1032–1037
11. Hajiakhondi-Meybodi Z, Mohammadi A, Abouei J (2021) Deep reinforcement learning for trustworthy and time-varying connection scheduling in a coupled UAV-based femtocaching architecture. *IEEE Access* 18(9):32263–81
12. Shrestha R, Omidkar A, Roudi SA, Abbas R, Kim S (2021) Machine-learning-enabled intrusion detection system for cellular connected UAV networks. *Electronics* 10(13):1549
13. Khan AA, Khan MM, Khan KM, Arshad J, Ahmad F (2021) A blockchain-based decentralized machine learning framework for collaborative intrusion detection within UAVs. *Comput Netw* 4(196):108217
14. Li Y, Pawlak J, Price J, Al Shamaileh K, Niyaz Q, Paheding S, Devabhaktuni V (2022) Jamming detection and classification in OFDM-based UAVs via feature-and spectrogram-tailored machine learning. *IEEE Access* 8(10):16859–70
15. Jung S, Yun WJ, Kim J, Kim JH (2021) Coordinated multi-agent deep reinforcement learning for energy-aware UAV-based big-data platforms. *Electronics* 10(5):543
16. Cai YD, Feng KY, Lu WC, Chou KC (2006)7 Using LogitBoost classifier to predict protein structural classes. *J Theor Biol* 238(1):172–6
17. Murphy KP (2006) Naive Bayes classifiers. *Univ Br Columbia* 18(60):1–8
18. Webb GI, Keogh E, Miikkulainen R (2010) Naïve Bayes. *Encyclopedia of machine learning*, vol 15, pp 713–714
19. Qi Y (2012) Random forest for bioinformatics. In: *Ensemble machine learning*. Springer, Boston, MA, pp 307–323
20. Probst P, Wright MN, Boulesteix AL (2019) Hyperparameters and tuning strategies for random forest. *Wiley Interdiscip Rev: Data Min Knowl Discov* 9(3):e1301
21. CrowdFlower. <https://www.crowdfunder.com/data-for-everyone/>. Accessed 10 Dec 2021
22. Squires M, Tao X, Elangovan S, Gururajan R, Zhou X, Acharya UR (2022) A novel genetic algorithm based system for the scheduling of medical treatments. *Expert Syst Appl* 14:116464
23. Sarosh P, Parah SA, Bhat GM (2022) An efficient image encryption scheme for healthcare applications. *Multimed Tools Appl* 25:1–8

Distributed and Hash-Based Mixers for User Anonymity on Blockchain



P. Guna Shekar, Raghwendra Singh, Debanjan Sadhya,
and Bodhi Chakraborty

1 Introduction

In its essence, a Blockchain is a distributed ledger technology enabling data immutability. It helps facilitate financial transactions in a decentralized and peer-to-peer manner without the involvement of a third party. While third parties like banks maintain the ledger of financial transactions in the traditional scenario, Blockchain acts as a public ledger, maintaining all the transactions on the network in a decentralized and distributed fashion. All the transactions on the network are collected and organized into *blocks*, with each block containing a limited number of transactions. As and when more transactions come into the network, new blocks get appended to the existing set of blocks, thus forming a chain of blocks. The blocks are connected using cryptographic methodologies like hashing, as shown in Fig. 1. Since Blockchain is a decentralized system, it is crucial to maintain consensus about the content of each block in the Blockchain among all the network participant nodes.

Anonymity in Blockchain refers to its property of not revealing the user's identity to the network. This property enables users to stay unidentified on the Blockchain network, thus enhancing user privacy. Though the user's identity needs not to be revealed to make transactions, de-anonymization inference attacks can be used to link a set of transactions to a user. Unlinkability refers to the property where a particular set of transactions cannot be related to a single entity with a high confidence

P. G. Shekar · R. Singh (✉) · D. Sadhya
ABV-Indian Institute of Information Technology and Management Gwalior, Madhya Pradesh,
Gwalior, India
e-mail: 202109@iiitm.ac.in

D. Sadhya
e-mail: debanjan@iiitm.ac.in

B. Chakraborty
ITM University, Madhya Pradesh, Gwalior, India
e-mail: bodhi.cse@itmuniversity.ac.in

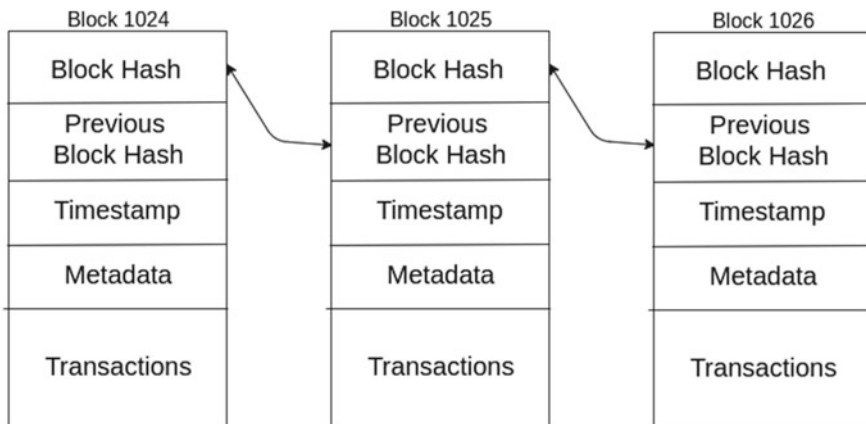


Fig. 1 Basic structure of a Blockchain

level. This property is essential despite anonymity because once a set of transactions can be linked to a particular entity, other details such as account balance, i.e., type of merchants/transactions and the frequency can be easily inferred. In turn, malicious parties can use these statistics to reveal the user’s true identity. For instance, a statistical analysis based on the origin IP Address and the physical location of transactions can be used to trace back the transactions and link a certain set of transactions to a particular user [1].

Mixing is a service that allows users to regain their privacy on Blockchain that was challenged by recent works on de-anonymization. Mixing refers to transferring coins or tokens from one account to another unrelated account before finally transferring them to the desired destination to prevent attackers from relating transactions to each other. In this work, we devise a solution for maintaining anonymity over a Blockchain network while using mixers over a Merkle Patricia Trie. The proposed solution has an interface for the users to deposit their coins to the mixer, which would then mix them to multiple addresses. The mixer would take note of the amount it owes to the user and return it to the desired destination as requested by the user.

2 Related Work

In this section, we discuss and analyze the existing studies on unlinkability and anonymity on the Blockchain network. Dupont and Squicciarini [2] provided a demonstration on obtaining a user’s real-life information based on their bitcoin transactions. To achieve this, they used publicly available Bitcoin transaction data. They performed a statistical analysis of the users’ spending habits to determine their physical location worldwide. The authors considered the timings during which the transactions were made from a particular user and were able to decode a user’s timezone

and, in some cases, their country with 75% accuracy. Jawaheri et al. [3] examined the possibility of de-anonymizing Tor hidden service users who pay with Bitcoin using public information obtained from online social networks, the Blockchain, and onion websites. This experiment successfully enabled to link of a user's Twitter handle to their Tor hidden service handle based on the transactions they made using Bitcoin. The authors conducted a real-world experiment simulating a passive, limited adversary to demonstrate the feasibility of this de-anonymization attack. Specifically, 1500 hidden services were crawled, and 88 different Bitcoin addresses were gathered. The authors subsequently crawled 5 billion tweets and 1 million BitcoinTalk forum pages, thus collecting 4.2 million and 41 thousand unique Bitcoin addresses, respectively. Each user's address was linked to the online identity and public profile information. A total of 125 individual users were linked to 20 Tor hidden services, including sensitive ones like The Pirate Bay and Silk Road.

Mixcoin, a protocol to enable users to have anonymity and unlinkability on the Blockchain, was proposed by Bonneau et al. [4]. The authors expanded on the emerging area of currency mixes by introducing an accountability mechanism to disclose thievery. It was demonstrated how the incentives of the mixes and the clients could be aligned to ensure that rational mixes do not steal. Gregory Maxwell developed a method to anonymize user transactions on the Blockchain using a joint transaction approach¹. In this approach, if a user wishes to send a transaction, they would find another user who also wants to make a transaction and send joint transactions to their respective destinations. This process improves anonymity as patterns of a user would be very difficult to identify when their transactions are combined with that of other users. CoinJoin also has a requirement that the users would have to find a pair for themselves to make a joint transaction.

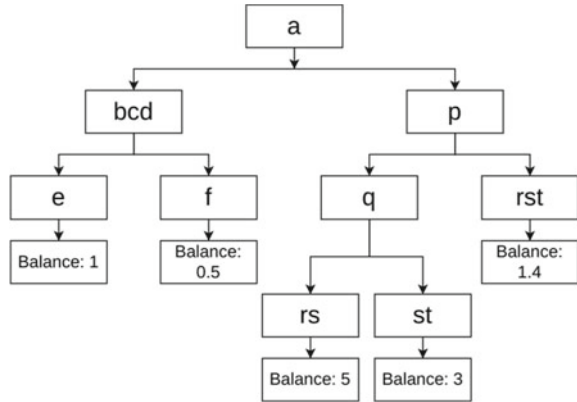
Some older mixing protocols, such as SharedCoin, proposed by Henrique et al. [5] used a centralized server to facilitate user matching. This process lessened the load on users to find a pair for themselves to make a joint transaction. However, they acted as a single point of failure for the entire system. The model also maintained the original addresses of users and transaction logs, leakage of which would cause de-anonymization of the users.

3 Merkle Patricia Trie

A Merkle Patricia Trie is an improved combination of a Merkle Tree and a Trie. An example of a Merkle Patricia Trie can be seen in Fig. 2. Each node has a hash, which is used as the identifier for the node. Starting from the leaf node, all children of a particular node are hashed to get the parent's hash; hence, the root of the tree can act as a cryptographic hash for the entire tree. A Merkle Patricia Trie has four different types of nodes.

1. **Empty Nodes:** These blank nodes do not contain any data.

Fig. 2 Example of a Merkle Patricia Trie with five hashes of user addresses—abcde, abcdf, apqrs, apqst, aprst; and their balances



2. **Leaf Nodes:** Lowermost nodes in the tree containing the final key-value pair data (in our case, the key is the hash of the user's address, and the value is the amount the user deposited in our mixer).
3. **Branch Nodes:** These are internal nodes, a list of characters that link to either other branch nodes or a leaf node. The list is the size of the alphabet used in the data structure.
4. **Extension Nodes:** These nodes contain the hash of another node as its value.

The Merkle Patricia Trie is much more efficient in terms of space than a Merkle Tree or a Trie because of the extension nodes that make it compact. Searching for a particular element in a Merkle Patricia Trie takes (1) time. While indexed databases can search for elements in (1) time, the indices take up extra storage space. At the same time, Merkle Patricia Trie does not utilize extra storage space for any form of indices. The Merkle Patricia Trie also provides a Merkle Proof via the cryptographic hash of the root node, using which other servers in the network can verify balances using (1) time without having to fetch all balances. In case of balance differences, comparison and update among a distributed set of Merkle Patricia Trie happen in $(\log n)$ time, where the maximum value of n is the number of branches from an element in the Merkle Patricia Trie. A Merkle Patricia Trie is beneficial when hashes are stored in it because hashes have a fixed domain; hence, the Trie width will always remain less than or equal to the size of the domain.

4 Proposed Methodology

With user anonymity being the primary objective of the implementation, we now discuss the design details of our model.

4.1 Model Overview

Our solution adopts a distributed approach to avoid such a single point of failure. Instead of having a single mixer handling all users' requests, the proposed solution will have multiple servers interacting with each other, as shown in Fig.3. The database too will be distributed across a cluster to ensure data replication. In this case, the data will be stored in a data structure called the Merkle Patricia Trie. In order to prevent location-based de-anonymization, the mixer servers are deployed on a cloud cluster with each node located in different time zones. Hence, a particular user's transactions will also occur randomly from different locations worldwide, thus preventing location-based de-anonymization.

The second aspect of our design is to provide a better user experience. Older mixers put the burden of finding a group/partner for mixing on the end user. Moreover, it would not be a viable solution if the problem had to be solved for the masses and scale to a massive number of users. A much more scalable approach would be to have the mixers take on the responsibility of creating one-time-use addresses for mixing, as shown in Fig. 4.

The final aspect of our solution is the protection of user anonymity. To efficiently mix the coins to ensure unlinkability, it is not necessary to store the user addresses



Fig. 3 Illustration of distributed servers to avoid a single point of failure

Fig. 4 Mixer server distributing the coins to one-time-use addresses

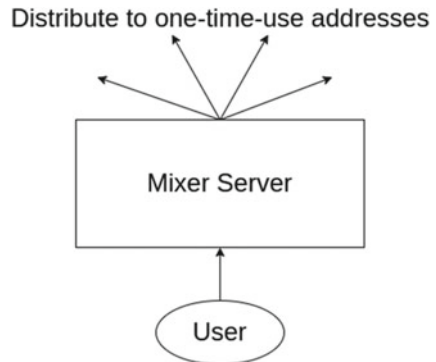
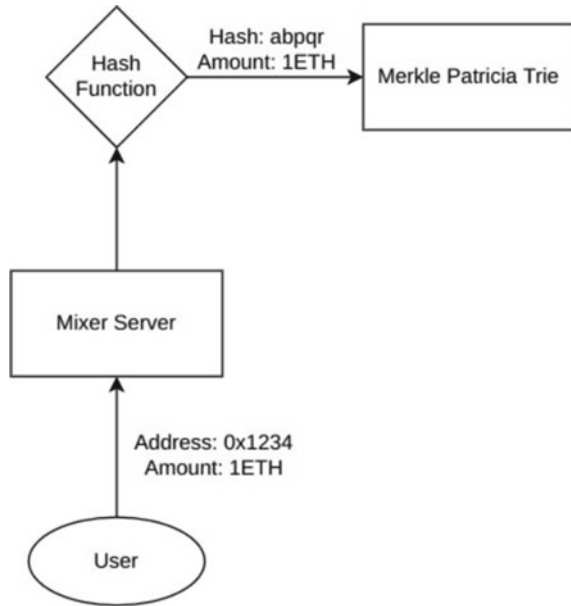


Fig. 5 Storing only the hash to protect user anonymity



directly. Instead, any other identifier that only the user would have access to would be sufficient to prove that the mixer owes a certain amount of coins to the user. For this purpose, instead of the user's original address, the address's hash will be stored in the database of the proposed solution, as shown in Fig. 5. Users who wish to withdraw their amount or send it to another account will send a signed transaction to the mixer. The signed transaction would come from the user's original address, which is enough proof that it is the original user requesting their coins back. The mixer would then hash the user's addresses, check if the mixer owes any coins to that user, and finally send the coins to the destination accordingly.

4.2 Detailed Architecture

There are three main components to consider for architecture design: (1) user interaction with the mixer while depositing coins, (2) the mixing protocol, and (3) user interaction with the mixer while withdrawing coins. Now we individually discuss these mechanisms.

4.2.1 Depositing Coins

For depositing the coins, the client chooses a random mixer from the available n mixers in the distributed system. The user would be presented with a user interface

using which they can enter the amount they want to deposit. They would then sign a transaction message to verify that they are the valid owner of the coins they are depositing. This message would then be sent to the mixer server. The mixer would then validate the request and hash the user's address for storage. The hashed address and the amount the mixer now owes to the user are stored in the Merkle Patricia Trie and eventually replicated over other mixer servers.

4.2.2 Mixing Protocol

Once the user sends the coins to the mixer, the mixer proceeds to carry out a protocol for mixing. The total number of coins is divided into three to five chunks with random amounts in each, totaling up to the original amount. The chunks are random and not uniform to ensure that an adversary cannot multiply the amount in a chunk to derive insights into what the original amount could be. The mixer then creates n one-time-use addresses, n being the number of chunks the original amount was divided into and transfers the amount in each chunk to each of the newly created addresses. At this point, the user's coins have reached new addresses, and future transactions cannot be linked to the user's original account. To increase the number of jumps before reaching the final destination and improve unlinkability, the coins will be shuffled between the newly created addresses once for a certain period (viz., one month).

4.2.3 Withdrawing Coins

When users wish to withdraw their coins, they first sign a message with the private key of the address whose coins are stored with the mixer. Along with this message, they also enter a new address to which they wish to transfer their coins and the number of coins they wish to transfer. This message and the details are then sent to the mixer. The mixer validates the message, verifies whether the mixer owes the amount to the given address, and proceeds to transfer the funds. For transferring funds, the mixer chooses three to five random addresses it previously created, whose total balance is at least as much as the amount the user wishes to withdraw. These n addresses are then made to transfer chunks of the amount, totaling up to the desired amount the user wants to withdraw, to the final destination where the user wishes to withdraw their coins. Finally, the destination address has now received the total sum requested by the user, but from n different and unrelated addresses, which cannot be linked back to the original address of the user. The process of withdrawing coins is demonstrated in Fig. 6.

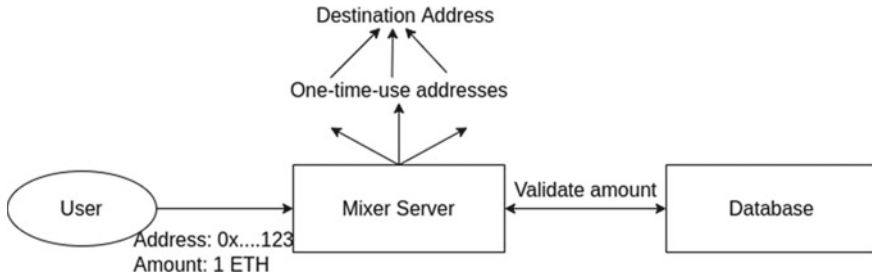


Fig. 6 The process of withdrawing coins by a user

5 Experimental Results

The developed model is a distributed and hash-based mixer server that provides user anonymity and transaction unlinkability on Blockchain. For simulation and testing purposes, the servers have been deployed on Microsoft Azure’s Virtual Machines, as shown in Fig. 7.

5.1 Usage of Merkle Patricia Trie

While providing a data storage solution for hashes, the Merkle Patricia Trie proved more efficient than a relational database. The Merkle Patricia Trie’s insertion time is significantly lesser than a relational or a document database. The insertion of a hash of a user address along with their balance took 21.755 ms in the case of a Merkle

Resource group (move) : 2018BCS-031-RTP-RG	Operating system : Linux (ubuntu 20.04)
Status : Running	Size : Standard B1s (1 vcpu, 1 GiB memory)
Location : Central India	
Subscription (move) : Visual Studio Enterprise Subscription	
Subscription ID : 6e3e7f69-8600-4d63-912c-484e00c02db4	
Resource group (move) : 2018BCS-031-RTP-RG	Operating system : Linux (ubuntu 20.04)
Status : Running	Size : Standard B1s (1 vcpu, 1 GiB memory)
Location : East US 2	
Subscription (move) : Visual Studio Enterprise Subscription	
Subscription ID : 6e3e7f69-8600-4d63-912c-484e00c02db4	
Resource group (move) : 2018BCS-031-RTP-RG	Operating system : Linux (ubuntu 20.04)
Status : Running	Size : Standard B1s (1 vcpu, 1 GiB memory)
Location : West Europe	
Subscription (move) : Visual Studio Enterprise Subscription	
Subscription ID : 6e3e7f69-8600-4d63-912c-484e00c02db4	

Fig. 7 Mixer servers are deployed at different data centers to prevent a single point of failure and location-based de-anonymization

Patricia Trie, while it took 49.262 and 45.278 ms in the case of a Document and Relational Database, respectively. Hence, Merkle Patricia Trie proves to be better in terms of computational memory and time.

5.2 Mixer Server

The Mixer Server is a Node.js-based server that contains the core logic of the mixer. This server is deployed on a cloud cluster on Microsoft Azure. The mixer has the following APIs.

1. **Deposit Funds:** The Deposit Funds API accepts a signed transaction from a user that contains a certain amount of cryptocurrencies. The signature on the transaction is done using the Metamask wallet on the client side. This signature is performed by the user's private key on the Blockchain network that can be used to ensure that it is the legitimate user who is sending the transaction. The API then hashes the user's address to provide anonymity on our server and then updates the user's balance in the Merkle Patricia Trie. Once the balance has been updated, the API distributes the funds to one-time-use addresses according to the mixing protocol described previously.
2. **Fetch Balance:** The Fetch Balance API accepts the user's address in the form of a signed transaction with zero funds to ensure the legitimacy of the user. The user's address is then hashed, and the record is fetched from the Merkle Patricia Trie. The balance fetched from the Merkle Patricia Trie is sent to the client, which then displays it to the user.
3. **Withdraw Funds:** The Withdraw Funds API accepts a destination address, the amount to be withdrawn, and a signed transaction with zero funds to ensure the legitimacy of the user. The first step is to hash the user's address and verify in the Merkle Patricia Trie whether the user has enough balance to be withdrawn. Once verified, the API proceeds to withdraw funds into the destination address provided by the user. The destination address will then receive funds from a random one-time address that cannot be linked back to the original user.

5.3 Client Side

A React.js-based client has been developed to help the mixer users interact with the mixer server. The client has the following routes:

- **Deposit Funds:** The Deposit Funds route contains a form that accepts the amount the user wishes to deposit with the mixer. Once the user enters the amount, they will be prompted to sign a transaction with the specified amount on their Meta-mask wallet, as shown in Fig. 8. This signed transaction will then be sent to the backend. Figure 8 demonstrates this process for a particular user.

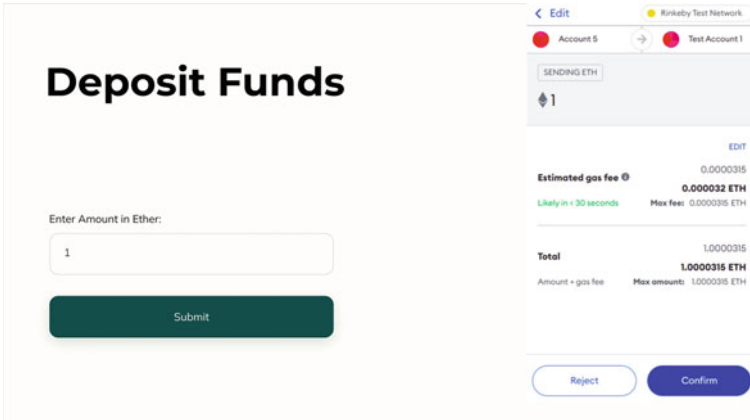


Fig. 8 An user trying to deposit 1 ETH to the mixer server

- **View Balance:** The View Balance route prompts the user to sign an empty transaction sent to the backend. Based on this, the route shows the balance of the user that will be returned by the server, as shown in Fig. 9.
- **Withdraw Funds:** Withdraw Funds contains a form that accepts the amount to be withdrawn and the destination address where the funds are to be received. The user will then be prompted to sign an empty transaction to verify their legitimacy, as shown in Fig. 10.

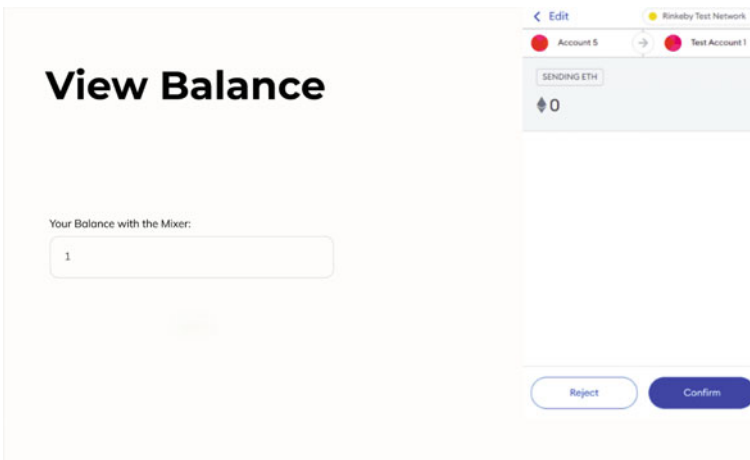
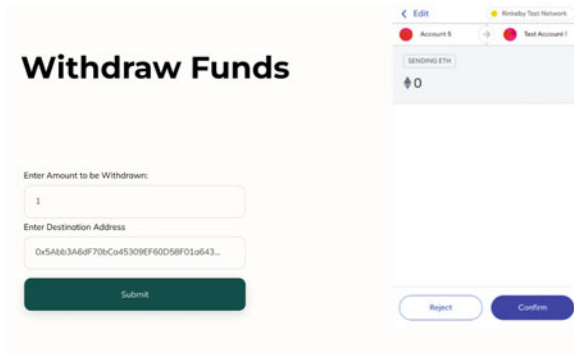


Fig. 9 An user is viewing the balance that was previously deposited into the mixer

Fig. 10 An user trying to withdraw 1 ETH to their destination address



5.4 Quantitative Evaluation

After the deployment of mixer servers, it was deliberately taken down to test the robustness of the system, as shown in the *Status* property in Fig. 11. A test user was then made to deposit funds mixed by the protocol. Finally, the user withdrew their funds to a destination address. Figure 12 shows the transactions that were sent to the destination address. Since these transactions were received from servers with different locations, the location of the original user cannot be decoded.

Importantly, our model protects user anonymity even in cases of data leaks. The mixer server only stores the irreversible hashes of the user addresses in the form of a Merkle Patricia Trie and not the user addresses directly. Fig. 2 shows what an adversary would see if they were to attack the server and leak the data. It can be seen that the original user addresses are hidden and the adversary cannot decode the identity of the user from the Trie that contains the hashes of user addresses. An Ethereum address consists of 64 characters and is a hexadecimal string, i.e., a domain size of 16. Therefore, on average, it would take an adversary attempt to crack a hash of a user address



Fig. 11 Mixer Server stopped to test the robustness of the system

Txn Hash	Method	Block	Age	From	To	Value	Txn Fee
0x974f450c366c9379d5...	Transfer	14745053	53 secs ago	0x555ee36293b612c52...	0x5AbB3A6dF70bCa453...	0.35 Ether	0.0029031686772
0xc0bc1e628189f1b34aa...	Transfer	14744715	1 min 23 secs ago	0x3c8cdd2e501Ac010F...	0x5AbB3A6dF70bCa453...	0.25 Ether	0.0029787779001
0x6a07e9bb1b00713078...	Transfer	14744715	1 min 34 secs ago	0x7765F238e08F97394...	0x5AbB3A6dF70bCa453...	0.4 Ether	0.0037400927732

Fig. 12 Destination address receiving funds from one-time addresses

6 Challenges and Limitations

While the problem of user anonymity is solved using the mixers, the solution lags on a few other facets. Following are the drawbacks of using a mixer instead of a direct transfer:

1. **Additional Gas Fee:** The proposed solution uses the Rinkeby Test Network, one of the testing environments for the Ethereum Blockchain. The average gas fee while making a transaction on this network is 0.000045ETH. Assuming an average hop size of 3, it would take the user an additional 0.000135ETH to stay anonymous while sending the transaction to their destination.
2. **Higher Latency:** It would take additional time for the mixer to send the desired amount to the destination address compared to a direct transfer of funds, owing to the mixing protocol that makes the funds hop among different addresses. However, since the mixing is done at a stage earlier than withdrawal and the final transactions to the destination are sent in parallel, the latency is lesser than the architecture proposed by the Mixcoin [5] protocol. The latency for the direct transfer, the proposed mixer-based model, and the Mixcoin protocol were 8.62, 1.85, and 39.46 s, respectively.

7 Conclusion

While anonymity and unlinkability are supposed to be one of the primary features of the Blockchain, studies have proved that this is not necessarily the case. Though there were attempts to resolve these issues in the past, they contained flaws due to which they could not become permanent solutions to these problems. The proposed solution attempts to fix these issues by mixing coins to improve user anonymity and transaction unlinkability on the Blockchain. The simulation results have shown that the proposed solution successfully solves the issues identified in existing solutions and provides a layer of user anonymity and transaction unlinkability on the existing Blockchain.

References

1. Zhang R, Xue R, Liu L (2019) Security and privacy on blockchain. *ACM Comput Surv* 52(3). <https://doi.org/10.1145/3316481>
2. DuPont J, Squicciarini AC (2015) Toward de-anonymizing bitcoin by mapping users location. *CODASPY '15*, pp 139–141. Association for Computing Machinery, New York, NY, USA. <https://doi.org/10.1145/2699026.2699128>
3. Jawaheri HA, Sabah MA, Boshmaf Y, Erbad A (2020) Deanonimizing tor hidden service users through bitcoin transactions analysis. *Comput Secur* 89(C). <https://doi.org/10.1016/j.cose.2019.101684>

4. Bonneau J, Narayanan A, Miller A, Clark J, Kroll JA, Felten EW (2014) Mixcoin: Anonymity for bitcoin with accountable mixes. In: Christin N, Safavi-Naini R (eds) *Financial cryptography and data security*, pp 486–504. Springer, Berlin, Heidelberg
5. Moniz H, Neves NF, Correia M, Verissimo P (2006) Experimental comparison of local and shared coin randomized consensus protocols. In: *2006 25th IEEE Symposium on Reliable Distributed Systems (SRDS'06)*, pp 235–244. <https://doi.org/10.1109/SRDS.2006.19>
6. Bonneau J, Narayanan A, Miller A, Clark J, Kroll J, Felten E (2014) Mix-coin: Anonymity for bitcoin with accountable mixes, pp 486–504. <https://doi.org/10.1007/978-3-662-45472-531>

Implementation and Analysis of Different Visual Cryptographic Schemes



Vanashree Gupta and Smita Bedekar

1 Introduction

Visual Cryptography (VC) scheme is an information security method. As with many cryptographic schemes, trust is the most difficult part whereas VC provides a very powerful technique by which one secret can be distributed into two or more shares. If all or threshold number of shares come together then secret can be recovered. It is also called as image secret sharing. VC uses idea of hiding secret within images. These images, text, picture or printed data are encoded into multiple shares and later decoded by superimposing transparencies without any computation. This allows the secret to be recovered. This is perfectly secure and very easily implemented. Here, many types of VC are examined starting from traditional VC right up to the latest developments. In traditional VC, sharing of single binary secret between number of participants is done. Extended VC has significant visual meaning. This detract from the suspicious looking encrypted share that are generated using traditional methods. VC has applications in copyright protection and watermarking domain. Here, comparison and analysis of different VC schemes are done. While designing these schemes, we have to consider certain parameters like number of pixels, relative difference in weight between the combined shares and the size of collection of subpixels. Pixel expansion and increased number of shares affect resolution. Therefore, optimum number of shares are used to hide the secret information.

V. Gupta (✉) · S. Bedekar

Department of Scientific Computing, Modeling and Simulation, Savitribai Phule Pune University, Pune, Maharashtra, India

e-mail: vanashreegupta@gmail.com

1.1 Organization

The rest of the paper is organized as follows: Sect. 2 covers related work based on different image encryption algorithms. Section 3 is about results and discussions. Section 4 concludes with final remarks. Section 5 discusses about the future scope.

2 Related Work

VC was first proposed by Naor and Shamir [1] in 1994. It is a cryptographic technique in which visual information like images, pictures, texts are encrypted in such a fashion that decryption can be done by human eye. When shares are stacked on one over other human eye can easily decrypt it. The mechanism is very secure and easy to implement. No prior knowledge of cryptography is required. Reconstruction can be done without performing any computation.

There are different types of Visual Cryptographic Schemes (VCS) like traditional VC, Size Invariant VC, Extended VC, Colour VC, Recursive Threshold VC, Random Grid-based VC, Probabilistic VC, Cheating Immune VC, OR and XOR VC, etc. Different factors such as contrast, security, pixel expansion, meaningful shares or meaningless shares, type of secret image (binary, coloured or grayscale image), number of secret images encrypted (single or multiple) are considered for categorization. Few of these techniques are discussed below.

In [1], they assumed image as collection of black and white (binary images) pixels. Here, each pixel is handled individually which leads to pixel expansion. White colour is the transparent colour. The limitation of this scheme is decryption process is lossy. It affects contrast (clarity). Ito et al. [2] first considered size invariant scheme in 1998. Pixel expansion leads to the difficulty in carrying shares and consume more storage space. Ito removes the need for pixel expansion. Till 1997 visual cryptographic schemes were applied only to black and white images. But Verheul and Van Tilborg [6] proposed first visual cryptographic scheme for colour images. Since majority of people use colour images and interact with them more frequently, it is one of the potentially useful type of scheme. VC has application in copyright protection and watermarking domain. Here comparison and analysis of different VC schemes are presented.

3 Results and Discussions

Following is the list of visual cryptographic algorithms with examples. All of them are (2, 2) Visual cryptographic secret sharing encryption techniques but can be extended to (n, n) schemes. The algorithms are applied to different types of images (grayscale, black and white (binary) and colour) as shown below. The observational results that

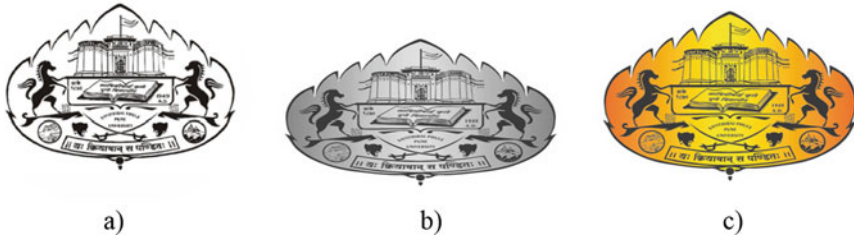


Fig. 1 a Binary input image b Grayscale input image c Colour input image

we could get from implementation of the algorithm are also shown. All these are the results for (2, 2) VC scheme image encryption. Two parameters Peak Signal to Noise Ratio (PSNR) and Mean Normalized Cross-Correlation are considered. PSNR is the common metric to check the appearance of a recovered image. It shows peak of error between original image and the recovered image. Ideally it should be infinity but practically as large as possible. PSNR value of reconstructed image greater than 30 dB is acceptable. Quality of cryptosystem is determined by correlation coefficient metric. Ideally it should be 1 for indistinguishable image and 0 for uncorrelated image. Practically, this value should be minimum (towards 0). Following are the different types input images used for encryption and decryption (Fig. 1).

3.1 Pixel Expansion Algorithm

In this, each pixel is divided into subpixels and two shares are generated from the original image. When the two shares are overlapped the final decrypted image is generated. Here, we are dividing each pixel into four subpixels which results in increase in size of shares as well as resultant decrypted image. Decryption with single share is impossible and for a brute force attack it will take (number of blocks)^(m*n) states to retrieve back the image, where m*n is the size of the original image. There are two ways for decryption. First one is overlap and the second is extraction. In overlap, the decrypted image is double the size of original image, to match with original image it is further resized. In Extraction, pixel with all its subpixels as black are marked as black otherwise white (Figs. 2 and 3).

Advantages

- Best and oldest algorithm.
- Applicable to binary images of any size.

Disadvantages

- Computation costs are high.
- Lossy method, original image is not same as decrypted image.
- The PSNR value is 29.096859139920994 dB and Mean NCORR value is -0.09613788418338697.

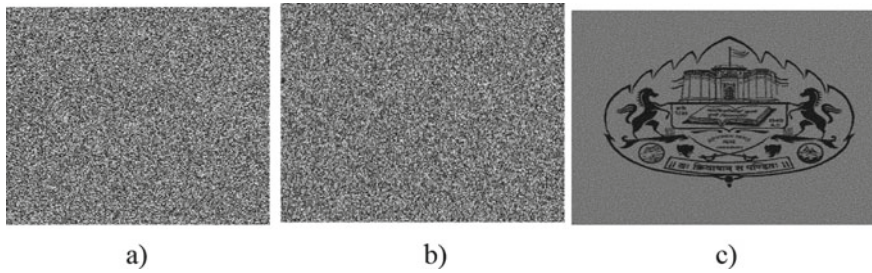


Fig. 2 a, b are binary shares c reconstructed (decrypted) binary image for Overlap

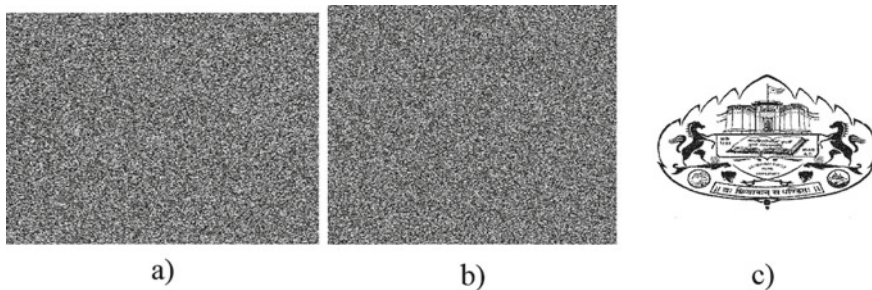


Fig. 3 a, b are binary shares c reconstructed (decrypted) binary image for Extraction

The PSNR value is 100 dB and Mean NCORR value is -0.09908745073681847 .

3.2 XOR Image Encryption Algorithm

It is similar to encrypting numbers using XOR with a secret number. This idea is used for each pixel value in an image. The original image is XORed with k shares of same size to get secret image. Retrieving of original image is done by XORing secret image with k shares. If one of the shares is missing it is hard to decrypt. For a brute force attack it will take $2^{(m*n)}$ states, where $m*n$ is the size of the original image (Figs. 4, 5 and 6).

Advantages

- Easy to build.
- Computation cost is low.
- Can be used to encrypt or decrypt any type of image.

Disadvantages

- Attacker can easily track the secret image.
- This is similar to one-time pad, hence less secure.
- The PSNR value is 100 dB, and Mean NCORR value is -0.12021190442825512 .

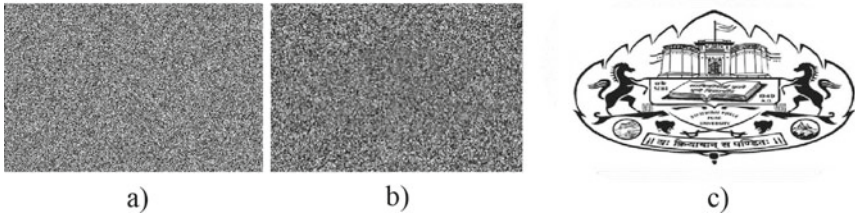


Fig. 4 a, b are binary shares c reconstructed (decrypted) binary image for XOR Image Encryption algorithm

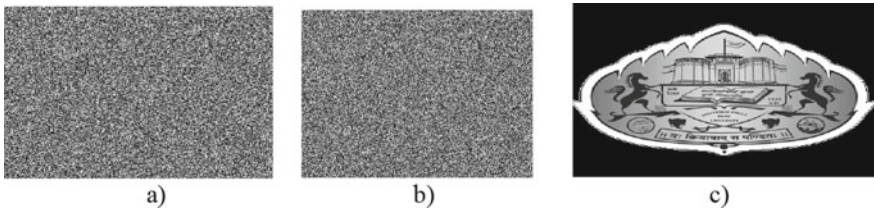


Fig. 5 a, b are Grayscale shares c reconstructed (decrypted) Grayscale image for XOR Image Encryption algorithm

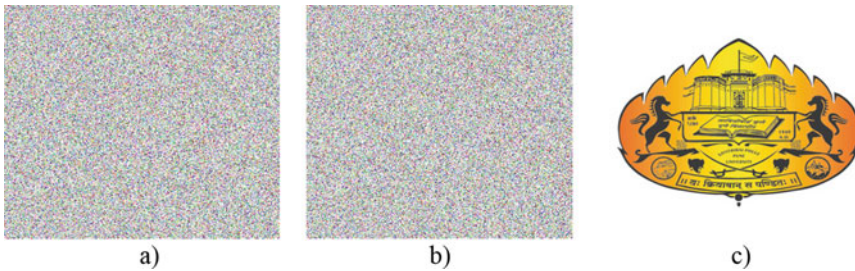


Fig. 6 a, b are shares c reconstructed (decrypted) Colour image for XOR Image Encryption algorithm

The PSNR value is 100 dB, and Mean NCORR value is 0.11484586182709378.
The PSNR value is 100 dB, and Mean NCORR value is 0.037898043706412705.

3.3 Modular Arithmetic Image Encryption

This is based on modular arithmetic and cyclic ring. Each pixel in an image has the value ranging between 0 and 255, it is written modulo 256. The original image is

added to k shares of same size and it's modulo 256 is taken which results in generation of secret image. For decryption, share images are subtracted from the secret image and their modulo 256 values result in decrypted image. This means all the shares and the secret image are required for decryption. If one of the shares is missing it is hard to decrypt. For a brute force attack, it will take $256^{(m*n)}$ states, where $m*n$ is the size of original image (Fig. 7, 8 and 9).

Advantages

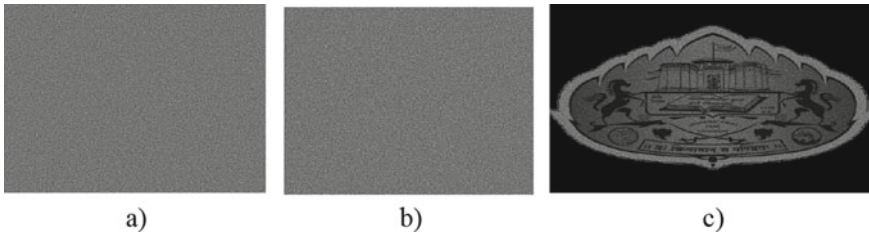


Fig. 7 a, b are binary shares c reconstructed (decrypted) binary image for modular arithmetic image encryption

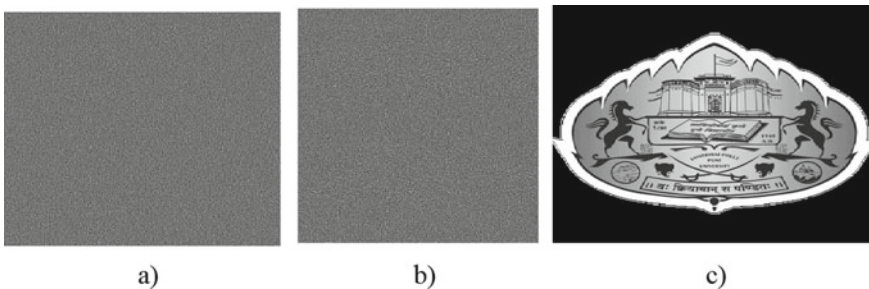


Fig. 8 a, b are Grayscale shares c reconstructed (decrypted) Grayscale image for modular arithmetic image encryption

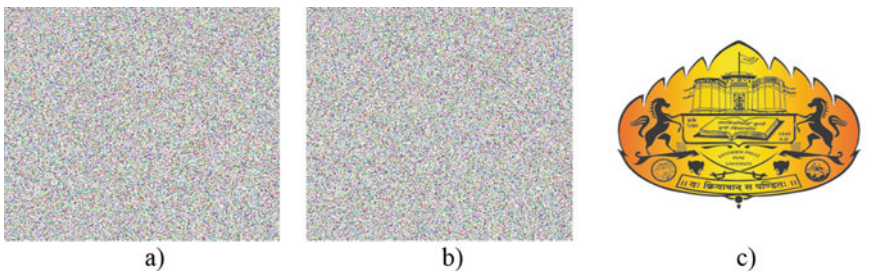


Fig. 9 a, b are shares C reconstructed (decrypted) Colour image for modular arithmetic image encryption

- Easy to build.
- Computation cost is low.
- Can be used to encrypt or decrypt any type of image.

Disadvantages

- Attacker can easily track the secret image.
- This is similar to one-time pad, hence not secure.
- The PSNR value is 100 dB, and Mean NCORR value is -0.09908745073681847 .
- The PSNR value is 100 dB and Mean NCORR value is 0.11484586182709378 .
- The PSNR value is 100 dB and Mean NCORR value is 0.037898043706412705 .

3.4 Bit Level Decomposition Algorithm

This is an extension of Pixel expansion applied on grayscale images. The input grayscale image is divided into binary images. Pixel expansion is applied on that. Resultant binary images are combined to form grayscale shares. There are two ways for decryption. First one is overlap and the second is extraction. In overlap, the decrypted image is double the size of original image, to match with original image it is further resized. In Extraction, pixel with all its subpixels as black are marked as black otherwise white (Figs. 10 and 11).

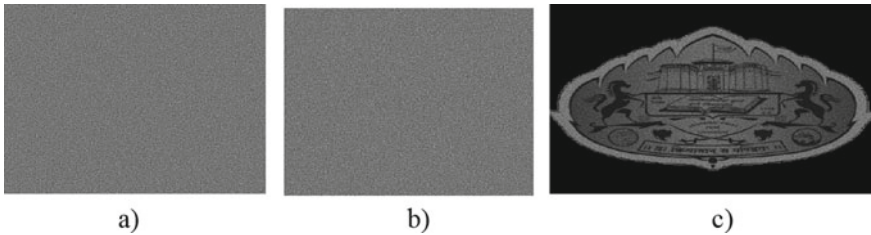


Fig. 10 a, b are Grayscale shares c reconstructed (decrypted) Grayscale image for overlap

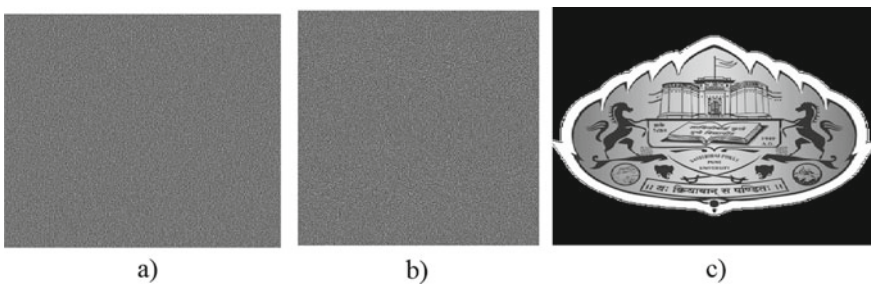


Fig. 11 a, b are Grayscale shares c reconstructed (decrypted) Grayscale image for extraction

Advantages

Applicable to grayscale images of any size.

It is extension of pixel expansion, efficiency is same as pixel expansion.

Disadvantages

Computation costs are high.

Lossy method, original image is not same as decrypted image.

The PSNR value is 30.64300933811834 dB and Mean NCORR value is 0.11574893914784816.

The PSNR value is 100 dB and Mean NCORR value is 0.11484586182709378.

3.5 CMYK Decomposition Algorithm

This is used in colour images. Colour image is decomposed to Cyan, Magenta, Yellow and Black. Three monotone images' conversion using halftoning is done. Using nearest neighbour interpolation, three shares are generated. If we combine all the three shares the original image is retrieved back (Fig. 12).

Advantages

Applicable to colour images of any size.

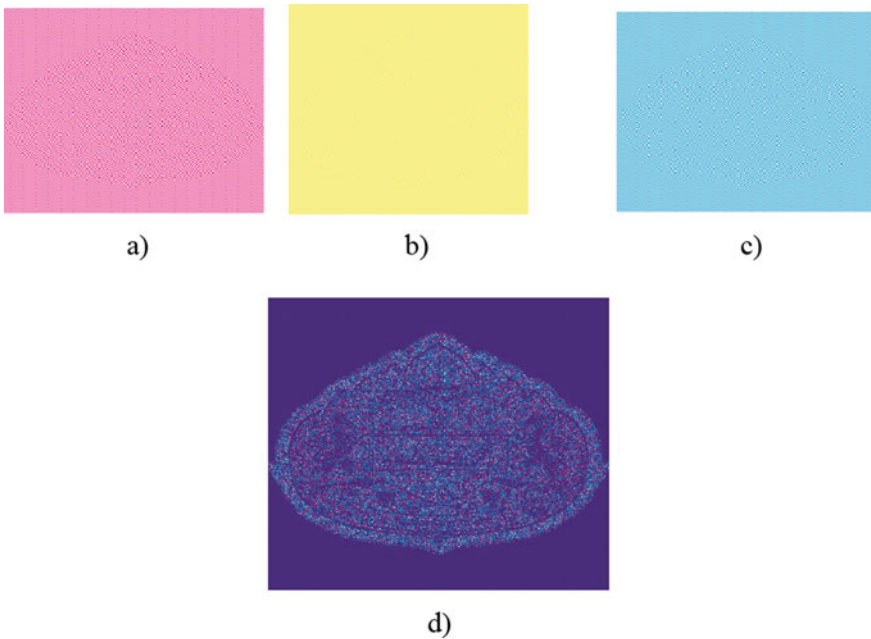


Fig. 12 a, b, c colour shares, d reconstructed (decrypted) colour image

Can be easily extended to n -shares.

Disadvantages

Computation costs are high.

Halftone is used so Lossy method and results in loss of information.

The PSNR value is 100 dB, and Mean NCORR value is 0.13645658129855046.

4 Conclusion

VC Scheme is image sharing method. This scheme is simple, secure, efficient and easy to implement. Decryption can be done by human visual system without doing any computation. This scheme has lots of applications such as bio-metric authentication, medical imaging, steganography, etc. In this paper, already existing algorithms along with their advantages and disadvantages are mentioned. Analysis and comparative study of these algorithm are also done based on two parameters PSNR and Mean Normalized Cross-Correlation. The study shows that an optimum number of shares are required to hide secret information since increased number of shares and pixel expansion affect resolution.

5 Future Work

Visual cryptography is a data security method. It has lots of applications in authentication, human identification, copyright protection, visual signature checking, mobile ticket validation etc. In future, we will try to apply to some real-life example like image security in smart and secure healthcare management as well as we will try to develop our own algorithm for VCS. In healthcare management any image can be shared among n number of clinicians using this technique and if k number of clinicians come together can recover the secret image back. This way sensitive image data or medical record can be protected.

Main open problem in visual cryptography is pixel expansion. It is affecting the quality of reconstructed image. The schemes that address this issue are still costly or inefficient. So optimization technique can be further explored. Development of multiple secret image techniques that are efficient and easy to use can also be done.

References

1. Naor M, Shamir A (1995) Visual cryptography, advances in cryptology—EUROCRYPT94. Lect Notes Comput Sci 950:1–12
2. Ito R, Kuwakado H, Tanaka H (1998) Image size invariant visual cryptography. IEICE Trans Funda: 2172–2177

3. Chang C, Tsait C, Chen T (2000) A new scheme for sharing secret colour images in computer network. In: Proceeding of international conference on parallel and distributed systems, pp 21–27
4. Chang C-C, Yu T-X (2002) Sharing secret gray image in multiple images. National Chung Cheng University, Taiwan
5. Droste S (1996) New results on visual cryptography, advances in cryptology–CRYPTO'96. Lect Notes Comput Sci 1109:401–415
6. Verheul E, Tilborg HV (1997) Constructions and properties of k out of n visual secret sharing schemes. Des Codes Crypt 11(2):179–196
7. Thien C-C, Lin J-C (2002) Secret image sharing. Comput Graph 26:765–770
8. Lukac R, Plataniotis KN (2004) Colour image secret sharing. Electron Lett 40:529
9. Lou D-C, Chen H-H, Wu H-C, Tsai C-S (2011) A novel authenticatable color visual secret sharing scheme using non-expanded meaningful shares. Displays 32:118–134
10. Tsai D-S, Horng G, Chen T-H, Huang Y-T (2009) A novel secret image sharing scheme for true-color images with size constraint. Inf Sci 179:3247–3254
11. Chen T-H, Tsao K-H (2009) Visual secret sharing by random grids revisited. Pattern Recogn 42:2203–2217
12. Alex NS, Anbarasi LJ (2011) Enhanced image secret sharing via error diffusion in halftone visual cryptography. In: Electronics computer technology (ICECT), 2011 3rd international conference on, 2011, pp 393–397
13. Yang C-N (2004) New visual secret sharing schemes using probabilistic method. Pattern Recogn Lett 25:481–494
14. Lin T-L, Horng S-J, Lee K-H, Chiu P-L, Kao T-W, Chen Y-H et al (2010) A novel visual secret sharing scheme for multiple secrets without pixel expansion. Expert Syst Appl 37:7858–7869
15. Sasaki M, Watanabe Y (2014) Formulation of visual secret sharing schemes encrypting multiple images, in Acoustics, Speech and Signal Processing (ICASSP). IEEE International Conference on 2014:7391–7395
16. He J, Lan W, Tang S (2016) A secure image sharing scheme with high quality stego-images based on steganography. Multimed Tools Appl

A New Data Communication Method Using RSA and Steganography



Varun Shukla, Manoj Kumar Misra, Shivani Dixit, and Himanshu Dhumras

1 Introduction

Information exchange is an inseparable part in modern life. The security of information is a key term. Cryptography is all about keeping data secure. Cryptography provides various security categories such as confidentiality, authentication, data integrity and non-repudiation, and they are called as goals of information security or classic goals of cryptography as shown below in Fig. 1 [1–5].

On the other hand, steganography is used to hide the presence of message. Steganography hides the text message in a carrier file. The carrier file can be an image, sound file or a video file. The aim of steganography is also to safeguard information and hence steganography is always seen as a supporting tool of cryptography [6, 7]. A basic comparison between steganography and cryptography is shown in Fig. 2.

The remaining part of this paper is organized as follows: The proposed method is given in Sect. 2. Security analysis and advantages are discussed in Sect. 3. Conclusion and future scope are given in Sect. 4.

V. Shukla · S. Dixit

Department of ECE, Pranveer Singh Institute of Technology, Kanpur, India

M. K. Misra (✉)

Department of CSE, Pranveer Singh Institute of Technology, Kanpur, India

e-mail: manojmisra12@gmail.com

H. Dhumras

Department of Mathematics, Jaypee University of Information Technology, Wakanaghat, India

Classic goals of cryptography



- **Confidentiality:**
 - Information is only accessible to an authorized party
- **Integrity:**
 - Correctness and completeness of information can be verified
- **Authenticity:**
 - Source of information can be verified by a receiving party
- **Non-Repudiation:**
 - Source of information can be verified by any third party

Fig. 1 Goals of information security

	Cryptography	Steganography
Application	Secret communication using scrambled information	Secret communication using hidden information
Supported data	Text	Digital medium (e.g., Text, audio, image, video)
Secret key type	Single (private) Double keys (public)	Single (private)
Key size importance	Critical	Moderate
Processing time	Part of the roundtrip delay	Add processing time to the roundtrip delay
Usage	All communications types	Dependent on payload capacity
Human perception	Visible but unreadable	Invisible/Inaudible
Machine based attack	Cryptanalysis	Steganalysis
Attack result	Secret information recovered	Secret communication detected

Fig. 2 Basic comparison between cryptography and steganography

2 Proposed Method

Step 1: In the first step of the proposed method, RSA is used to generate the cipher text [8, 9]. For RSA, we need two prime numbers p and q and $n = pq$. We calculate

$$\varphi(n) = (p - 1) \times (q - 1).$$

The public key is $\{e, n\}$ and the private key is $\{d, n\}$ where $ed \bmod \varphi(n) = 1$. If the plain text is represented by m and cipher text is given as c then $c = m^e \bmod n$ and $m = c^d \bmod n$. We show the readings in Table 1. These readings are only for illustration point of view but user can extend the values of p and q and then the corresponding parameters will also be changed. The security of RSA algorithm is based on the selection of p and q , and these prime numbers must be large enough for security. So the proposed method provides this flexibility that user can select the large prime numbers also.

Step 2: In the second step, the generated cipher text is kept inside the carrier image (the process of steganography). The selected cipher text and carrier image are shown in Figs. 3 and 4, respectively.

Now the cipher text is kept inside the carrier image so that intruder never knows the presence of cipher text. Since only receiver knows it, he or she will be able to extract the data from the embedded output [10–12]. The comparison of carrier image and embedded output is shown in Fig. 5, and the histogram comparison is also shown in Fig. 6 to prove that both the images look exactly the same.

3 Security Analysis and Advantages

- **Security of RSA:** The proposed method utilizes RSA algorithm for the generation of cipher text. RSA is the most trusted Public Key Cryptosystem (PKC) and its security is still trusted [13, 14]. The large values of prime numbers p and q will make sure that the generated cipher text remains secure from intruders.
- **Usage of steganography:** The use of steganography makes sure that intruders will have no idea about the cipher text. RSA secures the plain text but steganography makes the cipher text invisible. So steganography acts as a second layer of security for the proposed method. User can select any image of his choice as carrier image and generate the corresponding embedded output.
- **Hybrid method:** The proposed method is a combination of cryptography and steganography [15, 16]. Suppose the level of security provided by RSA is A and steganography is B then the overall level of security of the proposed method will be $A + B$. The benefit of using steganography is that the presence of cipher text remains unknown to intruders.

Table 1 Showing readings of the proposed method

Plain text	S.N	p	q	$n = p \times q$	$\varphi(n) = (p - 1) \times (q - 1)$	e	d	Cipher text
Hello Alice let us share a secret number	1	11	13	143	120	7	103	91 62 4 4 45 98 59 4 118 44 62 98 4 62 129 98 39 80 98 80 91 59 49 62 98 59 98 80 62 44 49 62 129 98 33 39 21 32 62 49
	2	17	19	323	288	5	173	168 271 109 109 42 223 241 109 22 131 271 223 109 271 165 223 53 115 223 115 168 241 190 271 223 241 223 115 271 131 190 271 165 223 230 53 181 319 271 190
	3	23	29	667	616	3	411	302 453 416 416 281 85 217 416 380 481 453 85 416 453 116 85 146 115 85 115 302 217 137 453 85 217 85 115 453 481 137 453 116 85 335 146 382 55 453 137

302 453 416 416 281 85 217 416 380 481 453 85 416 453 116 85 146 115 85 115 302 217 137 453 85 217 85 115 453 481 137 453 116 85 335 146 382 55 453 137
--

Fig. 3 Showing the selected cipher text

- Innovative method:** Many security methods have been presented till now using PKCs but the proposed method is a combination of PKC with steganography which is quite unique. This unique combination will show the new direction of research, and it will provide safe and reliable data communication in various applications.
- Customized method:** The proposed method is customized from user’s perspective. User can select prime numbers of his own choice. User can also select the carrier image. Any plain text message can be encrypted and kept inside the carrier image which will produce embedded output. User can select any prime numbers based on the required level of security.



Fig. 4 Showing the carrier image



Fig. 5 Showing comparison of carrier image (left) and embedded output (right)

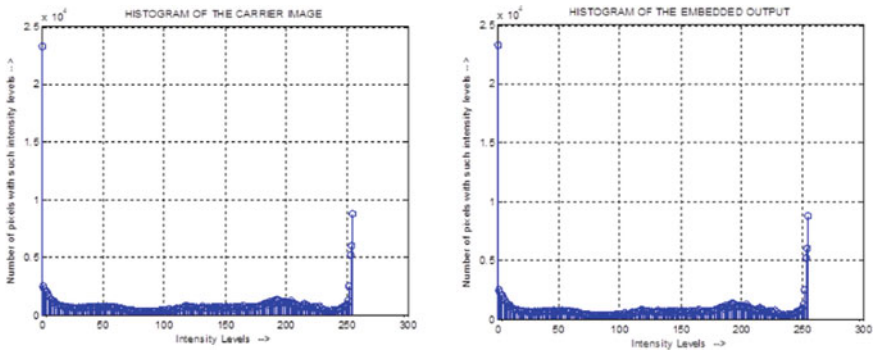


Fig. 6 Showing histogram comparison of carrier image (left) and embedded output (right)

- **Useful in various applications:** The proposed method is a general method that provides strong encryption and then hides the presence of the cipher text using steganography. The proposed method can be used in all those applications where data communication security is vital. The examples can be financial transactions, passing military messages, Electronic Health Records (EHR), e-commerce application, etc. [17–19]. Various password strategies can also be implemented to make the proposed method more secure [20, 21].
- **Resistive against brute force:** The proposed method uses RSA encryption which is resistive against brute force attack. Proper selection of key size makes sure that intruders will not be able to search all the possible combinations in feasible time duration. As an additional security measurement, the presence of cipher text is also hidden which enhances the insurance that intruders will not launch brute force because they don't know about the existence of cipher text.
- **Resistive against DoS:** In Denial of Service (DoS) attack, the intruder intentionally disrupts the ongoing services because they know that communication is going on in an encrypted fashion but in the proposed method, the presence of cipher text is hidden. So intruders will never think of launching DoS [22, 23].
- **Resistive against MITM:** Man in the Middle Attack (MITM) is very dangerous for communication protocols but it is not applicable in the proposed method. RSA itself is resistive against MITM and even if any possibility of MITM is there, it will be nullified by the generated embedded output which hides the cipher text. Since only transmitter and receiver know about the hidden cipher text, there is no question about MITM [24–28].
- **Easily implementable:** The proposed method is easily implementable in various platforms. Mobile apps can also be developed where user needs to select prime numbers and carrier image of his choice and the hidden encrypted message will be transmitted. No additional memory or hardware requirements are needed for the implementation of proposed method. Various hash mechanisms can also be incorporated specifically when the proposed method is used for financial transactions [29–31] or any other-related applications [32–41].

4 Conclusion and Future Scope

An innovative data communication method using RSA and steganography is presented in this paper. The proposed method generates the cipher text using RSA and the cipher text is kept inside the carrier image and embedded output is produced. The carrier image and embedded output look exactly the same, and intruder will not be able to find any difference. The method is an innovative hybrid method and resistive against various well-known security attacks such as brute force, DoS, MITM, etc. The method is easily implementable and can be used in a variety of applications. The future extension of the proposed method is also possible as other encryption algorithms instead of RSA can be used. Similarly, other innovative steganographic procedures can also be applied in order to increase difficulty for intruders.

Acknowledgements The authors thank the editor and the anonymous reviewers for reviewing this article and providing valuable and kind suggestions.

Conflict of Interest The authors declare no competing interests.

References

1. Menezes AJ, Oorschot PCV, Vanstone SA (2001) Handbook of applied cryptography, 5th edn. CRC Press Inc, USA, ISBN: 9780849385230
2. Stallings W (2005) Cryptography and network security, principles and practices, 7th edn. Prentice Hall, ISBN-13:978-0134444284, ISBN-10:0134444280
3. Shukla V, Chaturvedi A, Srivastava N (2019) Nanotechnology and cryptographic protocols: issues and possible solutions. *Nanomater Energy* 8(1):1–6. <https://doi.org/10.1680/jnaen.18.00006>
4. Shukla V, Chaturvedi A, Srivastava N (2019) A secure stop and wait communication protocol for disturbed networks. *Wirel Pers Commun* 110:861–872. <https://doi.org/10.1007/s11277-019-06760-w>
5. Shukla V, Srivastava N, Chaturvedi A (2016) A bit commitment signcryption protocol for wireless transport layer security (wtls). In: IEEE international conference on electrical, computer and electronics engineering, pp 83–86. <https://doi.org/10.1109/UPCON.2016.7894629>
6. Subramanian N, Elharrouss O, Maadeed SA, Bouridane A (2021) Image steganography: a review of the recent advances. *IEEE Access* 9:23409–23423. <https://doi.org/10.1109/ACCESS.2021.3053998>
7. Shukla V, Mishra A (2020) A new sequential coding method for secure data communication. In: IEEE international conference on computing, power and communication technologies, pp 529–533. <https://doi.org/10.1109/GUCON48875.2020.9231252>
8. Zhou X, Tang X (2011) Research and implementation of RSA algorithm for encryption and decryption. In: Proceedings of 6th international forum on strategic technology, pp 1118–1121. <https://doi.org/10.1109/IFOST.2011.6021216>
9. Kaabi SSA, Belhaouari SB (2019) Methods toward enhancing RSA algorithm: a survey. *Int J Netw Secur Appl* 11(3):53–70. <https://doi.org/10.5121/ijnsa.2019.11305>
10. Maniriho P, Ahmad T (2019) Information hiding scheme for digital images using difference expansion and modulus function. *J King Saud Univ Comput Inf Sci* 31(3):335–347. <https://doi.org/10.1016/j.jksuci.2018.01.011>
11. Attaby AA, Ahmed MFMM, Alsammak AK (2018) Data hiding inside JPEG images with high resistance to steganalysis using a novel technique: DCT-M3. *Ain Shams Eng J* 9(4):1965–1974. <https://doi.org/10.1016/j.asej.2017.02.003>
12. Chaturvedi A, Shukla V, Misra MK (2021) A random encoding method for secure data communication: an extension of sequential coding. *J Discr Math Sci Cryptogr* 24(5):1189–1204. <https://doi.org/10.1080/09720529.2021.1932902>
13. Kaur J, Ramkumar KR (2021) The recent trends in cyber security: a review. *J King Saud Univ Comput Inf Sci*. <https://doi.org/10.1016/j.jksuci.2021.01.018>
14. Hassan MA, Shukur Z, Hasan MK (2020) An efficient secure electronic payment system for e-commerce. *Computers* 9(3):1–13. <https://doi.org/10.3390/computers9030066>
15. Taha MS, Rahim MSM, Lafta SA, Hashim MM, Alzuabidi HM (2019) Combination of steganography and cryptography: a short survey. *IOP Conf Ser Mater Sci Eng* 518(5):1–13. <https://doi.org/10.1088/1757-899X/518/5/052003>
16. Jan A, Parah SA, Hussan M, Malik BA (2021) Double layer security using crypto-stego techniques: a comprehensive review. *Health Technol* 1–23. <https://doi.org/10.1007/s12553-021-00602-1>

17. Shukla V, Chaturvedi A, Srivastava N (2015) A new secure authenticated key agreement scheme for wireless (mobile) communication in an EHR system using cryptography. *Commun Appl Electron* 3(3):16–21. <https://doi.org/10.5120/cae2015651903>
18. Misra MK, Chaturvedi A, Tripathi SP, Shukla V (2019) A unique key sharing protocol among three users using non-commutative group for electronic health record system. *J Discr Math Sci Cryptogr* 22(8):1435–1451. <https://doi.org/10.1080/09720529.2019.1692450>
19. Chaturvedi A, Srivastava N, Shukla V, Tripathi SP, Misra MK (2015) A secure zero knowledge authentication protocol for wireless (mobile) ad-hoc networks. *Int J Comput Appl* 128(2):36–39. <https://doi.org/10.5120/ijca2015906437>
20. Shukla V, Chaturvedi A, Srivastava N (2019) A new one time password mechanism for client-server applications. *J Discr Math Sci Cryptogr* 22:1393–1406. <https://doi.org/10.1080/09720529.2019.1692447>
21. Zviran M, Haga WJ (1999) Password security: an empirical study. *J Manage Inf Syst* 15(4):161–185. <https://www.jstor.org/stable/40398409>
22. Yang ZC (2011) DOS attack analysis and study of new measures to prevent. In: *International conference on intelligence science and information engineering*, pp 426–429. <https://doi.org/10.1109/ISIE.2011.66>
23. Mahjabin T, Xiao Y, Sun G, Jiang W (2017) A survey of distributed denial-of-service attack, prevention, and mitigation techniques. *Int J Distrib Sens Netw* 13(12):1–34. <https://doi.org/10.1177/1550147717741463>
24. Aliyu F, Sheltami T, Shakshuki EM (2018) A detection and prevention technique for man in the middle attack in fog computing. *Proc Comput Sci* 141:24–31. <https://doi.org/10.1016/j.procs.2018.10.125>
25. Conti M, Dragoni N, Lesyk V (2016) A survey of man in the middle attacks. *IEEE Commun Surv Tutor* 18(3):2027–2051. <https://doi.org/10.1109/COMST.2016.2548426>
26. Chaturvedi A, Srivastava N, Shukla V (2015) A secure wireless communication protocol using Diffie-Hellman key exchange. *Int J Comput Appl* 126(5):35–38. <https://doi.org/10.5120/ijca2015906060>
27. Shukla V, Chaturvedi A, Misra MK (2021) On authentication schemes using polynomials over non commutative rings. *Wirel Pers Commun* 118(1):1–9. <https://doi.org/10.1007/s11277-020-08008-4>
28. Mallik A, Ahsan A, Shahadat MMZ, Tsou JC (2019) Man-in-the-middle-attack: understanding in simple words. *Int J Data Netw Sci* 3(2):77–92. <https://doi.org/10.5267/j.ijdns.2019.1.001>
29. Shukla V, Misra MK, Chaturvedi A (2021) A new authentication procedure for client-server applications using HMAC. *J Discr Math Sci Cryptogr* 24(5):1241–1256. <https://doi.org/10.1080/09720529.2021.1932908>
30. Shukla V, Chaturvedi A, Srivastava N (2019) Authentication aspects of dynamic routing protocols: associated problem & proposed solution. *Int J Recent Technol Eng* 8(2):412–419. <https://doi.org/10.35940/ijrte.B1503.078219>
31. Shukla V, Mishra A, Agarwal S (2020) A new one time password generation method for financial transactions with randomness analysis. *Innovations in electrical and electronic engineering (Part of the lecture notes in electrical engineering book series (LNEE, vol 661))*, pp 713–720. https://doi.org/10.1007/978-981-15-4692-1_54
32. Shukla V, Mishra A, Yadav A (2019) An authenticated and secure electronic health record system. In: *IEEE international conference on information and communication technology*, pp 1–5. <https://doi.org/10.1109/CICT48419.2019.9066168>
33. Chaturvedi A, Shukla V, Misra MK (2018) Three party key sharing protocol using polynomial rings. In: *5th IEEE Uttar Pradesh section international conference on electrical, electronics and computer engineering (UPCON)*, pp 1–5. <https://doi.org/10.1109/UPCON.2018.8596905>
34. Shukla V, Chaturvedi A, Srivastava N (2017) Secure wireless communication protocol: to avoid vulnerabilities in shared authentication. *Commun Appl Electron* 7(6):4–7. <https://doi.org/10.5120/cae2017652680>
35. Shukla V, Misra MK, Chaturvedi A (2022) Journey of cryptocurrency in India in view of financial budget 2022–23. *Cornell university arxiv*, pp 1–6. <https://doi.org/10.48550/arXiv.2203.12606>

36. Chaturvedi A, Shukla V (2020) Miracle of number theory. *Everyman's Sci* 50(3–4):131–134. http://www.sciencecongress.nic.in/pdf/e-book/august_nov_2020.pdf
37. Shukla V, Chaturvedi A (2018) Cryptocurrency: characteristics and future perspectives, vol 53, number 2, pp 77–80. <http://164.100.161.164/pdf/e-book/june-july-18.pdf#page=14>
38. Shukla V, Kushwaha A, Parihar SS, Srivastava S, Singh VP (2016) Authenticated wireless information display system using GSM module. *Commun Appl Electron* 5(3):7–11. <https://doi.org/10.5120/cae2016652251>
39. Shukla V, Chaturvedi A, Srivastava N (2017) Double layer cryptographic protocol for mobile ad-hoc networks (MANETs) by commitment scheme. *Commun Appl Electron* 7(9):32–36. <https://doi.org/10.5120/cae2017652716>
40. Shukla V, Dixit S, Dixit P (2022) An IoT based user authenticated soil monitoring system. *Adhoc Sensor Wirel Netw* 53(3–4):269–283. <https://doi.org/10.32908/ahsw.n.v53.9453>
41. Chaturvedi A, Shukla V, Srivastava N (2017) A secure wireless peer to peer authentication protocol using triple decomposition problem. *Asian J Math Comput Res* 22(2):63–69. <https://archives.biciconference.co.in/index.php/AJOMCOR/article/view/1167>

Some Computational Attacks on Threshold Secret-Sharing Scheme by Outside Adversaries



L. Sreenivasulu Reddy

1 Introduction

Secret-sharing schemes are designed to distribute a secret to a group of authenticated participants such that the secret will be retrieved later by the combiner who is also a member in the authenticated participants group. This scheme is named as threshold secret-sharing scheme when the combiner will be retrieved the total secret without all authenticated participants shares. The public parameters of any (t, n) -threshold secret-sharing schemes are represented by the pair of integers t and n . The parameter n denotes the number of shares in the secret distribution process and t denotes the minimum number of shares required to retrieve the secret. The threshold secret-sharing scheme is perfect if the total secret is not retrieved with less than threshold value t number of shares. When the participants reveal their shares at the same time to the combiner at secret retrieval process, the process is called synchronous. When the participants reveal their shares one at a time to the combiner at secret retrieval process, the process is called asynchronous.

Numerous mathematical concepts, including field theory, number theory, and numerical methods, have a substantial impact on (t, n) -threshold secret-sharing schemes, according to the literature on secret-sharing schemes. Many of these (t, n) -threshold secret-sharing systems produce a secret share for authenticated participants using a polynomial $f(x) = a_0 + a_1x + \dots + a_{t-1}x^{t-1}$ over a finite field F_p , where p is a predetermined prime and $t > 1$ is an integer. Shamir [1] proposed the first polynomial-based threshold secret-sharing system in 1979. It is well knowledge that polynomials are crucial to the theory of the algebraic structure of finite fields. A secret-sharing technique based on polynomials was later presented by Sun and Shieh [2]. They built their scheme using the Diffie-Hellman concept. In order to

L. Sreenivasulu Reddy (✉)

Department of Mathematics, Kalasalingam Academy of Research and Education, Krishnankoil, Tamil Nadu, India

e-mail: sreenivasulureddy.svu@gmail.com

build their secret-sharing mechanism, Hwang and Chang [3] also used polynomials. A recent secret-sharing technique with secure secret reconstruction was proposed by Jian Ding, Pinhui Ke, Changlu Lin, and Huaxiong Wang [4]. It uses two variables and is based on asymmetric polynomials. Çalkavur et al. [5] proposed a secret-sharing scheme based on polynomials over exploiting the structure of field extension of degree $d + 1$, and Sergey Bexxateev, Vadim Davydov, and Ometov [6] proposed Newton's Polynomial based secret-sharing scheme.

On the other hand, studies on secret-sharing schemes have been done in a variety of ways. For instance, Gharahi and Khazaei [7] investigated the best linear secret-sharing techniques for Graph Access Structures on Six Participants, and Oguzhan Erasoy, et al. [8] investigated homomorphic extensions of CRT-based sharing. Furthermore, a verifiable secret-sharing method with combiner verification and cheater identification was studied by Kndar and Dhara [9]. En Zhang, et al. [10] investigated a reputation-based outsourcing hierarchical threshold secret-sharing protocol, which only requires one round of operation and allows participants from different levels to fairly reconstruct the secret. A secret-sharing technique based on both (t, n) -threshold and adversarial structure was studied by Huawang Qin, Yuewei Dai, and Zhiquan Wang [11] to prevent participants from providing their genuine shadows when the shared secret is modified.

First and foremost, Shamir's approach [1] was demonstrated by Martin Tompa and Heather Woll [12] as a method for some types of cheating in the Lagrange interpolation polynomial-based secret-sharing scheme. They maintain Shamir's scheme's property that the security of the system is not based on any unproven hypotheses, such as the intractability of computing number-theoretic functions. When shares are revealed asynchronously, inside competitors can cheat, according to Tompa and Woll's method. An inside adversary can always release a false share last in the secret reconstruction of Shamir's secret-sharing scheme when shares are released asynchronously to ensure that the dishonest shareholder can only obtain the secret; nevertheless, other honest shareholders can also obtain a fake secret. Later, Harn, Lin, and Li [13] designed a secret reconstruction scheme that secures against both internal and external attackers and does not require an interactive dealer, challenging cryptographic primitives, or any assumptions on the number of truthful shareholders. In order to determine the potential for cryptanalysis in Shamir's secret-sharing scheme (SSSS) when secrets are revealed synchronously, Tieng and Nocon [14] first examined the scheme's vulnerability to both a single insider adversary as well as a group of inside adversaries under synchronous and asynchronous secret sharing.

All previous efforts on polynomial-based secret-sharing schemes lack any information on how to choose the degree value of the polynomial so that the scheme is perfect. Although no mathematical justification is given, the best value for t in a (t, n) -threshold interpolation polynomial-based secret-sharing system is one that prevents any adversary from discovering the secret or even the chance that they might. Regardless of the secret reveal process, adversaries in all of their research require a minimum number of secret sharing from authenticated participants to obtain the entire secret (synchronous or asynchronous). In this paper, we'll show how adversaries can reveal the complete secret when there are fewer secret shares available

than what's necessary (threshold number). In this work, the shares of the secret are revealed either synchronously or asynchronously.

The rest of this paper is organized as follows. In the next section, we introduce some preliminaries of polynomial-based secret-sharing schemes including Shamir's secret-sharing scheme along with the best bounds for assumptions of Shamir's secret-sharing scheme. In Sect. 3, Adversaries' possibilities on interpolation polynomial-based (t, n) -threshold secret-sharing schemes including Shamir's secret-sharing scheme. In Sect. 4, the Adversary is outside the threshold secret-sharing scheme. The conclusion is given in Sect. 5.

2 Preliminaries and Best Range for the Parameter t

This section will explore secret-sharing schemes, outline the Shamir secret-sharing premise, and present examples of secret reconstruction by many groups when each group has a minimum number of members. The optimum bound for parameter t in secret-sharing schemes, such as Shamir's secret-sharing scheme, should be found so that the secret can be reconstructed by a distinct group that includes the combiner.

Secret-Sharing Schemes

In a secret-sharing scheme, the n number of authenticated participants each has their own secret share, denoted by the notion s_i for $i = 1, 2, 3, \dots, n$. These shares s_i 's are created and distributed by the dealer using a special function with the input secret s , which allows the secret s to be retrieved from the entire set of shares s_1, s_2, \dots, s_n . Sometimes, a suitable subset of the entire set of shares can be used to retrieve the secret. The secret-sharing scheme in these situations is referred to as a threshold secret-sharing scheme. The secret-sharing scheme is referred to as perfect if no appropriate subset of the entire set of shares is capable of retrieving the secret. If t secret shares out of the total set of n secret shares are used to retry the secret, the secret-sharing scheme is known as a threshold (t, n) -secret-sharing scheme. If knowledge of $t - 1$ or fewer shares does not reveal any information regarding the total secret s , the threshold scheme is perfect. Shamir's secret-sharing scheme is interpreted in various ways by researchers; here, we focus on the interpretation that was taken from Stinson [15] and Trappe [16].

Shamir's (t, n) -threshold secret-sharing scheme serves as a model for secret-sharing schemes. Using the numerical algorithm known as Lagrange interpolation, this scheme is created as a perfect (t, n) -threshold scheme. The mathematical description of this scheme is that the dealer selects n randomly chosen distinct non-zero values $x_i \not\equiv 0 \pmod{p}$, where p is a positive prime integer, in the field F_p and also selects the polynomial $f(x) = a_0 + a_1x + \dots + a_{n-1}x^{t-1}$ of degree at most $t - 1$ over a field F_p such that the constant term of the polynomial $a_0 \in F_p$ is the secret s and other a_i 's for $i = 1, 2, \dots, t - 1$ are constants in the field F_p such that $a_{t-1} \not\equiv 0 \pmod{p}$. The order pairs (x_i, y_i) are then sent to a i^{th} participant who owns the share with the dealer computing $y_i = f(x_i)$ for $x_i \not\equiv 0 \pmod{p}$. At a later time,

combiner requires at least t shares of (x_i, y_i) in order to reconstruct the equation $f(x) = a_0 + a_1x + \dots + a_{t-1}x^{t-1}$, and as a result, discover the secret $s = a_0$ as well by using Lagrange's interpolation polynomial.

Assumptions of the Shamir's Secret-Sharing Scheme

Other than the trust of the dealer and combiner, Shamir's secret-sharing scheme has the following assumptions:

- The dealer should not be a participant in the scheme.
- The parameters t, n are public.
- One of the components x_i of each share is the public.
- The secret should be a member in the field F_p and it is a constant term of the polynomial.
- The system of t number of non-homogeneous linear equations must have solutions.

Best Bound for Parameter t in (t, n) -Threshold Secret-Sharing Scheme

Suppose the dealer chooses the polynomial of degree $t - 1$ such that t is less than or equal to $\lfloor \frac{n}{2} \rfloor$ where n is the total number of authenticated participants in the threshold secret scheme to share the secret. Later, the combiner collects t number of secret shares from the n number of participants including his share to reconstruct the secret. Since $\leq \lfloor \frac{n}{2} \rfloor$, so $2t \leq n$ because $\lfloor \frac{n}{2} \rfloor \leq \frac{n}{2}$. Thus, there is more than one t number of distinct authenticated participant groups. The combiner may be a member of one group only. So, if the other group of t members can contribute their secret shares, then they can construct the threshold polynomial without a combiner in the process. Therefore the best bound for the parameter t is greater than $\lfloor \frac{n}{2} \rfloor$.

For example, consider a field F_p where $p = 17$. Therefore, the coefficients of the polynomial and secret share pair values are in finite field F_p . For instance, the secret scheme has five authenticated parties $n = 5$, the degree of the polynomial is $t = 2$, and the public components say $x_i = i, 1 \leq i \leq 5$. Let's say that out of five two: P_1 and P_3 (one of them is a combiner) pool their shares, which are respectively 8 and 10. Consider the polynomial $a(x)$ as $a(x) = a_0 + a_1x$ and compute $a(1)$ and $a(3)$. Then we obtain the following two linear equations: $a_0 + a_1 = 8$ and $a_0 + 3a_1 = 10$ in Z_{17} . This system has a solution in the finite field Z_{17} : $a_1 = 1$ and $a_0 = 7$. Therefore, the secret S is $a_0 = 7$. Suppose there is another participants group say: P_2 and P_5 (no one is a combiner) different from the previous participants group such that they pool their shares, which are, respectively, 9 and 12. Compute $a(2)$ and $a(5)$ using the above polynomial $a(x) = a_0 + a_1x$. It yields the following two linear equations $a_0 + 2a_1 = 9$ and $a_0 + 5a_1 = 12$ in Z_{17} . This system has a solution in Z_{17} : $a_1 = 1$ and $a_0 = 7$. Thus, this group of participant finds the secret S is $a_0 = 7$.

Based on this disadvantage, choose the threshold parameter t such that there no more than one patrician subsets having the order t of the set of authenticated participants. Consequently, this is possible only when $t > \lfloor \frac{n}{2} \rfloor$.

3 Adversaries' Possibilities on Interpolation Polynomial-Based (t, n) -Threshold Secret-Sharing Schemes

Our contribution to the paper begins in this part with the potential adversarial mathematical attacks on any secret-sharing system for obtaining the secret unethically. The adversary in this case could not be a member of the scheming party and the dealer. The dealer doesn't need to perform any computations in order to obtain the secret because he already knows it. The primary chances of the scheme when he tackles any interpolation polynomial-based (t, n) -threshold secret-sharing schemes.

- (i) There is no any information for selecting degree t value to the polynomial such that the scheme becomes perfect.
- (ii) Let x'_i 's are public in a scheme. The secret $S = a_0$ can be computed by the adversary if he knows $k \neq 1$ number of shares (x_i, y_i) .
- (iii) Let x'_i 's are not public in a scheme. The secret $S = a_0$ can be computed by the adversary if he knows $k \neq 1$ number of shares (x_i, y_i) along with at least $t - k$ number of others x'_i 's values.
- (iv) The secret $S = a_0$ stored in the constant term of the polynomial $f(x)$ only and so the secret is get easily if other t coefficients a_1, a_2, \dots, a_t values are known to us.
- (v) If the field F_p size p is small, the internal/external adversary can able to reveal the secret with at most $(p - 1)(p - 2) \dots (p - t)$ chances.

The diagrammatic approach of the adversary possibilities to find threshold secret sharing polynomial based on combinatorial principle is shown in Fig. 1.

There is insufficient information regarding how to choose polynomial coefficients in the threshold secret-sharing methods. However, the polynomial is univariate and monic in the majority of research studies. The best and worst algorithms for adversaries to create threshold secret-sharing polynomials are presented below, respectively.

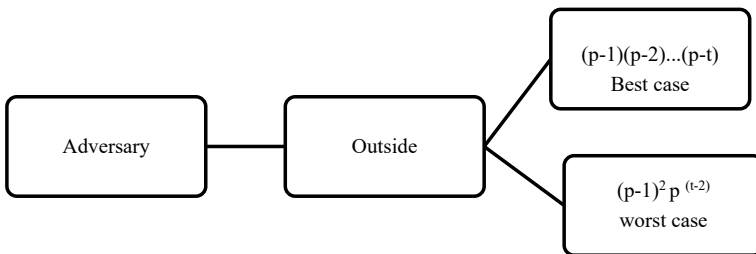


Fig. 1 Outside adversary computational possibility bounds

Worst-Case Algorithm

Choose the coefficients $a_i, i = 0, 1, 2, \dots, t - 1$ of the polynomial $f(x) = a_0 + a_1x + \dots + a_{n-1}x^{t-1}$ from the finite field F_p .

The constant term a_0 has $(p - 1)$ chances because a_0 is the total secret which must be non-zero.

1. The coefficient a_{n-1} must be nonzero because the polynomial has degree $t - 1$, so a_{n-1} is chosen in $(p - 1)$ ways.
2. For the other $t - 2$ coefficients $a_i, i = 1, 2, \dots, t - 2$, each one is chosen in p ways.
3. Thus the polynomial is chosen in $(p - 1).p.p \dots (p - 1) = (p - 1)^2 p^{t-2}$ ways.

Best-Case Algorithm

Choose the coefficients $a_i, i = 0, 1, 2, \dots, t - 1$ of the polynomial $f(x) = a_0 + a_1x + \dots + a_{n-1}x^{t-1}$ from finite field F_p . But in most of the cases, the polynomials have been chosen such that the coefficients are pair wise distinct.

1. The constant term a_0 has $(p - 1)$ chances because a_0 be the total secret which must be non-zero.
2. The coefficient a_{n-1} must be non-zero because the polynomial has degree $t - 1$ and chose differently from a_0 , so a_{n-1} is chosen in $(p - 2)$ ways.
3. For the other $t - 2$ coefficients $a_i, i = 1, 2, \dots, t - 2$, each one is chosen in $(p - 3), (p - 4), \dots, (p - t)$ ways, respectively.
4. Thus, the polynomial is chosen in $(p - 1).(p - 2).(p - 3) \dots (p - t)$ ways.

4 Role of Outside Adversary

Throughout this section, the adversary is an outsider who is different from the dealer. He knows the public information about the scheme only. Therefore, he has no secret share (x_i, y_i) but he has filed F_p and polynomial of degree t . This section discusses the above two possible ways of attacking an interpolation polynomial based secret-sharing scheme.

The adversary possibilities based on x_i 's are represented diagrammatically as shown in Fig. 2.

Case 1: x_i 's are public.

Consider a filed F_p where $p = 17$. Therefore, the coefficients of the polynomial and secret share pair values (the first component is public and the second is private) are in F_p . For example, the secret-sharing scheme has five authenticated participants $n = 5$, the degree of the polynomial is $t = 3$, and public components say $x_i = i, 1 \leq i \leq 5$. Suppose that among the five participants, the following three participants P_1, P_3, P_5 pool their shares, which are, respectively, 8, 10, and 11. Choose the polynomial $a(x)$ as $a(x) = a_0 + a_1x + a_2x^2$ because we have three

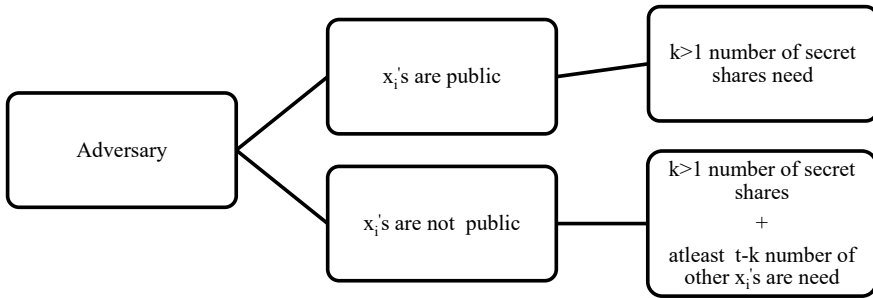


Fig. 2 Outside adversary computational value depending on public key information

shares only. Then compute $a(1), a(3)$, and $a(5)$. These yields the following three linear equations: $a_0 + a_1 + a_2 = 8, a_0 + 3a_1 + 9a_2 = 10$ and $a_0 + 5a_1 + 8a_2 = 11$ in Z_{17} . This system has a unique solution in Z_{17} : $a_0 = 13, a_1 = 10$, and $a_2 = 2$. Therefore, the key S is $a_0 = 13$.

The aforementioned example explains how an outside adversary could attack a secret-sharing scheme.

First way: Suppose the outside adversary gets the secret share of any one of the authenticated participants, say $(3, 10)$. Adversaries need two secret sharing pairs of others because the degree of the polynomial is $t = 3$. Here $x_i = i$'s are public and so in each pair, the second component is only a varying quantity. In the field we have $(p - 1) = (17 - 1) = 16$ number of non-zero elements and among them one element 10 (the second component of the adversary) is for the adversary. After Adversary's second element, the remaining elements are $(p - 2) = (17 - 2) = 15$. Now the adversary has the chance to get the other two pairs of secretes is in $(p - 3).(p - 4) = 14.13 = 182$ ways to get secretes share b_0 such that $b(1) = 8$. Here, the second component of secret share pairs is only varying because they are private.

Second way: Suppose the outside adversary gets two secret shares of any two authenticated participants, say $(3, 10)$ and $(5, 11)$ of other authenticated party. Then he chose a polynomial $b(x) = b_0 + b_1x + b_2x^2$ of degree two such that $b(3) = 10$ and $b(5) = 11$ over the field Z_{17} , the following two equations are $b_0 + 3b_1 + 9b_2 = 10$ and $b_0 + 5b_1 + 8b_2 = 11$.

Subtract one equation from the other, we get an equation with two variables b_1, b_2 : $2b_1 + 16b_2 = 1$. For choosing one value either b_1 or b_2 , we get the other. Suppose b_1 is chosen first from F_p , then we get p number of solutions (b_1, b_2) and among them, remove $b_2 = 0$ because the polynomial degree is two. Thus the total number of possible number of solutions is $p - 1$. Substituting each (b_1, b_2) value in $b_0 + 3b_1 + 9b_2 = 10$ or $b_0 + 5b_1 + 8b_2 = 11$, we get the secret b_0 with $p - 1$ possible values. That is there are $p - 1 = 17 - 1 = 16$ chances to find the secret b_0 such that $b(1) = 8$.

Third way: Suppose the outside adversary gets the two secret shares of any two authenticated participants, say (3, 10) and (5, 11). Then he chose a polynomial $b(x) = b_0 + b_1x + b_2x^2$ of degree two such that $b(3) = 10$ and $b(5) = 11$ over the field Z_{17} , the following two equations are $b_0 + 3b_1 + 9b_2 = 10$ and $b_0 + 5b_1 + 8b_2 = 11$.

Since $x_i = i$'s are public, so he can pick any one of $x_i = i$'s other than above two $x_i = i$'s. For example he picks $x_1 = 1$ then substitutes in $b(x)$, we get $b(1) = b_0 + b_1 + b_2$ and this value is not known, say $b(1) = a$. Find a such that the system of equations $b_0 + 3b_1 + 9b_2 = 10$, $b_0 + 5b_1 + 8b_2 = 11$, and $b_0 + b_1 + b_2 = a$ has unique solution such that $a \neq 0, 10, 11$ modulo p . Because every authenticated party has their own different shares and each has at least some contribution, the value a must be different from other existing share values 10, 11 modulo p and also non-zero. Thus, there are $p - (t + 1) = 17 - (2 + 1) = 14$ possible chances to get $b(1) = 8$. Find all $b(x)$ values, so for $x_1 = 1$ then $b(1) = 8$.

Fourth way: Suppose the outside adversary gets a secret shares of two authenticated participants, say (3, 10) and (5, 11). Then, he chose a polynomial $b(x) = b_1 + b_2x$ of degree one such that $b(3) = 10$ and $b(5) = 11$ over the field Z_{17} , the following two equations are $b_1 + 3b_2 = 10$ and $b_1 + 5b_2 = 11$. Then solving the above equation, we get $b_1 = 0$ and $b_2 = 9$ over the field Z_{17} . Thus the polynomial equation $b(x) = b_1 + b_2x$ becomes $b(x) = 0 + 9.x = 9.x$. Since $x_i = i$'s are public, so he pick any one of $x_i = i$'s other than the above two $x_i = i$'s. For example he pick $x_1 = 1$ then substitute in $b(x)$, we get $b(1) = 0 + 9.1 = 9$ this is the approximate secrete value, but actual secret value is $a_0 = 13$. The existing secret is near to the original secret, so the adversary can get the exact secret with fewer chances (in the worst case with in $(p - 2) = 15$). If the degree of the polynomials is more, then the adversary can get good approximate value to the secret.

Case 2: x_i 's are not public.

First way: Suppose the outside adversary gets a secret share of any one of the authenticated participants, say (3, 10). Adversaries need two secret sharing pairs of others because the degree of the polynomial is $t = 3$. Here $x_i = i$'s are not public and so in each pair both the first and second components are varying quantity. In the field, we have $(p - 1) = (17 - 1) = 16$ non-zero elements and so we have $(p - 1).(p - 1) = 16.16 = 256$ number of order pairs. The secret share order pair (3, 10) is known to him, so he needs two other secret sharing order pairs from the remaining $256 - 1 = 255$ order pairs. He can choose two order pairs among 255 order pairs in $255.(255 - 1) = 255.254 = 64770$ ways. Thus the adversary have the chances to get the other two pair of secretes is in 64770 ways to get secrete share a_0 such that $a(1) = 8$. Here, the second components of secret share pairs are only varying because they are private.

Second way: Suppose the outside adversary gets two secret shares of authenticated participants, say (3, 10) and (5, 11). Then he chose a polynomial $b(x) = b_0 + b_1x + b_2x^2$ of degree two such that $b(3) = 10$ and $b(5) = 11$ over the filed Z_{17} , the following two equations are $b_0 + 3b_1 + 9b_2 = 10$ and $b_0 + 5b_1 + 8b_2 = 11$.

Subtract one equation from the other we get an equation in two variables b_1, b_2 : $2b_1 + 16b_2 = 1$. For choosing one value either b_1 or b_2 , we get the other. Suppose b_1 is chosen first from F_p , then we get p number of solutions (b_1, b_2) and among them remove $b_2 = 0$ because the polynomial degree is two. Thus the total number of possible solutions is $p - 1$. Substituting each (b_1, b_2) value in $b_0 + 3b_1 + 9b_2 = 10$ or $b_0 + 5b_1 + 8b_2 = 11$, we get the secret b_0 with $p - 1$ possible values. That is there are $p - 1 = 17 - 1 = 16$ chances to find the secret b_0 such that $b(3) = 10$ and $b(5) = 11$. In this way no need to guess others x_i 's. The computational value is same as the computational value in the case of x_i 's are public.

Third way: Suppose the outside adversary gets two secret shares of any two authenticated participants, say $(3, 10)$ and $(5, 11)$. Then, he chooses a polynomial $b(x) = b_0 + b_1x + b_2x^2$ of degree two such that $b(3) = 10$ and $b(5) = 11$ over the field Z_{17} , the following two equations are $b_0 + 3b_1 + 9b_2 = 10$ and $b_0 + 5b_1 + 8b_2 = 11$.

Since $x_i = i$'s are not public, he can't pick any exact share with the first component x_i but he can guess in $(p - 1) - 2 = 14$ ways (other than known shares and zero) from the field F_p , so he has 14 equations. For example, he picks one value from 14 values in the field F_p , say $x_1 = 1$ then substitute in $b(x)$, we get $b(1) = b_0 + b_1 + b_2$ and this value is not known say $b(1) = a$. Find a such that the system of equations $b_0 + 3b_1 + 9b_2 = 10$, $b_0 + 5b_1 + 8b_2 = 11$, and $b_0 + b_1 + b_2 = a$ has a unique solution such that $a \neq 0, 10, 11$ modulo p . Because every authenticated party has its own different shares, and each has atleast some contribution, the value of a must be different from other existing share values 10, 11 modulo p and also non-zero.

Thus, there are $p - (t + 1) = 17 - (2 + 1) = 14$ possible chances to get $b(3) = 10$ and $b(5) = 11$. Repeat this process for each value from the other 13 values of the field. For each one, we have 14 possible chances to get $b(3) = 10$ and $(5) = 11$. Therefore, he finds all $b(x)$ values, so that the secret $a_0 = 13$ in $14 \cdot 14 = 196$ ways.

Fourth way: Suppose the outside adversary gets two secret shares of any two authenticated participants, say $(3, 10)$ and $(5, 11)$. Then, he chose a polynomial $b(x) = b_1 + b_2x$ of degree one such that $b(3) = 10$ and $b(5) = 11$ over the field Z_{17} , the following two equations are $b_1 + 3b_2 = 10$ and $b_1 + 5b_2 = 11$. Then, solving the above equation, we get $b_1 = 0$ and $b_2 = 9$ of x_i over the field Z_{17} . Thus, the polynomial equation $b(x) = b_1 + b_2x$ becomes $b(x) = 0 + 9 \cdot x = 9 \cdot x$. Since $x_i = i$'s are not public, he can't pick any one of the exact shares value with the first component x_i , but he can guess the x_i value in $(p - 1) - 2 = 14$ ways (other than known shares and zero) from the field F_p , so he has 14 equations. For example, he picks one value from 14 values of the field F_p , say $x_1 = 1$ then substitute in $b(x)$, we get $b(1) = 0 + 9 \cdot 1 = 9$ this is the approximate secrete value but actual secret value is $a_0 = 13$. Repeat this process for each value from the other 13 values of the field. For each one he has one possible chances to get $b(3) = 10$ and $b(5) = 11$. Therefore, he finds the secret $a_0 = 13$ in 14 approximation ways. Here also, the existing secret is near to the original secret, so the adversary can get the exact secret with fewer chances (in the worst case with in $(p - 2)(p - 3) = 15 \cdot 14 = 210$ ways). If the degree of the polynomials exceeds one, then the adversary can get good approximate value to the secret.

5 Conclusion

We see in this work the significant contribution of the adversary, who is an external member of the scheme who is different from the dealer, made by the computational way of finding some weaknesses of the secret-sharing schemes in the form of certain computational attacks. In this connection, we have studied computational attacks on secret-sharing schemes, especially threshold secret-sharing schemes. Throughout this paper, the adversary is not the internal participant. Based on the sharing of information components, the attacks were divided into two cases. In each case, four different ways of adversary attacks have been studied and found to have the same computational numerical value for the same problem with five secret shares.

This work can be considered as the first attempt of this kind to study the combinatorial ways of the outside adversary in threshold secret-sharing schemes to find secrets. This approach can be extended in various directions to find secrets in threshold secret schemes, for instance, the combinatorial ways of the inside adversary in a threshold secret-sharing scheme and the possible combinatorial ways of the inside/outside adversary in algebraic homomorphism-based threshold secret-sharing schemes. One can also relate the methods to bivariate polynomial-based threshold secret-sharing schemes.

References

1. Shamir A (1979) How to share a secret. *Commun ACM* 22(11):612–613
2. Sun HM, Shieh SP (1994) Construction of dynamic threshold schemes. *Electron Lett* 30:2023–2026
3. Hwang S, Chang C (1993) A dynamic secret sharing scheme with cheater detection. In: *Lecture Notes in Computer Science 1172, ACISP'96*. Springer, Berlin, Germany, pp 136–146
4. Ding J, Ke P, Lin C, Wang H (2022) Bivariate polynomial-based secret sharing schemes with secure secret reconstruction. *Inf Sci* 593:398–414
5. Çalkavur S, Solé P, Bonnetaze A (2020) A new secret sharing scheme based on polynomials over finite fields. *J Math* 8:1200. <https://doi.org/10.3390/math8081200>
6. Bezzateev S, Davydov V, Ometov A (2020) On secret sharing with Newton's polynomial for multi-factor authentication. *J Cryptogr* 4:34. <https://doi.org/10.33390/cryptography4040034>
7. Gharahi M, Khazaei S (2018) Optimal linear secret sharing schemes for graph access structures on six participants. *Theoret Comput Sci*
8. Ersoy O, Pedersen TB, Anarim E (2020) Homomorphic extensions of CRT-based secret sharing. *Discr Appl Math* 285:317–329
9. Kandar S, Dhara BC (2020) A verifiable secret sharing scheme with combiner verification and cheater identification. *J Inf Secur Appl* 51:102430
10. Zhang E, Zhu J-Z, Li G-L, Chang J, Li Y (2019) Outsourcing hierarchical threshold secret sharing scheme based on reputation. *Hind Secur Commun Netw* 2019, Article ID 6989383, 8 p
11. Qin H, Dai Y, Wang Z (2009) A secret sharing scheme based on (t, n) threshold and adversary structure. *Int J Inf Secur* 8:379–385
12. Woll TMH (1989) How to share a secret with cheaters. *J Cryptol* 1(3):133–138
13. Harn L, Lin C, Yong Li (2015) Fair secret reconstruction in (t, n) secret sharing. *J Inf Secur Appl* 23:1–7

14. Tieng DG, Nocon E (2016) Some attacks on Shamir's secret sharing scheme by inside adversaries. In: DLSU research congress 2016, De La Salle University, Manila, Philippines, March 7-9
15. Stinson D (2006) Cryptography theory and practice, 3rd edn
16. Trappe W, Washington LC. Introduction to cryptography with coding theory. Prentice-Hall Inc, Upper Saddle River, New Jersey

Influence of COVID-19 Pandemic on Digital Payment Market Growth



Mohammed Kamran Siddiqui and Krishan Kumar Goyal

1 Introduction

As of the publishing of this article in late May 2020, more than 200 nations and territories around the world had been hit by the Coronavirus pandemic. This applied to both metropolitan centres and more rural areas.

Nearly everywhere as the virus has spread, lockdowns have been enacted, with public spaces including schools, malls, temples, workplaces, airports, and train stations all closed. Because of the lockdown, the vast majority of people are using the internet and related services to stay in touch with one another and carry out their regular duties from the comfort of their own homes. Increases in utilisation of online services range from 40 to 100% when compared to activity before the shut-down. There has been a tenfold growth in the use of video conferencing services like Zoom, and a 30% increase in the use of content delivery services like Akamai. The amount of people using the internet in places like Bangalore has increased by 100%.

An increase in the use of information systems and networks has resulted from the lockdowns throughout countries, along with significant shifts in usage habits and behaviour. Workers are adjusting to the new “normal,” wherein all meetings are held virtually and work is done from home. These shifts can be seen in just about every institution today, from businesses to nonprofits to even governments. In addition, the pace of these shifts has been extremely rapid, leaving little time for institutions and individuals to make the necessary preparations and adjustments, forcing them to improvise and invent solutions where none existed previously [1].

M. K. Siddiqui (✉)

Bhagwat University, Sikar Road, Ajmer, Rajasthan 305009, India

e-mail: kamransiddiquidba@gmail.com

K. K. Goyal

Raja Balwant Singh Management Technical Campus, Agra 282002, India

1.1 Background of Digital Payment System

Digital technologies and digital payment systems can alleviate many difficulties and improve daily living during a pandemic. As the COVID-19 epidemic continues, more individuals are choosing to stay indoors out of their own free will. In the meanwhile, they've been able to work from home thanks to digital innovations that have also helped them deal with the effects of isolation. Lockdowns have been enacted over the whole country of India as the disease spreads. As a result of the CoVH19 epidemic, there has been an uptick in the demand for various forms of digital service provision at all levels of society. The speed with which digital services may be provided is, however, contingent upon the resources required and the quality of the service provided. In a similar vein, the maturity and adaptability of a country's digital infrastructure are crucial. The epidemic has had a devastating effect on the economy. The lockdown and subsequent isolation of the population have resulted in numerous noticeable shifts in food shopping and consumption patterns. The actions and routines of customers during this time period can be better understood by gaining insight into their food-buying habits [2].

The economic impacts of the COVID-19 pandemic measures were split into two. Two types of global impacts are evaluable: constraints on economic operations, and changes in expectations and behaviour. Many people's spending habits have been shown to have been influenced by "panic purchasing," as seen in the second effect. So, we have witnessed a phenomenal growth in the use of online retailers and shipping methods. The use of these technologies made it simple for both merchants and consumers to make and receive financial transactions. Precautions made before the epidemic start are anticipated to enhance the demand for digital banking services. The result has been a rise in the usage of both virtual currency and credit card purchases made online. It has moved away from conventional cash and toward electronic payment networks as a result of widespread use of digital currencies [3].

1.2 Digital Currencies

It is feasible that pre-pandemic and post-pandemic processes might benefit significantly from the use of digital payment networks and digital currencies. As an alternate instrument, digital money offers a number of benefits for contactless transactions. Blockchain-based cryptocurrencies, which are well-known for their credit cards, bank transfers, electronic wallets, and digital value transfer methods, have gained significance in light of the fact that the use of contactless payment methods has made it possible to reduce the risk of the spread of viruses. Credit cards, which are the most used payment method after banknotes, are reported to have the potential to spread viruses at least as much as banknotes do if they are hand-picked and used by entering a password through the panel. This information is based on reports made by professionals in the healthcare industry. The usage of mobile wallet software, such

as PayPal, Cash App, Android Pay, or Apple Pay, which are accessed through mobile devices, is a technologically advanced alternative technique that is often utilised in industrialised nations. Therefore, fresh insights into conventional currencies have been presented in light of any probable increase in the use of digital currencies. Governments need to do more to support and encourage this system.

1.3 Credit Cards

The more conventional methods of payment, such as paying with cash, have been rendered largely obsolete by the advent of the modern method of payment known as the credit card. The financial sector of the economy was impacted as a result of the widespread spread of the COVID-19 epidemic and increasing social isolation. As a result of the coronavirus pandemic, many banks have redirected their consumers to digital prospects. As a result, the percentage of total banking transactions that take place in branches has plummeted to the single digits. The push that banks have made to their clients to “exploit digital opportunities” has resulted in a drop in the percentage of transactions performed from branches. However, the share of payments made digitally has climbed throughout this time. During the epidemic, the digital payment systems in which banks have been investing for years have become increasingly popular. Many banking transactions may be completed swiftly and securely using alternative distribution channels, which eliminates the need for consumers and branch personnel to visit the physical location of the bank [4].

1.4 The Rising of Online Shopping

The quick growth of e-commerce and global connectivity has made it possible for customers to change their buying habits. E-commerce volume in India climbed to 31.5 billion, while retail international operations (holiday travel and online legal betting) were revealed as 28.4 billion. The rising rate of GDP in India has remained steady at 42% (Deloitte and Digital TUSIAD, 2019).

In the midst of the COVID-19 epidemic, many consumers have turned to buying online because they are confined to their homes and shopping centres are closed. However, many items are still being purchased and eaten in conventional market-places. The statistics from the Reserve Bank of India (RBI) indicates that when this time period is compared to March 2019, it was discovered that the panic scenario during the COVID-19 procedure had also grown in the number of contactless payments, but the overall payment on a monthly basis had around 28 million. Some people are afraid that they won't be able to get food because of the widespread COVID-19 outbreak. As a result, more people have started using contactless payment methods, and as a result, the demand for shopping among individuals has soared [5].

2 Increasing Digitalization

Businesses will boost their IT resources to accommodate the growing demand for video and audio conferencing capabilities. As a result, more money will be spent on increasing network bandwidth, purchasing networking hardware, and developing applications that take use of cloud computing. As workers become more accustomed to WFH, it will become the standard rather than the exception for businesses to have meetings and conduct transactions digitally. Many businesses are adopting this since they already have the necessary digital infrastructure in place to support the increased traffic and storage needs.

Another sector experiencing a rapid transition to online transactions is the educational sector. Since the lockdown began, educational institutions around have begun holding lessons using video conferencing apps like Zoom and Google Meet [6].

2.1 *Work-from-Home and Gig Workers*

A major force behind the rise of the “gig economy” was the rise of internet platforms that hired employees on a casual, temporary, and as-needed basis. Uber and Ola are two of the most well-known companies in this category worldwide; in India, Ola and Swiggy are two of the most well-known companies in this sector. Since smartphones became widely available in 2010, the popularity of these platforms has skyrocketed. Taxi drivers, Ola hosts, and those who undertake skilled work have all seen a significant drop in demand for their services during the shutdown. Furthermore, since these employees were not protected by a contract, their pay plummeted.

Telecommuting, digital nomads, and virtual teams are all areas that have been explored in IS research in relation to the rise of freelance and remote work. Work allocation and cooperation is a major challenge that spans teams and projects. As the number of WFH and gig workers grows in the post-pandemic era, this problem will become increasingly widespread and consequential. Aspects of the design of work standards, work contracts, trust-building, and team-building may be investigated [7].

2.2 *Workplace Monitoring and Techno Stress*

The ability to keep tabs on the office at all times and work without stopping is another benefit of digital technology that has attracted the attention of the working population. All interactions are “hyper-focused,” putting remote workers under a microscope when they utilise video conferencing equipment from home. Managers and upper-level executives may more easily track down and contact their staff members at any time, thanks to the widespread availability of contact information made possible by digital tools. Anecdotal evidence suggests this has enhanced productivity, but it has

also raised technostress among workers who are now expected to constantly adapt to new technologies, be in constant contact with their digital gadgets, and juggle many tasks at once.

As a means of escaping the stress of their jobs, employees may band together after the epidemic to lobby for “no digital hours.” The issues of fairness in the workplace, a healthy work–life balance, and stress management may be the subject of future study [8].

2.3 Online Fraud

Since more people are using digital technology, so too have instances of fraud, scams, intrusions, and security breaches occurred online. The uncertainty brought on by the epidemic makes it easy for scammers to take advantage of the crisis for their own ends, whether it be the theft of funds or personal data, or the introduction of new security flaws. More and more people are relying on internet resources, and some of these individuals are easy prey for con artists. Aware of the danger, businesses and governments are taking steps to counter it; for instance, certain governments have taken a firm stance against the use of Zoom in the classroom, prompting the platform provider to increase security.

After the epidemic ends, it’s probable that these cons and frauds will pick up speed. Massive security measures will be implemented by businesses, and the government will launch significant awareness efforts. The number of security-focused startups and service providers will increase. Studies will probably analyse the causes of security breaches and the resulting economic and societal costs [9].

3 Internet Access and Digital Divide

Even after a pandemic has ended, the internet and other forms of information technology will continue to play a crucial role. The internet’s own administration and control will be crucial to the success of this uptick. Despite the fact that the internet is a global resource over which no one country has sway, ensuring that its citizens have access to it where they live remains a domestic problem. Access to the internet has been banned in several countries due to the outbreak.

Those who aren’t online are at danger of being left behind as the epidemic spreads over the world. In order to maintain the new routines imposed by social and physical barriers, it is necessary to use the internet for the vast majority of these tasks. Those who aren’t able to bridge the digital gap, therefore, have no way to participate. The chasm has multiple root causes: gadget availability, cost, Internet availability, content relevancy, user expertise, and government-mandated shutdowns of the Internet are all factors. The problem is even more severe in underdeveloped nations. Given this, it’s crucial to investigate how connectivity might be ensured. These topics have been

studied and discussed before, but the events of COVID-19 have made it clear that having access to the internet is now essential for human life. Some research has found that having or not having access to ICTs can exacerbate existing societal divides, and the post-pandemic environment may further exacerbate this trend. With so much reliance on technology to get necessities like health and education, it's crucial to examine how the digital divide affects fairness in society. Research on the effects of connectivity is needed to pique the interest of politicians and, perhaps, provide suggestions for improving connectivity to promote greater social inclusion [10].

3.1 Internet Governance: Net Neutrality and Zero-Rating

People's data needs have increased as a result of their increased reliance on the internet during the epidemic. This increase in Internet data needs has resurrected the debate about zero-rating strategies, which is particularly relevant given the widening digital gap in today's nations.

With zero-rating plans, businesses may offer their services and sites to customers without charging them for the data they consume. This goes against the principles of net neutrality, which state that all data transmitted over the internet should be treated equally in terms of priority and cost.

In terms of controlling zero-rating schemes, India, for example, has a stellar track record. Although the government did not approve such plans, the telecom regulatory body of India (TRAI) allowed data and voice prices to be waived for specific websites following the outbreak. Organizations like the World Health Organization and the Indian Ministry of Health and Family Welfare were prioritized on the list because of their relevance to COVID-19. Some private entities were also included on the list. The primary objective was to ensure access to information on COVID-19 for people of all income levels [11].

3.2 Internet Governance: Shutdowns

Now more than ever, a loss of internet access may have devastating effects on modern civilizations, whose productivity depends more on the global information infrastructure. Even in these circumstances, though, internet shutdowns are not unheard of. Kashmir, a union region of India, had the longest ever mandated internet blackout in a democracy on August 5, 2019, and it did not end until May 2020. Locals in Kashmir took use of a train called the Internet Express to go to the next town with internet connectivity and perform tasks like applying for driver's licenses online. The internet blackout has cost businesses in Kashmir an estimated \$1.4 billion, according to the Kashmir Chamber of Commerce. The Arab Spring served as a key beginning point, but similar occurrences have since been observed in a number of other nations.

The impact of internet outages has become direr in the wake of the epidemic, when the internet has become the most crucial tool available to individuals. There are numerous questions that needs investigation because of the profound effects shutdowns have on society as a whole. Foreign investors may be put off by the resultant atmosphere of uncertainty, which might have repercussions across many industries, including the ones dealing with education, healthcare, the press, the media, and online commerce. In the current context, it is especially crucial to appreciate the far-reaching consequences of internet shutdowns on human rights. Many of the repercussions of a shutdown can't be predicted because of the underlying political reasons they're caused. The domino effect that can occur, leading to major political crises, can be the subject of study [12].

4 Retail Industry to Drive the Market Growth

The payment industry is evolving to meet the needs of modern shoppers. The rise in the influence of regulatory bodies, the shift toward a cashless economy, mobile banking, rapid payments, digital commerce, and the rise of these technologies are all developments that are influencing the payment sector.

Customers like the ease and convenience of contactless payments since they mean fewer steps in the payment process, less time waiting in line, and no need to worry about whether or not they have enough cash on hand.

The e-commerce industry is booming as more and more people turn to the internet to purchase necessities like food and clothing. The Reserve Bank of India (RBI) created the digital payments index (DPI) in January 2021 to measure the level of digitalization of payments in India; the index for September 2021 was 304.06, up from 270.59 in March. This points to a nationwide trend toward and immersion in cashless transactions [13].

5 Different Digital Payments Apps

5.1 Google Pay

In 2015, Google Inc. released the software for public use. Google Pay, accessible for both Android and iOS, is the digital payment app that is now the most widely used. Only by using the Google Pay app can a person create a Unified Payment Interface (UPI) id, which is required for making bank-to-bank transfers and utility bill payments. The software provides users with not just one, but two layers of protection, one of which is biometric fingerprint scanning, relieving them of the anxiety that comes with the risk of their sensitive information being stolen or lost.

To accept or send payments, it may be utilised by merchants of all sizes, from mom-and-pop stores to wholesalers and multinational conglomerates. There are already more than 100 million users of the app, with more than 67 million users located in India alone. These users together transact more than \$110 billion annually.

5.2 *Paytm*

Paytm was developed as a third-party mobile and desktop application for digital payment services and has its roots in India. In 2010, Paytm was founded as a private company that processes payments. The app is tailored at the e-commerce, financial technology, and digital wallets industries. It is a popular software because it caters to Indians of many linguistic backgrounds by providing support for 11 official languages. The app not only allows users to send and receive money, but also provides them with a number of other services, such as Paytm Mall, Gamepind, Paytm Money, Paytm smart retail, and Paytm Payments bank. There are currently more than 350 million active users of the app, and it is estimated that these individuals are responsible for bringing in more than \$360 million in annual revenue for the company.

5.3 *PhonePe*

PhonePe, like many other payment apps, was created in India and is launching in 2015 as a privately owned multilingual mobile and PC software. The company's headquarters are located in the Indian city of Bangalore, in the state of Karnataka. In order to make a purchase or pay a bill using the PhonePe app, users must first link their bank account and create a UPI id using the app's UPI service. It's similar to Paytm in that it supports 11 different Indian languages. PhonePe has more over 280 million users as of this writing. The "PhonePe ATM" service, which will be available to customers starting in January of 2020, will be the first of its kind offered by the firm. It brings in about \$60,000,000 annually in income.

5.4 *Internet Banking*

Users also commonly use the terms "web banking" and "online banking" to refer to the same notion of banking conducted through the internet. Internet banking allows account holders to transfer funds electronically from one party's bank account to another party's bank account without physically visiting either party's bank. This provides a wide range of banking services for both businesses and individuals, including the ability to transfer funds, view details of previous transactions, create

statements, pay bills, and more. So that its users may use it without worrying about potential security risks, the software employs a dual layer of protection. Over the past several years, internet banking has expanded rapidly, and today more than 45 million people in urban India utilise these services [14].

6 Methodology

The introduction of online shopping has fundamentally changed the way people purchase, yet the electronic goods and clothes industries have profited the most from this change. The purpose of the research is to investigate the relationship between the use of credit cards and the expansion of the GDP. The Granger causality test was utilised in order to analyse the findings of this study, which investigated whether or not the usage of credit cards contributed to the expansion or contraction of GDP. The conclusions of the study were compiled using monthly data taken from the website of the World Bank, beginning in January 2016 and continuing through April 2020.

7 Results

This research looks at how digital payment methods and credit card use have evolved in India [15–20]. There is a spike in epidemic-era internet buying in India; nonetheless, consumer purchasing habits per individual should be investigated. Keeping customers happy with what they buy is essential for the growth of the online retail industry. While there has been a notable uptick in the value of digital currencies, online payment spending overall has grown, notably with debit cards and contactless payments. Some numbers on people's perspectives and preferences regarding the use of debit or credit card spending by industry have been taken from the Reserve Bank of India (RBI). During this time, their popularity has increased mostly in the electronics and fashion sectors. However, a key result of pandemic process is often a drop in India's GDP. There has also been an increase in spending as a result of people panic-buying groceries. All banks have maintained electronic and telephone banking services throughout the shutdown. Thus, our research has determined how credit card purchases affect GDP growth.

For each set of data, the Granger causality test estimates the likelihood that one set of changes will cause the next set of changes. Linear Granger causality tests, a modification of the Granger (1969) model, were used for this purpose.

Table 1 displays the changes in the 2-year period involving debit and credit card numbers. Data for May 2020 have been issued by the Reserve Bank of India (RBI) As of the end of May, 242.7 million cards have been discovered in India, with 71.4 million credit cards and debit cards making up the 171.3 million units. This information comes from the International Credit Card Association. As of May 2019, credit card usage was up 6% yearly and debit card usage was up 12% annually.

Table 1 Developments of cards number

Number of cards	2019 May (million)	2020 May (million)	Percentage
Debit card	154.6	171.3	12
Credit card	67.8	71.4	6
Total	222.4	242.7	8

The growth or decline in the average amount paid by debit and credit card over the course of 2 years is displayed in Table 2. There were 76.2 billion TL worth of card payments processed in May, per RBI statistics. Of this total, 63.3 billion TL came from credit card payments and 12.9 billion TL came from debit card transactions. Accordingly, the amount of payments has made with credit cards declined by 12% compared to the same month of the previous year, while the amount of payments has made with debit cards climbed by 20% yearly.

Credit card expenditures in monthly data between 2016 and 2020 are shown in Fig. 1. After rising steadily since 2016, expenditure on credit cards fell during the first 3 months of 2020. People are being urged to make the transition to using online contactless debit card payment alternatives during this epidemic.

In Table 3, we can see the growth in online card payments and their percentage of total automobile payments from May 2019 to May 2020.

It's only in recent years that the internet has become a reliable source of information. Because of this, we now have a domestic problem with access and availability. As people stay indoors owing to the epidemic, their spending habits and consumption patterns have shifted. The use of credit cards for making purchases online has grown in popularity and significance in recent years.

May saw a 13% increase in the total amount of money spent via credit card transactions online, with total sales hitting 19 billion India. That's why May saw a historic high in online credit card transactions. There has also been a record-breaking increase in the percentage of card payments that originate from online transactions.

For the period of May 2019–May 2020, Table 4 displays the percentage of all online card payments made by industry. The “electronic goods” and “clothing” sectors, which have undergone the most rapid transition to online shopping, have been revealed to have the highest proportion of card payments made online relative to all card purchases in May. From May 2019 to May 2020, the proportion of online card purchases in the “electronic goods” industry rose from 42 to 71%, while the proportion rose from 13 to 41% in the “clothing and accessories” industry and the

Table 2 Developments of card payment amount

Cards payment amount	2019 May (billion)	2020 May (billion)	Percentage
Debit card	10.9	12.9	20
Credit card	73.2	63.3	–12
Total	84.1	76.2	–8

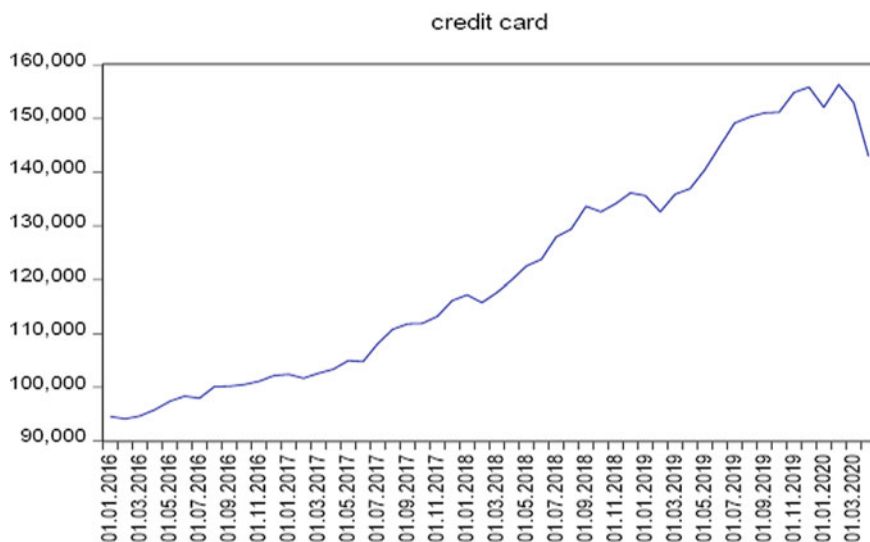


Fig. 1 Credit card spending

Table 3 Development of online card payment amount

Internet card payment amount	2019 May	2020 May	Percentage
Card payment amount from internet	16.8	19.1	13
Share of total card payments (%)	18.9	25.%	

“education/stationery” industry, respectively. It has risen from 5 to 13% in the “food” area, from 10 to 23% in the “furniture and decorating,” and from 13 to 29% overall. It’s one of the areas of expertise that has grown the most popular over the past few years (Table 5).

Although Table 6 reveals that the hypothesis that credit cards are not a Granger cause of GDP cannot be rejected, the inverse hypothesis that GDP does not cause credit cards may be rejected. Therefore, it would appear that the Granger causality test only works in one direction, from GDP to credit card.

Table 4 Sectoral share of online card payments

Sectoral share of online card payments	2019 May (%)	2020 May (%)
Electronic stuff	42	71
Clothing and accessories	13	41
Education/stationery	13	29
Furniture and decoration	10	23
Food sector	5	13

Table 5 Unit root tests of model variables at level and first difference

	Null hypothesis: CREDIT_CARD has a unit root	Null hypothesis: D(CREDIT_ CARD) has a unit root	Null hypothesis: GDP has a unit root	Null hypothesis: D(GDP) has a unit root
Augment Dickey-Fuller test statistic	-1.438115	-4.418368	-2.951701	-3.074613
Prob.*	0.5559	0.0009	0.4366	0.0001
Test critical values				
1% level	3.574446	-3.568308	-3.568308	-3.568308
5% level	-2.923780	-2.921175	-2.921175	-2.921175
10% level	-2.599925	-2.598551	-2.598551	-2.598551

Table 6 Granger causality test

Null hypothesis	Obs	F-statistic	Prob.*
CREDIT_CARD does not Granger Cause GDP	34	-14.3102	1.0000
GDP does not Granger cause CREDIT_CARD		5.68352	0.0083

The history and forecast of online transactions in India

The past trends refer to an analysis of previous years related to digital payments in India, and the future projections show an idea of some upcoming years. In order to have more stronger conclusion research has considered some graphs which shows trend of digital payments in India from financial year 2012 to financial year 2023 (Fig. 2).

The graph shows an increasing trend year by year that is in financial year share of digital payment was less than 5% in an economy which increased up to 15% approximately in financial year 2019 that is before pandemic. In year 2020, the share experienced a growth between 15 and 20%, which has further increased to nearby 20% in current year that is financial year 2021. It is being predicted that by the year 2023 the Indian economy will have approximately 25% of its share, which will lead to US\$1 trillion (Fig. 3).

8 Conclusions

During this time period, they have been given a higher priority of preference in the electrical goods and garment sectors. On the other hand, the most significant effect that pandemic processes often have is a negative impact on the GDP of India. Another impact, the widespread panic purchasing in the food industry, has led to an increase in

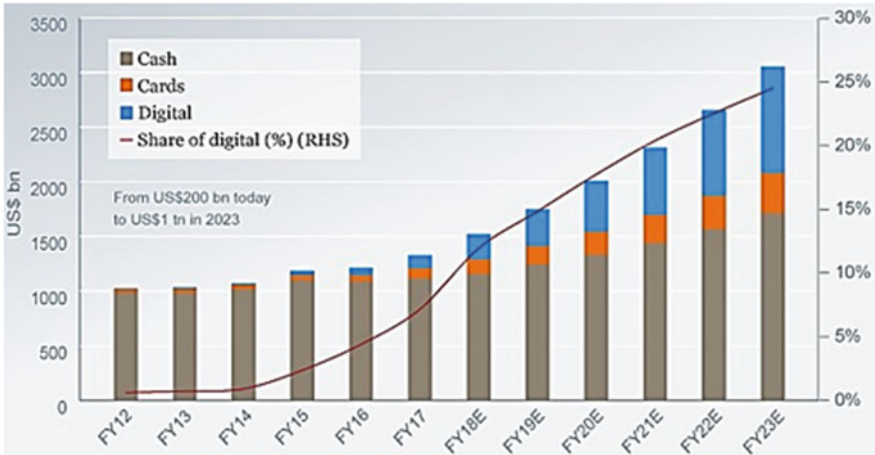


Fig. 2 Growth trend in India's digital payment systems

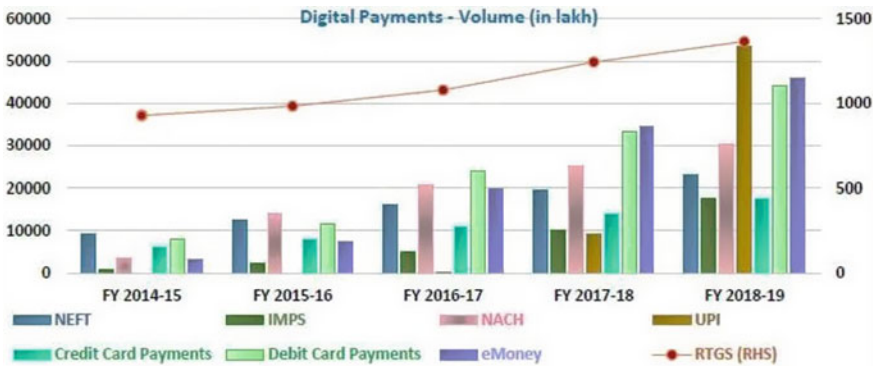


Fig. 3 Different digital payments' volume graph

expenditure as a direct result. The Granger causality test was utilised in this research project in order to conduct an analysis of the effect that spending using credit cards has on the development of GDP. Digital services or digital payment applications are a significant contributor to the development of the Indian economy's gross domestic product (GDP) as well as the standards of society. Prior to the introduction of digital services or digital payment applications, many members of Indian society were either not very aware of their surroundings or were struggling with trade-related issues because they relied solely on cash in their daily lives [21–24].

References

1. Agarwal S, Sengupta D, Kulshrestha A, Anand S, Guha R (2017) The economic times. internet users to touch 420 million by June
2. Aghasian E, Garg S, Gao L, Yu S (2017) Montgomery J. Scoring users' privacy disclosure across multiple online social networks. *IEEE Access* 5:13118–13130
3. Akala A (2020) More big employers are talking about permanent work-from-home positions. *CNBC*
4. Aker JC, Boumniel R, McClelland A, Tierney N (2016) Payment mechanisms and antipoverty programs: Evidence from a mobile money cash transfer experiment in Niger. *Econ Dev Cult Change* 1–37
5. Tarafdar M, Tu Q, Ragu-Nathan BS, Ragu-Nathan T (2007) The impact of technostress on role stress and productivity. *J Manage Inf Syst* 24(1):301–328. Taylor & Francis
6. Pollach I, Treiblmaier H, Floh A (2005) Online fundraising for environmental nonprofit organizations. *Hawaii Int Conf Syst Sci*
7. *New York Times* (2020) Virus forces Cambridge to hold most classes online next year—The New York
8. Misa TJ, Brey P, Feenberg A (2003) *Modernity and technology*. MIT Press
9. Khetarpal S (2020) Post-COVID, 75% of 4.5 lakh TCS employees to permanently work from home by. *Bus Today* 25
10. Digital payments and its types. *RashiSood* (2018)
11. Balaji RP, Vijayakumar T (2020) Diffusion of digital payment system in rural India. *Glob J Manage Bus Res*
12. Bashir R, Mehboob I, Bhatti WK (2015) An empirical study of Pakistan. Effects of online shopping trends on consumer-buying behavior. *J Manage Res* 2
13. De R, Pandey N, Pal A (2020) Impact of digital surge during Covid-19 pandemic: a viewpoint on research and practice. *Int J Inf Manage*
14. Laudon KC, Traver CG (2009) *E-commerce business. Technology, Society*, 5th edn. Prentice Hall, New Jersey
15. Shukla V, Chaturvedi A, Srivastava N (2019) Nanotechnology and cryptographic protocols: issues and possible solutions. *Nanomater Energy* 8(1):1–6. <https://doi.org/10.1680/jnaen.18.00006>
16. Misra MK, Chaturvedi A, Tripathi SP, Shukla V (2019) A unique key sharing protocol among three users using non-commutative group for electronic health record system. *J Discr Math Sci Cryptogr* 22(8):1435–1451. <https://doi.org/10.1080/09720529.2019.1692450>
17. Shukla V, Chaturvedi A, Misra MK (2021) On authentication schemes using polynomials over non commutative rings. *Wirel Pers Commun* 118(1):1–9. <https://doi.org/10.1007/s11277-020-08008-4>
18. Chaturvedi A, Shukla V, Misra MK (2018) Three party key sharing protocol using polynomial rings. In: 5th IEEE Uttar Pradesh section international conference on electrical, electronics and computer engineering (UPCON), pp 1–5. <https://doi.org/10.1109/UPCON.2018.8596905>
19. Shukla V, Misra MK, Chaturvedi A (2022) Journey of cryptocurrency in India in view of financial budget 2022–23. *Cornell university arxiv*, pp 1–6. <https://doi.org/10.48550/arXiv.2203.12606>
20. Shukla V, Srivastava N, Chaturvedi A (2016) A bit commitment signcryption protocol for wireless transport layer security (wtls). In: IEEE international conference on electrical, computer and electronics engineering, pp 83–86. <https://doi.org/10.1109/UPCON.2016.7894629>
21. Shukla V, Chaturvedi A, Srivastava N (2019) A secure stop and wait communication protocol for disturbed networks. *Wirel Pers Commun* 110:861–872. <https://doi.org/10.1007/s11277-019-06760-w>
22. Chaturvedi A, Shukla V, Misra MK (2021) A random encoding method for secure data communication: an extension of sequential coding. *J Discr Math Sci Cryptogr* 24(5):1189–1204. <https://doi.org/10.1080/09720529.2021.1932902>

23. Shukla V, Chaturvedi A, Srivastava N (2019) A new one time password mechanism for client-server applications. *J Discr Math Sci Cryptogr* 22:1393–1406. <https://doi.org/10.1080/09720529.2019.1692447>
24. Shukla V, Misra MK, Chaturvedi A (2021) A new authentication procedure for client-server applications using HMAC. *J Discr Math Sci Cryptogr* 24(5):1241–1256. <https://doi.org/10.1080/09720529.2021.1932908>

Two-Level Security of Color Image Using 9D-Hyperchaotic System and DWT



Sonali Singh and Anand B. Joshi

1 Introduction

In this digital world, the security of digital images is required during the communication process. Digital images are frequently used to exchange information as images speak louder than words. These digital images may contain lots of important and confidential information therefore secure communication of digital images over an unsecured channel is a subject of major concern.

Several techniques have been developed to provide secure communication and protect the significant information from unauthorized users. These methods include cryptography, steganography, and digital watermarking. Cryptography is a technique of converting readable information into unreadable one so that only an authorized person can read and process it. It is mostly used to protect the passwords, e-mail security, banking transaction details, e-commerce transaction, etc. Steganography is a technique of concealing secret information within a cover file in order that any unauthorized person can not detect the embedded information, only the intended one can detect and reveal the secret information. This technique also has several uses, such as business, defense, military, etc. Digital watermarking is a branch of steganography which is used in copyright protection, content authentication, etc.

Many researchers have worked on the security of digital images using cryptography [17–23] and steganography [24]. Cryptography is based on keys in which encryption process converts the data into encrypted form to which only the intended person can read and process and nobody else can know the secret content without knowing the decryption key. But still, anyone can see that something is there in encrypted form and that one can make an attempt to break the encryption to get the secret content. To overcome this type of risk, steganography has been used in

S. Singh (✉) · A. B. Joshi

Department of Mathematics and Astronomy, University of Lucknow, Lucknow 226007, Uttar Pradesh, India

e-mail: singhsonali09192@gmail.com

which secret information is embedded in cover file and transmitted from one place to another and no one will be able to know about the secrecy of the data. But this embedding should also be done in a very strong manner so that no one else can extract the secret information other than the intended one. Hence, a strong technique is needed to protect the secret information and also a better communication.

This research paper proposes a novel technique to protect digital color images by using 9D-hyperchaotic system and discrete wavelet transform (DWT). In this technique, the secret color image is decomposed into its three components (Red, Green, and Blue) and converted into row vectors. Secret keys generated by 9D-hyperchaotic system, are used with these row vectors to form three encrypted images which are concatenated to obtain the final encrypted image. After getting the encrypted images, our next goal is to embed them in any other cover image. For this, we apply six-level DWT on encrypted image as well as on cover image and then decomposing the six-level approximation coefficients into their three components and perform the embedding operation on these components using the secret key obtained from 9D-hyperchaotic system. After embedding, inverse six-level DWT is performed and stego image is obtained. This technique provides double security for digital images, as it used both encryption and steganography techniques. The performance of the proposed scheme has been checked by performing computer simulations and experimental results. This scheme is robust against brute-force and cropping attacks too.

The rest of this paper has been organized in the following manner: Sect. 2 presents the information about related works. Section 3 presents the preliminary knowledge of the 2D discrete wavelet transform (DWT) and 9D-chaotic system. Section 4 describes the techniques used in the proposed scheme. Section 5 provides simulation and experimental results. Performance of the proposed scheme is evaluated in Sect. 6 by using metrics MSE, PSNR, and SSIM. Section 7 provides a security analysis of the scheme. Comparison with related work has been done in Sect. 8. Finally, Sect. 9 concludes the proposed technique.

2 Related Works

In paper [12], Baluja et al. proposed a technique of steganography to embed any color image into another image, where both the images are of same size. In this scheme, there are three components: preparation of network, hiding image, and reveal image. This scheme has a higher bit rate compared to the traditional methods. In paper [13], Rehman et al. proposed a technique in which CNN-based encoder and decoder architectures are used for embedding the images as payload. This method has used a color image as a cover image and a grayscale image as a secret image but many problems are still addressed. References [14, 15] also worked on steganography techniques. In these schemes, color images of the same size have been used for the cover image as well as for the secret image. Li et al. [16] proposed a method in which an encrypted secret image is hiding in a cover image of the same size. The encryption of secret image has been done by using the chaos encryption technology

before hiding it. Then encrypted and cover images have been transformed into the stego image by using a convolutional neural network (CNN).

3 Preliminaries

3.1 2D Discrete Wavelet Transform

Wavelet Transform (WT) plays an important role in many areas of image processing [1, 2, 23], pattern recognition [3], document analysis, etc. In the 1980s, WT [4, 5] (such as Daubechies, Haar, and Meyer wavelets) was developed.

In the image processing, an image is formed by discrete samples called pixels. DWT is an implementation of the discretely sampled wavelet. Here in our paper, we have used Haar wavelet. Haar sequence was first proposed by Alfréd Haar in 1910 [6]. A sequence of square-shaped functions is formed by this wavelet which together form a wavelet basis. Matrix form of the Haar wavelet transform [4] is expressed as $I' = H I H^T$, where I is an image matrix, H is the Haar transform matrix, and I' is the resulting transformed matrix, each of order $N \times N$ that contains the Haar basis functions, $h_n(z)$ defined in $z \in [0, 1]$ where $n = 0, 1, 2, \dots, N - 1$ can be decomposed uniquely as

$$n = 2^{t_1} + t_2,$$

where t_1 is the highest power of 2 contained in n and t_2 is the remainder, i.e., $t_2 = n - 2^{t_1}$. The Haar Basis function is defined by Eq. (1);

$$h_n(z) = \begin{cases} 1 & \text{if } n = 0 \text{ and } 0 \leq z < 1, \\ 2^{t_1/2} & \text{if } n > 0 \text{ and } t_2/2^{t_1} \leq z < (t_2 + 0.5)/2^{t_1}, \\ -2^{t_1/2} & \text{if } n > 0 \text{ and } (t_2 + 0.5)/2^{t_1} \leq z < (t_2 + 1)/2^{t_1}, \\ 0 & \text{otherwise .} \end{cases} \quad (1)$$

The transformation matrix of the 2D discrete Haar wavelet transform (2D-DHWT) is obtained by substituting the inverse transformation kernel which is given by the Eq. (2):

$$h'(z, n) = \frac{1}{\sqrt{N}} h_n(z/N) \text{ for } z = 0, 1, 2, \dots, N - 1. \quad (2)$$

For $n = 0, 1, 2, \dots, N - 1$, the transformation matrix is given by Eq. (3).

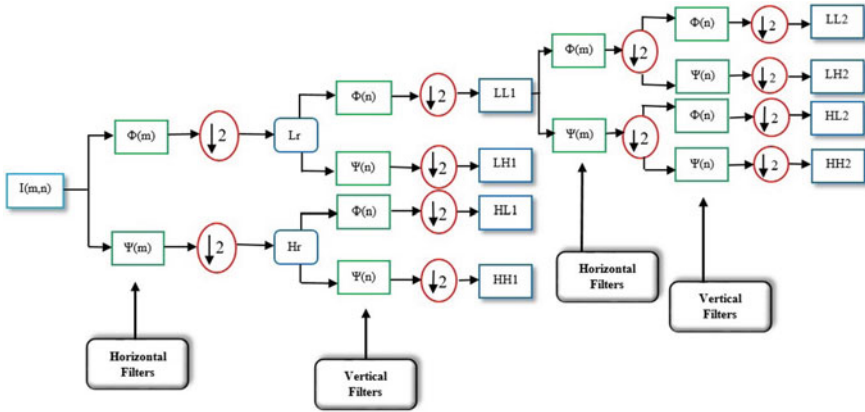


Fig. 1 Block diagram of 2D-DWT up to second-level decomposition

$$H' = \begin{bmatrix} h_0(0/N) & h_0(1/N) & \dots & h_0(N-1/N) \\ h_1(0/N) & h_1(1/N) & \dots & h_1(N-1/N) \\ h_2(0/N) & h_2(1/N) & \dots & h_2(N-1/N) \\ \vdots & \vdots & \ddots & \vdots \\ h_{N-1}(0/N) & h_{N-1}(1/N) & \dots & h_{N-1}(N-1/N) \end{bmatrix},$$

$$H = \frac{1}{\sqrt{N}} H'. \tag{3}$$

In the case of two-dimensional discrete signals $I(m, n)$, i.e., image of size $M \times N$, the basic operation of 2D-DHWT is as follows: each row of the two dimension signal $I(m, n)$ is analyzed by low-pass filter $\phi(m)$ and high-pass filter $\varphi(m)$ in the horizontal direction and the output of each filter is down-sampled by a factor of 2 and then intermediate signals L_r and H_r are obtained. Then, each column of these new signals L_r and H_r is examined by low-pass filter $\phi(m)$ and high-pass filter $\varphi(m)$ in the vertical direction, and the output of each filter is down-sampled by a factor of 2 and then produce four new sub-bands LL_1, LH_1, HL_1 , and HH_1 which are shown in Fig. 1. All these sub-bands contain all information of the original signal. If 2D-DHWT is applied to the sub-band LL_1 again, then we get four new sub-bands: LL_2, LH_2, HL_2 , and HH_2 . If we repeat this process again and again for t times then 2D-DHWT gives the sequences of sub-bands LL_t, LH_t, HL_t , and HH_t . In this paper, we have repeated this process six times as six-level DWT has been used here.

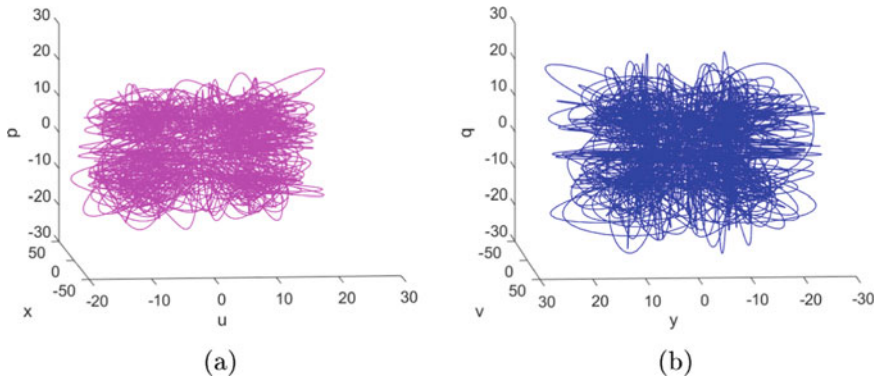


Fig. 2 9D-hyperchaotic attractors of the system (4) when $a_1 = 10$, $a_2 = 28$, $a_3 = 8/3$, and $r_1 = r_2 = r_3 = 0.05$: **a** in the (u, x, p) space and **b** in the (y, v, q) space

3.2 9D-Chaotic System

A continuous hyperchaotic structure has a minimum dimension of four. Grassi et al. [7] used three same 3D Lorenz chaotic systems and modeled an eight-wing hyperchaotic system which is represented by Eq. (4):

$$\begin{cases} \dot{u} = a_1(v - u) \\ \dot{v} = a_2u - v - uw + r_1(p - q) \\ \dot{w} = uv - a_3w \\ \dot{x} = a_1(y - x) \\ \dot{y} = a_2x - y - xz + r_2(u - v) \\ \dot{z} = xy - a_3z \\ \dot{p} = a_1(q - p) \\ \dot{q} = a_2p - q - ps + r_3(x - y) \\ \dot{s} = pq - a_3s, \end{cases} \quad (4)$$

where a_1, a_2, a_3 and r_1, r_2, r_3 are the positive and coupling parameters, respectively. When $a_1 = 10$, $a_2 = 28$, $a_3 = 8/3$, and $r_1 = r_2 = r_3 = 0.05$. Eight wing attractors have been generated by Eq. (4) as shown in Fig. 2. In this proposed technique, parameters $a_1, a_2, a_3, r_1, r_2, r_3$ and initial conditions u, v, w, x, y, z, p, q and s are considered as the secret keys, to form the keystream.

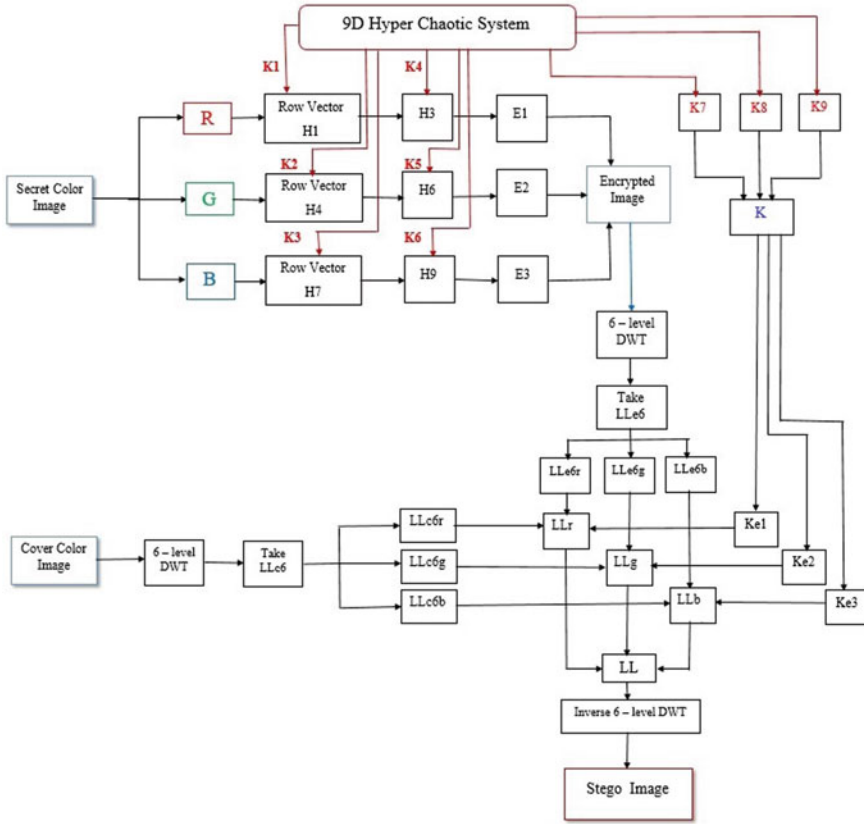


Fig. 3 Block diagram of embedding process of proposed technique

4 Proposed Scheme

This section describes the proposed scheme. Figure 3 represents the pictorial form of embedding process along with the encryption of secret image. Figure 4 gives the pictorial description of extraction process along with the decryption of secret image.

4.1 Key Generation Process

Step 1: To eliminate the transitory response, iterate system (4) 1000 times and generate these random sequences:

$$u = \{u_1, u_2, u_3, \dots, u_{mn}\}, v = \{v_1, v_2, v_3, \dots, v_{mn}\}, w = \{w_1, w_2, w_3, \dots, w_{mn}\},$$

$$x = \{x_1, x_2, x_3, \dots, x_{mn}\}, y = \{y_1, y_2, y_3, \dots, y_{mn}\}, z = \{z_1, z_2, z_3, \dots, z_{mn}\},$$

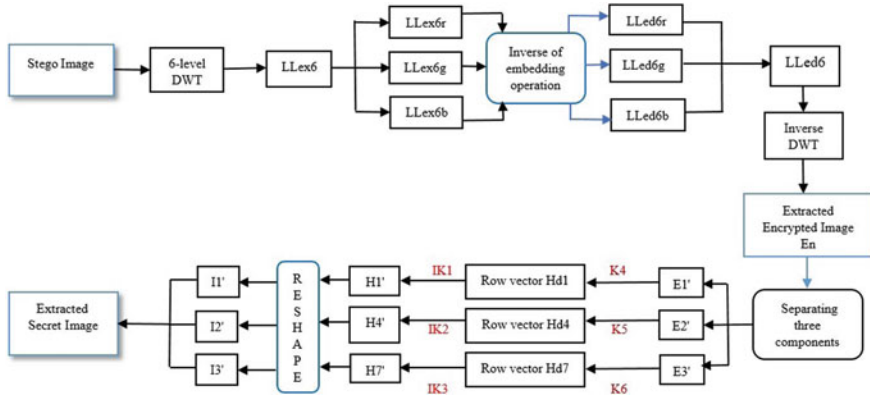


Fig. 4 Block diagram of extracting process of proposed technique

$p = \{p_1, p_2, p_3, \dots, p_{mn}\}$, $q = \{q_1, q_2, q_3, \dots, q_{mn}\}$, and $s = \{s_1, s_2, s_3, \dots, s_{mn}\}$, respectively, each of size $1 \times mn$.

Step 2: All these sequences u, v, w, x, y, z, p, q and s are converted into integers as

$$U = \text{floor}(u \times 10^{15}) \bmod mn, \quad (5)$$

$$V = \text{floor}(v \times 10^{15}) \bmod mn, \quad (6)$$

$$W = \text{floor}(w \times 10^{15}) \bmod mn, \quad (7)$$

$$X = \text{floor}(x \times 10^{15}) \bmod mn, \quad (8)$$

$$Y = \text{floor}(y \times 10^{15}) \bmod mn, \quad (9)$$

$$Z = \text{floor}(z \times 10^{15}) \bmod mn, \quad (10)$$

$$P = \text{floor}(p \times 10^{15}) \bmod mn, \quad (11)$$

$$Q = \text{floor}(q \times 10^{15}) \bmod mn, \quad (12)$$

$$S = \text{floor}(s \times 10^{15}) \bmod mn, \quad (13)$$

where $\text{floor}(t)$ gives the greatest integer less than or equal to t and \bmod defines modulo function.

Step 3: After sorting these sequences (5)–(13), we get sorted sequences $\overline{U}, \overline{V}, \overline{W}, \overline{X}, \overline{Y}, \overline{Z}, \overline{P}, \overline{Q}$, and \overline{S} . Now, find the positions of the values of $\overline{U}, \overline{V}, \overline{W}, \overline{X}, \overline{Y}, \overline{Z}, \overline{P}, \overline{Q}$, and \overline{S} in U, V, W, X, Y, Z, P, Q , and S and note down the transform positions, i.e., $A = \{A(i) : i = 1, 2, 3, \dots, mn\}$, $B = \{B(i) : i = 1, 2, 3, \dots, mn\}$, $C = \{C(i) : i = 1, 2, 3, \dots, mn\}$, $D = \{D(i) : i = 1, 2, 3, \dots, mn\}$, $E = \{E(i) : i = 1, 2, 3, \dots, mn\}$, $F = \{F(i) : i = 1, 2, 3, \dots, mn\}$, $G = \{G(i) : i = 1, 2, 3, \dots, mn\}$, $H = \{H(i) : i = 1, 2, 3, \dots, mn\}$, $I = \{I(i) : i = 1, 2, 3, \dots, mn\}$, where $U(A(i)) = \overline{U}(i)$,

$V(B(i)) = \overline{V}(i)$, $W(C(i)) = \overline{W}(i)$, $X(D(i)) = \overline{X}(i)$, $Y(E(i)) = \overline{Y}(i)$, $Z(F(i)) = \overline{Z}(i)$, $P(G(i)) = \overline{P}(i)$, $Q(H(i)) = \overline{Q}(i)$ and $S(I(i)) = \overline{S}(i)$.

Step 4: Now, position sequences A , B , and C are converted into row vectors M_1 , M_2 , and M_3 , respectively, where each vector is of size $1 \times mn$. D , E , F , G , H , and I are converted into matrices M_4 , M_5 , M_6 , M_7 , M_8 , and M_9 , respectively, each matrix is of size $m \times n$ and generate keys K_1 , K_2 , K_3 , K_4 , K_5 , K_6 , K_7 , K_8 , and K_9 as given below:

$$\begin{aligned} K_1 &= M_1, \\ K_2 &= M_2, \\ K_3 &= M_3, \\ K_4 &= (M_4) \bmod 256, \\ K_5 &= (M_5) \bmod 256, \\ K_6 &= (M_6) \bmod 256, \\ K_7 &= (M_7) \bmod 256, \\ K_8 &= (M_8) \bmod 256, \\ K_9 &= (M_9) \bmod 256. \end{aligned}$$

Step 5: Now combine last three keys K_7 , K_8 , and K_9 to form a three dimension array L . Then by applying six-level DWT on L , we generate key K .

4.2 Proposed Encryption and Embedding Algorithm

The proposed embedding technique uses both permutation and substitution methods along with six-level DWT to achieve high security of data. A pictorial representation of embedding process is shown in Fig. 3. The step-by-step process is described below:

Step 1: Take a color image I of size $m \times n$ and separate all its three components Red (R), Green (G), and Blue (B).

Step 2: In this step, R , G , and B are converted into row vectors. Then these row vectors are scrambled using keys K_1 , K_2 , and K_3 , respectively. After scrambling, these row vectors are again converted into matrices each of size $m \times n$, namely, H_3 , H_6 , and H_9 , respectively.

Step 3: In this step, substitution method is being used in H_3 , H_6 , and H_9 with the help of keys K_4 , K_5 , and K_6 . The substitution method is as below: Find the minimum values of the matrices H_3 , H_6 , and H_9 . Let these values are δ_1 , δ_2 , and δ_3 , respectively. Further, calculate the minimum value δ of δ_1 , δ_2 , and δ_3 . Then let $\alpha = -\delta + \mu$, where $1 \leq \mu \leq 5$ and calculate the pixel values of the encrypted image using the equation given below:

$$\begin{cases} E_1(i, j) = \text{mod}((H3(i, j) + \alpha) \oplus K_4(i, j), 256) \\ E_2(i, j) = \text{mod}((H6(i, j) + \alpha) \oplus K_5(i, j), 256) \\ E_3(i, j) = \text{mod}((H9(i, j) + \alpha) \oplus K_6(i, j), 256) \end{cases} \quad (14)$$

where $1 \leq i \leq m$, $1 \leq j \leq n$, and \oplus is the bitwise XOR operator. Thus, encrypted components E_1 , E_2 , and E_3 are obtained. After combining these components, we get the final encrypted image E .

Step 4: Applying six-level DWT on E then separating red ($LLe6r$), green ($LLe6g$), and blue ($LLe6b$) parts of approximation component $LLe6$.

Step 5: Separate all three red, green, and blue components of key K , namely, $Ke1$, $Ke2$, and $Ke3$.

Step 6: Take an RGB cover image. Applying six-level DWT on it and separating its sixth-level approximation component ($LLc6$) into its red, green, and blue components $LLc6r$, $LLc6g$, and $LLc6b$, respectively.

Step 7: Using the following process, embed encrypted image components into cover image components:

$$\begin{cases} LLr = LLc6r + Ke1 * LLe6r \\ LLg = LLc6g + Ke2 * LLe6g \\ LLb = LLc6b + Ke3 * LLe6b \end{cases} \quad (15)$$

Step 8: Combining these three components LLr , LLg , and LLb , we get LL . Stego image S is obtained after applying inverse six-level DWT on LL .

4.3 Proposed Decryption and Extraction Algorithm

Figure 4 shows the pictorial representation of the extraction process along with decryption of extracted encrypted image. The process is started with stego image. Using the embedding and encryption processes in reverse order, original secret image can successfully be obtained.

Step 1: Applying six-level DWT on stego image S , separate all three channel components of its sixth-level approximation component $LLex6$, namely, $LLex6r$, $LLex6g$, and $LLex6b$.

Step 2: Finding inverses of matrices $Ke1$, $Ke2$, and $Ke3$ which are named $Ke1_{inv}$, $Ke2_{inv}$, and $Ke3_{inv}$, respectively. By using the following extraction process, obtain RGB components $LLed6r$, $LLed6g$, and $LLed6b$

$$\begin{cases} LLed6r = Ke1_{inv} * (LLe6r - LLc6r) \\ LLed6g = Ke2_{inv} * (LLe6g - LLc6g) \\ LLed6b = Ke3_{inv} * (LLe6b - LLc6b) \end{cases} \quad (16)$$

After combining these three components, $LLed6$ is obtained.

Step 3: By applying inverse DWT, secret image is obtained in encrypted form.

Step 4: In this step, decryption process is applied over encrypted image to obtain the original secret image. Following equations are used to calculate the pixel values by using the same keys K_4 , K_5 , and K_6 as in encryption process:

$$\begin{cases} D2d(i, j) = \text{mod}((Z4e(i, j) \oplus K_4(i, j)) - \alpha, 256) \\ D4d(i, j) = \text{mod}((Z5e(i, j) \oplus K_5(i, j)) - \alpha, 256) \\ D6d(i, j) = \text{mod}((Z6e(i, j) \oplus K_6(i, j)) - \alpha, 256) \end{cases} \quad (17)$$

where $1 \leq i \leq m$ and $1 \leq j \leq n$. And, $Z4e$, $Z5e$, and $Z6e$ are red, green, and blue components of extracted image, respectively.

Step 5: Now, $D2d$, $D4d$, and $D6d$ are converted into row vectors $Hd1$, $Hd4$, and $Hd7$ and also scrambled by using keys IK_1 , IK_2 , and IK_3 which are the inverses of keys K_1 , K_2 , and K_3 , respectively.

Step 6: After scrambling, convert these vectors into matrices $Hd3$, $Hd6$, and $Hd9$, respectively, each of size $m \times n$.

Step 7: In this step, combine these three components and obtain final secret image in original form.

5 Simulation and Experimental Results

This proposed algorithm is implemented in the computer system having Windows 10, Intel(R), Core(TM) i3-1005G1 CPU with a clock speed of 1.20 GHz, and 8 GB RAM, using MATLAB R-2015a software. For the experimental results, we have taken two color images: secret image and cover image each of size 256×256 shown in Fig. 5a, b. The initial values and control parameters of the

9D-chaotic system are $u_0 = 2.543210007543721$, $v_0 = 3.674515623875401$, $w_0 = 1.235685120036054$, $x_0 = 1.6758174322200155643$, $y_0 = 4.785400011325467$, $z_0 = 2.3576335564327899543$, $p_0 = 3.463200032157649$, $q_0 = 1.8473569860087554321$, $s_0 = 4.143205412568765$, $a_1 = 10.000102302324532$, $a_2 = 28.004238236135784$, $a_3 = 2.424874598634125$, $r_1 = 0.036793421834854$, $r_2 = 0.040002249502146$ and $r_3 = 0.055622474001536$. Figure 5c, d shows the encrypted secret image and stego image. Figure 5e, f shows the extracted encrypted image and then extracted secret image.

5.1 More Experimental Results

We have performed our proposed scheme on some other images to get more experimental results. Images, used for it, are shown in Fig. 6. No changes are made in secret keys, all keys and initial values are same as given in Sect. 5. Figure 6a, b display the original images, Fig. 6c displays the encrypted secret image, Fig. 6d displays stego image, and Fig. 6e, f display extracted encrypted image and extracted secret image, respectively.

6 Performance Evaluation Metrics

6.1 Mean Squared Error (MSE)

MSE defines the error between two like images. We calculate MSE between extracted image and its corresponding original secret image and also between stego image and original cover image. Value of MSE should be less, i.e., 0 for a high secured algorithm.

MSE [8] is calculated by following Eq. (18):

$$\text{MSE} = \frac{1}{MN} \sum_{m=1}^M \sum_{n=1}^N [I_1(m, n) - I_2(m, n)]^2, \quad (18)$$

where I_1 is the original secret image and I_2 is the extracted image of size $M \times N$. In the same manner, we take I_1 as the original cover image and I_2 as the stego image in our second evaluation of MSE.

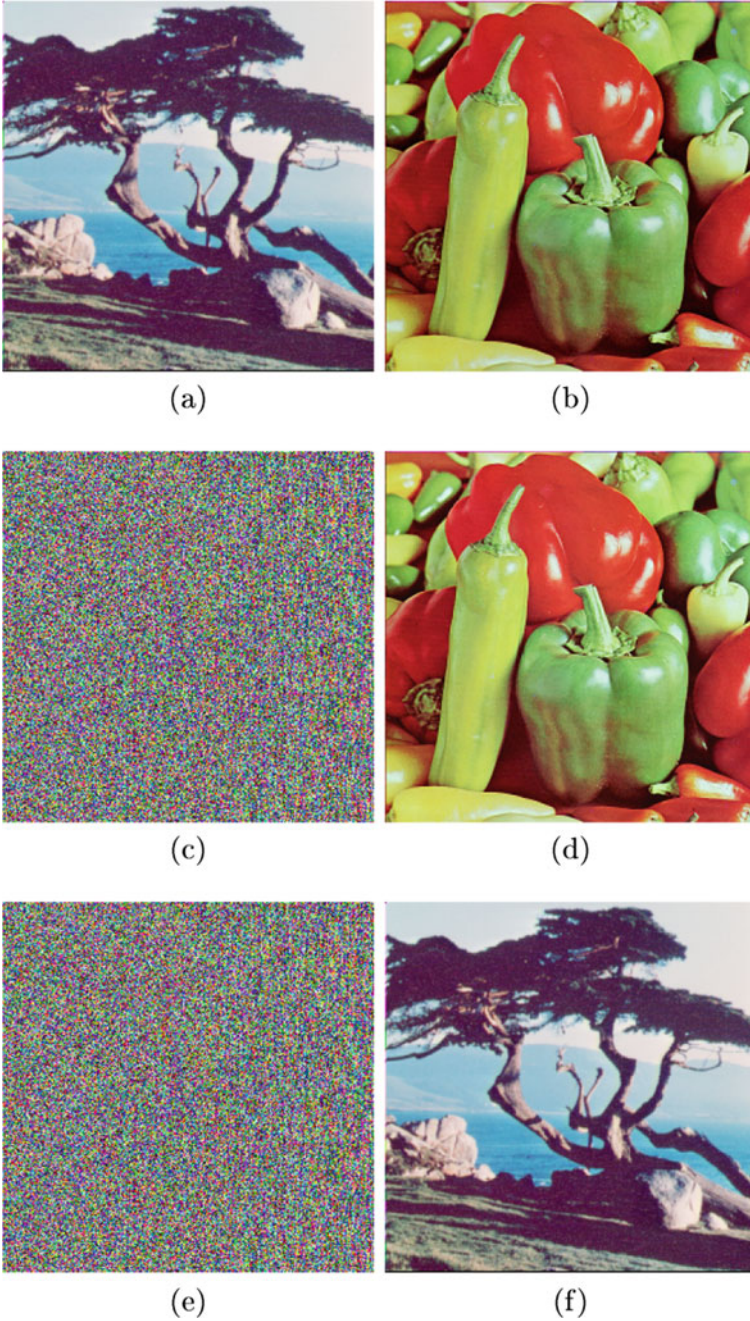


Fig. 5 Experimental results : **a** Secret image, **b** cover image, **c** encrypted secret image, **d** stego image, **e** extracted encrypted image, and **f** extracted secret image

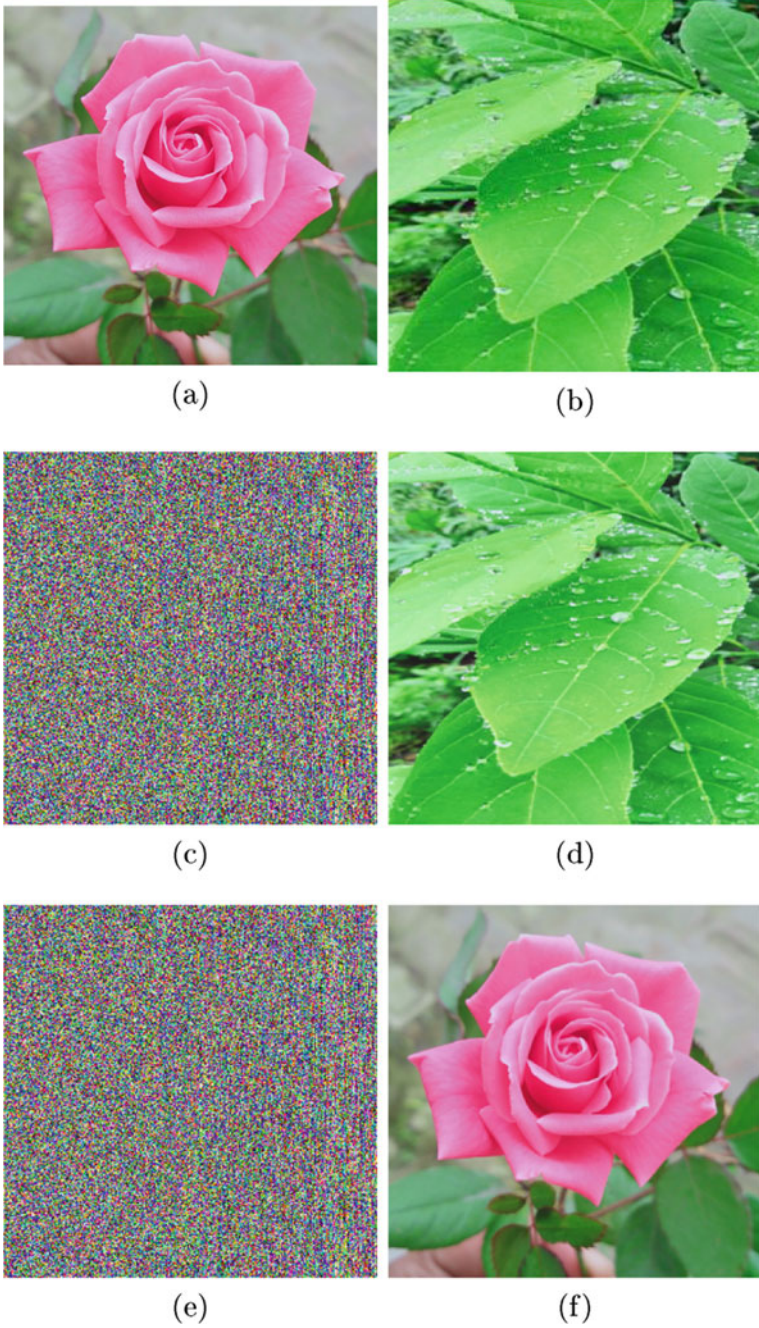


Fig. 6 Experimental results : **a** Secret image, **b** Cover image, **c** Encrypted secret image, **d** Stego image, **e** Extracted encrypted image, and **f** Extracted secret image

6.2 Peak Signal-To-Noise Ratio (PSNR)

“Peak Signal-to-Noise Ratio” (PSNR) is used to evaluate the quality of the extracted secret image. This ratio is used as a quality measurement between original secret image and extracted secret image. PSNR value for extracted image should be high as higher value of PSNR shows the better quality of extracted secret image. PSNR [8] is calculated by the following Eq. (19):

$$PSNR = 10 \log_{10} \frac{(255)^2}{MSE}. \quad (19)$$

Here, 255 is the maximum possible pixel value of the image. We also calculate the PSNR value between cover and stego images.

6.3 Structural Similarity Index Measure (SSIM)

SSIM [9] is a metric which is used to measure the image quality degradation because of some processing like image compression and image transmission. This SSIM value always lies between 0 and 1. SSIM gets its value 1 only when two datasets are identical and therefore it indicates the perfect structural similarity. The SSIM value between two images I_1 and I_2 is calculated by given Eq. (20):

$$SSIM = \frac{(2\mu_{I_1}\mu_{I_2} + J_1)(2\sigma_{I_1I_2} + J_2)}{(\mu_{I_1}^2 + \mu_{I_2}^2 + J_1)(\sigma_{I_1}^2 + \sigma_{I_2}^2 + J_2)}, \quad (20)$$

where μ_{I_1} and μ_{I_2} are the mean of the original secret image and extracted secret image, respectively. σ_{I_1} and σ_{I_2} are the standard deviation of the original and extracted secret images, respectively. $\sigma_{I_1I_2}$ is the covariance between the original and extracted secret images. $J_1 = (k_1L)^2$, $J_2 = (k_2L)^2$ and $k_1 = 0.01$, $k_2 = 0.03$ and $L = 2^{\text{number of bits per pixel}} - 1$.

SSIM value is also calculated between cover and stego images, in the similar manner as between secret and extracted images.

Calculated MSE, PSNR, and SSIM values between cover and stego images and again between secret and extracted images are shown in Table 1.

7 Security Analysis

In this section, we have analyzed some security parameters to check the validity and robustness of the proposed scheme.

Table 1 Experimental values of MSE, PSNR, and SSIM

Cover image	Stego image	MSE	PSNR	SSIM
Figure 5b (peppers)	Figure 5d	120.9988	27.3370	0.9884
Figure 6b (leaves)	Figure 6d	121	27.3369	0.9935
Secret image	Extracted image			
Figure 5a (tree)	Figure 5f	0	∞	1
Figure 6a (rose)	Figure 6f	0	∞	1

7.1 Key Space Analysis

In the proposed technique, 9D-chaotic system have control parameters $a_1, a_2, a_3, r_1, r_2, r_3$ and initial values $u_0, v_0, w_0, x_0, y_0, z_0, p_0, q_0$ and s_0 as secret keys. For the control parameters and initial values of 9D-chaotic system, if the precision is 10^{-15} , then the key space will be $10^{(15+15+15+\dots+15)_{15\text{-times}}} = 10^{225}$, which is much more sufficient to resist the brute-force attacks and makes data more secure.

7.2 Key Sensitivity Analysis

For key sensitivity analysis, we make some slight changes in secret keys. For this, we add a very small value $\Delta = 10^{-14}$ in control parameters and initial values. Due to the chaotic nature, we see that a very small change in key shows an unexpected and dramatic change in extracted image.

From Fig. 7, we can observe that extracted image is totally different from original image.

Figure 7a represents the extracted image of Fig. 5 by making a small change in control parameter a_1 , i.e., $a'_1 = a_1 + \Delta$.

Figure 7b represents the extracted image of Fig. 5 by making a small change in control parameter a_2 , i.e., $a'_2 = a_2 + \Delta$.

Figure 7c represents the extracted image of Fig. 5 by making a small change in control parameter a_3 , i.e., $a'_3 = a_3 + \Delta$.

Figure 7d represents the extracted image of Fig. 5 by making a small change in control parameter r_1 , i.e., $r'_1 = r_1 + \Delta$.

Figure 7e represents the extracted image of Fig. 5 by making a small change in control parameter r_2 , i.e., $r'_2 = r_2 + \Delta$.

Figure 7f represents the extracted image of Fig. 5 by making a small change in control parameter r_3 , i.e., $r'_3 = r_3 + \Delta$.

Figure 7g represents the extracted image of Fig. 5 by making a small change in initial value u_0 , i.e., $u'_0 = u_0 + \Delta$.

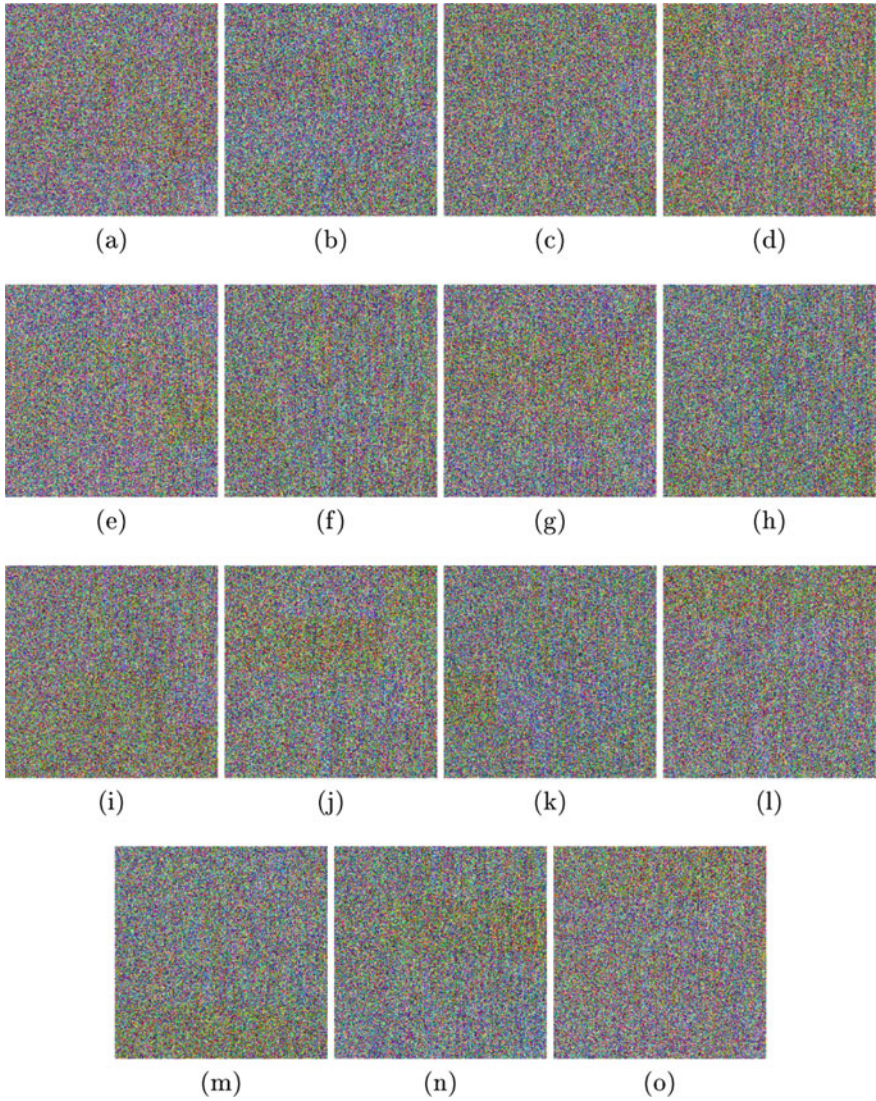


Fig. 7 Key sensitivity analysis : extracted images of Fig. 5 with wrong keys

Figure 7h represents the extracted image of Fig. 5 by making a small change in initial value v_0 , i.e., $v'_0 = v_0 + \Delta$.

Figure 7i represents the extracted image of Fig. 5 by making a small change in initial value w_0 , i.e., $w'_0 = w_0 + \Delta$.

Figure 7j represents the extracted image of Fig. 5 by making a small change in initial value x_0 , i.e., $x'_0 = x_0 + \Delta$.

Figure 7k represents the extracted image of Fig. 5 by making a small change in initial value y_0 , i.e., $y'_0 = y_0 + \Delta$.

Figure 7l represents the extracted image of Fig. 5 by making a small change in initial value z_0 , i.e., $z'_0 = z_0 + \Delta$.

Figure 7m represents the extracted image of Fig. 5 by making a small change in initial value p_0 , i.e., $p'_0 = p_0 + \Delta$.

Figure 7n represents the extracted image of Fig. 5 by making a small change in initial value q_0 , i.e., $q'_0 = q_0 + \Delta$.

Figure 7o represents the extracted image of Fig. 5 by making a small change in initial value s_0 , i.e., $s'_0 = s_0 + \Delta$.

7.3 Histogram Analysis

In digital images, histogram represents the relationship between the number of pixels and their intensity. Corresponding to Fig. 5, histogram of secret and cover images are shown in Fig. 8a, b. Figure 8c, d shows the histogram of encrypted secret image and stego image. Histogram of extracted encrypted image and extracted secret image are in Fig. 8e, f. Similarly, corresponding to Fig. 6, histogram of secret and cover images are shown in Fig. 9a, b. Figure 9c, d shows the histogram of encrypted secret image and stego image. Histogram of extracted encrypted image and extracted secret image are in Fig. 9e, f. This evaluation shows that our proposed technique is resistant to histogram analysis and protects data very efficiently.

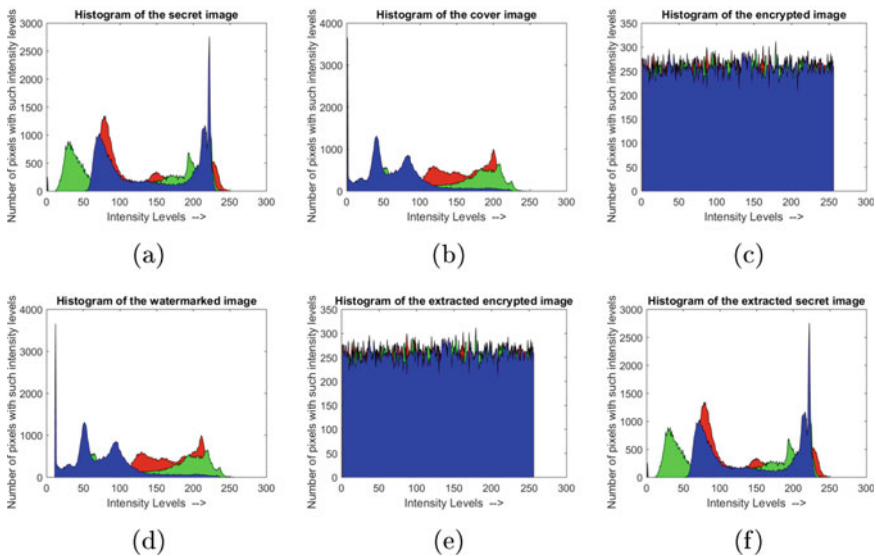


Fig. 8 Histogram analysis of Fig. 5: **a** Histogram of secret image, **b** Histogram of cover image, **c** Histogram of encrypted secret image, **d** Histogram of stego image, **e** Histogram of extracted encrypted image, and **f** Histogram of extracted secret image

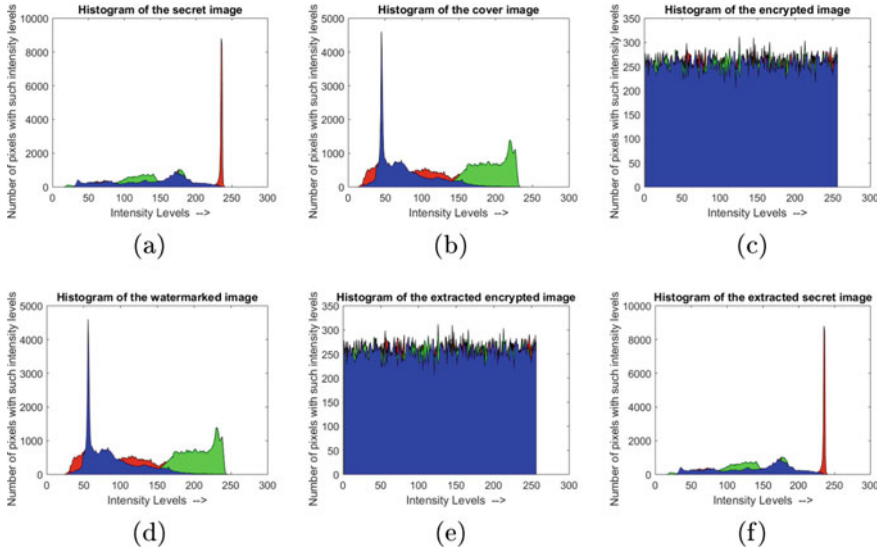


Fig. 9 Histogram analysis of Fig. 6: **a** Histogram of secret image, **b** Histogram of cover image, **c** Histogram of encrypted secret image, **d** Histogram of stego image, **e** Histogram of extracted encrypted image, and **f** Histogram of extracted secret image

7.4 Correlation Analysis

In this section, we have analyzed the pixel intensity distribution of adjacent pixels in horizontal(H), vertical(V), and diagonal(D) direction by calculating the correlation coefficient (CC) [10] using the given Eq. (21):

$$CC_{IE} = \frac{\sum_{i=1}^u \sum_{j=1}^v (I_{i,j} - \bar{I})(E_{i,j} - \bar{E})}{\sqrt{[\sum_{i=1}^u \sum_{j=1}^v (I_{i,j} - \bar{I})]^2 [\sum_{i=1}^u \sum_{j=1}^v (E_{i,j} - \bar{E})]^2}}, \quad (21)$$

where \bar{I} and \bar{E} are the mean of original and encrypted images, respectively.

Calculated CC values are displayed in Table 2. Correlation coefficients are measured for each direction (H, V, and D) and also for each color component (R, G, and B). From Table 2, it is clear that CC values of original images are near about 1 and that of encrypted image are close to 0 which shows that adjacent pixels in the original image are more correlated rather than encrypted image. Adjacent pixels in encrypted image have a slight correlation.

Figure 10 shows the pixel intensity distribution of adjacent pixels in H, V, and D directions for original and encrypted images.

Table 2 Experimental values of CC for original and encrypted images

Images	Component	Direction		
		H	V	D
Figure 5a (tree)	R	0.9590	0.9361	0.9159
	G	0.9687	0.9457	0.9318
	B	0.9612	0.9406	0.9265
Figure 5c (encrypted)	R	0.0022	0.0400	-0.0007
	G	0.0001	0.0326	0.0079
	B	-0.0028	0.0407	-0.0091
Figure 6a (rose)	R	0.9920	0.9922	0.9854
	G	0.9683	0.9665	0.9455
	B	0.9816	0.9805	0.9663
Figure 6c (encrypted)	R	-0.0085	0.0513	-0.0016
	G	0.0008	0.0378	-0.0048
	B	0.0024	0.0457	0.0006

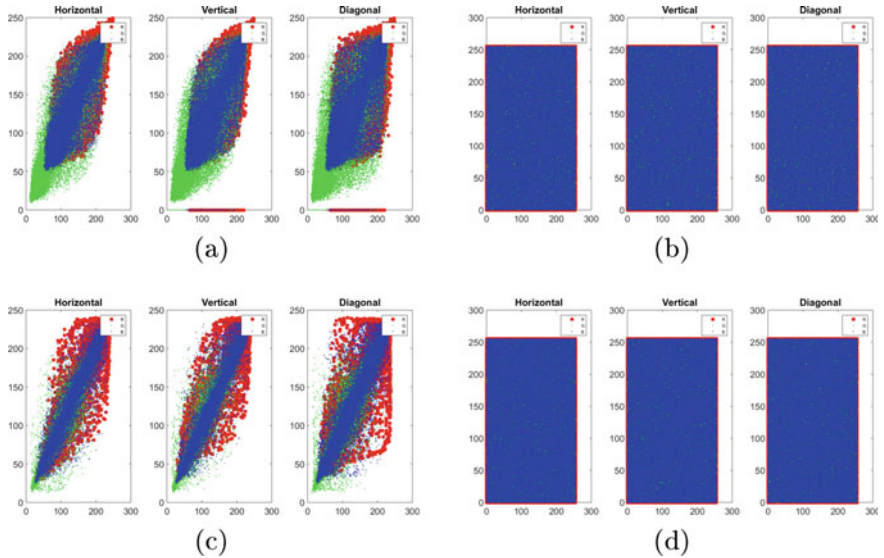


Fig. 10 Graphical analysis of pixel intensity distribution: **a** pixel intensity distribution of Fig. 5a at H, V, and D directions, **b** pixel intensity distribution of Fig. 5c at H, V, and D directions, **c** pixel intensity distribution of Fig. 6a at H, V, and D directions, and **d** pixel intensity distribution of Fig. 6c at H, V, and D directions

Table 3 Experimental values for entropy

Secret image	Entropy for secret image	Encrypted image	Entropy for encrypted image
Tree Fig. 5a	7.1816	Figure 5c	7.9971
Rose Fig. 6a	6.9914	Figure 6c	7.9972

7.5 Entropy Analysis

Entropy is the measurement of the degree of randomness in the image. The mathematical method to calculate entropy is given by Shannon [11], which is given below:

$$H(x) = - \sum_{i=1}^N P(x_i) \log_2 P(x_i), \quad (22)$$

where $P(x_i)$ is the probability of occurrence of x_i .

Entropy values for original and encrypted images are shown in Table 3. Entropy value for encrypted image is close to its ideal value 8, which shows that our proposed technique is secure to resist the entropy-based attacks.

7.6 Cropping Attack Analysis

Cropping attacks have been performed on stego image in different formats to check the robustness of the proposed scheme. The stego image Fig. 5d is cropped with block size of 32×32 , 64×64 , 128×128 from left top corner, 32×32 , 64×64 from middle, and 256×64 , 256×128 from left, which are shown in Fig. 11a, c, e, g, i, k, m, respectively. The corresponding extracted images are shown in Fig. 11b, d, f, h, j, l, n, respectively. This analysis deduces that our proposed scheme is robust and prevents images from cropping attacks.

8 Comparison with Related Work

We have compared our proposed scheme with some existing methods. Comparison is based on payload capacity, SSIM, and PSNR for both stego and extracted secret images. Comparison results are shown in Table 4, from where we can deduce that

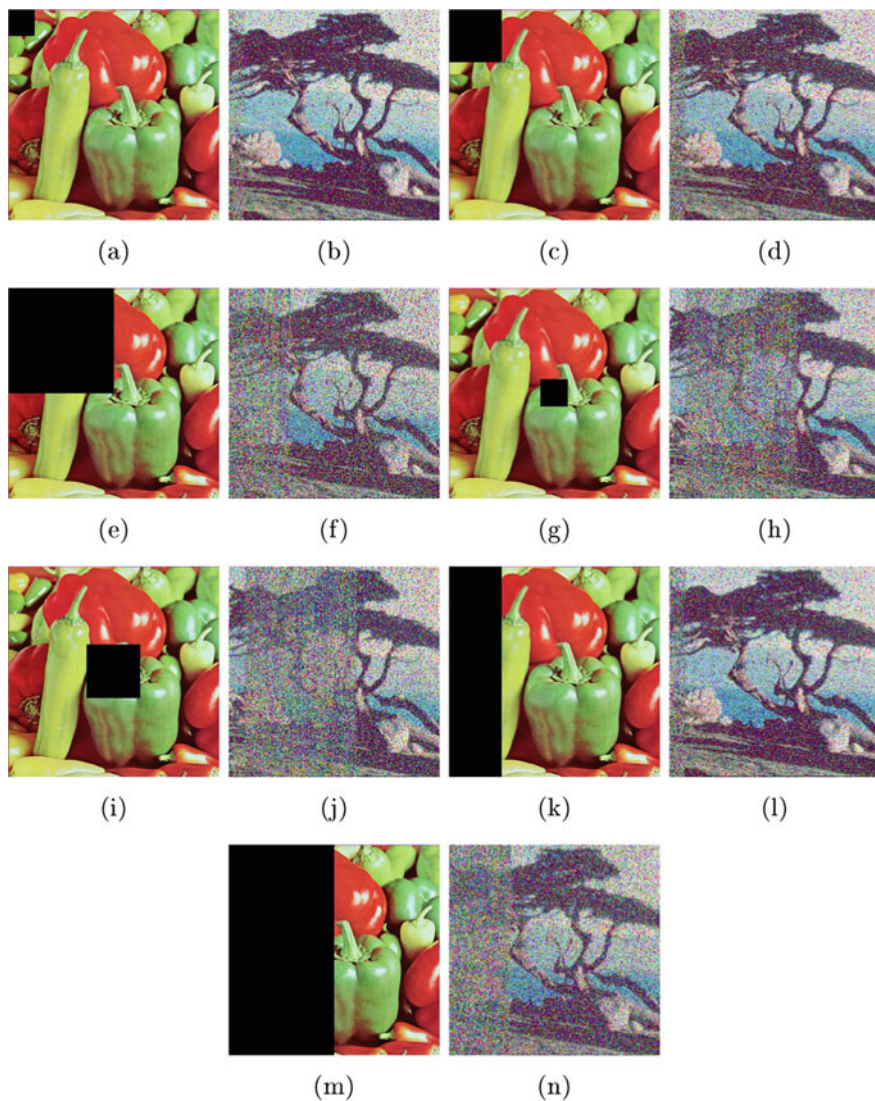


Fig. 11 Cropping attack analysis: extracted images under different levels of cropping attack

our proposed scheme has a very high payload (embedding) capacity and better SSIM as compared to other existing techniques. It also has high PSNR for extracted secret image.

Table 4 Comparison of our technique with other steganography techniques on the basis of Payload, SSIM, and PSNR

Steganography scheme	Payload (%)	Stego image (SSIM)	Extracted secret (SSIM)	Stego image (PSNR)	Extracted secret (PSNR)
[12]	100	0.98	0.97	41.2	37.6
[13]	33.3	0.937	0.93	32.5	34.76
[14]	100	0.985	0.981	40.45	37.32
[15]	100	0.97	0.984	40.47	36.92
[16]	100	0.987	0.953	42.3	38.45
Ours	100	0.9884	1	27.3370	∞

9 Conclusion

This paper proposes a new steganography technique for digital images. The 9D-hyperchaotic system is used to generate keys for the proposed permutation and substitution encryption scheme and DWT is used for embedding process. Performance evaluation metrics like MSE, PSNR, and SSIM confirm the robustness of the proposed scheme. It has a very large key space, which shows that it is resistant to brute-force attacks. Keys generated in the proposed scheme are highly sensitive with respect to the bit change. Histogram, correlation, and entropy analysis are also used to define the robustness of the proposed scheme against the static attacks. The proposed scheme is capable to resist the cropping attacks too, which is being analyzed here.

References

1. Shi S, Zhang Y, Hu Y (2010) A wavelet-based image edge detection and estimation method with adaptive scale selection. *Int J Wavelets Multiresolut Inf Process* 8(3):385–405. <http://dx.doi.org/10.1142/S0219691310003547>
2. Ehler M, Koch K (2010) The construction of multiwavelet biframes and applications to variational image denoising. *Int. J. Wavelets Multiresolut. Inf Process* 8(3):431–455. <http://dx.doi.org/10.1142/S0219691310003560>
3. Tang YY (2009) *Wavelets theory approach to pattern recognition*, 2nd edn, vol 74. World Scientific
4. Gonzalez R, Woods R (2008) *Digital image processing*, 3rd edn. Prentice Hall, Upper Saddle River
5. Jain N, Singh M, Mishra B (2018) Image compression using 2D-discrete wavelet transform on a light-weight reconfigurable hardware. In: 31st international conference on VLSI design and 2018 17th international conference on embedded systems. Pune, pp 61–66. <http://dx.doi.org/10.1109/VLSID.2018.38>
6. Haar A (1910) Zur theorie der orthogonalen funktionen-systeme. *Math Ann* 69(3):331–371
7. Grassi G, Severance FL, Miller DA (2009) Multi-wing hyperchaotic attractors from coupled Lorenz systems. *Chaos, Solitons Fractals* 41(1):284–291. <https://doi.org/10.1016/j.chaos.2007.12.003>

8. Wang Z, Bovik AC (2006) Modern image quality assessment. *Synth Lect Image Video Multimed Process* 2(1):1–156. <http://dx.doi.org/10.2200/S00010ED1V01Y200508IVM003>
9. Wang Z, Bovik AC, Sheikh HR, Simoncelli EP (2004) Image quality assessment: from error visibility to structural similarity. *IEEE Trans Image Process* 13(4):600–612. <http://dx.doi.org/10.1109/TIP.2003.819861>
10. Khan JS, Boulila W, Ahmad J, Rubaiee S, Rehman AU, Alroobaea R, Buchanan WJ (2020) DNA and plaintext dependent chaotic visual selective image encryption. *IEEE Access* 8:159732–159744. <https://doi.org/10.1109/ACCESS.2020.3020917>
11. Shannon CE (1948) A mathematical theory of communication. *Bell Syst Tech J* 27(3):379–423. <http://dx.doi.org/10.1002/j.1538-7305.1948.tb01338.x>
12. Baluja S (2017) Hiding images in plain sight: deep steganography. In: *Neural information processing systems*, pp 2066–2076
13. Rehman AU, Rahim R, Nadeem MS, Hussain SU (2018) End-to-end trained cnn encoder-decoder networks for image steganography. In: *European conference on computer vision*, pp 723–729
14. Duan X, Jia K, Li B, Guo D, Zhang E, Qin C (2019) Reversible image steganography scheme based on a U-net structure. *IEEE Access* 7:9314–9323
15. Duan X, Liu N (2019) Hide the image in FC-densenets to another image. [arXiv: Multimedia](https://arxiv.org/abs/1908.08114)
16. Li Q et al (2020) A novel grayscale image steganography scheme based on chaos encryption and generative adversarial networks. *IEEE Access* 8:168166–168176. <https://doi.org/10.1109/ACCESS.2020.3021103>
17. Joshi AB, Kumar D, Mishra DC (2020) Triple color image encryption based on 2D multiple parameter fractional discrete fourier transform and 3D arnold transform. *Opt Lasers Eng (Elsevier)* 133:106139. <https://www.sciencedirect.com/science/article/abs/pii/S0143816619319864>
18. Joshi AB, Kumar D, Mishra DC, VandanaGuleria, Color-image encryption based on 2D discrete wavelet transform and 3D logistic chaotic map. *J Modern Opt (Taylor and Francis)* 67(10):933–949. <https://doi.org/10.1080/09500340.2020.1789233>
19. Joshi AB, Kumar D, Mishra DC (2020) Security of digital images based on 3D arnold cat map and elliptic curve. *Int J Image Graph* 21(1):2150006. <https://doi.org/10.1142/S0219467821500066>
20. Joshi AB, Kumar D (2019) A new method of multi color image encryption. In: *2019 IEEE conference on information and communication technology*, pp 1–5. <https://ieeexplore.ieee.org/abstract/document/9066198>
21. Kumar D, Joshi AB, Mishra VN (2020) Optical and digital double color-image encryption algorithm using 3D chaotic map and 2D-multiple parameter fractional discrete cosine transform. *Results Opt* 1:100031
22. Kumar D, Joshi AB, Singh S (2021) 6D-chaotic system and 2D fractional discrete cosine transform based encryption of biometric templates. *IEEE Access* 9:103056–103074
23. Kumar D, Joshi AB, Singh S (2021) A novel encryption scheme for securing biometric templates based on 2D discrete wavelet transform and 3D Lorenz-chaotic system. *Results Opt* 5:100146
24. Gaffar A, Joshi AB, Singh S et al (2022) A high capacity multi-image steganography technique based on golden ratio and non-sampled contourlet transform. *Multimed Tools Appl* 81:24449–24476. <https://doi.org/10.1007/s11042-022-12246-y>

A Novel Approach For Secure Data Aggregation Scheme in Battlefield Surveillance Using Elliptic Curve Cryptography



Abhishek Bajpai and Anita Yadav

1 Introduction

Wireless sensor network (WSN) is a fast-expanding area of research that has attracted considerable research attention and has found tons of applications in the real world. The technological advancements in recent times have made it possible to create low-cost sensor nodes with wireless network interfaces that are quite efficient to be deployed in the real world. The number of nodes deployed in WSN can range from a few hundred to thousands; the number could reach even millions in IoT. These nodes can be deployed densely or in an uncrowded fashion depending upon the application for which they are used. WSNs are often deployed to monitor area that is unapproachable to humans, e.g., forest fire detection, flood detection, battlefield surveillance, etc. The sensor nodes that operate in WSN are fragile and are restricted by low computational power, less memory, and short battery lifetime. Also, energy is consumed both in transmitting and receiving data. Transmission of large data over the network considerably consumes the energy of sensors. Also in comparison to processing transmission depletes the majority of the energy of small nodes, according to research. As a result, it is critical to keep WSN communication to a minimum. Data aggregation refers to the process of eliminating the redundant data collected by the sensor nodes. Data aggregation plays a major role in enhancing the network lifetime by eliminating redundancy in data. The number of sensor nodes deployed in WSN can be large at times. The data collected by sensor nodes is huge, highly correlated, and repetitive. As the sensor nodes are resource-constrained, data aggregation is a vital technique that should be used to utilize the resources adequately. In the process of data aggregation, the data collected from source nodes are aggregated using a proper

A. Bajpai (✉)

Computer Science and Engineering, Rajkiya Engineering College, Kannauj, India
e-mail: abhishek.neophyte@gmail.com

A. Bajpai · A. Yadav

Computer Science and Engineering, Harcourt Butler Technical University, Kanpur, India

aggregation algorithm, generating data that is small in size. WSN can be deployed in critical areas to collect some sensitive information, such types of areas are prone to attacks and battlefields are one such type of area [1]. Attackers may attempt to take control of the WSN or intercept data transmissions from source nodes to sink in order to obtain access to sensitive data. As a result, securely storing and transmitting data in such type of network becomes vital. Also, there can be critical data generated by nodes in WSN which should reach the base station with no delay. Attackers may attempt to steal sensitive data as it is being transmitted from source to aggregator nodes to sink, so data secrecy must be maintained. Encryption can be used to ensure data confidentiality [2]. Prior to transmission, the data should be encrypted. Adversaries may attempt to introduce fraudulent information into the network; hence data authenticity must be ensured. Integrity is another crucial aspect to consider when transmitting data from nodes to sink as adversaries may attempt to alter the data which could have negative consequences for the base station's final outcome. There are various techniques that have been applied for securely transmitting the data. Several schemes have been proposed, some use encryption while some don't. Schemes that use encryption can be further classified into symmetric encryption and asymmetric encryption. Much Public key cryptography is usually heavy-weight and hence would require a lot of computational power [3, 4]. Symmetric cryptography on the other hand doesn't provide that much security. In this paper, we propose a novel approach for battlefield surveillance.

Our scheme is divided into four phases: in the first phase cluster head selection is done based on network parameters, and in the next phase the sensed data is encrypted using ECC, and signatures are calculated using the HMAC-based scheme. In the next phase, the cluster head (CH) aggregated the data and forwarded it to the sink node. In the final phase, the base station verifies the received data and decrypts the received data. We choose ECC-based scheme as our encryption algorithm because it involves low computational and communication overheads as compared to other asymmetric homomorphic schemes [5].

The rest of the paper is organized as follows Sect. 2 presents the literature review for the research study, Sect. 3 describes the preliminaries needed to understand the approach, Sect. 4 presents the proposed approach, Sect. 5 gives its security analysis and Sect. 6 gives its power analysis and Sect. 7 concludes the paper.

2 Literature Review

Castelluccia et al. [3] suggested an additive polymorphic scheme using symmetric key cryptography, a modified version of the Vernam cipher. Replacing X-OR in Vernam with an addition operation, making it ideal for resource-constrained WSN. The keys are chosen at random from key space in the Vernam cipher. However, in CMT, keys are generated using the stream cipher. Privacy is improved by using a unique message identifier for generating keys between the node and base station. The CMT cryptosystem also aids in the distribution of communication overhead across all sen-

sensor nodes. This load balancing across WSNs extends the lifetime of WSNs. Sruthi et al. [6] added multi-path topology, slicing, and combination techniques to DB-MAC. The proposed algorithm suffers from high computational overhead as data slicing and combining is used. Though the scheme improves data security it is not computationally vital. Harn et al. [7] proposed a lightweight data encryption scheme for WSNs. A pair-wise shared key is established between every pair of sensors. Collected data is combined with these pair-wise shared keys over and over until it reaches the receiver node (by addition and subtraction only). Any intermediate sensor node will be unable to recover the data. At the end of transmission, the data can only be recovered by the receiver node. The scheme is ideal to be used in resource-constrained WSNs but does not provide much security. Khashan et al. [8, 9] designed an automated lightweight cryptographic scheme for WSNs called FlexCrypt. The FlexCrypt scheme works in three phases in the first phase cluster head selection mechanism is harnessed to rotate the cluster head, which balances power consumption and improves the quality of service. The second step proposes automatic lightweight encryption to secure wireless sensor data transmitted. For the connection between the sensors, the cluster head, and the neighborhood AP, a lightweight key management method is added in the final phase. The scheme proposed is both energy efficient and secure but involves a lot of computational overheads. Reshma et al. [10] proposed VEDSDA protocol in the research to reduce data redundancy, and data length, and provide data transfer security. The VEDSDA protocol used compression techniques to shorten data, resulting in lower energy use. Leveling, encoding, and decoding are all aspects of the data compression technique. The leveling step turns data into logical data, and the encoding and decoding phases compress and decompress data sizes at the source and destination, respectively. To encrypt and decrypt aggregated data, the Voronoi diagram concept is used. The proposed scheme provides data security and is energy efficient as data was compressed before transmission. In their research, Fang et al. [11] enhanced the CPDA protocol for WSNs and added intrusion detection to prevent WSNs from sinkholes and selective forwarding attacks. The cluster head increases the data security on the WSNs. The method, which slices data using a tree-structure network, can save energy and transmission time. The proposed method is practical, secure, and precise. For secure data transmission in WSN, Elhoseny et al. [12] present a unique encryption architecture based on Elliptic Curve Cryptography (ECC) and Homomorphic Encryption. The suggested encryption scheme is based on the GASONeC algorithm which employs a genetic process to create the best network topology in the form of clusters. The suggested encryption key is 176 bits long and is generated by combining the ECC key, node identification number, and cluster head distance (CH). Also, the encryption process uses permutation and combination. The proposed scheme is secure, energy efficient, and increases network lifetime. Boudia et al. [13] designed a scheme that provided end-to-end privacy and does hop-by-hop verification thus rejecting the malicious data at an early stage. This scheme extends the tiny ECC algorithm and uses HMAC for verification purposes. This scheme is very secure, fast, and reliable. Gopikrishnan et al. [14] proposed a scheme called HSDA which ensures data integrity, confidentiality, and authenticity while consuming less energy, computing overhead, and fewer collisions. For encryption, the HSDA

Table 1 Comparison of various schemes

Previous work	Technique	Merits	Demerits
Castelluccia et al. [3]	Efficient aggregation of encrypted data in wireless sensor networks	(1) Increase in network lifetime (2) Polymorphic homomorphism enhances network security	Uses symmetric encryption (3) Not fault tolerant
Sruthi et al. [6]	An efficient secure data aggregation technique for Internet of things network	Improves data security	(1) Not computationally vital (2) Has high computational overhead involved at the nodes. Hence reduces network lifetime
Elhoseny et al. [12]	A secure data routing scheme for WSN using elliptic curve cryptography and homomorphic encryption	(1) Energy efficient (2) Improved network lifetime (3) Proper key management	Not fault tolerant
Boudia et al. [13]	Secure and efficient verification for data aggregation in wireless sensor networks	(1) Less computational overheads (2) Secure (3) Data integrity satisfied	Network lifetime not improved

secure algorithm employs both public and private keys. This technique does not execute any decryption when the intermediate nodes aggregate the data from the leaf nodes. The aggregation tree created in the proposed work significantly reduces delay and collision problems.

3 Preliminaries

This section describes two concepts ECC and HMAC which are very important to understand the proposed scheme.

3.1 Elliptic Curve Cryptography

Elliptic curves are plane algebraic curves that have all points x, y and are defined by the equation:

$$Ax^3 + Bx^2y + Cxy^2 + Dy^3 + Ex^2 + Fxy + Gy^2 + Hx + Iy + J = 0$$

Cryptography uses elliptic curves in a simplified form (Weierstrass form), which is defined as

$$Y^2 = X^3 + Ax + B$$

Here A and B are constants such that

$$4A^3 + 27B^2 \neq 0$$

Elliptic Curve Cryptography (ECC) employs elliptic curves over a finite field F_p (where p is prime and $p > 3$) or F_{2m} (where the size of the field is $p = 2m$). This means that the field is a $p \times p$ square matrix, and the points on the curve can only have integer locations within the field. Within the field, all algebraic operations (such as point addition and multiplication) result in another point [15]. The modular form that the elliptic curve equation takes over the finite field F_p is

$$Y^2 = (X^3 + Ax + B) \text{ mod } p$$

ECC is based on Elliptic Curve Discrete Logarithmic which is a computationally hard problem that allows ECC to use a smaller key size than other cryptographic techniques while maintaining a high level of security [16]. The advantages of employing elliptic curves are well known: they provide smaller key sizes and more efficient implementations while maintaining the same level of security as other extensively used schemes like RSA also they can easily achieve homomorphic properties. Let the set (Y, p, E, B, n) be the elliptic curve parameters where Y is the public key generated, p is the prime field over which the elliptic curve is defined, E is the elliptic curve, B is the base point of order n . K is generated by the scalar multiplication of B with an integer x this integer x is kept private while the parameter set is kept public. To encrypt a plain text m the sender selects a random integer k and encrypts the message as $(kB, M + kY)$ in order to decrypt the receiver calculates $(M + kY - xkB)$. The aggregated cipher text can be computed as $M_{agg} = (k_1B, M_1B + k_1Y) + (k_2B, M_2B + k_2Y) = ((k_1 + k_2)B, (k_1 + k_2)Y, (M_1 + M_2)B)$.

3.2 Hash Based Message Authentication Code (HMAC)

HMAC can be used to certify a message’s data integrity and authenticity at the same time. An HMAC can be calculated using any cryptographic hash algorithm, such as SHA-2 or SHA-3. The HMAC’s cryptographic strength is determined by the underlying hash function’s cryptographic strength, the size of its hash output, and the size and quality of the key. SHA-1 is used in the proposed scheme due to the low hash digest size in output.

4 Proposed Scheme

The proposed scheme is divided into four phases. In the first phase clustering of the sensor nodes is done the sensor nodes are nominated for cluster head selection based on their residual power, distance to the base station, and the number of nodes in the vicinity of the cluster head. In the second phase key establishment and plaintext, encryption is done using ECC. In the next phase, data is verified using HMAC at the cluster heads and the data from cluster members are aggregated at the cluster head. In the final phase, decryption is done at the base station and the data is checked for authenticity also using HMAC.

4.1 Cluster Head Selection Phase

Since the nodes in WSN are low-powered and have very less computational power so we proposed a dynamic cluster head selection algorithm to save the battery power of the network. A major battery is consumed in transmitting the data, the more the data bits the more energy will be used in transmitting it. So, the very first parameter that we used to select the cluster head is the number of nodes in the vicinity. Also, the residual power of a node should also be considered while nominating it for cluster head selection if a node has residual power less than a threshold it should not be nominated for cluster head selection. The third and very last parameter is the distance of a node to the base station. The node should be close to the base station to reduce the transmission cost. Each node sends its node ID and residual power to all its neighbors by using data packets which are used to update the adjacent node table of each sensor node.

- If the number of nodes in the vicinity of a node (N_V) is higher than a threshold, its chances to be nominated as a cluster head get reduced.
- If the distance between a node and base station (d_b) is higher than a threshold, the chances of that node being nominated as cluster head gets reduced.
- If the residual power of a node (P) is less than a threshold, its chances of getting selected as a cluster head reduces to a minimum.

4.2 Data Encryption Phase

We assume that all the curve parameters (Y, p, E, B, n) and a master key K are preloaded in the sensor nodes. Each node in the network computes its shared keys using the master key. After every node has computed its shared key master key is deleted from the memory. A new encryption key pair is generated in each transmission round. Different underlying elliptic curves can be used in ECC cryptographic

Algorithm 1 Cluster Head Selection Phase

```

1: Input:  $n, d_T, N_T, P_T$ 
2: for each node  $i \in (1, n)$  do
3:   Compute  $d_{Bi}$ 
4:   Compute  $N_{vi}$ 
5:   if  $i \neq CH$  and  $d_{Bi} > d_T$  and  $N_{vi} > N_T$  then
6:     find  $CH$  that is closest to  $i$ 
7:     if such  $CH$  is found then
8:        $i$  sends join request to  $CH$ 
9:        $i$  sends release request to previous  $CH$ 
10:      update adjacent node tables
11:     else
12:       Select a new  $CH$ 
13:       New  $CH$  sends announcement
14:       Closest members send join request to  $CH$ 
15:       Closest members send release request to previous  $CH$ 
16:       Update adjacent nodes table
17:     end if
18:   end if
19:   if  $i = CH$  and  $CM_i > N_T$  and  $P_i > P_T$  then
20:     send release request to all  $CM$ 
21:     convert  $CH$  to ordinary node
22:     find  $CH$  that is closest
23:     send join request
24:     update adjacent node table
25:   end if
26: end for

```

techniques. Different curves give varying levels of security (cryptographic strength), performance (speed), and varying key size. In this scheme, we use the *secp224k1* curve which uses a 224-bit key. All the nodes in the network agree upon a curve and a generator or base point. Y is the public key of the base station, E is the elliptic curve, B is the generator or base point, and x is the private key of the base station. The scheme also uses two other functions $map()$ and $ramp()$. $map()$ is a deterministic function that converts a plain text message into a point on the plain. The $rmap()$ is the aggregate recovery function. Each sensor node in the network senses data m in the network, the data is passed into the $map()$ function then. After that, the data is encrypted using the ECC algorithm random integer k is also used in the process. The random integer k makes sure that a different cipher text is generated for the same plain text each time. The signatures are calculated using the shared keys generated using the master key and node N . A is a randomly generated number that can only be used once in a cryptographic transmission. It is used to prevent replay attacks. If the sensed data is of high importance and it is extremely important for it to reach the base station as early as possible a critical bit is introduced for this purpose. A critical bit set equal to 1 is appended to the left of the message packet which is being transmitted to the cluster head otherwise if the data is not critical it is set to 0. The compressed data along with the signatures are transmitted to the cluster head.

Algorithm 2 Data Encryption Phase

```

1: Input:  $m_i, Y, B, k_i, N_i, K_i$ 
2: Compute  $M_i = \text{map}(m_i)$ 
3: Compute  $C_i = \{K_i B, M_i + K_i Y\} = \{T_i, S_i\}$ 
4: Compress  $T_i$  and  $S_i$ 
5: Compute  $\text{signature}_{1i} = \text{HMAC}(K_i^{\text{CHi}}, T_i, N_i)$ 
6: Compute  $\text{signature}_{2i} = \text{HMAC}(K_i^{\text{CHi}}, S_i, N_i)$ 
7: if data is critical then
8:   Append bit=1 on left of  $T_i$  and  $S_i$ 
9: else
10:  Append bit=0 on left of  $T_i$  and  $S_i$ 
11:  Transmit  $(T_i, \text{signature}_{1i})$  and  $(S_i, \text{signature}_{2i})$  to CH
12: end if

```

Table 2 Variables used in the proposed scheme

Variable	Meaning
N	Total number of sensor nodes in network
N_{vi}	Number of nodes in the vicinity of node i
d_{Bi}	Distance between basic station and node i
d_r	Threshold distance
N_r	Number of cluster members allowed per cluster head
CH	Cluster head
CM_i	Cluster members of node i
P_i	Residual power of node i
P_T	Threshold power
m_i	Sensed data of node i
CM_{CH}	Cluster members of a cluster heads
n_{CH}	Set of all the cluster heads
D_{agg}	The final decrypted aggregated data

4.3 Data Aggregation Phase

When the data reaches CHs for each cluster member the CH first checks for the critical bit if the critical bit is set the data along with the signature is transmitted to the base station. If the critical bit is not set the received signatures are verified and the points are decompressed. If the signatures do not match the packet is rejected and the node is removed from the cluster member list and added to the malicious node list.

Algorithm 3 Data Aggregation Phase

```

1: Input:  $CM_{CH}, T_i, S_i$ 
2: for each  $i \in (1, CM_{CH})$  do
3:    $b = \text{leftmost bit of } T_i \text{ and } S_i$ 
4:   remove leftmost bit from  $T_i$  and  $S_i$ 
5:   if  $b=1$  then
6:     Transmit  $(T_i, \text{signature}_{1i})$  and  $(S_i, \text{signature}_{2i})$  to base station
7:   else
8:     Compute  $\text{signature}'_{1i} = \text{HMAC}(K_i^{CH}, T_i, N_i)$ 
9:     Compute  $\text{signature}'_{2i} = \text{HMAC}(K_i^{CH}, S_i, N_i)$ 
10:    if  $\text{signature}'_{1i} = \text{signature}_{1i}$  and  $\text{signature}'_{2i} = \text{signature}_{2i}$  then
11:      Decompress  $T_i$  and  $S_i$ 
12:    else
13:      Reject data packet, remove a node from CM and add a node to the malicious node list
14:    end if
15:  end if
16: end for
17: Compute  $T_{agg}$  and  $S_{agg}$  using ECC
18: Transmit  $(T_{agg}, \text{signature}_{1CH})$ ,  $(S_{agg}, \text{signature}_{2CH})$  and malicious node list to base station

```

4.4 Decryption and Verification at Base Station

When the data reaches the base station the message authenticity and integrity are checked at the base station. Signatures are compared if the signatures don't match the data packet is rejected. If the signatures match the points are decompressed. The aggregated data is decrypted using base station private key x . To recover the aggregated data the elliptic curve discrete logarithmic problem needs to be solved but since this needs to be solved at the base station, the base station has enough power and computational resources so it can be easily computed.

Algorithm 4 Decryption and Verification at Base Station

```

1: Input:  $n_{CH}, T_{agg}, S_{agg}, M_{agg}, N_{CHi}, K_{CH}, x$ 
2: for each  $i \in (1, n_{CH})$  do
3:   Compute  $\text{signature}'_{1i} = \text{HMAC}(K_{CH}^{BS}, T_{aggi}, N_{CHi})$ 
4:   Compute  $\text{signature}'_{2i} = \text{HMAC}(K_{CH}^{BS}, S_{aggi}, N_{CHi})$ 
5:   if  $\text{signature}'_{1i} = \text{signature}_{1i}$  and  $\text{signature}'_{2i} = \text{signature}_{2i}$  then
6:     Decompress  $T_{aggi}$  and  $S_{aggi}$ 
7:   else
8:     reject the message packet
9:   end if
10: end for
11:  $M_{agg} = (S_{agg} - xT_{agg})$ 
12:  $D_{agg} = \text{rmap}(M_{agg})$ 

```

5 Threat Model

The wireless sensor network, which is critical for monitoring areas where human intervention is difficult, is vulnerable to active and passive attacks. In active attacks, intruders may attempt to alter the originality of the aggregated data by inserting, deleting, or altering the message to be transmitted. In a passive attack, intruders attempt to eavesdrop on confidential information. Some of the attacks are:

- **Cipher text analysis:** In this type of attack intruder tries to obtain confidential data by illuminating cipher texts.
- **Known plain text attack:** The known-plaintext attack (*KPA*) is a cryptanalysis attack paradigm in which the attacker has access to both the plaintext (also known as a crib) and the encrypted version of the plaintext (cipher text). These can be used to uncover more classified information.
- **Replay attack:** In such type of attack a packet that is transmitted is retransmitted again by the adversary. This may lead to false or incorrect data collection at the base station.
- **Malleability:** Some cryptographic algorithms have the property of malleability. An encryption technique is malleable if it can convert one cipher text into another cipher text that decrypts to the same plaintext. That is, given an encryption of a plaintext m , another ciphertext can be generated that decrypts to $g(m)$, for a known function g , without knowing or learning m .

The proposed scheme maintains end-to-end data confidentiality. The scheme uses ECC for encrypting the data, the security of ECC relies on the elliptic curve discrete logarithmic problem, which is again very hard to solve. It may take years to solve this problem if the secret key is not known. We have considered a key size of 224-bit which can provide security much higher than many state-of-the-art techniques. Even if an adversary has a huge set of known plain text and cipher text, he cannot infer the confidential data and shared keys since the encryption scheme is probabilistic. Also, the cluster head does not decrypt the data in order to aggregate it, so compromising the CH does not reveal any confidential information. Data integrity and authenticity are satisfied by the use of HMAC. If data integrity is not satisfied the hop-by-hop verification fails and the data packet gets rejected. The scheme can also handle replay attacks. If an intruder tries to inject an old packet into the network, the use of nonce prevents the attack.

6 Performance Analysis

In this section, we first discuss the simulation environment and the parameters used. Contiki is an operating system designed for small IoT devices with limited memory, power, bandwidth, and processing capacity. It has a simple look but has all of the essential capabilities found in modern operating systems. It has management features for programs, processes, resources, memory, and communication. The Cooja

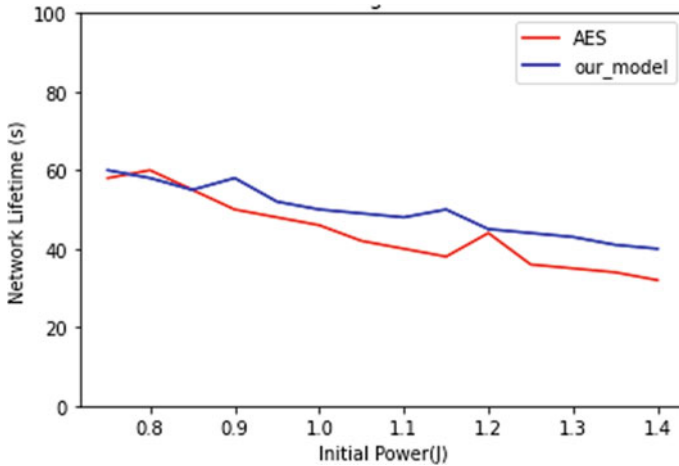


Fig. 1 Analysis of network lifetime

simulator is a piece of software that runs on the Contiki operating system and is used to model *6LoWPAN*, *RPL*, *CoAP* architecture, and networks. The simulation is run on a PC equipped with Intel i3 with 8 GB RAM and a 1.8 GHz processor. In the WSN 100 sensors are initially distributed in a grid fashion. The proposed is compared with the AES block cipher. AES is a common cryptographic algorithm that has a 128-bit block size and a variable number of rounds for each of the defined key sizes. It employs 10, 12, and 14 rounds, with 128-bit, 192-bit, and 256-bit key lengths, respectively.

We have analyzed the network lifetime of our proposed scheme and the previous research AES scheme in Fig. 1. The result shows that our scheme has considerably improved the network lifetime. There was an increase in network lifetime by about 12%. The total residual power percentage of the network against various simulation times was analyzed for both schemes in Fig. 2. The results show that our scheme considerably improved the total residual power left as compared to the AES block cipher scheme.

7 Conclusion

Security remains a challenging issue in the applications related to WSN and when it comes to battlefield surveillance secure transmission of data is of at most importance. Any adversary would try to steal sensitive data or try to manipulate it. Also, improving the network lifetime is a big issue in WSNs because of the fragile and tiny nature of the nodes it is essential to save battery power whenever possible. In this approach, we have given our very own cluster head selection algorithm, which aims at increasing

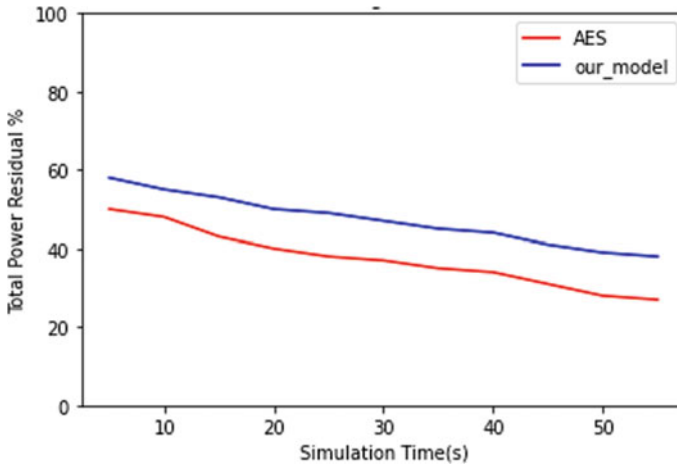


Fig. 2 Analysis of total residual power percentage

the network lifetime. Also, we have used Elliptic Curve-based cryptography for encrypting the data packets since it supports shorter key lengths and provides a much higher level of security. The concept of the critical bit is also introduced in this paper. If the data need to reach the base station as early as possible its critical bit is set so that data is not aggregated and sent directly to the base station. HMAC is used to maintain data authenticity and integrity. Security analysis of the approach shows that the method can resist various active and passive attacks.

References

1. Jaigirdar FT, Islam MM (2016) A new cost-effective approach for battlefield surveillance in wireless sensor networks. In: 2016 international conference on networking systems and security (NSysS). IEEE, pp 1–6
2. Vinodha D, Anita EM (2019) Secure data aggregation techniques for wireless sensor networks: a review. *Arch Comput Methods Eng* 26(4):1007–1027
3. Castelluccia C, Mykletun E, Tsudik G (2005) Efficient aggregation of encrypted data in wireless sensor networks. In: The second annual international conference on mobile and ubiquitous systems: networking and services. IEEE, pp 109–117
4. Wang H, Sheng B, Li Q (2006) Elliptic curve cryptography-based access control in sensor networks. *Int J Secur Netw* 1(3–4):127–137
5. Parmar PV, Padhar SB, Patel SN, Bhatt NI, Jhaveri RH (2014) Survey of various homomorphic encryption algorithms and schemes. *Int J Comput Appl* 91(8)
6. Sruthi SS, Geethakumari G (2016) An efficient secure data aggregation technique for internet of things network: an integrated approach using DB-MAC and multi-path topology. In: 2016 IEEE 6th international conference on advanced computing (IACC). IEEE, pp 599–603
7. Harn L, Hsu C-F, Xia Z, He Z (2021) Lightweight aggregated data encryption for wireless sensor networks (WSNS). *IEEE Sens Lett* 5(4):1–4

8. Khashan OA, Ahmad R, Khafajah NM (2021) An automated lightweight encryption scheme for secure and energy-efficient communication in wireless sensor networks. *Ad Hoc Netw* 115:102448
9. Medileh S, Laouid A, Euler R, Bounceur A, Hammoudeh M, AlShaikh M, Eleyan A, Khashan OA et al (2020) A flexible encryption technique for the internet of things environment. *Ad Hoc Netw* 106:102240
10. Reshma S, Shaila K, Venugopal K (2021) Vedsda: Voronoi encryption and decryption for secure data aggregation in WSNS. *Wirel Person Commun* 119(3):2675–2694
11. Fang W, Wen X, Xu J, Zhu J (2019) CSDA: a novel cluster-based secure data aggregation scheme for WSNS. *Clust Comput* 22(3):5233–5244
12. Elhoseny M, Elminir H, Riad A, Yuan X (2016) A secure data routing schema for WSN using elliptic curve cryptography and homomorphic encryption. *J King Saud Univ-Comput Inf Sci* 28(3):262–275
13. Merad Boudia OR, Senouci SM, Feham M (2018) Secure and efficient verification for data aggregation in wireless sensor networks. *Int J Netw Manag* 28(1):2000
14. Gopikrishnan S, Priakanth P (2016) HSDA: hybrid communication for secure data aggregation in wireless sensor network. *Wirel Netw* 22(3):1061–1078
15. Singh LD, Singh KM (2015) Implementation of text encryption using elliptic curve cryptography. *Procedia Comput Sci* 54:73–82
16. Bos JW, Halderman JA, Heninger N, Moore J, Naehrig M, Wustrow E (2014) Elliptic curve cryptography in practice. In: *International conference on financial cryptography and data security*. Springer, pp 157–175

A Recent Survey of Reversible Data Hiding Techniques for 2D and 3D Object Models



Amit Verma, Ruchi Agarwal, and Bhogeswar Borah

1 Introduction

Over the past decades, digitization has increased rapidly, and digital communication has become a part of daily life. With the help of the internet, transferring data in any digital form over the globe within seconds becomes easier. But, due to the ease of accessing digital information, its security has become an important concern. During the transfer of any digital content over any type of network or channel, there are lots of attacks that may be performed by illegal users. Many information security techniques have been performed to secure digital information. Information security is the soul of digital communication. The main techniques of information hiding are steganography [1], cryptography [2–4] and watermarking [5–8].

Steganography takes a binary form of information and hides it in the digital cover medium, which may be in the form of audio, video, images, or 3D objects. The main objective of cryptography is to use different encryption techniques and hide the information in a different way, which may be in a buzzed or scrambled form. The unauthorized party cannot detect the image content during the peer-to-peer transmission. On the other hand, watermarking is the process of secretly hiding information in multimedia content in such a way that perceptually its effect is negligible. In Fig. 1, the basic tree structure of a general data security system for digital content is presented. Watermarking makes users that the extraction and manipulation of secret information cannot be done by the third party. Using the watermarking techniques, security of the original data can be guaranteed, and no one can find whether the received media has been manipulated by the information hider or not. Although the

A. Verma · B. Borah
Tezpur University, Napaam, Sonitpur, Assam 784 028, India
e-mail: bgb@tezu.ernet.in

R. Agarwal (✉)
Babasaheb Bhimrao Ambedkar University, (A Central University), Vidya Vihar, Raebareli Road,
Lucknow 226025, India
e-mail: ruchi7777agarwal@gmail.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2024
B. K. Roy et al. (eds.), *Cryptography and Network Security with Machine Learning*,
Algorithms for Intelligent Systems, https://doi.org/10.1007/978-981-99-2229-1_24

279

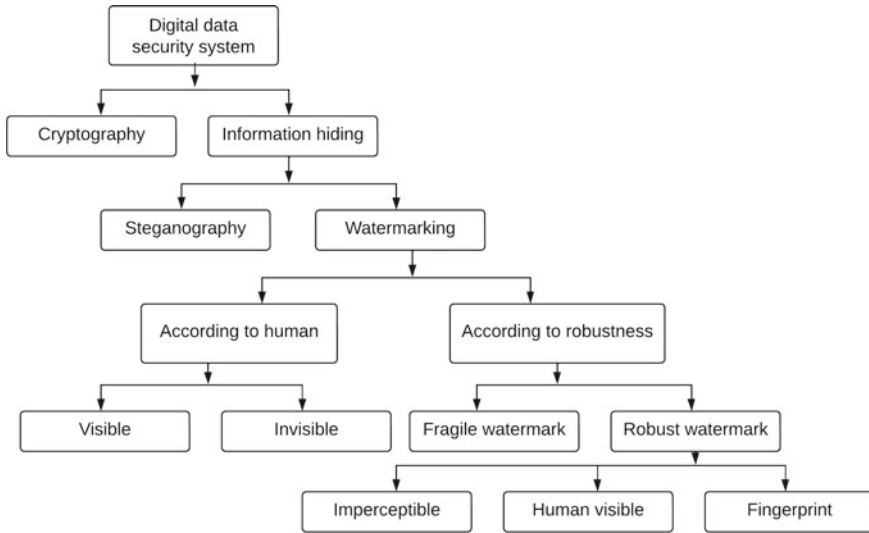


Fig. 1 Tree diagram of digital data security

watermarking may distort the crucial information present in the original cover media, thus at the receiver end, accurate recovery may not be possible.

For example, if conventional watermarking is used in the medical image, then the cover image may lose some important information, and after extraction of the watermark on the receiver side, the specialist may diagnose the wrong thing due to the loss of information. Thus, to solve this problem, reversible watermarking was introduced [8, 9]. Reversible watermarking is also called lossless watermarking and allows the user to extract full information from the embedded medium and recover the original cover media losslessly [9].

Reversible watermarking has caught the attention of the researchers in the last few years because of its use in different areas like medical and healthcare, military organization, legitimate documentation, and industrial communication. In Fig. 2, the basic block diagram of the reversible watermarking has been shown. There are two most important component of reversible watermarking, first one embedding capacity, that is, to embed the maximum payload in terms of bits (information) in the cover media; and imperceptibility, which is measured by the localization similarity between the cover media and the watermarked image. Honsinger et al. [10] introduced the methodology by manipulating the patchwork algorithm and performing modulo addition 256. However, the reversibility was achieved by the modulo 256 watermark, but the imperceptibility was not very impressive. In another work, [9, 11] techniques experienced salt and pepper noise. These techniques use the 256 modulo algorithm in the embedding system. Fridrich et al. [12] created the space between the cover image and the embedded information using compression techniques in the

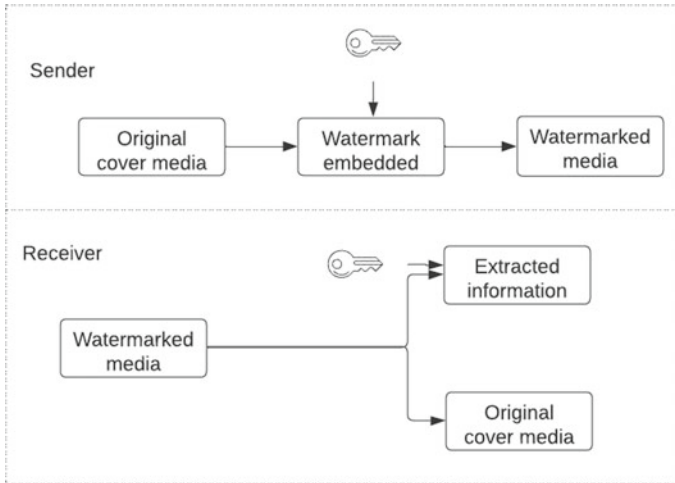


Fig. 2 Block diagram of reversible watermarking

least-significant bit plane, but the embedding capacity was very limited. To enhance embedding capacity and imperceptibility, many methodologies have been proposed by the research community in the last few years. In recent years, an amazing amount of state-of-the-art techniques and some enhanced versions of the previous versions have been proposed for transmission security, which is the main concern in the modern era.

Although there are many different types of reversible watermarking techniques available, most of the techniques focus on increasing the payload capacity and minimizing the distortion in cover media and also making it robust against different attacks by using various techniques. The previously reported reversible watermarking techniques mostly used 2D images as a cover medium, but the data-carrying capacity of 2D images is limited in comparison to 3D images. Due to the large capacity of data storage in 3D objects, it attracts the research communities to explore it more. Since 3D models display high information-carrying capacity in the digital world, the potential of work in different organizations like humans, architecture, organ structure, and cinematic characters can grow drastically. During the transmission of the 3D models, the cloud administrator is allowed to embed the information in the 3D models for content authentication. So, the requirement for object authentication and copyright protection is raised, and here comes the role of reversible watermarking in 3D models. To the best of my knowledge, there is few such survey reported in the field of RDH for 3D objects. Specifically in an encrypted domain and spatial domain there is no such work has been reported, which could describe the recent development in RDH in a 3D object and help researchers and academicians explore more in this area. Thus, a recent review on emerging reversible data hiding in 3D models is highly desirable.

With this motive, in this paper, a brief summary of emerging reversible data hiding techniques by using block diagrams, tabular representation, and comparison of available the state-of-the-art techniques has been presented. This paper is divided into the following sections: section II describes reversible data hiding and its various techniques. In section III, the RDH for 3D models and their various properties are explained. Section IV performance metrics and section V concludes and gives the future work of emerging techniques in the fields of RDH on 3D models.

2 Reversible Data Hiding

The aim of RDH is to secure the data in the transmission channel over any network. It works by embedding the secret information in a binary form in the source media, which is fully extracted at the receiver end. The receiver uses the extraction process and retrieves the source media and the secret message. Reversible data hiding has been exponentially investigated due to its ability and highly desirable application in the field of secure transmission, which may be in the form of copyright protection, source tracing, document tampering, photograph tampering, and content distribution in different organizations. RDH was first introduced by Barton in 1997 [9]. This development gives a boost to the traditional data hiding techniques. Due to this, ample amount of work on information hiding focused on RDH techniques in the last few decades. The main focus of the researcher while using RDH is to recover lossless cover media along with the embedded watermarking while maintaining a high payload. Reversible data hiding can be classified mainly into two domains: spatial domain and encrypted domain, shown in Fig. 3.

RDH mainly categories into three parts, lossless compression [9], difference expansion [13] and histogram sifting [14]. In spatial domain, RDH mainly works by developing the techniques to improve the payload capacity in the cover media and at the same time minimizing the distortion. In spatial domain, original cover media pre-processed first and then space is created for data embedding. After that, marked cover media is obtain by embedding secret information into cover media using lossless embedding techniques. Several noteworthy works have been offered by scholars in recent years that demonstrate the significance in the field of RDH in the spatial domain.

Jung and Hyun [20] proposed a method based on sorting and prediction. In this method, sub-block is divided into two groups—max and min. The pixel pair first predicted for both max and min groups and then modified for embedding the secret. It has an embedding capacity of 15684 bits and a PSNR of 51.44 dB. The proposed approach lacks resistance to malicious attacks. Qiu et al. [21] suggested an improved RDH scheme which basically used images texture to reduce insignificant moving of pixels intensity in histogram shifting. To assess the degree of smoothness, the flutter value of each pixel in each component is calculated and sorted in ascending order. It achieves good payload capacity and better visual quality of cover image with significantly improved PSNR. Hou et al. [18] propose a dynamical multiple

histogram generation using a deep neural technique. For minimization of distortion, multiple histograms are established according to the image content which helps to find the optimal bins using the proposed algorithm. It achieves 112,767 bits embedding capacity which was optimal in comparison of previous methods. RDH in neural networks still has the potential to explore in terms of embedding capacity.

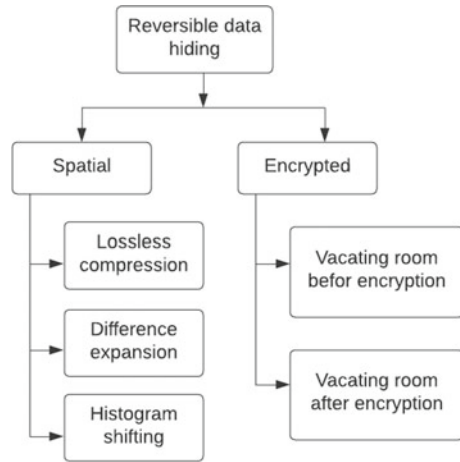
He and Cai et al. [17] introduced a dual pairwise prediction-error expansion technique that uses double-layered embedding to expand the original embedding in order to find the best way to define a spatial position for each block. It achieves 30,000 bits embedding capacity. To significantly enhance the capacity-distortion performance, a novel local complexity computation and multi-peak embedding (MPE) were suggested by Fu et al. [16]. In order to increase accuracy, the pixel local complexity (PLC) of every set of group is calculated using the pixel correlation techniques. This method finds the good embedding capacity of 110,725 bits and the PSNR of 62.56 dB. RDH schemes in spatial domain and their embedding capacity along with different methods are listed in Table 1. Spatial techniques are less resistant against different types of attacks and the computational complexity is also higher. Also, spatial methods have low embedding capacity in comparison to other domains. Further, the information hiding was investigated in the encrypted domain to achieve high-capacity data embedding and privacy protection for the secure transmission of the information. RDH in the encrypted domain has been introduced in a two-way fashion; one is vacating room before encryption (VRBE) [15]. In vacating room before encryption the data owner encrypts the cover media after creating space in the original cover data to hide the secret information by the data hider. On the other, original cover media is encrypted using the vacating room after encryption (VRAE) [16] technique without any preprocessing on the part of the content owner. Additionally, the data hider modifies some pieces of the encrypted cover media to incorporate the hidden information. The high entropy of the encrypted image limits this approach's ability to embed data. However, this strategy works quickly and simply for the end user.

All these types of techniques are used in the cloud computation and signal processing of healthcare systems. Some significant work has been proposed by the researchers in the field of encrypted domain in recent years. Weng et al. [19] proposed an improved multi-histogram point selection based Criss-cross Optimization techniques to find prominence solution in the global solution space also K-means clustering technique is employed in this study to divide all prediction errors into various groups based on the local complexity. The selection of embedding locations for multiple sub-histograms is reduced to a classic multi-choice knapsack issue. It achieves payload capacity of 45,000 bits and the PSNR of 60.76 dB. These techniques have the potential to improve in the field of designing better features for each image block also, automatically determining optimal clustering numbers, when a large payload has been embedded. Xiang et al. [20] proposed an RDH scheme for redundant space transfer. Also created the encryption processes using control variables derived from the original image and introduced the 2D Logistic Adjusted—Sine Map (2D-LASM). The hidden data was removed by outer processing while the associated plain-text pixel intensity was kept unaltered via lossless data hiding. To

Table 1 Some significant work on reversible data hiding

Ref.	Techniques	ER (bits), PSNR (dB)	Domain
[17]	Sorting and prediction	15684, 51.44	Spatial
[18]	Reducing invalid shifting of pixels	25038, 63.48	Spatial
[19]	Multi-histogram point selection, criss-cross optimization	10000, 60.03	Encrypted
[20]	Redundant space transfer, chosen-plain text attack	~1.7114 bpp, ~49.43	Encrypted
[21]	Dual pairwise prediction-error expansion	30000, 53.21	Spatial
[22]	Deep neural networks for dynamical multiple histograms generation	10000, 59.97	Spatial
[23]	Pixel prediction and multi-MSB planes rearrangement	~2.797 bpp, –	Encrypted
[24]	Multiple data hidings, cipher-feedback secret sharing	2.91 bpp, –	Encrypted
[25]	Combining encryption, singular value decomposition, mersenne twister and thilox counter-based	~1 bpp, ~ 64.17	Encrypted
[26]	Two side histogram shifting	~2 bpp, –	Encrypted
[27]	Local complexity calculation and multi-peak embedding (MPE)	30000, 48.75	Spatial
[28]	High-capacity RDH-EI Pixel, prediction and entropy encoding	~3.14 bpp, –	Encrypted
[29]	Random element substitution	~5 bpp, –	Encrypted
[30]	Vector quantization encoded images	23171, 31.60	Encrypted
[31]	Adaptive gradient prediction scheme	2.9860 bpp, 9.2231	Encrypted
[32]	Multi-MSB prediction and Huffman coding, homogeneity index modification algorithm	6.19, ~7.03	Encrypted
[33]	Multi—most significant bit prediction	0.9896, 9.56	Encrypted
[34]	RDH in encrypted images, bit-plane partition	2.50 bpp, 16.9	Encrypted
[35]	Combines image encryption and data hiding	0.1134 bpp, –	Encrypted
[36]	High-capacity RDH-EI, gradient edge detection predictor	3.065 bpp, –	Encrypted
[37]	Partitioned into non-overlapping blocks	0.0312 bpp, 30.83	Encrypted
[38]	Integer wavelet transform and chaotic system	7.98 bpp, 60.2628	Encrypted
[39]	Separable, histogram shifting and the paillier cryptosystem	20000, ~50.17	Encrypted
[40]	Combining median prediction and bit plan cycling-XOR	1.280 bpp, ~15	Encrypted
[41]	Scalable secret reference matrix	4 bpp, ~34	Encrypted

Fig. 3 Tree diagram for reversible data hiding



address this issue Wu et al. [29] proposed a random element substitution method. It used the pallier cryptosystem and attained low computational complexity. Yin et al. [23] suggested to use a high-capacity, fully reversible RDH-EI technique which is basically based on pixel prognostication and multi most significant bit planes displacement. It produced a higher embedding rate (ER) when used in place of a direct pixel prediction approach on grayscale images (Table 1).

With the accelerated growth on the 3D application and their intrinsic long capacity on the Internet mesh models are frequently employed. Cloud storage has grown in popularity and attention over the past few years and various privacy-preserving applications, particularly for multimedia files like videos, audio files, images, and three-dimensional (3D) mesh models that need a lot of storage space. Because there are more points (or vertices) available in 3D mesh models than in 2D cover media, they are capable of carrying more hidden bits than 2D cover media. Reversible data hiding for 2D images laid the foundation of 3D mesh models.

3 RDH For 3D Object Models

In this section, a brief discussion of RDH on 3D object models is presented. 3D models provide mostly three types of components: vertices, faces, and edges as $M = \{V, F, E\}$ where, 3D triangular mesh M can be recognized as a vertices $(V) = \{v_1, \dots, v_m\}$ and faces $(F) = \{f_1, \dots, f_n\}$, edges $E = \{e_{ij}\}, 1 \leq i, j \leq m$ here, e_{ij} pair of connected vertices. A face can be a combination of three vertices, five vertices, or with n number of vertices. Vertex normals are displayed as $V_i = (V_{i,x}, V_{i,y}, V_{i,z})$ shows in Table 3. The purpose of taking 3D object models is to increase the embedding payload because the information storage capacity is larger than 2D images. The mesh model contains different types of information like geometric structure, topo-

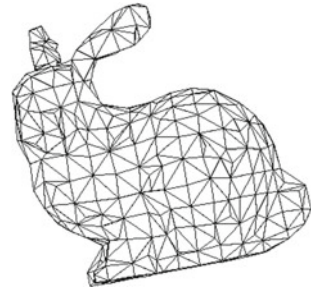
Table 2 Various RDH approaches proposed in recent years for 3D object models

Ref.	ER (bpv), SNR (dB)	Methodology	Robust
[42]	1.4, 58.63	Optimal three-dimensional modification, prediction-error histogram, optimal probability transition matrix, recursive construction coding	–
[43]	6, –	Two-tier RDH-ED framework, homomorphic Paillier cryptosystem	–
[44]	3.38, 25.88	Vacating room after encryption, least-significant bits	–
[45]	1.58, 48	Hybrid prediction scheme and multilayer strategy	–
[46]	3, –	Affine invariant difference shifting and logistic map	RST, Reordering
[47]	6, –	Shortest distances between neighboring vertices, fifth decimal places of the cartesian coordinate	Reordering, Rotation, Translation
[48]	0.972, 34.14	Prediction error expansion, homomorphic Paillier cryptosystem	–
[49]	13.5, –	Two-tier reversible data hiding paillier cryptosystem	–
[50]	1.07, 143.56	Integer mapping, most significant bit prediction, reserving room after encryption	–
[51]	7.75, –	Separable reversible data hiding	–
[52]	21.76, –	Multiple most significant bit	–
[53]	0.9, –	Variable direction double modulation	RST, Reordering
[54]	6.94, 96.40	Separable reversible, shortest distance between neighbor	Reordering

Fig. 4 Graphical representation of 3D model of bunny



Fig. 5 Mesh structure of 3D model



logical information, texture information, etc. The Bunny mesh model and its triangle mesh structure are shown in Figs. 4 and 5, respectively, which is provided by [59]. These types of models are available in such formats, .off, .obj, .pmd, .ply, .amf, .curl, etc. which are used for analysis by researchers according to the use of method and process applied to the models [59–61].

To enhance the embedding capacity and robustness, many RDH techniques for 3D objects have been reported in the last few years. RDH for 3D object models classification according to the domain, mainly three types, which are the spatial domain [47], the transform domain [46], and the compressed domain [54]. The spatial domain has some advantages, like being easy to construct algorithms in and having low computational complexity. Wu et al. [56] proposed a manipulation between the face centroid and the object gravity. In this technique, the original cover medium has minimal distortion during the reconstruction after the extraction of secret data. Further, to resolve the distortion of the cover media, Zhu et al. [57] suggested a prediction-error expansion-based method. The spatial domain has a higher embedded capacity of data with the minimum complexity, but the algorithm is fully fragile, which means in the transmission phase there may be hidden information loss. Thus, some flaws remain to be improved in the future. To resolve the robustness problem, 3D transform-domain proposed by Luo et al. [48], they performed a discrete cosine transform by modifying the vertex cluster and secret data embedded by a slight change in the high-frequency coefficients. Further DCT transformation with feature segmentation based on a double reversible data hiding proposed by Peng et al. [53], gives better robustness in comparison to the previous methodologies [50, 53, 54, 57]

Table 3 3D mesh file format

V index (i)	X-axis	Y-axis	Z-axis	F
1	V_1, x	V_1, y	V_1, z	1, 2, 8
2	V_2, x	V_2, y	V_2, z	7, 8, 1
3	V_3, x	V_3, y	V_3, z	5, 7, 1
4	V_4, x	V_4, y	V_4, z	1, 5, 4
5	V_5, x	V_5, y	V_5, z	3, 4, 1
6	V_6, x	V_6, y	V_6, z	2, 1, 3
...
...
36	V_{36}, x	V_{36}, y	V_{36}, z	36, 17, 133
...
...
90	V_{90}, x	V_{90}, y	V_{90}, z	90, 26, 177

stated work shows that these techniques transmit an ample amount of information to the receiver end. The applicability of the methods was substantially constrained by this limitation. Hence, before the transmission of information data compression is required. To address this issue compressed domain is introduced in RDH and it is a very challenging task for the 3D object models to merge with compressed RDH methodologies. Girdhar and Kumar [46] proposed affine invariant on a 3D mesh model and use difference shifting to hide the secret information inside the vertex and a Chaotic logistic Map was used to decide the coordinates. It achieves up to 3 bits per vertex (bpv) embedding capacity and this technique also withstands rotation, scaling, translation, and vertex reordering attacks. Bhardwaj [54] proposed efficient RDH for compressed 3D mesh models. It uses the shortest distance between neighbor vertices for removing the ambiguity of the traversal order of the mesh vertices. It achieves a higher embedding rate of 6.94 bpv and PSNR of 96.40 dB than [46]. Recent development in RDH techniques for 3D object models is represented in Table 2.

However, it is clearly noticed that the RDH method cannot be directly implemented in the 3D object models because of its complex geometric structure. Researcher faces some issues and challenges during the implementation of RDH on 3D mesh models. Multimedia security-based attacks are very common these days and due to this, it is very difficult to find whether the attacks are intentional or unintentional. Attacks on 3D object models are majorly categorized into three types, 3D mesh preserving, edge connectivity, and edge connectivity adjusting [63–67]. The mesh preserving attacks incorporate file attacks that try to reorder mesh vertices and a combination of vertices (faces) due to this digital visualization changed. Also, rotation, scaling, and translation of vertices lie under the similarity transformation attacks. Under connectivity preserving attacks generally adding some noise to each and every vertex. Further, the most challenging attack in the 3D mesh model to connectivity altering attacks. It creates the synchronization issue and reduces the

approximate ratio of vertices. It requires during the implementation of blind techniques for 3D models. In recent years, the animated video games are highly explored by people so, they should be robust against the pose changing attacks. 3D models have some different attacks and noises like cropping, low pass filter, etc.

4 Performance Metrics

For the evaluation of the RDH techniques, different type of metrics has in proposed [41, 46, 53–55, 63–67]. One of the significantly used metrics is PSNR and some other metrics are like payload, structure similarity, correlation factor, etc.

• Peak Signal to Noise Ratio

In the PSNR, when the data hider hides the information in the cover media then the original information of the cover media will affect. So, the output of the stego media has to measure by the mean squared error value to check the similarity between cover media and stego media. PSNR is calculated as in Eq. 1.

$$PSNR = 10 \log_{10} \left(\frac{R^2}{MSE} \right) \quad (1)$$

where, R is the maximum variations in the marked media.

$$MSE = \frac{\sum_{i=1}^N (M_i - S_i)^2}{N} \quad (2)$$

here, M_i is the frequency of the i th pixels, stego cover media intensity define by S_i and N is the intensity of pixels in the original cover media. For 3D data SNR measured as in Eq. 3.

$$SNR = 10 \log \left(\frac{\sum_{i=1}^N [(v_{i,x} - \bar{v}_x)^2 + (v_{i,y} - \bar{v}_y)^2 + (v_{i,z} - \bar{v}_z)^2]}{\sum_{i=1}^N [(c_{i,x} - v_{i,x})^2 + (c_{i,y} - v_{i,y})^2 + (c_{i,z} - v_{i,z})^2]} \right) \quad (3)$$

here, $\bar{v}_x, \bar{v}_y, \bar{v}_z$ are the mean intensity of the embedded model and $v_{i,x}, v_{i,y}, v_{i,z}$, are the original cover media coordinates. $c_{i,x}, c_{i,y}, c_{i,z}$ are the modified intensity of the 3D model.

• Payload Capacity

The embedded information to the cover media measured by the payload capacity basically it represent as bit per pixel (BPP), where

$$BPP = \frac{\text{Number of information hide}}{\text{Total pixel in the cover media}} \quad (4)$$

- **Structural Similarity Index Measure**

SSIM is used to compare the structural information after the data extraction done by the data receiver. Mainly it is used to check the identical between the original cover data and reconstructed data as in Eq. 5.

$$SSIM(i, j) = \frac{(2\mu_i\mu_j + m_1)(2\sigma_{ij} + m_2)}{(\mu_i^2 + \mu_j^2 + m_1)(\sigma_i^2 + \sigma_j^2 + m_2)} \quad (5)$$

now,

$$m_1 = (k_1 Q)^2$$

$$m_2 = (k_2 Q)^2$$

here, μ_i and μ_j are the mean values of the image data i, j . σ_i^2 and σ_j^2 are the covariance of i, j , respectively. m_1, m_2 are the constant used as stabilizing parameters. Q is the variation of pixel value and constant k_1 and k_2 are 0.01, 0.03 respectively

- **Robustness**

Robustness is measured basis of the correlation coefficient (ρ) between the embedded information and the extracted information from the cover media and it lies between range $[-1, 1]$, where higher values refer to low degradation.

$$\rho = \frac{\sum_i (V_i - \bar{v}) \cdot (V'_i - \bar{v}')}{\sqrt{\sum_i (V_i - \bar{v})^2 \cdot \sum_i (V'_i - \bar{v}')^2}} \quad (6)$$

here, \bar{v} and \bar{v}' are the mean of the inserted data sequence (V) and recovered data sequence (V'), respectively.

- **Tamper detection**

The 3D mesh model is altered if at least one of the vertices coordinate is modified or we can say the face is modified by inserting one or more vertices or moving the sequence of vertices in the face of 3d mesh models. Tamper detection can be assessed as in equations 7 and 8, respectively,

$$\text{False Positive} = \frac{N_{\text{correct mesh model}}}{N_{\text{tampered mesh model}}} \quad (7)$$

$$\text{False Negative} = \frac{N_{\text{incorrect}}}{N_{\text{total}} - N_{\text{tampered}}} \quad (8)$$

here, rejected tampered 3D mesh models is $N_{\text{incorrect}}$ and total number of 3D mesh models N_{total} .

5 Conclusion and Future Scope

This survey briefly describes several reversible data hiding techniques for 2D and 3D objects. Payload, robustness, and imperceptibility are the preliminary specifications for any reversible data hiding techniques. However, studies show that meeting these specifications at the same time is difficult. Additionally, this survey covers nearly every recent advancement in reversible data hiding techniques in different domains like spatial domain, and encrypted domain. We have also summarized some recently proposed states of the work in terms of their robustness toward various attacks and embedding capacity. Future work includes improving the schemes to achieve, high payload capacity, and recover secret information from the tampered 3D model without using prior information. Tamper detection accuracy has to maximize false negative detection and minimize false positive detection. Reversible data hiding techniques can be extended further using deep learning and artificial intelligence methods to develop robust solutions and generalized frameworks for both spatial and encrypted domains.

References

1. Pfitzmann B (1996) Information hiding terminology. In: Proceedings of first international workshop information hiding, Lecture notes in computer science, vol 1,174. Springer, Berlin, pp 347–356
2. Shannon C-E (1949) Communication theory of secrecy systems. *Bell Syst Tech J* 28:656–715
3. Goldwasser S, Micali S (1984) Probabilistic encryption. *J Comput Syst Sci* 28:270–299
4. Diffie W, Hellman M-E (1976) New directions in cryptography. *IEEE Trans Inf Theory* 22:644–654
5. Request for Proposals-Embedded Signalling Systems Issue 1.0. International Federation of the Phonographic Industry, U.K., London (1997)
6. Abraham D-G (1991) Transaction security system. *IBM Syst J* 30(2):206–229
7. Miller ML, Cox IJ, Bloom JA (1998) Watermarking in the real world: an application to DVD. In: Proceedings of multimedia and security-workshop at ACM multimedia '98, pp 71–76
8. Bender W, Gruhl D, Morimoto N, Lu A (1996) Techniques for data hiding. *IBM Syst J* 35(3):313–336
9. Barton J-M (1997) Method and apparatus for embedding authentication information within digital data, U.S. Patent 5, pp 646–997
10. Honsinger CW, Jones PW, Rabbani M, Stoffel JC (2001) Lossless recovery of an original image containing embedded data, US Patent, 6, pp 278, 791
11. Delaigle JF, Vleeschouwer CD, Macq B-M-M (1996) Digital watermarking. In: Proceedings of SPIE 2659, optical security and counterfeit deterrence techniques
12. Fridrich J, Goljan M, Du R (2002) Lossless data embedding for all image formats. In: Proceedings of SPIE 4675, security and watermarking of multimedia contents IV
13. Ni N, Shi Y-Q, Ansari N, Su W (2006) Reversible data hiding. *IEEE Trans Circ Syst Vid Technol* 16(3):354–362
14. Tian J (2003) Reversible data embedding using a difference expansion. *IEEE Trans Circ Syst Vid Technol* 13(8):890–896
15. Ma K, Zhang W, Zhao X, Yu N, Li F (2013) Reversible data hiding in encrypted images by reserving room before encryption. *IEEE Trans Inf Forensics Secur* 8(3):553–562

16. Li M, Xiao D, Zhang Y, Nan H (2015) Reversible data hiding in encrypted images using cross division and additive homomorphism. *Signal Process Image Commun* 39:234–248
17. Jung K-H (2016) A high-capacity reversible data hiding scheme based on sorting and prediction in digital images. *Multimed Tools Appl* 76
18. Jia Y, Yin Z, Zhang X, Luo Y (2019) Reversible data hiding based on reducing invalid shifting of pixels in histogram shifting. *Signal Process* 163:238–246
19. Weng S, Tan W, Ou B, Pan J-S (2021) Reversible data hiding method for multi-histogram point selection based on improved criss-cross optimization algorithm. *Inf Sci* 549:13–33
20. Xiang Y, Xiao D, Zhang R, Liang J, Liu R (2021) Cryptanalysis and improvement of a reversible data-hiding scheme in encrypted images by redundant space transfer. *Inf Sci* 545:188–206
21. He W, Cai Z (2021) Reversible data hiding based on dual pairwise prediction-error expansion. *IEEE Trans Image Process* 30:5045–5055
22. Hou J, Ou B, Tian H, Qin Z (2021) Reversible data hiding based on multiple histograms modification and deep neural networks. *Signal Process: Image Commun* 92:116118
23. Yin Z, She X, Tang J, Luo B (2021) Reversible data hiding in encrypted images based on pixel prediction and multi-MSB planes rearrangement. *Signal Process* 187:108146
24. Zhongyun H et al (2021) Secure reversible data hiding in encrypted images using cipher-feedback secret sharing. [arXiv:2106.14139](https://arxiv.org/abs/2106.14139)
25. Agarwal R, Kumar M (2021) A two side histogram shifting based reversible data hiding technique in encrypted images. In: *Computer vision and image processing*, vol 1376. Springer, Singapore, CVIP (2021)
26. Agarwal R, Kumar R (2021) Block-wise reversible data hiding in encrypted domain using SVD. *Optik* 247:16801
27. Fu Z, Gong M, Long G et al (2022) Efficient capacity-distortion reversible data hiding based on combining multippeak embedding with local complexity. *Appl Intell*
28. Qiu Y, Ying Q, Yang Y, Zeng H, Li S, Qian Z, High-capacity framework for reversible data hiding in encrypted image using pixel prediction and entropy encoding. *IEEE Trans Circuits Syst Video Technol*
29. Wu H-T, Cheung Y-M, Zhuang Z, Xu L, Hu J, lossless data hiding in encrypted images compatible with homomorphic processing. *IEEE Trans Cybern*
30. Fang C, Yujie F, Heng Y, Mian Z, Jian I, Chuan Q (2022) Separable reversible data hiding in encrypted VQ-encoded images. *Secur Commun Netw* 16
31. Qin J, He Z, Xiang X, Tan Y (2022) Reversible data hiding in encrypted images based on adaptive gradient prediction. *Secur Commun Netw* 12
32. Tsai Y-Y, Liu H-L, Kuo P-L, Chan C-S (2022) Extending multi-MSB prediction and huffman coding for reversible data hiding in encrypted HDR images. *IEEE Access* 10:49347–49358
33. Xu D (2022) Reversible data hiding in encrypted images with high payload. *IET Inf Secur* 16(4):301–313
34. Arai E, Imaizumi S (2022) High-capacity reversible data hiding in encrypted images with flexible restoration. *J Imaging* 8:176
35. Yang C-H, Weng C-Y, Chen J-Y (2022) High-fidelity reversible data hiding in encrypted image based on difference-preserving encryption. *Soft Comput* 26:4
36. Rai A-K, Om H, Chand S (2022) High capacity reversible data hiding in encrypted images using prediction error encoding. *Multimed Tools Appl*
37. Panchikkil S, Manikandan V-M, Zhang Y-D (2022) A pseudo-random pixel mapping with weighted mesh graph approach for reversible data hiding in encrypted image. *Multimed Tools Appl*. 81(12):16279–16307
38. Meng L et al (2022) Reversible data hiding in encrypted images based on IWT and chaotic system. *Multimed Tools Appl* 81(12):16833–16861
39. Tsai C-S, Zhang Y-S, Weng C-Y (2022) Separable reversible data hiding in encrypted images based on Paillier cryptosystem. *Multimed Tools Appl* 81(13):18807–18827
40. Fengyong L, Zhu H, Qin C (2022) Reversible data hiding in encrypted images using median prediction and bit plane cycling-XOR. *Multimed Tools Appl* 1–20

41. Lin J et al (2022) A large payload data hiding scheme using scalable secret reference matrix. *Symmetry* 14(4):828
42. Jiang R, Zhang W, Hou D et al (2018) Reversible data hiding for 3D mesh models with three-dimensional prediction-error histogram modification. *Multimed Tools Appl* 77:5263–5280
43. Shah M, Zhang W, Hu H et al (2018) Homomorphic encryption-based reversible data hiding for 3D mesh models. *Arab J Sci Eng* 43:8145–8157
44. Jiang R, Zhou H, Zhang W, Yu N (2018) Reversible data hiding in encrypted three-dimensional mesh models. *IEEE Trans. Multimed* 20(1):55–67
45. Zhang Q, Song X, Wen T et al (2019) Reversible data hiding for 3D mesh models with hybrid prediction and multilayer strategy. *Multimed Tools Appl* 78:29713–29729
46. Girdhar A, Kumar V (2019) A reversible and affine invariant 3D data hiding technique based on difference shifting and logistic map. *J Ambient Intell Human Comput* 10:4947–4961
47. Farrag S, Alexan W (2020) Secure 3D data hiding technique based on a mesh traversal algorithm. *Multimed Tools Appl* 79:29289–29303
48. Luo T, Li L, Zhang S, Wang S, Gu W (2021) A novel reversible data hiding method for 3D model in homomorphic encryption domain. *Symmetry* 13(6):1090
49. Rensburg B-J-V, Puteaux P, Puech W, Pedebay J-P (2021) Homomorphic two tier reversible data hiding in encrypted 3D objects. In: *IEEE International conference on image processing (ICIP)*, pp 3068–3072
50. Xu N, Tang J, Luo B et al (2022) Separable reversible data hiding based on integer mapping and MSB prediction for encrypted 3D mesh models. *Cogn Comput* 14:1172–1181
51. Tsai YY (2021) separable reversible data hiding for encrypted three-dimensional models based on spatial subdivision and space encoding. *IEEE Trans Multimed* 23:2286–2296
52. Iyu W, Cheng L, Yin Z (2022) High-capacity reversible data hiding in encrypted 3d mesh models based on multi-MSB prediction volume 108686:201
53. Peng F, Liao T, Long M (2022) A semi-fragile reversible watermarking for authenticating 3D models in dual domains based on variable direction double modulation. *IEEE Trans Circuits Syst Video Technol* 31:11
54. Bhardwaj R (2022) Efficient separable reversible data hiding algorithm for compressed 3D mesh models. *Biomed Signal Process Control* 73:103265
55. Girdhar A, Kumar V (2017) Comprehensive survey of 3D image steganography techniques. *IET Image Proc* 12(1):1–10
56. Wu X, Xie Z, Gao Y, Xiao Y (2020) SSTNet: detecting manipulated faces through spatial, steganalysis and temporal features. In: *IEEE international conference on acoustics, speech and signal processing (ICASSP)*, pp 2952–2956
57. Zhu J, Kaplan R, Johnson J, Fei-Fei L (2018) Hidden: Hiding data with deep networks. In: *Proceedings of the European conference on computer vision*, pp 657–672
58. Lavoué G (2009) A local roughness measure for 3d meshes and its application to visual masking. *ACM Trans Appl Percept (TAP)* 5(4), 1–23
59. Silva S, Madeira J, Ferreira C, Santos B-S (2007) Comparison of methods for the simplification of mesh models using quality indices and an observer study. In: *Human vision and electronic imaging XII*, vol 6492, pp 64921L
60. Cho J-W, Prost R, Jung H-Y (2006) An oblivious watermarking for 3d polygonal meshes using distribution of vertex norms. *IEEE Trans Signal Process* 55(1):142–155
61. <http://graphics.stanford.edu/data/3Dscanrep/>
62. Zhou Q, Jacobson A (2016) Thingi10k: a dataset of 10,000 3d-printing models, p 04797. [arXiv: 1605](https://arxiv.org/abs/1605.04797)
63. Cho J-W, Prost R, Jung H-Y (2006) An oblivious watermarking for 3-d polygonal meshes using distribution of vertex norms. *IEEE Trans Signal Process* 55(1):142–155
64. Nader G, Wang K, Hétyou-Wheeler F, Dupont F (2015) Just noticeable distortion profile for flat-shaded 3d mesh surfaces. *IEEE Trans Visual Comput Graph* 22(11):2423–2436
65. Bors A-G, Luo M (2012) Optimized 3d watermarking for minimal surface distortion. *IEEE Trans Image Process* 22(5):1822–1835

66. Mun S-M, Jang H-U, Kim D-G, Choi S, Lee H-K (2015) A robust 3d mesh watermarking scheme against cropping. In: 2015 international conference on 3D imaging (IC3D), pp 1–6
67. Seo Y-S, Joo S, Jung H-Y (2003) An efficient quantization watermarking on the lowest wavelet subband. *IEICE Trans Fundam Electron Commun Comput Sci* 86(8):2053–2055

Demystifying Facial Expression Recognition Using Residual Networks



Pratyush Shukla and Mahesh Kumar

1 Introduction

After the evolution of deep learning techniques, there is an exponential growth and development in the field of computer vision and pattern recognition. Convolution Neural Networks (CNNs) have emerged as pioneer in the field of visual imagery specifically image and object detection tasks. Facial Expression Recognition (FER) is one such domain which has CNN as its foundation. FER deals with identifying an individual's emotions and intentions and also has a huge impact over several disciplines such as human–computer interaction (HCI), mental disease diagnosis, psychology, neuroscience, and data-driven animations. The authors, Wang and Gu [9], explain various application of FER in the HCI field. The Residual Network (ResNet) is one of the phenomenal CXNN architecture which is widely used for FER purposes. In this paper, we have presented several ResNet composition [1, 2] and variants such as 3D Inception-ResNet Layer Composition [3], Dynamic Geometric Image Networks with ResNet [4], ResNet and Atrous Convolutions [5], Conditional Random Fields and ResNet [6], FER using ResNet and Heart Rate Variability Observation [7], and ResNet escorted by Squeeze and Exception Networks [8].

2 Explicating ResNet

Deep neural networks often produce satisfactory output but are difficult to train. Research indicates that deep convolutional neural networks are leading in the image recognition domain, but it is not always the case that adding more layers to the network improves learning. The convergence of deeper networks leads to the prob-

P. Shukla (✉) · M. Kumar

Department of Computer Science and Engineering, Jaypee University of Engineering and Technology, Guna 473226, India

e-mail: pratyush19shukla@gmail.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2024
B. K. Roy et al. (eds.), *Cryptology and Network Security with Machine Learning*,
Algorithms for Intelligent Systems, https://doi.org/10.1007/978-981-99-2229-1_25

295

lem of degradation [11], the problem of vanishing gradients, and can result in low confidence in the network.

The residual networks consist of shortcut connections, also known as skip connections, which provide good generalization ability [12] to the network and produce superior results without the degradation problem. The vanishing gradient problem, or the exploding gradient problem, is caused when the gradient becomes either very small, close to zero, or very large as the number of layers increases. It affects the accuracy, and as the number of layers increases, the training and test error simultaneously increase. The technique of using skip connections links the activation of one layer to further layers by skipping some intermediate layers. These residual blocks, when stacked together, form ResNets.

3 Explicating Facial Expression Recognition via Distinguishable Residual Networks

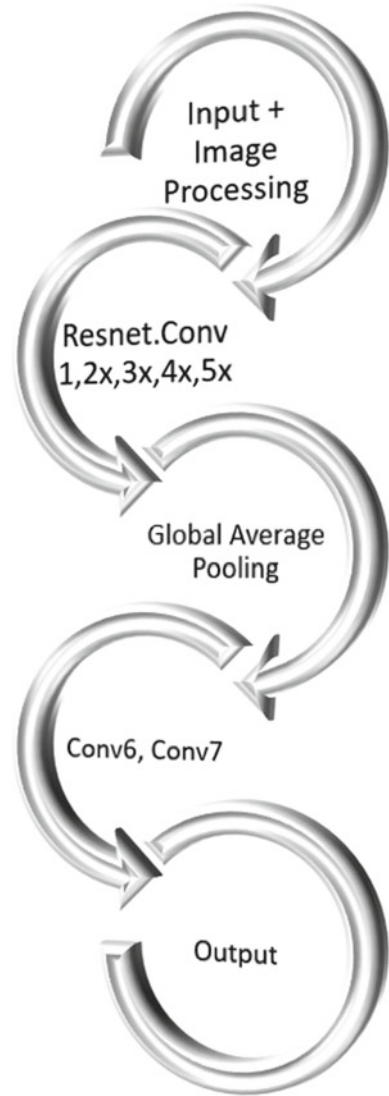
3.1 *ResNet-18*

Zhou et al. [1] proposed a new method for recognizing facial expressions using the ResNet-18 model is introduced. The study's main objective is to classify six basic emotions (happiness, sadness, surprise, anger, fear, and disgust) from static facial images. To accomplish this, the ResNet-18 model is employed to extract features from facial images, and a fully connected layer is trained to classify the emotions. The accuracy of the proposed method is evaluated using the publicly available Japanese Female Facial Expression (JAFFE) dataset, which contains 213 grayscale images. The proposed method achieved an accuracy of 87.56% on this dataset.

To test the proposed method's effectiveness, the authors compared it with other state-of-the-art techniques, such as Support Vector Machine (SVM) and Convolutional Neural Network (CNN). The proposed approach outperformed both SVM and CNN in terms of accuracy and computational efficiency. The authors also tested the proposed method on two additional datasets, the Cohn–Kanade (CK) dataset, containing 2105 grayscale images of 123 subjects, and the Extended Cohn–Kanade (CK+) dataset, with 593 grayscale images of 123 subjects. The proposed method achieved higher accuracies of 91.36 and 92.81% on the CK and CK+ datasets, respectively, than other state-of-the-art techniques.

The results demonstrate that the proposed method is highly effective in recognizing facial expressions, achieving high accuracy with relatively low computational cost. The study highlights the potential of using ResNet-18 models for facial expression recognition, which could have significant applications in various fields such as human–computer interaction, health care, and security. Nonetheless, the study's use of grayscale images and relatively small datasets imposes certain limitations. The authors suggest that future research should examine the use of larger and more diverse datasets to evaluate the effectiveness of the proposed method in recognizing facial expressions in real-world scenarios. The Fig. 1 depicts the ResNet 50 architecture.

Fig. 1 ResNet-18 model



3.2 ResNet-50

Li and Lima [2] demonstrate a method for detecting facial expressions using the ResNet-50 deep learning model. Recognizing facial expressions is a challenging task because of the intricate and subtle nature of human facial expressions. In the past, facial expression recognition relied on manually crafted features, which have limitations in capturing the nuances of facial expressions. However, deep learning models have shown potential in capturing complex patterns in data. The authors [2]

proposed a method that employs the ResNet-50 deep learning model, which is a cutting-edge convolutional neural network. The model is trained on a dataset of facial images tagged with the corresponding facial expressions. To evaluate their approach's performance, the authors used various metrics, such as accuracy, precision, recall, and F1-score. The Facial Expression Recognition and Analysis Challenge (FERA2017) dataset, which includes images of facial expressions in various positions, lighting conditions, and backgrounds, was used in the study.

The findings show that the ResNet-50 model achieves high accuracy in recognizing facial expressions, surpassing traditional machine learning approaches and other deep learning models. The authors also conducted experiments to examine the impact of various factors on the model's performance, including the size of the training dataset, the number of training epochs, and the learning rate. The results show that increasing the training dataset's size and the number of training epochs can enhance the model's performance while decreasing the learning rate can prevent overfitting. Overall, The method has potential applications in emotion recognition in human-computer interaction and psychology research. Additionally, the authors [2] provide insights into the factors that can affect the model's performance, which can be beneficial for researchers and practitioners in the facial expression recognition domain. Fig. 2 depicts the ResNet 50 architecture.

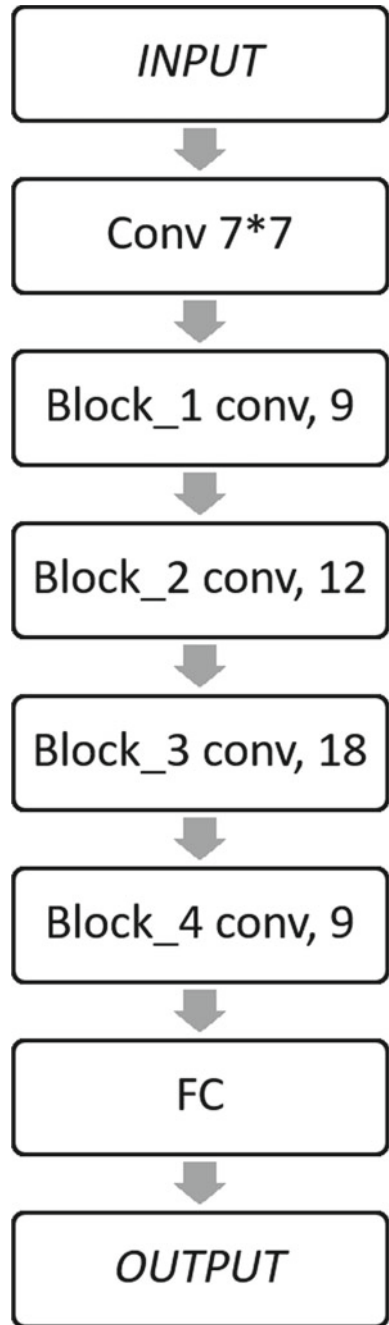
3.3 3D Inception-ResNet Layer Composition

The authors Hasani and Mahoor [3] proposed an approach for recognizing facial expressions using enhanced deep 3D convolutional neural networks (CNNs). The difficulty in recognizing facial expressions arises from the various facial expressions and poses. CNNs are promising in this area, but their performance is limited by the lack of depth perception in 2D images. To overcome this limitation, the authors suggest using 3D CNNs, which can extract spatiotemporal features from video data. The suggested method involves preprocessing the data by detecting and aligning the face and then extracting features through a pre-trained 3D CNN. The extracted features are then inputted into an enhanced deep 3D CNN, which consists of multiple layers of 3D convolutional, pooling, and fully connected layers. The enhanced network utilizes skip connections and residual learning to improve training and avoid overfitting. The authors evaluated the proposed method on three publicly available datasets, CK+, Oulu-CASIA, and AffectNet. The results show that the proposed method outperforms state-of-the-art methods on all three datasets, achieving accuracies of 98.32

3.4 Dynamic Geometric Image Networks with ResNet

The authors of the paper, Li et al. [4], have introduced a unique model for automatic 4D Facial Expression Recognition (FER) called the Dynamic Geometrical Image

Fig. 2 ResNet-50 model



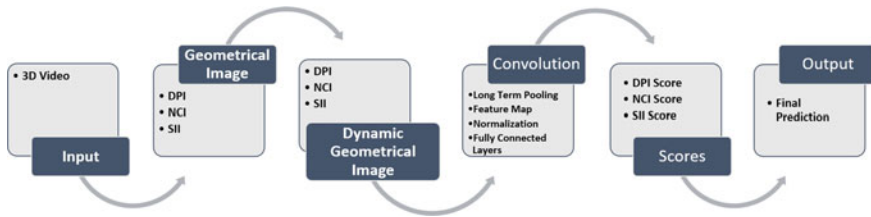


Fig. 3 Model of dynamic geometric image networks with ResNet

Network (DGIN). DGIN includes a short-term temporal pooling layer for generating dynamic geometric images, followed by multiple combinations of convolution, ReLU, and pooling layers that are beneficial for determining facial spatial features. Additionally, a long-term temporal pooling layer is used for dynamic feature map fusion, succeeded by a fully connected layer and a joint loss layer. The joint loss function includes cross-entropy loss and triplet loss. The 4D data accounts for short-term as well as long-term motion patterns, both of which are significant attributes for building an effective FER system.

As shown in Fig. 3, this model also engages with Geometrical Images and Dynamic Geometrical Images. Furthermore, there are significant applications of surface differential quantities such as normals and curvature in 3D face descriptors. Normal Component Image (NCI) is the first-order differential geometry quantity that contains much more information than original surface coordinates. Shape Index Images (SII) constitute the shape index, which is a homogenized value of two principal curvatures and is thus known as a second-order differential geometry quantity. The application of dynamic geometrical images is the recognition of video-based action, which epitomizes the dynamic and appearance of the whole video. The Depth Image (DPI), Normal Component Image (NCI), and Shape Index Image (SII) are the obligatory inputs for the DGIN model. With the help of the approximated rank pooling method, the overall computation of Dynamic Geometrical Images (DGI) is confirmed with DGIN. This architecture is now capable of extracting temporal deformities. The Two-Stage Sliding Window (TSSW) is also introduced since data augmentation is imperative. The sliding window is used for the fragmentation of input video into distinct frames or sub-videos. In TSSW, the first stage assimilates splitting the entire 3D video into segments with large window size and large stride. Through this stage, the number of samples is increased. In the subsequent second stage, it contains a small-sized window and smaller stride, which further produces multiple segments of short sub-videos. These segments are then served to the DGIN architecture.

ResNet was preferred for the CNN module because of its expertise in extracting spatial information. The dataset used for training and evaluation was BU-4DFE, which consists of six universal facial expressions, including surprise, sadness, disgust, anger, fear, and happiness. The dataset comprises 3D video sequences of 56 females and 44 males. The facial expressions were captured at a rate of 25 frames per second with a total time period of 3–4 s for each such sequence. The DGIN

model with ResNet as its CNN module exceeded other approaches with an accuracy of 86.67% in this BU-4DFE dataset.

3.5 *ResNet and Atrous Convolutions*

The facial micro-expressions constitute reflexive facial movements that divulge a human's hidden emotions in a high-stakes environment. These emotions appear for a very short interval of time, generally between 0.04 and 0.5 s. There are innumerable applications of such expressions, including psychotherapy, criminal investigations, customer interest discovery, lie detection, and communication. The recognition of such trivial expressions is only achievable through convolutional neural networks. A contemporary algorithm was composed by Lai et al. [5], which constitutes a distinct combination of kernels of atrous convolution and automatic facial correction that helps in the extraction of minute and fine features of micro-expressions. This novel algorithm is known as MACNN (Multi-scale convolution neural network model).

The convolution of atrous convolutions is assembled by enhancing the usual convolution kernel up to dilation-scale constraints and padding the unoccupied region of the original convolution with zeros. The atrous convolution is also known as dilated convolution. The dilation rate is appended to the convolution layer, which is actually used to introduce the interlude between the values during the processing of data. The MACNN model is associated with the VGG-16 network as the fundamental network framework. The overall network consists of 13 convolution layers and atrous convolution kernels. The residual blocks are eventually added to the network, which is dealing with the gradient disappearing problem, and hence these blocks simulate the effect of fast convergence in the model training process. In order to validate the newly introduced approach, CASME and CASMEII datasets were used. In the experiment, 70% of the dataset was used in training, and the rest 30% was used for testing purposes. CASMEII is the extension of CASME with additional image marking. These datasets constitute the samples of 22 males and 13 females, each one producing micro-expressions while watching a video, and hence the dataset is enriched with distinct and diverse micro-expressions.

There are a total of seven types of micro-expressions, including sad, surprise, fear, happy, disgust, neutral, and angry. The overall accuracy obtained with this methodology was much higher than the existing solutions. In the case of CASME, the accuracy was 70.16%, and in the CASMEII dataset, the accuracy reached 72.26%. In order to obtain a high accuracy score, some measures could be initiated in this approach. The traditional residual block could be optimized, and the deeper model may produce astounding results. Also, the feedback network enhancement will definitely strengthen the network.

3.6 *Conditional Random Fields and ResNet*

Traditional machine learning techniques such as Bayesian classifiers and Support Vector Machines (SVM) often provide insightful results but in a controlled and uncomplicated environment. They fail to consider any temporal links between successive frames in a video and hence cannot be deployed for the same. Deep Neural Networks (DNNs) have the exclusive ability to extract visual attributes even in videos and hence are widely accepted as frame-based techniques. The authors Hasani and Mahoor [6] initiated a coherent two-step learning approach for facial expression recognition. The first part constitutes residual networks, and in the second part, Conditional Random Fields (CRF) are deployed. This two-step model outperforms already established approaches since it is capable of extracting both spatial and temporal relations in sequences of the image.

The first section of the model is essentially a Dense Neural Network (DNN) composed of three combined Inception-v4 and ResNet layers succeeded by Average Pooling, Dropout, and two fully connected layers, respectively. The convolution layers are accompanied by batch normalization, and ReLU is used as an activation function. After the DNN portion, Conditional Random Fields (CRF) are initiated, which are responsible for extracting temporal relations of the input sequences, and hence this model provides extraordinary results in sequence labeling tasks. The CRF model was trained in a one-shot manner, whereas the DNN part was trained in multiple batches. The main advantage of using CRF over any other method, including Long Short-Term Memory (LSTM), is that the CRF entrusts the most probable and necessitous sequence of labels from the entire available sequence. The model was examined over three benchmark datasets. In the CK+ dataset, the obtained accuracy was 93.04%, in the MMI dataset, the accuracy is 78.68%, and in the case of the FERA dataset, the accuracy gained was 66.66%. These accuracies were superior to many state-of-the-art approaches. Figures 4 and 5 demonstrate the idea of conditional random fields and Resnet model.

3.7 *FER Using ResNet and Heart Rate Variability Observation*

The authors Singson et al. [7] presented a new method for recognizing emotions. This method utilizes short-term analysis of heart rate variability (HRV) and Residual Neural Network (ResNet) architecture, in order to address the limitations of traditional emotion recognition methods that rely on audio and visual cues, which may not accurately capture internal physiological states.

To evaluate this new method, 30 participants watched videos designed to elicit specific emotional responses, and their HRV data was collected, pre-processed, and segmented into short-term windows. Various HRV features were extracted from the segmented data, and a ResNet model was trained using these features to recognize six basic emotions: happiness, sadness, fear, anger, surprise, and disgust.

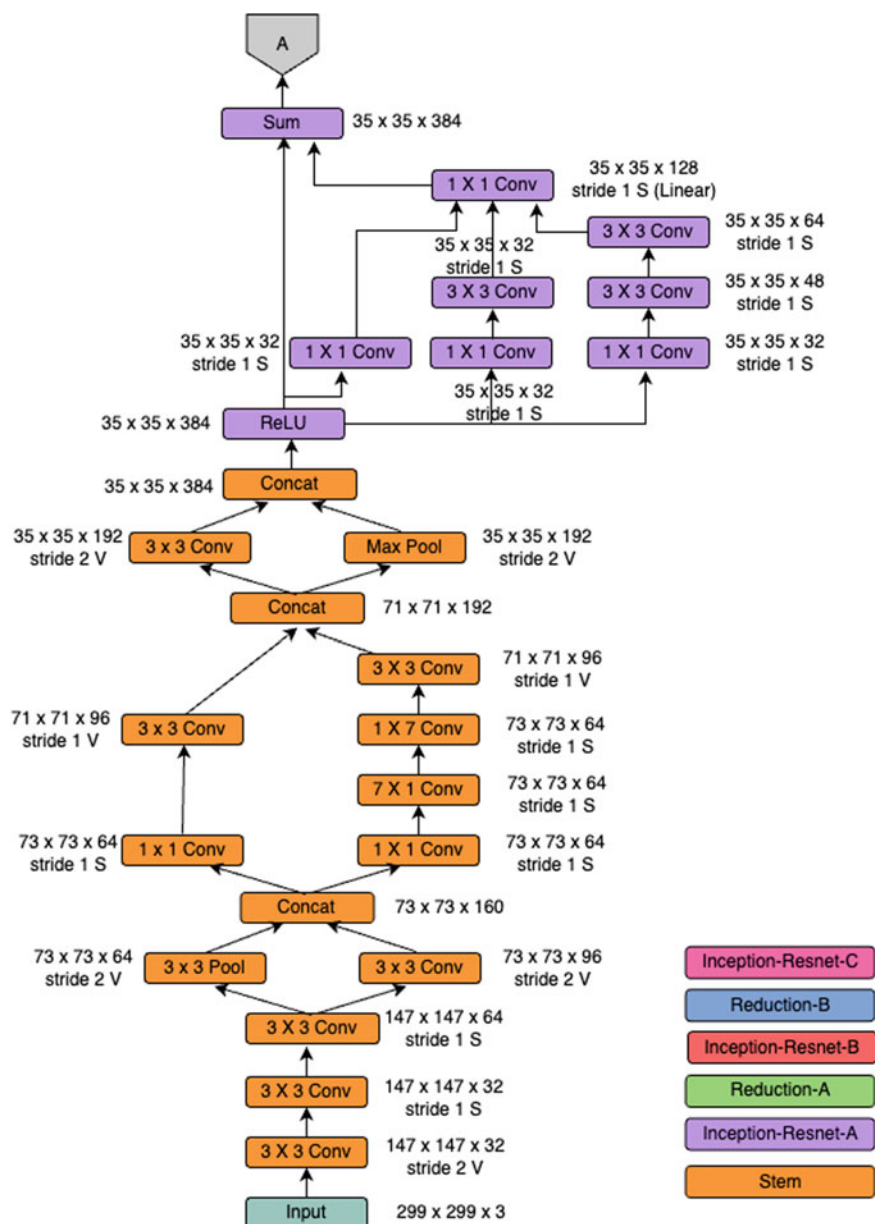


Fig. 4 Proposed conditional random fields and Resnet model part-1

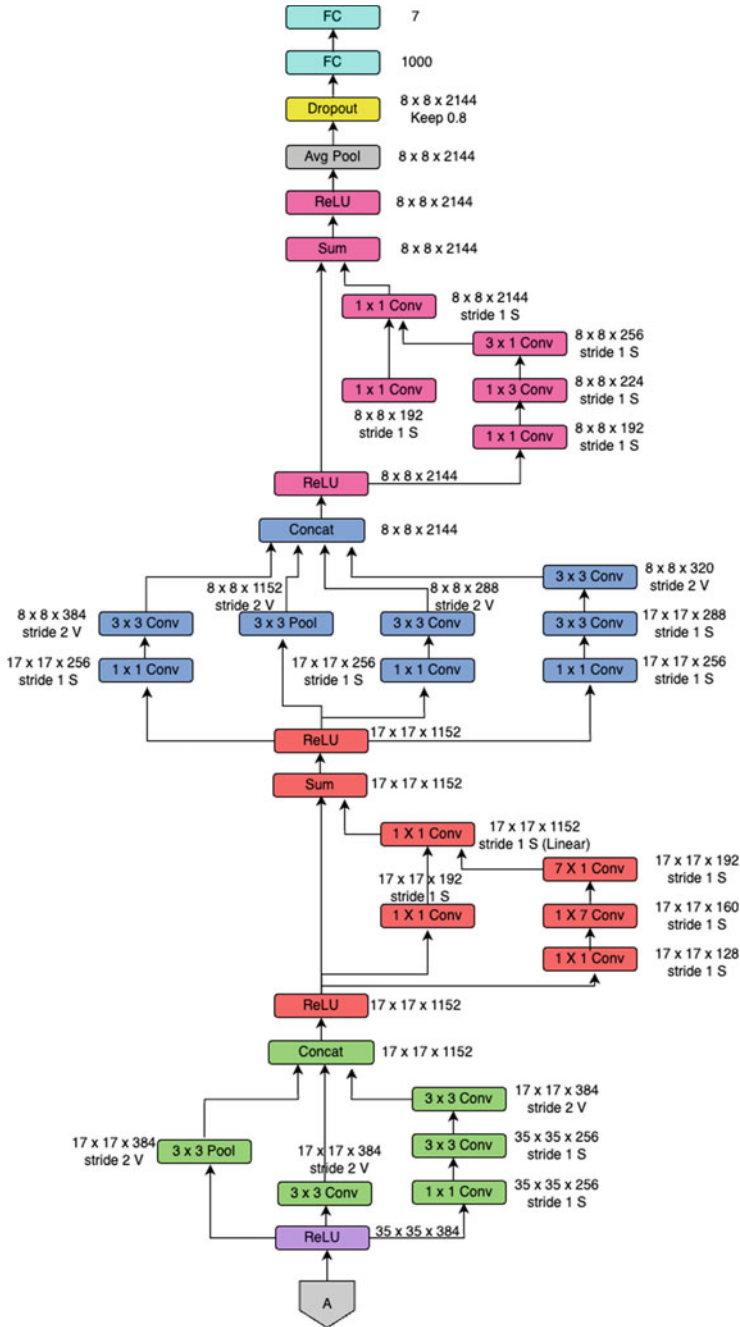


Fig. 5 Proposed conditional random fields and Resnet model part-2

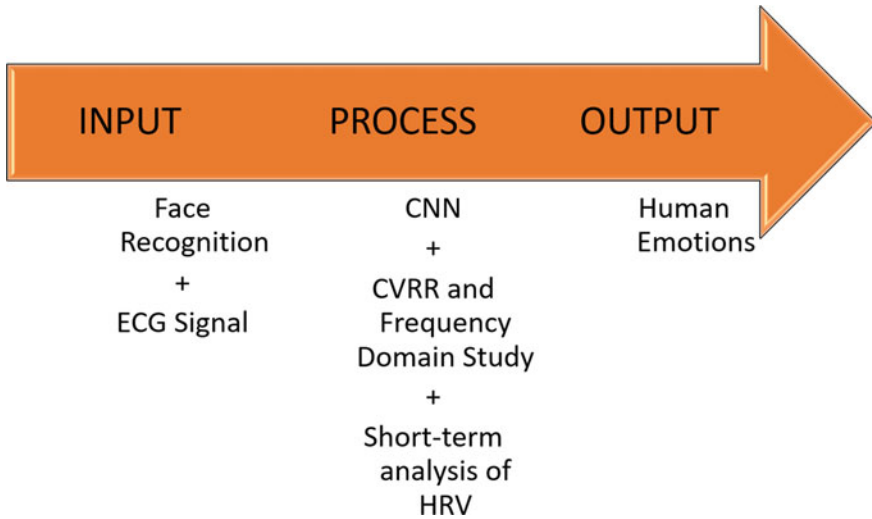


Fig. 6 Model of FER using ResNet and heart rate variability observation

The results of the study demonstrate that the proposed approach achieved high accuracy in recognizing emotions, with an average classification accuracy of 87.96%. This outperformed other machine learning methods, such as SVM, decision tree, and K-nearest neighbor. The study also revealed that certain HRV features, such as root mean square of successive differences (RMSSD) and low-frequency power (LF), were particularly useful in distinguishing between different emotions.

The proposed approach has several potential applications in psychology, psychiatry, and human–computer interaction. It could be used to develop emotion recognition systems that accurately capture internal physiological states and provide real-time feedback to users. Additionally, the approach could be used in clinical settings to monitor and track changes in emotional states in patients with mood disorders. Therefore, the paper [7] introduces a new approach to emotion recognition using short-term HRV analysis and ResNet architecture. The study demonstrates the effectiveness of the proposed approach in accurately recognizing different emotions and highlights its potential applications in various fields (Fig. 6).

3.8 *ResNet Escorted by Squeeze and Exception Networks*

Deep learning models provide exceptional results for FER, yet deploying such architectures for real-world applications and embedded systems is not common. The primary reason is the large size of such networks. Zhong et al. [8] have shown a network that provides good accuracy with fewer parameters. They introduced an elementary and effective composition containing ResNet and Squeeze and Excitation (SENet).

Squeeze and Excitation Networks (SENeTs) can be easily attached to any module of the network without increasing any of the computation load. To facilitate interdependence between feature channels, the SENet model consists of two key squeeze and excitation operations. The network extracts the meaningful attributes from each channel, and these SE blocks eventually utilize a global average pooling operation followed by two compact fully connected layers. This approach substantially decreases the number of parameters required and their calculations. It also provides more nonlinearity and stability to the network by efficiently maintaining complex correlations between channels. A simplified version of ResNet-18 (SResNet-18) was installed with the SENet block, and hence a new variant known as SE-SResNet-18 emerged with great capabilities, improved training speed, and accuracy.

The SE-SResNet-18 network was evaluated over two distinct datasets, named FER2013 and CK+ dataset. The inputs were resized to 44x44 pixels from the original size of 48×48 . The accuracy obtained in the FER2013 dataset was 74.143%, and the accuracy obtained over the CK+ dataset was 95.253%.

4 Comparative Inspection of Discussed ResNet Compositions

The chosen papers explore the use of deep learning techniques for facial expression recognition, with a focus on various ResNet models to improve classification accuracy. The following comparison examines similarities and differences in approaches, datasets, experimental setups, and results.

Zhou et al. [1] propose a ResNet-18 model for facial sentiment classification, achieving an accuracy of 95.06% on the CK+ dataset of 327 labeled images of seven facial expressions, outperforming other state-of-the-art methods. In contrast, Li and Lima use a ResNet-50 model for facial expression recognition on both CK+ and JAFFE datasets, reporting accuracy rates of 99.29 and 98.64%, respectively, and surpassing other methods. Although both papers use similar ResNet models, Li and Lima achieve higher accuracy on both datasets.

Hasani and Mahoor [3] presented an enhanced 3D CNN model for facial expression recognition, achieving an accuracy of 91.22% on the BU-3DFE dataset of 100 labeled videos of six facial expressions, surpassing other methods. They also compare their model with other 2D CNN models, showing that the 3D CNN model is more effective. In comparison, Li et al. [4] propose a dynamic geometrical image network for 4D facial expression recognition, reporting an accuracy of 90.7% on the BU-4DFE dataset of 101 labeled videos of six facial expressions, which is comparable to other methods. Both papers use different types of deep learning models and datasets, with Hasani and Mahoor achieving higher accuracy.

Lai et al. [4] introduced a real-time micro-expression recognition system based on ResNet and atrous convolutions, achieving an accuracy of 90.68% on the CASME II dataset of 247 labeled videos of three micro-expressions, outperforming other meth-

ods. They also compare their model with other deep learning models and demonstrate that the proposed model is more efficient. Hasani and Mahoor [6] propose a spatio-temporal facial expression recognition system based on ResNet and conditional random fields, achieving an accuracy of 59.9% on the AffectNet dataset of 1 million labeled images of seven facial expressions, which is comparable to other methods. Both papers use ResNet models but different datasets and approaches, with Lai et al. [4] achieving higher accuracy.

Overall, these papers illustrate the efficacy of deep learning models, particularly ResNet models, for facial expression recognition. However, the use of different datasets and experimental setups makes it challenging to compare results directly. Nonetheless, they all demonstrate that ResNet models can achieve high accuracy.

5 Conclusion

The use of ResNet models and 3D convolutional neural networks for facial expression recognition was explored in several papers, which demonstrated that ResNet models were effective in identifying emotions from both static and dynamic images with high accuracy. Furthermore, the incorporation of conditional random fields and enhanced deep 3D convolutional neural networks showed improvements in spatio-temporal and real-time micro-expression recognition. One study even proposed using short-term analysis of heart rate variability and ResNet architecture to recognize emotions, which has potential applications in healthcare, human–computer interaction, and security.

Despite these promising results, limitations exist, such as the need for large amounts of labeled data and the bias present in training data. To improve the models' robustness and generalization ability, researchers should focus on using more diverse and inclusive datasets. Moreover, further research should aim to enhance the accuracy of real-time micro-expression recognition.

Overall, the proposed approaches in the papers provide innovative solutions for facial expression recognition, which has the potential to transform various industries. Addressing the limitations and challenges in facial expression recognition will lead to more accurate and reliable results in the future.

References

1. Zhou Y, Ren F, Nishide S, Kang X (2019) Facial sentiment classification based on Resnet-18 model. *Int Conf Electron Eng Inform (EEI) 2019*:463–466. <https://doi.org/10.1109/EEI48997.2019.00106>
2. Li B, Lima D (2021) Facial expression recognition via ResNet-50. *Int J Cogn Comput Eng* 2:57–64. <https://doi.org/10.1016/j.ijcce.2021.02.002>

3. Hasani B, Mahoor MH (2017) Facial expression recognition using enhanced deep 3D convolutional neural networks. *IEEE Conf Comput Vis Pattern Recognit Work (CVPRW)* 2278–2288. <https://doi.org/10.1109/CVPRW.2017.282>
4. Li W, Huang D, Li H, Wang Y (2018) Automatic 4D facial expression recognition using dynamic geometrical image network. In: 2018 13th IEEE international conference on automatic face and gesture recognition (FG 2018), pp 24–30. <https://doi.org/10.1109/FG.2018.00014>
5. Lai Z, Chen R, Jia J et al (2020) Real-time micro-expression recognition based on ResNet and atrous convolutions. *J Ambient Intell Human Comput*. <https://doi.org/10.1007/s12652-020-01779-5>
6. Hasani B, Mahoor MH (2017) Spatio-temporal facial expression recognition using convolutional neural networks and conditional random fields. In: 2017 12th IEEE international conference on automatic face and gesture recognition (FG 2017), pp 790–795. <https://doi.org/10.1109/FG.2017.99>
7. Singson LNB, Sanchez MTUR, Villaverde JF (2021) Emotion recognition using short-term analysis of heart rate variability and ResNet architecture. In: 2021 13th international conference on computer and automation engineering (ICCAE), pp 15–18. <https://doi.org/10.1109/ICCAE51876.2021.9426094>
8. Zhong Y, Qiu S, Luo X, Meng Z, Liu J (2020) Facial expression recognition based on optimized ResNet. In: 2020 2nd world symposium on artificial intelligence (WSAI), pp 84–91. <https://doi.org/10.1109/WSAI49636.2020.9143287>
9. Wang H-H, Gu J-W (2018) The applications of facial expression recognition in human-computer interaction. *IEEE Int Conf Adv Manuf (ICAM)* 2018:288–291. <https://doi.org/10.1109/AMCON.2018.8614755>
10. Zhang Y-D et al (2016) Facial emotion recognition based on biorthogonal wavelet entropy, fuzzy support vector machine, and stratified cross validation. *IEEE Access* 4:8375–8385. <https://doi.org/10.1109/ACCESS.2016.2628407>
11. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: 2016 IEEE conference on computer vision and pattern recognition (CVPR). Las Vegas, NV, USA, pp 770–778. <https://doi.org/10.1109/CVPR.2016.90>
12. Antol et al S (2015) VQA: visual question answering. In: 2015 IEEE international conference on computer vision (ICCV). Santiago, Chile, pp 2425–2433. <https://doi.org/10.1109/ICCV.2015.279>
13. He K, Zhang X, Ren S, Sun J (2015) Deep residual learning for image recognition. <https://doi.org/10.48550/arXiv.1512.03385>
14. Szegedy C et al (2015) Going deeper with convolutions. In: 2015 IEEE conference on computer vision and pattern recognition (CVPR). Boston, MA, USA, pp 1-9. <https://doi.org/10.1109/CVPR.2015.7298594>

Block Farm: Blockchain-Based Platform for the Agriculture Supply Chain



Udai Bhan Trivedi, Manoj Srivastava, and Manish Kumar

1 Introduction

Blockchain is a peer-to-peer decentralized ledger that collects a growing number of transaction records from a hierarchically growing blockchain and uses cryptographic technology to secure each block to ensure transaction data integrity [1]. New blocks are added to the global blockchain only after the decentralized consensus system is completed. More specifically, a block stores a hash (the value of an entire block that can be considered a cryptographic image) and a hash value (the previous block that acts as a cryptographic relationship to the previous block in the blockchain). The network uses a secure blockchain and (III) a decentralized consensus method that monitors the acquisition of new blocks in a blockchain learning protocol to ensure the consistency of data records in each copy of the blockchain stored on each node. As a result, the blockchain ensures that transaction records cannot be modified or altered once a transaction record is added to a block and the block is successfully created and canceled on the blockchain. Blockchain also ensures the integrity of the data in each block of the chain, and the blocks created on the blockchain cannot be altered in any way. Blockchain acts as a secure, distributed ledger that records all transactions between two participants in an efficient, reliable, and verifiable manner in an open network system [2].

A block is added to the issuing node's blockchain containing the transaction. A block consists of a block header and data. The block header contains metadata about the block. Block data consists of a list of verified and legitimate transactions sent to

U. B. Trivedi (✉)

Pranveer Singh Institute of Technology, Kanpur, Uttar Pradesh, India

e-mail: udaibhantrivedi@gmail.com

M. Srivastava

Dr. Gaur Hari Singhanian Institute of Management and Research, Kanpur, Uttar Pradesh, India

M. Kumar

Galgotias College of Engineering and Technology (GCET), Greater Noida, Uttar Pradesh, India

Table 1 Component of block in blockchain

<ul style="list-style-type: none"> • Block header
(a) Block number: Indicate the version number of the block
(b) Hash of the previous block: Use SHA 256 (Algorithm)
(c) Hash of current block: Use SHA 256 (Algorithm)
(d) Timestamp: indicates when the block was created
(e) Difficulty target to adjust the difficulty of mining
(f) Nonce
<ul style="list-style-type: none"> • Block data
(a) Consist of all transactions and other data that may be present

the blockchain network. By cryptographically signing a transaction, the issuer of the digital asset verifies the correctness, legality, and validity of the transaction form. Ensure that the party presenting the digital asset for a transaction has access to the private key used to sign the existing digital asset. Each transaction in a published block is verified by additional full nodes to ensure correctness and validity. If the transaction is not valid, the lock is rejected. The following data fields are used by many blockchain implementations [3] (Table 1).

1.1 Chain of Blocks

Blocks are the building blocks of the blockchain. The header from the preceding block’s hash digest is included in each block. A new hash would be generated if a block that had already been published had been modified. As a consequence, since they incorporate the hash of the previous block, each successive block will have a unique hash. This makes it easier to spot corrupted data and reject it [4] (Fig. 1).

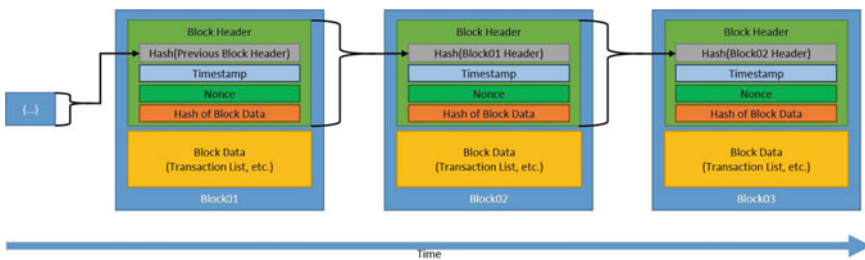


Fig. 1 Generic chain of blocks

1.2 Hash Chained Storage

The two main building blocks for building a blockchain using hash chain storage are the Hash index and the Merkel tree.

A hash pointer is a cryptographic hash of data that directs the user to where the data are stored. You can use a hash pointer to check if the data have changed. Hash tokens are used to combine blocks of data into a blockchain structure. Each block specifies a hash pointer to that block, specifying the address where the information from the previous block is stored. Users can publicly verify data hashes to prove that stored data have not been altered. If an attacker tries to change the data in each block of the entire chain to hide the damage, the attacker would have to change the hash of each previous block, which is almost impossible [5].

1.3 Digital Signature

A digital signature verifies data using cryptographic techniques. Also, it is a way to ensure that the data have not been changed. The digital signature system consists of three main parts. The first element of the key generation algorithm is the generation of two keys: one to sign messages and preserve their confidentiality. This key is called the private key. The other is available to the public. This key is called the public key, used to verify that the message was signed with the private key. The signature algorithm is the second important factor. Sender sign incoming messages using the provided private key. The validation algorithm is the third important element. It accepts three inputs: signature, message, and public key then uses a public key that verifies the signature of a message and returns a Boolean result [6].

1.4 Elliptic Curve Digital Signature Algorithm (ECDSA)

One of the most advanced public key encryption systems is Elliptic Curve Digital Signature Algorithm (ECDSA). Elliptic curve encryption keys are smaller than intermediate keys obtained using digital signature technology. Elliptic curve cryptography is a type of public key encryption based on the algebraic arrangement of an elliptic curve in a bounded field. Elliptic curve cryptography is often used to create fake numbers, digital signatures, and other similar data. An electronic digital signature is an authentication mechanism that uses a pair of public keys and a digital certificate as a signature to verify the identity of a recipient or sender of information [7].

1.5 Nonce

A random number that is used only once is called a cryptographic nonce. The data can be parsed into different values of the encrypted nonce generating multiple hash digests each time.

$$\text{Digest} = \text{hash}(\text{data} + \text{nonce})$$

There are ways to get different summary results by changing the unknown values while keeping the same data. The consensus model POW uses this method [5].

1.6 Consensus Protocol

The techniques through which all users within a distributed ledger concur on the accuracy of the underlying data are known as consensus protocols. The fact that all parties acknowledge a single “real” version of the data makes a distributed ledger one of its major features. An event known as a fork happens when current Blockchain participants choose to incorporate data in a way that is incompatible with established protocols.

The ledger splits as a result of forks, resulting in the formation of two groups, each of which validates its version of the ledger. Participants must stick to the same fork of the ledger to continue interacting with one another. As Block Farm is based on Ethereum PoS (Proof of Stake) consensus protocol has been used in the proposed system [8] (Table 2).

2 Literature Review

The various have been reviewed and considered to get an idea about the main stakeholders and their role in the traditional model. The various process where trust is the main deficit area. Try to find out the module where transparency and tracking are required. The literature review enabled us to understand Blockchain applications in agriculture, blockchain-based agricultural modeling, and the benefits and obstacles of blockchain deployment in the agricultural supply chain (Table 3).

Table 2 Consensus protocol

Consensus protocol	Overview
Proof of work	Validate new blocks of data using calculations. Participants in this plan must group transactions into a single block, apply a hash function along with some additional information, and then join the plan
Proof of stake	To participate in the verification process, validators (special nodes) must deposit collateral and cast votes on legitimate blocks. In contrast to Proof of Work, which focuses on proving the user has a significant amount of processing power, Proof of Stake is based on proving that users have invested in tokens of value in the network
Ripple protocol	The server confirms new transactions by combining open transactions into a “candidate list.” Then all participants vote on which legitimate transactions should be recorded in the ledger The final state of the closed ledger will include all transactions that have received at least 80% of “yes” votes
Proof of elapsed time	As part of the Intel Ledger Concept, Intel used the processor’s capability to provide the hardware with a cryptographically signed timestamp to create a validation lottery. The transaction that will enter the next block in the chain is determined by the transaction in the chain with the next timestamp. Compared to proof-of-work, this consensus mechanism consumes much less energy, so it is more suitable for IoT devices

3 Platforms for the Agriculture Supply Chain: Block Farm

Block Farm will help supply chain and market players by reducing inefficiencies. A dynamic and smooth software solution from Block Farm serves conventional, well-established agricultural supply networks. Block Farm will promote farmers and the companies that assist them to connect with customers to expand and strengthen their supply chains by enabling one-to-one global trade. Block Farm seeks to have a good influence on both societies at large and the world’s agricultural industry. Block Farm will offer a cutting-edge digital strategy and the necessary industry expertise to give farmers and the vital companies that support them more value. Block Farm can meet the demands of the sector as more farmers switch from bulk handling to residential storage. Farmers may sell their goods directly to local or foreign customers from their home storage systems by utilizing blockchain technology in a secure, low-risk setting [13] (Table 4).

Table 3 Literature review

Number	Year	Key findings of the paper specified in reference number
1	2021	Srivastava, R, etc. discussed emergent concerns for further blockchain research that have been found, with a focus on supply chain and agriculture management
2	2021	Yadav, V.S, etc. in his study identified and modeled the key drivers of long-term food security in India using a multi-criteria decision-making (MCDM) strategy
3	2021	Valoppi, F, etc. presented a description of cellular agriculture and other alternative methods of food production
4	2021	Zhu, L., etc. built a blockchain-based big data sharing solution after carefully analyzing the advantages of blockchain-based large data sharing
5	2021	Alkahtani, M., etc. in his research paper incorporated the blockchain impact into agricultural supply chain management using web design elements
6	2021	Liu, W., etc. presented the report which examines the body of research on blockchain technology and ICTs in agriculture from 2011 to 2020
7	2021	Tan, H.Y., etc. examined rural green credit utilizing a hierarchical blockchain model
8	2021	Yang, X., etc. presented popular methods for developing smart agriculture that covered: precision agriculture, convenience agriculture, and disciplined agriculture
9	2021	Rijanto, A.in his paper looked at how businesses finance themselves and how blockchain is used in agriculture
10	2021	VanWassenaer, L., etc. used a reference framework to identify key requirements for fresh agrifood use cases and better comprehend various blockchain applications
11	2021	Awan, S.H., etc. proposed a hybrid smart model that incorporates both Internet of Things (IoT) and blockchain capabilities, with a novel approach for transforming conventional agriculture into smart agriculture
12	2021	Vangala, A., etc. define the key needs for smart agriculture have been defined, and generic blockchain-based security architecture has been proposed
13	2020	Kamble, S.S., etc. This study identifies drivers for the implementation of blockchain technology in agricultural supply chains to increase the traceability
14	2020	Chen, Y.Y., etc. investigated agriculture's democratization, autocratization, centralization, and decentralization in terms of their ramifications and logical connections with other institutional and technological concepts
15	2020	Sharma, R., etc. research has shown the particular importance of how some machine learning application is used in agricultural supply chains (ASC) and accelerate ASC stability and performance
16	2020	Osmanoglu, M., etc. in their study recommended a blockchain-based approach to enhance agricultural product yield estimates
17	2020	Zhang, X.H., etc. in his study explored the opportunities and challenges of ensuring agricultural data quality through edge computing with other blockchain-based technologies
18	2020	Yadav, V.S, etc. Obstacles to the blockchain were identified, and an integrated model of those barriers' interrelationships and the strength of their relationships was introduced with a special focus on the Indian agriculture supply chain's adaption

(continued)

Table 3 (continued)

Number	Year	Key findings of the paper specified in reference number
19	2020	The Paper by Ferrag, M.A., etc. discussed the privacy issues and security challenges of IoT-based agriculture and outlined several risks affecting the sector, such as privacy attacks, authentication, confidentiality, availability, and integrity aspects
20	2020	Iqbal, R., etc. discussed the effectiveness of precision agriculture by utilizing blockchain and IoT systems in safe agricultural applications, such as tracking animal attacks, primarily through a Repelling and Notifying System
21	2020	Xiong, H., etc. presented an overview of blockchain applications in the food supply chain, agricultural insurance, smart agriculture, and agricultural transactions, including a discussion of data security issues in small farms

Table 4 Stakeholders of block farm

1. Farmers	2. Manufactures requiring agricultural products (dairies, flour mills, breweries, etc.)
3. Farmer representatives (brokers)	4. Smaller farming cooperatives
5. Trucking and logistics companies	6. Suppliers of agricultural inputs (chemicals, fertilizers, etc.)
7. Logistics brokers	8. Importers and exporters of bulk commodities
9. Buyers of agricultural products	10. Customers and end users

3.1 Operational Efficiency

3.1.1 Tracking and Automation (Supply Chain)

Combining user-friendly mobile applications for agricultural operations and transportation with powerful online business management, Block Farm provides an end-to-end view of the agricultural supply chain. Block Farm provides data to farmers, brokers, and logistics companies to automate the farm-to-consumer transportation process. Data are collected at each stage of the SCM and each shipment is time-stamped upon receipt and delivery. Real-time updates are sent to all parties after each transaction is completed. Block Farm will improve productivity and visibility, simplifies inventory, automates inventory orders, and removes tedious paperwork from the system (Fig. 2; Table 5).

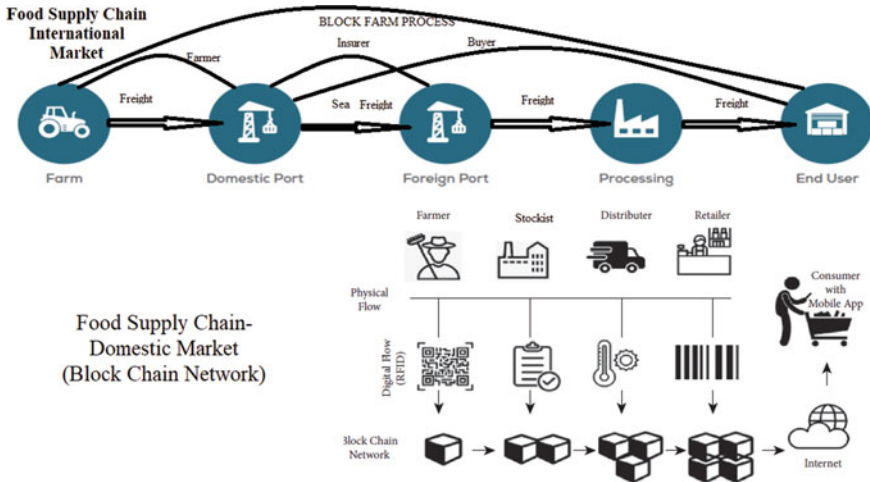


Fig. 2 Blockchain-based (Block Farm) process of the agriculture supply chain

Table 5 Advantages of block farm

FARM-domestic port/stockist	Domestic port-processing/distributer	Processing/retailer-end-user/customer
One computer program	Supply-chain management	Complete stock tracing
Exclusive origin stock	Decreased dependence on banks	Evidence of origin
An increase in stock visibility	De-risked business deals	More precise product recalls
Decreased capital costs	There are no currency concerns	Certainty of supply
Less handling	Exact insurance rates	
Reduced storage costs	Fewer manual documents	
E-commerce cash boards	Decreased administrative costs	
Under the control of the farmer		

3.1.2 Information and Data Transparency

Users will have more options with Block Farm to make a fact-based decision based on Market information and data. By viewing inventory, consumers can see exactly how many items they have, where they are going, when they will arrive and how much they cost. Users can be sure that improving profits is important at every level of the supply chain.

Buyers and end users will benefit from increased data quality and completeness as they can monitor and track the entire supply chain using a single system across all industries [8].

3.1.3 Smart Contract

Block Farm will offer simple contract options for farmers, brokers, buyers, and logistics companies. Block Farm will provide the largest automated software for creating products and shipping contracts. The blockchain system allows buyers, sellers, and goods to communicate in a blockchain system, ensuring the security of both sides of the transaction. Using blockchain technology, Block Farm will create the first global marketplace for authentic agricultural products [9, 10].

3.1.4 Record Keeping/Proof of Origin

By creating an immutable blockchain record of the journey from paddock to plate, proof of origin can be provided as goods move through the supply chain. At each level, the major supply chain actors are listed along with the supply chain nodes that each of these organizations owns and controls (e.g., fields, silos, vehicles, and delivery locations). Additionally, Block Farm tracks key product information including weights, species, grades, specifications, and inputs. Blockchain enables a fully transparent and traceable supply chain, giving all consumers a clear view of the goods, they buy [11].

3.2 Financial Benefits

3.2.1 Reducing the Cost of Capital

Block Farm users will pay FARMCOIN to create a contract to buy and sell agricultural products. Any subsequent changes to the business contract to add new data to the public blockchain will be charged as a FARMCOIN transaction fee. FARMCOIN payment is required to complete the transaction of goods between seller and buyer if the contract is suspended, closed, or suspended. On settlement day, both seller and buyer can trade using FARMCOIN tokens to reduce complexity and risk. This is especially important for international commodity contracts because it eliminates many burdensome transaction fees, including many foreign exchange contracts, bank guarantees, and letters of credit. Blockchain allows sellers and buyers around the world to use a single currency, eliminating the hassle and dependence on financial intermediaries.

Block Farm aims to make credit more affordable and accessible to farmers by providing greater visibility into farm inventory and reducing lender risk. With Block Farm, banks may view the farmer's current "position" and history of custody, both of which are recorded and certified by the blockchain. In addition, investors consider any guarantee of payments made by the farmer (future sales supported by an acceptable contract). Financial institutions can use Block Farm's data to reduce the risk of their financing by gaining access to previously unobtainable information about the farmers' real circumstances [14, 15].

3.2.2 Improving Cash Flow

Most small and medium-sized businesses fail due to poor cash flow, and agricultural supply companies are no different. Participants in the agricultural SCM suffer from low liquidity in their daily operations. Companies try to reconcile accounts receivable and accounts payable to maintain a stable and predictable cash flow. Unfortunately, as agricultural products move through the supply chain, the number of delinquent loans increases. Blockchain allows companies to raise capital and improve liquidity for supply chain participants using future contracts with existing customers [10, 12].

3.2.3 Insurance

Authorized insurance companies will have unprecedented access to farm inventory, storage, and transportation through Block Farm. In addition, new Internet of Things (IoT) solutions for heavy agricultural equipment and storage offer more opportunities to optimize the data stored in the Block Farm. With access to up-to-date, accurate, and irreversible data, insurance companies can more effectively assess claims, provide customized solutions for customers, pay claims, manage risk, detect suspicious activity, and improve fraud assessment. On the other hand, farmers, consumers, and logistics companies will have access to a wider range of insurance solutions with more accurate and affordable risk premiums [13].

3.2.4 Peer-To-Peer Lending

A farm of any size can connect to Block Farm in a secure (P2P) lending environment that allows companies with excess cash to provide short-term loans to those who need access to short-term financing. With Block Farm's platform, like-minded farmers can get better credit offers than conventional overdraft loans and reduce the cost of agricultural financial products.

4 FARMCOIN Tokens

FARMCOIN coins will use by Block Farm on the Ethereum public blockchain. The FARMCOIN token becomes a trading instrument token to pay for access to the Block Farm system through a payment gateway. Block Farm generates FARMCOIN tokens that will be profitably sold on the open market as revenue. In addition, a portion of FARMCOIN token sales will be transferred to Block Farm to expand the commodity pool. The price of the Block Farm system (such as fees and rewards) will be determined using the value of fiat currency to reduce volatility in crypto markets. However, users are charged with FARMCOIN tokens.

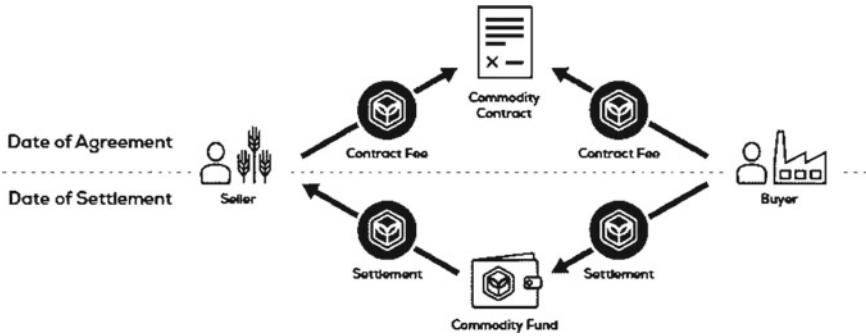


Fig. 3 Commodity contract between seller and buyer

4.1 Commodity Contract

Block Farm users will pay FARMCOIN to create a contract to buy and sell agricultural products. Any subsequent changes to the business contract to add new data to the public blockchain will be charged as a FARMCOIN transaction fee. FARMCOIN payment is required to complete the transaction of goods between seller and buyer if the contract is suspended, closed, or suspended. On settlement day, both seller and buyer can trade using FARMCOIN tokens to reduce complexity and risk. This is especially important for international commodity contracts because it eliminates many burdensome transaction fees, including many letters of credit, bank guarantees, and foreign exchange contracts. FARMCOIN will allow sellers and buyers around the world to use a single currency, which will eliminate the hassle and dependence on financial intermediaries (Fig. 3).

5 Technological Architecture of Block Farm

In the Block Farm Model of Blockchain, the application layer consists of various modules which can access directly through Public Block Farm Layer or indirectly through the Private Block Farm Layer [8] (Fig. 4).

5.1 Public Block Farm Layer

Block Farm uses the Ethereum blockchain (main contract, public blockchain data record, FARMCOIN token trading, and network rights) to create smart FARMCOIN tokens that can be traded on third-party token exchanges. Commodity smart contracts, inventory smart contracts, raw data sources, and FARMCOIN transactions are all available through the public blockchain layer. The public Ethereum blockchain and

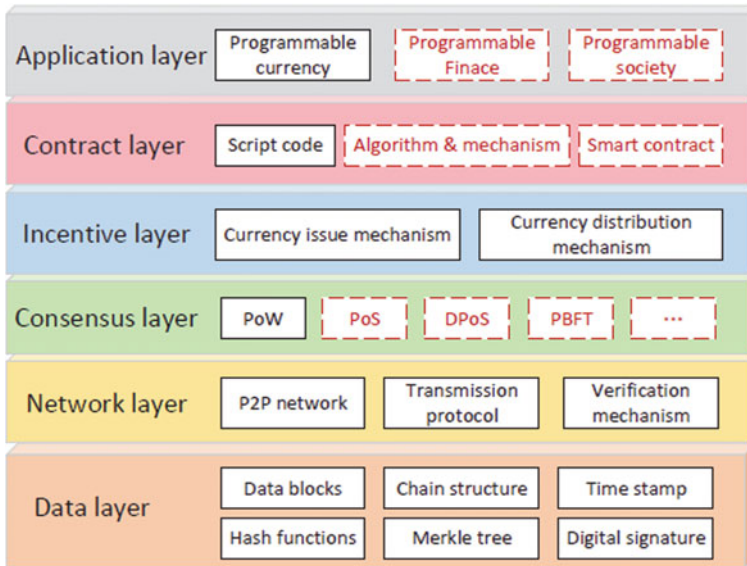


Fig. 4 Blockchain architecture [8]

third-party exchanges operate independently of Block Farm, providing secure and decentralized smart contracts and FARMCOIN tokens.

5.2 Private Block Farm Layer

Because agriculture and supply chain transactions occur frequently and in large numbers, a private blockchain layer can be used to better manage data volumes and reduce transaction costs and delays associated with public blockchains. The private layer of the Ethereum blockchain (main network) communicates with the public blockchain. Block Farm’s private blockchain will be used to store private data on commodity contracts, shipping contracts, supply chains, warehouse management, agriculture, and more. Block Farm uses this technique to achieve an optimal match to the underlying application, expected traffic, and load.

5.3 Payment Platform

Block Farm payment platform accepts direct payments for access to applications. Payment is accepted in the form of FARMCOIN tokens and fiat money (credit card) through payment gateways. As many consumers in Block Farm’s target market are

unfamiliar with blockchain technology and tokens, credit card payments will be allowed as a temporary solution, reducing barriers to adoption. Credit card payments can be used to purchase FARMCOIN tokens on public exchanges, which are then processed in the same way as live FARMCOIN payments. As Block Farm’s private blockchain evolves, the credit card payment gateway will be removed and FARMCOIN tokens will become the exclusive payment method for accessing the system (Fig. 5; Tables 6 and 7).

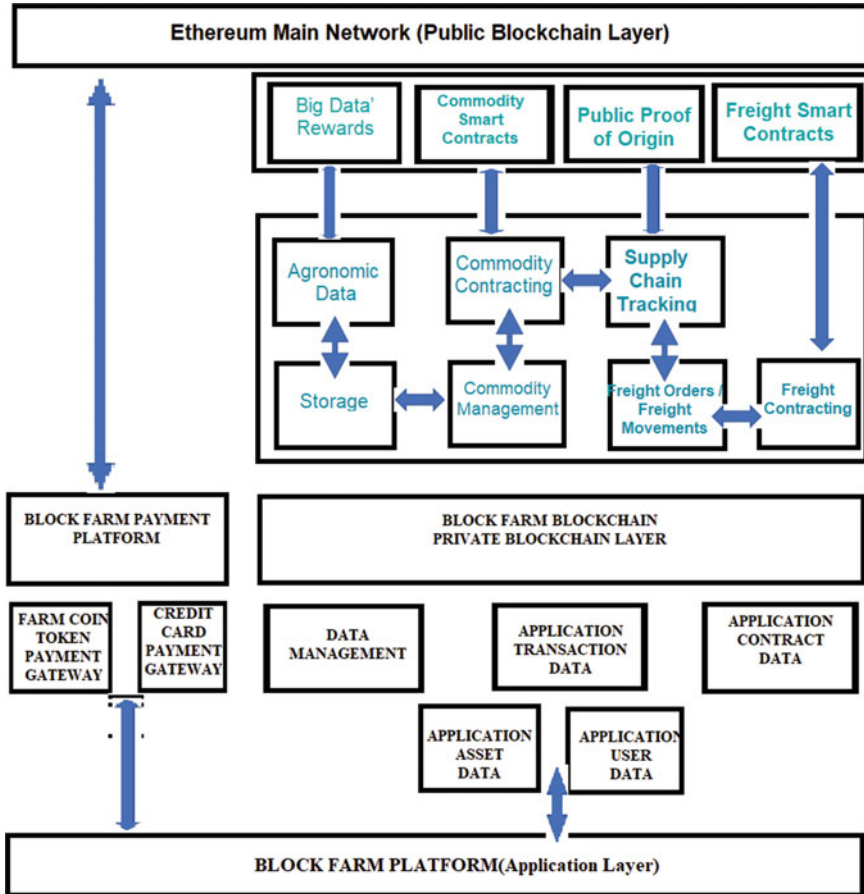


Fig. 5 Suggestive architecture of block farm (Application Layer)

Table 6 Proposed modules in application layers/users of block farm and platform

Module in application layer		Users	Platforms
Commodity contracts	Freight tendering	Growers (Farmers)	Web browser (PC/Mac)
Freight orders	Insurances	Brokers	Android (Mobile)
Freight movements	Online marketplace	Logistics	iOS (Mobile)
Stock management	Invoicing	Buyers	
Reporting	Production estimates	Receival sites	
Admin/Config	Cash boards		

Table 7 Module of application layer and its functionality

Module name	Functionality
Block Farm marketplace	Block Farm receives a commission when Resource Resellers join using the marketplace in exchange for making the link. FARMCOIN tokens are used to pay market commissions
Freight contract	Similarly, to freight contracts, freight orders and movement contracts will be created and managed where the parties in the supply chain pay in FARMCOIN to send the data to the public blockchain. Contracts for the supply of goods can be concluded between the seller or buyer of goods and an intermediary or supplier of goods. FARMCOIN will also receive compensation for the award of specific transport contracts to subcontractors for transport orders managed by a transport intermediary
Freight tender	Merchants in need of transportation services can register their needs using Block Farm’s freight bidding module. This feature promotes competitive pricing by enabling potential suppliers to offer the best prices to win the contract. FARMCOIN tokens are used to pay Block Farm commissions for connecting vendors and logistics providers
Supply chain tracking	Each point of the transaction must be recorded in the supply chain to track the product and provide a complete record of the journey from paddock to plate. Parental authority and ownership may change multiple times between farms and end consumers as products move from one member of the supply chain to another. Each stage of loading, noise, or change of owner is paid by the user and paid in FARMCOIN
Commodity management	Goods are frequently inspected as they are assembled and moved through the supply chain to confirm that they meet the specified quality, grade, and weight. Product data are entered into the blockchain at each checkpoint to create an end-to-end record and provide a complete view. In addition, as goods move through the supply chain, any treatment or use of gas is tracked. The cost of collecting product data is charged in FARMCOIN tokens

(continued)

Table 7 (continued)

Module name	Functionality
Share-farming contracts	Collective contracts provide an opportunity for farmers to work together and increase the profitability of their farms. Cooperative farming contracts give landowners and farmers the opportunity to grow their businesses by bringing together farmers who want to grow. Collaborative farm contracts are promoted by Block Farm, which also stores this data on Block Farm’s private blockchain. Agricultural contract structure, conversion, and termination fees are valued in FARMCOIN tokens
FARMCOIN as a reward token	Agronomic and agriculture data are collected on field activities and conditions during the growing season. These details include soil analysis, production forecast, fertilizer and nutrient content, chemicals, pesticides, and more. This data has a variety of uses, as it not only helps identify trends and patterns in agricultural “big data” but also influences consumer purchasing decisions. Agricultural data are collected and published on Block Farm’s private blockchain to ensure full transparency of each product. Block Farm plans to reward users who process data with FARMCOIN tokens generated directly from data revenue to encourage additional contributions

6 Conclusion

All the players in the agricultural supply chain, including farmers, transport businesses, intermediaries, traders, buyers, and consumers will be able to do so fast and risk-free transactions with the Block Farm platform. Block Farm connects stakeholders of the agricultural supply chain as a global platform, promoting transparency and traceability from the paddock to the fork. Agricultural supply chain internal and external complementary platforms can be customized by the Block Farm ecosystem, which is secure, scalable, and open. Customers can use Block Farm as: stand-alone solution or add-on to current software. A FARMCOIN token works as digital money that facilitates communication between supply chain participants and the Block Farm platform. FARMCOIN will evolve as multifunctional utility coin that trades on exchanges and serves as the “fuel” of the Block Farm system. This paper tries to outline the conceptual model of the agriculture supply chain system for the local and international market and help all stakeholders in their business processes.

References

1. Narayanan A, Bonneau J, Felten E, Miller A, Goldfeder S (2016) Bitcoin and cryptocurrency technologies: a comprehensive introduction
2. Zheng Z, Xie S, Dai H-N, Chen X, Wang H (2017) An overview of blockchain technology: architecture, consensus, and future trends, June. <https://doi.org/10.1109/BigDataCongress.2017.85>
3. Nakamoto S (2009) Bitcoin: a peer-to-peer electronic cash system. In: Cryptography mailing list, March. <https://metzdowd.com>

4. UK Government Office for Science (2016) Distributed ledger technology: beyond blockchain
5. Trivedi UB, Sharma S (2023) Digitally signed document chain (DSDC) blockchain. In: Singh PK, Wierzchoń ST, Tanwar S, Rodrigues JJPC, Ganzha M (eds) Proceedings of third international conference on computing, communications, and cyber-security. Lecture Notes in Networks and Systems, vol 421. Springer, Singapore
6. Srivastava R, Zhang JZ, Eachempati P (2021) Blockchain technology and its applications in agriculture and supply chain management: a retrospective overview and analysis. *Enterp Inf Syst* 94
7. Steem: An incentivized, blockchain-based, public content platform (2017)
8. Zhang R, Xue R, Liu L (2019) Security and privacy on blockchain. *ACM Comput Surv (CSUR)*
9. Mavilia R, Pisani R (2021) Blockchain for agricultural sector: the case of South Africa. *Afr J Sci Technol Innov Dev*
10. Liu ZL, Wei H, Wang DB (2021) Functional agricultural monitoring data storage based on sustainable block chain technology. *J Clean Prod*
11. Friha O, Ferrag MA, Shu L, Maglaras L, Wang XC (2021) Internet of Things for the future of smart agriculture: a comprehensive survey of emerging technologies. *IEEE-CAA J Autom Sin*
12. Torky M, Hassanein AE (2020) Integrating blockchain and the internet of things in precision agriculture: analysis, opportunities, and challenges. *Comput Electron Agric*
13. Salah K, Nizamuddin N, Jayaraman R, Omar M (2019) Blockchain-based soybean traceability in agricultural supply chain. *IEEE Access*
14. Kamilaris A, Fonts A, Prenafeta-Boldu FX (2019) The rise of blockchain technology in agriculture and food supply chains. *Trends Food Sci Technol*
15. Lin YP, Petway JR, Anthony J, Mukhtar H, Liao SW, Chou CF, Ho YF (2017) Blockchain: the evolutionary next step for ICT E-Agriculture. *Environments*

A Lattice-Based Key Exchange Protocol Over NTRU-NIP



Sonika Singh and Sahadeo Padhye

1 Introduction

In the world of rapid developments in communication technology and widely distributed systems, the security of our communicated the information relies heavily on the chosen key exchange (KE) protocol. It is common practise to exchange keys securely using the key exchange protocol created by Diffie and Hellman in 1976 [3]. The security of this scheme relies on DLP, a computationally hard mathematical problem, discrete logarithm problem. Computationally hard means no polynomial-time algorithm exists which can solve this problem. If anyone can solve this hard problem, then he can break this protocol. Now, there is an algorithm, developed by Shor in 1994 [14, 15], which can solve DLP and factoring problem in polynomial time. But, this algorithm will work only with quantum computers. Quantum computers theoretically exist, but not in practice. So, in the future, if someone can build a quantum computer that can be implemented in realistic situations, then it will be a very great threat to DLP & factoring-based cryptosystem and so for Diffie-Hellman key exchange protocol. We need a hard mathematical problem that is resistant to quantum effects as well as a protocol that can securely exchange keys based on this hard problem in order to counter this threat. Hoeffstein, Pipher, and Silverman developed a cryptosystem (NTRU cryptosystem) in 1998 [5] by using lattice's hard problems. Since hard instances of lattices are still considered to be quantum-resistant [13, 16], Lie et al. presented the NTRU-KE key exchange protocol, which is based on the NTRU cryptosystem [5] in 2013 [9]. Valluri conducted an analysis of the NTRU-KE protocol in 2018 using a man-in-the-middle attack [18]. With trun-

S. Singh (✉)

Department of Mathematics, Chaudhary Mahadeo Prasad Degree College, University of Allahabad, Prayagraj 211002, Uttar Pradesh, India
e-mail: sonikasinghcool1@gmail.com

S. Padhye

Department of Mathematics, Motilal Nehru National Institute of Technology Allahabad, Prayagraj 211004, Uttar Pradesh, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2024
B. K. Roy et al. (eds.), *Cryptology and Network Security with Machine Learning*, Algorithms for Intelligent Systems, https://doi.org/10.1007/978-981-99-2229-1_27

325

cated polynomials, the NTRU encryption and key exchange protocol functions. In its key generation process, we choose polynomials from $R = \frac{Z[x]}{(x^N - 1)}$ having their inverses under modulo two different integers p and q , respectively. If the inverses of chosen polynomial do not exist, then we discard it, and we choose a new one that has inverses. So, the problem is whether there exist enough polynomials from a given space that have their inverses. In NTRU encryption and NTRU-KE, there do not exist enough invertible polynomials for terminating the key generation algorithm successfully. Thus, our motivation in this article is to remove this issue by proposing a key exchange protocol that will assure us that we can find enough invertible polynomials, and hence key generation algorithm will generate keys successfully. The NTRU-NIP (NTRU with non-invertible polynomials) cryptosystem [1] serves as the foundation for the proposed key exchange protocol and we call it NTRU-NIP-KE. We refer to the article [1] for more information on the NTRU-NIP cryptosystem due to page limits. Recently a key exchange protocol that is quantum-resistant, MaTRU-KE [17] was introduced using the MaTRU cryptosystem. MaTRU-KE has a substantially greater computation cost than NTRU-NIP-KE because it uses the ring \mathcal{M} of $k \times k$ matrices of elements from ring $R = \frac{Z[x]}{x^N - 1}$ as its base ring. Additionally, NTRU-KE and MaTRU-KE are both susceptible to a lattice attack via SVP, whereas NTRU-NIP-KE is not. In addition to this, the parameter selection is more flexible in NTRU-NIP-KE in comparison to NTRU-KE and MaTRU-KE.

The article is structured as follows: In Sect. 2, we present a hardness assumption on which the security of the proposed protocol is based, and in Sect. 3, we present our key exchange protocol. The security evaluation of the given protocol is covered further in Sect. 4. In Sect. 5, a comparison of the proposed framework with NTRU-KE and MaTRU-KE is provided. The article is concluded at Sect. 6.

2 Parameter Sets and Hardness Assumption

Parameter Sets

The parameters' sets for the proposed protocol are the same as the NTRU-NIP scheme. We have three subsets $(L_f, L_g, L_\phi) \subset R = \frac{Z[x]}{x^N - 1}$ and (N, p, q) in parameter sets, where N, p, q are integers with $\gcd(N, q) = 1$ and p and q should be prime numbers. The sets L_f and L_g are used for choosing private keys and set L_ϕ is used to choose random polynomials. Assume d_f, d_g and d_ϕ all are positive integers. Here,

$$L(d_1, d_2) = \{f \in R : d_1 \text{ coefficients of } f \text{ are } 1, d_2 \text{ coefficients of } f \text{ are } -1 \text{ and rest coefficients are } 0\}$$

Then, $L_f = L(d_f + 1, d_f)$, $L_g = L(d_g, d_g)$, and $L_\phi = L(d_\phi, d_\phi)$. The symbol \odot denotes the multiplication of ring elements. We define the term 'width' of a polynomial: width = max coeff. – min. coeff.

Any public key cryptosystem’s security is based on some hard computational assumptions. The security of the NTRU-NIP encryption scheme is based upon the NTRU-NIP assumption, as described below.

Definition 1 The NTRU-NIP Inversion Problem: Determine m , for given security parameter k specifying (N, p, q) and spaces L_f, L_g, L_m and L_ϕ , a random public key pair (h, H) and $e = p\phi \odot h + H \odot m \pmod q$ as ciphertext, where $m \in L_m$ and $\phi \in L_\phi$.

Definition 2 The NTRU-NIP assumption: The NTRU-NIP inversion problem is computationally difficult for appropriate parameters.

3 Proposed Protocol—NTRU-NIP-KE

Here, the NTRU-NIP-KE protocol is proposed that works on the NTRU cryptosystem with non-invertible polynomials [1]. In Lie et al.’s key exchange protocol, there are three data flows. Similar to that, the proposed approach provides three data flows. All of the notations used in this article are from NTRU-NIP. For example, Alice (user A) and Bob (user B) need to exchange a secret key. The proposed protocol has the following steps:

Step-1

First of all, user A initiates with key generation. He generates his public and private keys in the following way. First of all, he selects $f_A \in L_f, g_A \in L_g$ and $G_A \in R$ in a manner such that G_A has its inverse under modulo q . Let, inverse of G_A under modulo q is G_{q_A} . That gives

$$G_{q_A} \odot G_A = 1 \pmod q$$

Now, she computes

$$\begin{aligned} h_A &= G_{q_A} \odot g_A \pmod q \text{ and} \\ H_A &= G_{q_A} \odot f_A \pmod q. \end{aligned}$$

The pair $((h_A, H_A), (f_A, g_A, G_A))$ is public and private key of user A. Then, A sends (h_A, H_A) to B.

Step-2

In same manner as A, B obtains his key pair. He chooses $f_B \in L_f, g_B \in L_g$ and $G_B \in R$ such that G_B has its inverse under modulo q . Let G_{q_B} is inverse of G_B under modulo q . That gives:

Table 1 NTRU-NIP-KE protocol

User A	User B
<p style="text-align: center;">Step 1:</p> $f_A \xleftarrow{r} L_f, g_A \xleftarrow{r} L_g$ $G_A \xleftarrow{r} R$ $h_A = G_{q_A} \odot g_A \text{ mod } q$ $H_A = G_{q_A} \odot f_A \text{ mod } q$ <div style="text-align: right; margin-top: 10px;"> $\xrightarrow{(h_A, H_A)}$ </div>	<p style="text-align: center;">Step 2:</p> $f_B \xleftarrow{r} L_f, g_B \xleftarrow{r} L_g$ $G_B \xleftarrow{r} R, \phi_B \xleftarrow{r} L_\phi$ $h_B = G_{q_B} \odot g_B \text{ mod } q$ $H_B = G_{q_B} \odot f_B \text{ mod } q$ $e_B = p\phi_B \odot h_A + H_A \odot f_B \text{ mod } q$ <div style="text-align: left; margin-top: 10px;"> $\xleftarrow{(h_B, H_B, e_B)}$ </div>
<p style="text-align: center;">Step 3:</p> $\phi_A \xleftarrow{r} L_\phi$ $e_A = p\phi_A \odot h_B + H_B \odot f_A \text{ mod } q$ <div style="text-align: right; margin-top: 10px;"> $\xrightarrow{e_A}$ </div>	<p style="text-align: center;">Step 4 :</p> $a_B = G_B \odot e_A \text{ mod } q$ $K_B = a_B \text{ mod } p$ $= f_B \odot f_A \text{ mod } p$
$a_A = G_A \odot e_B \text{ mod } q$ $K_A = a_A \text{ mod } p$ $= f_A \odot f_B \text{ mod } p$	

$$G_{q_B} \odot G_B = 1 \text{ mod } q.$$

After choosing f_B, g_B and G_B , he computes pair (h_B, H_B) as follows

$$h_B = G_{q_B} \odot g_B \text{ mod } q \text{ and}$$

$$H_B = G_{q_B} \odot f_B \text{ mod } q.$$

(f_B, g_B, G_B) is the user B's private key and pair (h_B, H_B) is his public key. Now he selects a random polynomial $\phi_B \in L_\phi$ and computes

$$e_B = p\phi_B \odot h_A + H_A \odot f_B \text{ mod } q.$$

and send triplet (h_B, H_B, e_B) to user A (Table 1).

Step-3

After obtaining (h_B, H_B, e_B) from user B, user A chooses a random polynomial $\phi_A \in L_\phi$ in order to calculate

$$e_A = p\phi_A \odot h_B + H_B \odot f_A \bmod q$$

and then passes e_A to user B . Now, user A has e_B and user B has e_A .

Step-4

User A obtains $a_A = G_A \odot e_B \bmod q$ by choosing coefficients of a_A from interval $(-q/2, q/2]$. Thus

$$\begin{aligned} a_A &= G_A \odot [(p\phi_B \odot h_A + H_A \odot f_B) \bmod q] \bmod q \\ &= [p G_A \odot \phi_B \odot h_A + G_A \odot H_A \odot f_B] \bmod q \\ &= [p G_A \odot \phi_B \odot (G_{q_A} \odot g_A) + G_A \odot (G_{q_A} \odot f_A) \odot f_B] \bmod q \\ &= [p \phi_B \odot (G_A \odot G_{q_A}) \odot g_A + (G_A \odot G_{q_A}) \odot f_A \odot f_B] \bmod q \\ &= [p \phi_B \odot g_A + f_A \odot f_B] \bmod q \end{aligned}$$

Now, if all of a_A 's coefficients lie within the range of $(-q/2, q/2]$, it implies user A can treat a_A as with integer coefficients in place of $\bmod q$ coefficients. So, he computes shared key by $K_A = a_A \bmod p$ and obtains $K_A = f_A \odot f_B \bmod p$.

Step-5

User B gets $a_B = G_B \odot e_A \bmod q$ by choosing coefficients of a_B from interval $(-q/2, q/2]$. Thus

$$\begin{aligned} a_B &= G_B \odot [(p\phi_A \odot h_B + H_B \odot f_A) \bmod q] \bmod q \\ &= [p G_B \odot \phi_A \odot h_B + G_B \odot H_B \odot f_A] \bmod q \\ &= [p G_B \odot \phi_A \odot (G_{q_B} \odot g_B) + G_B \odot (G_{q_B} \odot f_B) \odot f_A] \bmod q \\ &= [p \phi_A \odot (G_B \odot G_{q_B}) \odot g_B + (G_B \odot G_{q_B}) \odot f_B \odot f_A] \bmod q \\ &= [p \phi_A \odot g_B + f_B \odot f_A] \bmod q \end{aligned}$$

Now, if all of the coefficients of a_B lie within the range of $(-q/2, q/2]$, it means in a similar way as discussed above, user B can treat a_B as with integer coefficients in place of $\bmod q$ coefficients. So, he computes shared key by $K_B = a_B \bmod p$ and obtains $K_B = f_B \odot f_A \bmod p$.

Since, f_A and f_B are polynomials, therefore $f_A \odot f_B = f_B \odot f_A$, i.e., $K_A = K_B$. Hence, key exchange algorithm terminates successfully.

Following is a toy example to verify the proposed NTRU-NIP key exchange protocol.

Parameters are $(N, p, q, d) = (7, 3, 41, 2)$ satisfying the condition $q = 41 > (6d + 1) \cdot p = 39$.

$$\begin{aligned}
f_A &= x^6 - x^4 + x^3 + x^2 - 1 \\
f_B &= x^6 + x^4 - x^3 + x^2 - x \\
g_A &= x^6 + x^4 - x^2 - x \\
g_B &= x^6 - x^4 + x^2 - x \\
\phi_A &= x^6 - x^5 + x - 1 \\
\phi_B &= -x^6 + x^5 - x + 1 \\
G_A &= -x^6 + 2x^5 - x^4 + x^2 + 10x - 1 \\
G_B &= x^6 - 3x^5 + x^3 - x^2 - x + 12 \\
G_{q_A} &= 38x^6 + 27x^5 + 38x^4 + 11x^3 + 34x^2 + 21x + 32 \pmod{41} \\
G_{q_B} &= 3x^6 + 10x^5 + 23x^4 + 18x^3 + 2x^2 + 30x + 28 \pmod{41} \\
h_A &= G_{q_A} \odot g_A \pmod{q} \\
&= x^6 + 10x^5 + 14x^4 + 21x^3 + 26x^2 + 2x + 8 \pmod{41} \\
H_A &= G_{q_A} \odot g_A \pmod{q} \\
&= 9x^6 + 35x^5 + 12x^4 + x^3 + 20x^2 + 40x + 2 \pmod{41} \\
h_B &= G_{q_B} \odot g_B \pmod{q} \\
&= 39x^6 + 9x^5 + 7x^4 + 7x^3 + 6x^2 + 36x + 19 \pmod{41} \\
H_B &= G_{q_B} \odot f_B \pmod{q} \\
&= 25x^6 + 26x^5 + 33x^4 + 26x^3 + 23x^2 + 31x + 32 \pmod{41} \\
e_A &= p\phi_A \odot h_B + H_B \odot f_A \pmod{41} \\
&= 25x^6 + 30x^5 + 7x^4 + 36x^3 + x^2 + 38x + 18 \pmod{41} \\
e_B &= p\phi_B \odot h_A + H_A \odot f_B \pmod{q} \\
&= 35x^6 + 27x^5 + 9x^4 + 12x^3 + 40x^2 + 12x + 25 \pmod{41} \\
a_A &= G_A \odot e_B \pmod{q} \\
&= 39x^6 + 37x^5 + 5x^4 + 2x^2 + 2x + 39 \pmod{41} \\
a_B &= G_B \odot e_A \pmod{q} \\
&= 39x^6 + 2x^5 + 40x^4 + 37x^2 + 8x + 39 \pmod{41}
\end{aligned}$$

For coefficients of a_A and a_B to be lie in range $(-q/2, q/2]$, we do center-lifting. After center-lifting, a_A and a_B are

$$\begin{aligned}
a_A &= -2x^6 - 4x^5 + 5x^4 + 2x^2 + 2x - 2 \\
a_B &= -2x^6 + 2x^5 - x^4 - 4x^2 + 8x - 2 \\
K_A &= a_A \pmod{p} \\
&= x^6 + 2x^5 + 2x^4 + 2x^2 + 2x + 1 \\
K_B &= a_B \pmod{p} \\
&= x^6 + 2x^5 + 2x^4 + 2x^2 + 2x + 1.
\end{aligned}$$

Hence, $K_A = K_B = 1220221$.

In decimal and binary representation, the shared key is 1402 and 10101111010 respectively.

4 Security Analysis

The key exchange protocol proposed here is based on NTRU-NIP-KE problem. We will show how our proposed algorithm relates to the hardness of lattice problems and how an attacker can mount a lattice attack on this algorithm to get the secret key by solving hard problems of lattices.

Definition 3 NTRU-NIP KE Problem-Find $f_A \odot f_B \pmod p$ for a specified security parameter specifying (N, p, q) and spaces L_f, L_g and L_ϕ together with $e_A, e_B, (h_A, H_A)$ and (h_B, H_B) .

Definition 4 NTRU-NIP KE Assumption-The NTRU-NIP KE problem is computationally complex to solve with the appropriate parameters.

Relationship between NTRU-NIP encrypt assumption and NTRU-NIP KE assumption-If NTRU-NIP encrypts inversion problem can be solved by an efficient algorithm X , then X can be transformed into an efficient algorithm Y , which can calculate f_A from e_A and f_B from e_B . Algorithm X may then calculate the key $K = f_A \odot f_B \pmod p$. As a result, NTRU-NIP-KE Assumption is stronger than NTRU-NIP Encrypt Assumption.

4.1 Lattice Attack

The hard problems of lattices, SVP, and CVP have a significant impact on the security of our proposed protocol. $L_{\overline{H}}$ where $\overline{H} = h - H$ is used to represent the NTRU-NIP lattice. The rows of $2N \times 2N$ matrix made up of four $N \times N$ blocks are used to create this lattice $L_{\overline{H}}$ where the blocks are

$$\left[\begin{array}{c|cccc} I_{n \times n} & h_0 - H_0 & h_1 - H_1 & \dots & h_{N-1} - H_{N-1} \\ & h_{N-1} - H_{N-1} & h_0 - H_0 & \dots & h_{N-2} - H_{N-2} \\ & \dots & \dots & \dots & \dots \\ & h_1 - H_1 & h_2 - H_2 & \dots & h_0 - H_0 \\ \hline 0_{n \times n} & & & & qI_{n \times n} \end{array} \right]$$

Lattice $L_{\overline{H}}$ can be re-expressed as $L_{\overline{H}} = \{(G, f - g)|(h - H) \odot G = (g - f) \pmod q\}$

4.1.1 Lattice Attack Using SVP

Suppose a positive attacker, having the user's public key (h, H) , tries to retrieve his private key (f, g, G) . Since, he knows relations $h = G_q \odot g \bmod q$ and $H = G_q \odot f \bmod q$. By these relations, he can get $(h - H) = G_q \odot (g - f) \bmod q$ that is $(h - H) \odot G = (g - f) \bmod q$. Thus, \exists a polynomial $u \in R$ such that $g - f = (h - H) \odot G - qu$. It is evident that $[G, -u].L_{\overline{H}} = [G, -u] \begin{bmatrix} I & h - H \\ 0 & qI \end{bmatrix} = (G, G \odot (h - H) - qu) = (G, g - f)$. That means lattice $L_{\overline{H}}$ has the vector $(G, g - f)$ as part of its structure. Since g and f are ternary polynomials, i.e., having coefficients 0, 1 and -1 , therefore, polynomial $g - f$ will also have coefficients equal to 0, 1 and -1 except in two cases. In the first scenario, g has a coefficient of 1 and f has a corresponding coefficient of -1 , and second case is, g has a coefficient of -1 and f has a corresponding coefficient of 1. In these two cases, we will not get a ternary coefficient. We discard those cases and change those coefficients according to our requirement and repeat the whole process from the beginning. Since vector $(G, g - f)$ belongs to lattice $L_{\overline{H}}$, thus for a suitable choice of g and f , by lattice reduction algorithms such as LLL algorithm [10], adversary can obtain target vector $g - f$. Consequently, by this attack, he can get information about the difference $g - f$ of private keys g and f , not exactly private key pair (g, f) . So, it is a partial key exposure attack in place of a fully lattice attack. Thus lattice attack is much more robust here in comparison to NTRU-KE. Our polynomial G is an element of ring $R = \frac{\mathbb{Z}[x]}{x^{N-1}}$, hence coefficients of G range all over \mathbb{Z} instead of ternary polynomials. That is why the target vector $(G, g - f)$ is not among short vectors of lattices. So, we can not say that if we run any shortest vector algorithm to lattice $L_{\overline{H}}$, we will surely get vector $(G, g - f)$ or any rotation of this vector. Hence, the conclusion is, using the shortest vector problem, lattice attack is not possible on our key exchange protocol. This is an additional advantage of the proposed protocol over NTRU-KE.

4.1.2 Lattice Attack Using CVP

Since $e = p\phi \odot h + H \odot f \bmod q$, therefore it can be re-expressed in vector form as $[0, e] = [\phi, \phi \odot (ph) \bmod q] + [-\phi, H \odot f]$. By lattice $L_{\overline{H}}$. Clearly, vector $[\phi, \phi \odot (ph) \bmod q]$ belongs to lattice L_{ph} . Thus, the distance between a vector point $[0, e]$ and a vector point of a lattice L_{ph} is merely $[-\phi, H \odot f]$. The width condition is satisfied by the vector $[-\phi, H \odot f]$ if the proper parameters are selected. Here, since H contains G_q , therefore vector $[-\phi, H \odot f]$ will not be short with respect to the Euclidean norm as required for solving the closest vector problem (norm of distance vector should be minimized). So, this problem is not exactly based on the closest vector problem of lattices. But we can observe this in the approximate CVP version of the CVP of lattices. Approximate CVP is defined as, for a given lattice L and a non-lattice vector $w \notin L$ as a target vector, $\|(v - w)\| \leq \gamma \|v_{shortest}\|$ where $v_{shortest}$ is shortest vector of lattice L and γ is an approximate factor constant. Here, we have a lattice L_{ph} and a non-lattice target vector $[0, e]$. By relation, $[0, e] =$

$[\phi, \phi \odot (ph) \bmod q] + [-\phi, H \odot f]$, we can see that $\| [0, e] - [\phi, \phi \odot (ph) \bmod q] \| = \| [-\phi, H \odot f] \|$. Hence, by running an approximate closest vector algorithm to vector $[0, e]$ of the lattice L_{ph} , distance vector $[-\phi, H \odot f]$ can be recovered. We can see that $\| [-\phi, H \odot f] \| = \sqrt{\|\phi\|^2 + \|(H \odot f)\|^2} \leq \sqrt{\|\phi\|^2 + \|H\|^2 \|f\|^2}$. We know that f and ϕ are truncated polynomials with ternary coefficients (norm of these vectors will be very small and we can assume f or ϕ as a short vector for lattice L_{ph}) and $f \in \tau(d+1, d)$ and $\phi \in \tau(d, d)$, we can say $\|f\| \approx \|\phi\|$. Hence $\| [-\phi, H \odot f] \| \leq \sqrt{\|\phi\|^2 + \|H\|^2 \|f\|^2} = \sqrt{\|f\|^2 + \|H\|^2 \|f\|^2} = \sqrt{\|f\|^2(1 + \|H\|^2)} = \|f\| \sqrt{(1 + \|H\|^2)}$. Hence, $\| [0, e] - [\phi, \phi \odot (ph) \bmod q] \| \leq \gamma \|f\|$ where $\gamma = \sqrt{(1 + \|H\|^2)}$. In this manner, we may use the approximate closest vector algorithm to locate the user's private key f . As a result, the security of the proposed key exchange protocol is heavily dependent on how difficult the approximate CVP is to break.

5 Comparison of NTRU-NIP-KE Over NTRU-KE and MaTRU-KE

Since choosing secure parameters for [9, 17] and NTRU-NIP-KE is still remaining hard. Therefore, an experimental comparison of performance among these is currently not possible. Here, we are listing some theoretical comparisons as

- Lattice attack is not possible on NTRU-NIP-KE using SVP, whereas it is possible on others.
- NTRU-NIP-KE delivers higher flexibility in the choice of parameters compared to [9, 17].
- Computation cost of NTRU-NIP-KE and NTRU-KE is much smaller than MaTRU-KE as MaTRU-KE uses matrix ring \mathcal{M} whereas the other two uses polynomial ring R as their underlying ring.
- Compared to NTRU-KE and NTRU-NIP-KE, MaTRU-KE is more efficient since it generates k polynomial at a time, whereas NTRU-NIP-KE generates a single key. But in the case of MaTRU-KE, to store generated keys secretly and securely is itself a challenging task.

6 Conclusion

In this paper, we proposed a quantum-safe key exchange protocol that provides enough space to choose invertible polynomials for generating keys. An additional advantage of the proposed KE protocol over NTRU-KE and MaTRU-KE is that an attacker cannot attack it by using the shortest vector algorithm.

References

1. Banks WD, Shparlinski IE (2002) A variant of NTRU with non-invertible polynomials. In: Proceeding of INDOCRYPT 2002. LNCS, Springer, pp 62–70
2. Coppersmith D, Shamir A (1997) Lattice attacks on NTRU. In: Proceeding of EUROCRYPT 1997. LNCS, Springer, pp 52–61
3. Diffie W, Hellman ME (1976) New directions in cryptography. *IEEE Trans Inf Theory* 22:644–654
4. Hankerson D, Menezes A, Vanstone S (2004) Guide to elliptic curve cryptography. Springer, New York
5. Hoffstein J, Pipher J, Silverman JH (1998) NTRU: a ring based public key cryptosystem. In: Proceedings of the ANTS. LNCS, Springer, pp 267–288
6. Hoffstein J, Silverman JH (2000) Optimizations for NTRU. *Public-key cryptography and computational number theory*. DeGruyter
7. Hermans J, Vercauteren F, Preneel B (2010) Speed records for NTRU. *Topics in cryptography-CT-RSA*. Springer, pp 73–88
8. Koblitz N (1987) Elliptic curve cryptosystems. *Math Comput* 48:203–209
9. Lei X, Liao X (2013) NTRU-KE : a lattice-based public key exchange protocol. *Cryptology ePrint Archive*
10. Lenstra AK, Lenstra HW, Lovsz L (1982) Factoring polynomials with polynomial coefficients. *Math Ann* 261:513–534
11. Lochter M, Merkle J (2010) Elliptic curve cryptography (ECC) brainpool standard curves and curve generation
12. Maurer U, Wolf S (1999) The relationship between Breaking the Diffie-Hellman protocol and computing discrete logarithm. *SIAM J Comput* 28:1689–1721
13. Perlner R, Cooper D (2009) Quantum resistant public key cryptography: a survey. In: Proceedings of the 8th symposium on identity and trust on the internet. ACM New York, USA, pp 85–93 (2009)
14. Shor P (1994) Algorithms for quantum computation: discrete logarithms and factoring. In: Proceedings of the 35th annual IEEE symposium on foundations of computer science. IEEE Press, Piscataway, pp 124–134
15. Shor P (2006) Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM J Comput* 26:1484–1509
16. Stehle D, Steinfeld R (2011) Making NTRU as secure as WorstCase problems over ideal lattices. In: *Advances in cryptology—Eurocrypt’11*. LNCS, Springer, pp 27–47
17. Singh S, Padhye S (2018) MaTRU-KE: a key exchange protocol based on MaTRU cryptosystem. *Int J Commun Syst* 32(4):e3886
18. Valluri MR (2018) Cryptanalysis of Xinyu et al.’s NTRU-lattice based key exchange protocol. *J. Inf. Optimiz. Sci* 39(2):475–479

Blockchain Within the Insurance Industry: A Bibliometric Analysis



Lalit Garg , Luca Bugeja, Corinne Marie Formosa, and Varun Shukla

1 Introduction

Blockchain and distributed ledgers have been attracting attention for several years now. This technology started becoming a solution to problems related to the ownership of an asset, especially in the financial industry. The financial crisis showed that even in financial services, pinpointing the owner of an asset is not always possible [42]. In computer science, many articles have been published revolving around blockchain-related topics. These have analysed consensus algorithms and proposed new concepts on how to deal with issues regarding the privacy of smart contracts [51]. Together with the significant benefits that this technology brings, there are also drawbacks discussed in various literature pieces [35]. The vast number of benefits blockchain brings shows its great potential to permanently affect and change aspects of societies' different operations. Swan [50] highlights how blockchain and its disruptive potential may be broken into three main categories. These include Blockchains 1.0, 2.0 and 3.0 [50].

Blockchain 1.0 refers to the currency aspect of blockchain. This denotes the use of cryptocurrencies within the application for things such as currency transfers, modern digital payment infrastructure and fee payments [50].

L. Garg (✉) · L. Bugeja · C. M. Formosa
University of Malta, Msida, Malta
e-mail: lalit.garg@um.edu.mt

L. Bugeja
e-mail: luca.bugeja.18@um.edu.mt

C. M. Formosa
e-mail: corinne.formosa.18@um.edu.mt

V. Shukla
Pranveer Singh Institute of Technology, Kanpur, India

Blockchain 2.0 started with the introduction of Ethereum, facilitating developers to develop decentralised apps (dApps). It refers to smart contracts and their potential to heavily impact the entire economy. Blockchain contracts enable better secure, trustless transactions with respect to bonds, stocks, loans, smart property and smart contracts [50]. However, Blockchain 2.0 suffers from scalability, slow processing, heavy-weight applications and astronomically high energy consumption.

Blockchain 3.0 refers to all other use cases apart from currency, finance and markets. Thus, this would include health, art, science and government implementations [50]. It facilitates ‘Smart everything’ and claims to solve the issues of its previous generation.

Gatteshci et al. [25] defined blockchain as a distributed ledger heavily maintained by network nodes, continuously recording transactions completed by the nodes. Information found within the blockchain is of public interest and accessible by anyone, and it cannot be modified or erased. The insurance sector was at the forefront of this new technology and is widely being investigated by small and large companies and consultancy firms. Moreover, in 2016, a B3i was founded—a company that was the first blockchain-centred insurance firm [25].

2 Blockchain Background

Many of the world’s biggest revolutions started from a quiet disruptor [29]. One can say that we are currently during another quiet disruptor—blockchain. Blockchain may be defined in various ways. However, IBM defines it as a ‘shared, immutable ledger for recording transactions, tracking assets and building trust’ [37]. One can say it is a decentralised distributed database of records. In other words, it is a public ledger responsible for storing digital events implemented by partaking groups [12]. Blockchain, though still relatively new to the industry of technologies, is on its way to being the ‘bedrock’ of record-keeping systems [29]. The idea of blockchain was first thought of by two scientists, S. Haber and W. Stornetta, back in the eighteenth century. The idea included a practical computational solution that enabled digital documents to be timestamped, thus hindering the backdating or tampering of any documents [29]. Later, in 2004, computer scientist Hal Finney introduced the ‘Reusable proof-of-work system’. This system worked by receiving a non-exchangeable hash cash-based proof-of-work token and, in return, created an RSA-signed token that could then be transferred from person to person. This system helped in solving the double-spending problem. This was possible by storing the ownership data of tokens on a secure and trusted server. Moreover, the server allowed various types of users to verify the integrity and overall correctness of the data. Moreover, in the late stages of 2008—a white paper titled ‘Bitcoin: A Peer-to-Peer Electronic Cash System’ by Satoshi Nakamoto mainly highlighted the inner working of blockchain architecture. Essentially, Bitcoins are ‘mined’ or executed for a reward using the proof-of-work mechanism by independent miners and then authenticated by the decentralised nodes found within a network. In early 2009, Bitcoin came into existence—having been

mined by Satoshi Nakamoto—having a reward of 50 bitcoin—later, the first transaction took place with Hal Finney being the recipient of 10 Bitcoin. Some years later, in 2013, smart contracts were created due to the scripting language—Ethereum [37]. Although blockchain is mainly known for its insinuation with Bitcoin—it has undoubtedly come a long way since then. Nowadays, blockchain is gaining tremendous attention in various applications and industries, such as health care [10, 23], finance and insurance.

2.1 *Blockchain Technology*

Blockchain is a distributed database of records. It can also be described as a public ledger of digital events (or transactions) executed [12]. Every transaction needs to be verified, and data cannot be erased. This series of records have a timestamp. These are regulated by a group of computers not possessed by one person or organisation. All these blocks of data are stored securely and connected using cryptographic principles. Since this network does not have a central authority figure, information is open and available to everyone. Anything uploaded to a blockchain is transparent, and people can be held accountable for their actions [46]. Information that is stored on a blockchain exists as a shared database. The database is held in several locations to ensure security. The records ‘are truly kept public and easily verifiable’ [46]. The blockchain contains three crucial components: blocks, nodes and miners. Each chain has numerous blocks, and, in turn, each of the blocks contains the following basic features: the data within the block, a nonce and the hash. A nonce is a 32-bit number which is generated randomly when a block is initially created and then generates a block header hash. A hash is a 256-bit number that is linked to the nonce. It is tiny and thus begins with a lot of zeros. Once an initial block of a chain is made, the nonce produces a cryptographic hash. The block data is presumed as signed and tied to the nonce and hash unless it gets mined [5]. Blockchain is referred to as a Distributed Ledger Technology (DLT), where transactions are registered using an unchangeable cryptographic signature defined as a hash. The DLT refers to the decentralised database managed by several participants [16].

Miners are the function that creates the new blocks on the chain. This process is called ‘mining’. Mining a block is not an easy task, especially given large chains. This is because while every block in the blockchain has its nonce and hash, it also refers to the previous block’s hash. Special software is used to solve the problem of finding a nonce that generates an accepted hash due to its very complex nature. Since the nonce is 32 bit and the hash is 256 bit, billions of combinations possible must be mined before the right one is found. This is said to be ‘the golden nonce’ [5]. It is so complex to manipulate blockchain technology because altering a block demands the re-mining of not only that block getting alterations but even those following it. This is ‘safety in math’. Nodes are any electronic device that contains copies of the blockchain and keep the network functioning [5]. Decentralisation is one of the most important aspects of blockchain. No computer or organisation owns the blockchain.

Public information is combined with a system of checks-and-balances which aids in the blockchain maintaining integrity and creating a sense of assurance among its users.

From a technical perspective, blockchain has the following features: decentralisation, traceability, transparency and immutability. Decentralisation refers to the fact that the information is not stored in one area. Everyone in the network owns the information. There is no governing authority. Transactions using blockchain can be conducted anywhere in the world between various users. Decentralisation is the process of ‘verifying, storing, maintaining and transmitting on the blockchain’ based on a distributed system’s structure. Mathematical methods build trust between nodes rather than the centralised organisation [28]. Immutability is permanence. Once something has been inputted into the blockchain, it cannot be altered. This property comes from cryptographic hash work [33]. Data is only included in a block once everyone on the network approves it. This ensures confidential transactions. Transparency refers to the ‘straightforwardness’ of blockchain technology. While the person’s actual personality remains secure, one can observe every one of the exchanges conducted within their location. All transactions are stored chronologically in the blockchain. Since a block relates to another two blocks through the cryptographic hash function, every transaction can be tracked (hence traceability) [28]. These features tie in with blockchain’s important feature, increased security. Without a central authority, one cannot alter the characteristics of the network as one like and tailor it to their benefit [32]. Adding encryption also ensures a layer of security for the system.

Blockchain is also consensus-driven. Each block on the chain is verified independently using Consensus models that provide rules for validating a block [49]. They also use proof to show that effort was made, such as computing power. This works without a central authority figure or a third party (like a trust-granting agent). An important part of the blockchain is the fact that it allows the use of smart contracts. This is an agreement between two parties in the form of computer code. These run on the blockchain, so they are kept on a public database, making them unchangeable. The blockchain is responsible for processing transactions that occur in a smart contract. There is no need for third-party involvement because they can be sent automatically. This also means there is no need to rely on anyone [7]. Furthermore, the transactions only happen when conditions in the agreement are met. This means there is no third party, so there are no issues with trust.

2.2 Blockchain Concepts

The blockchain is a chain of blocks. Initially, the main objective of blockchain was to facilitate the secure transaction of bitcoin. Within the context of bitcoin, the blockchain’s role was to act as the ledger, recording transactions and ensuring the transfers went to the correct end-user. This was done with the use of cryptographic signatures.

Gatteschi et al. [25] state that blockchain has six core concepts: transactions, blocks, nodes, majority consensus, mining and wallet.

Transactions: Transactions within the blockchain are permanently recorded and cannot be erased.

Blocks: A block comprises a group of transactions primarily grouped based on the period in which they took place.

Nodes: Rather than being kept in a centralised database server, the blockchain is stored over several computers connected to a network, these computers are referred to as nodes, and they contain an updated copy of the blockchain.

Majority Consensus: This refers to the fact that since there is no official authority for decision-making, blockchain follows a consensus or majority-based approach. Each node is responsible for modifying the blockchain's local copy based on the status of many of the other nodes.

Mining: Nodes within the blockchain could participate in the blockchain in two ways. They can inactively store the local copy or participate in the upkeep and updating of the blockchain, often referred to as mining. During the mining process, nodes check for transaction verification; each time a new block is added to the blockchain, a complex mathematical problem must be solved. The mathematical problem was designed to limit the possible chances of an entity attempting to fabricate previous transactions.

Wallet: In the case of a transfer using cryptocurrencies, digital wallets are used. Unlike a tangible wallet, a digital wallet does not store memory of your current balance but a history of previous transactions; thus, only your specific and complex credentials are stored. Each wallet is associated with a unique address, and the recipient's address must be known for a transaction to take place.

2.3 *Smart Contracts*

Smart contracts are increasing in popularity. This innovation might eliminate the need for certain lawyers and banks to be involved in contracts dealing with selling and purchasing assets, which they have been a part of for years [42]. Smart contracts are a mechanism involving many digital assets between two or more groups of people. In a smart contract, these digital assets are reallocated among the people groups according to a code instruction structured around data readily available to the people when the contract was written [35]. Contract terms and assets to be transferred are specified and may be coded mathematically into open-sourced, consensus-based chains, and their execution is verified autonomously. The code cannot be tampered with by either party involved. This means that a smart contract enabled by blockchain will instantly conclude, irrespective of how long the process is to take or whether either person from the group has doubts or changes about the contract being executed [35].

A self-ruling organisation managed by smart contracts whose operation is as far removed as possible from the day-to-day input of managers, employees or owners is called a Decentralised Autonomous Organisation (DAO). These are said to tackle a widespread problem of governance that scientists and economists refer to as the leading-agent dilemma [36]. This problem occurs when an organisation's principal agent has the decision-making power on behalf of, or even impacts, the principal (another person or entity in the organisation) [2, 3]. DAOs provide an operating system for people and institutions that do not know, and hence do not trust each other. They might also not be in the same geographical area, differ in languages and be subject to varying jurisdictions. DAOs involve a group of people that are collaborating on a self-enforcing open-source protocol. Blockchains and smart contracts, therefore, decrease transaction costs of management at an increased level of transparency, connecting the interests of all stakeholders by the consensus rules tied to the native token [2, 3]. These tokens are sometimes referred to as 'protocol tokens'. They are a part of the incentive scheme of blockchain infrastructure. Their main purpose is to encourage a diverse group of people who do not know each other to coordinate around the purpose of a specific blockchain [2, 3].

2.4 Public, Private and Hybrid Blockchains

2.4.1 Public Blockchains

A blockchain is considered public if all participants can read and use a blockchain to carry out transactions. Furthermore, on a public blockchain, everyone is integrated into the process of creating the consensus. In this type of blockchain, 'nodes on the network validate choices discussed and initiated by the developers by deciding whether to integrate the proposed modifications' [13]. This is based on 'crypto-economics', the combination of economic incentives and verification mechanisms that use cryptography. This system has shown its strength and resilience because it is based on a community approach to the economy [13].

2.4.2 Private Blockchains

A blockchain is considered private if the consensus process may only be reached by a limited and pre-established number of participating parties. Accessing to writing is only possible if given by the organisation, and read permissions may be made public, or they can also be restricted. A preselected group of nodes restricts the consensus process. This type of blockchain 'does not necessarily use mechanisms based on cryptography' [13]. Moreover, private blockchains have no mining, working proof or remuneration. This distinguishes the 'two types of storage and transmission technologies' [13].

2.4.3 Hybrid Blockchains

A hybrid blockchain serves as an in between of the two extremes, which are public and private blockchains. This type of blockchain can therefore enjoy characteristics from both private and public blockchains. The members of this blockchain or a particular entity with higher power can establish which transactions remain publicly available and which must be restricted to fewer people. The hybrid blockchain that consists of the public and private state of the network ensures that transactions are private but still verifiable by an immutable record on the public state of the blockchain. In a public state, ‘every transaction gets approved by a large network and is secure and trustworthy’ [34]. This means a central governing body or an exhaustive chain of intermediate parties to supervise processes is not required. The hybrid blockchain technology ‘can be used to build enterprise-grade implementations of the open-source technology across different trades leading to real-world use cases’ [34].

3 Blockchain in Insurance

3.1 Blockchain Merits

Though still a relatively new and up-and-coming technology, it is already quite apparent the instant and great benefits blockchain can bring when implemented within the insurance industry. This section will further emphasise essential drivers to increase efficiency and usability [41]. The implementation of blockchain may be matched with the use of smart contracts in the day-to-day running of business activities. Such business activities may include identity authentication, validation, data management, document formations and payment options. Thus, as a result of this implementation, additional personalised products can be offered to potential and current customers [28]. These new and improved products reap various advantages by being better priced and transparent, benefitting increased reliability, improving automated processes and preventing fraud.

Insurance companies are well known for not being transparent and hiding consumer data. Consumers have little to no information about what data insurance companies have, how it is used and which third parties it is shared with. This, in turn, creates a sense of distrust between clients and insurance agencies, especially when the time comes to open a claim or receive compensation [36]. However, with the implementation of blockchain technology, various benefits emerge for the insurance company and the potential client. First, insurance companies would now be able to build a more reliable and complete customer profile, thus eliminating duplicated data. Since data within a blockchain cannot be modified, there is no doubt that the data entered is reliable and authentic. Secondly, clients can now visibly see what data insurance companies have about them and what is being done with them. Thirdly, blockchain enables third-party claims to be automatically verified and

processed through personal devices. More importantly, the insurance company can see the transactions in the blockchain transparently [36].

Moreover, another advantage of implementing blockchain within the insurance industry is the prevention of fraudulent activities. In America, it is said that fraudulent insurance claims accumulate to over 40 million dollars. Thus, it is no surprise that why opening up an insurance claim can seem daunting [45]. However, when employing a blockchain infrastructure, these issues are solved. This is because the data stored on the blockchain is secured using a cryptographic signature and consent settings. Different groups can share, verify and authenticate client data without revealing personal data. Using a decentralised ledger enables companies to foresee suspicious patterns and prevent activities such as clients attempting to open multiple claims, insurance being sold by unlicensed brokers or policyholders manipulating ownership. Also, insurance companies are now providing their clients with specifically encrypted electronic ID Cards [36]. Blockchain can also facilitate decentralised patient IDs (generation) [8].

3.2 Blockchain Applications in Insurance

Blockchain is likely to change the way insurance companies do business. This technology can help carriers save time, reduce costs, enhance transparency, comply with regulations and produce better products and markets [11]. The technology of blockchain has potential uses that are apparent in every part of the insurance value chain. This means from the financing and valuation of their services and goods, their sales and supply, to the continuous product control and processing of claims [43]. Blockchain is a promoter and stimulus to quicken digitisation, encourage change and transformation and increase a sense of innovation. While reading through different types of literature, it was apparent that insurance companies are so eager to start implementing this new emerging technology to its close-to-instant benefits with regard to fraud detection, pricing, cost savings and growth opportunities [53]. Caitlyn Long, Chairman and President of Symbion, states that blockchain and DLT technologies greatly aid insurers and financial institutions in managing security claim settlements [30]. In the following, attention will be given explicitly to the potential and current uses of blockchain within the insurance sector and the benefits that can be achieved.

3.2.1 Better Customer Experience and Decreased Operating Costs

In the case of creating better experiences for customers, Valentina Gatteschi et al. [25] mention how the combination of smart contracts and blockchain could be used to speed up claim processing times, as well as reduce the costs which are often associated with human error. A simple use case scenario includes programming the smart contract to issue an automatic refund if the customer gets his, for example,

injuries examined by a certified doctor. The refund would occur once the doctor sends the transaction to the smart contract to verify his identity. Moreover, a more complex use case would be with the involvement of oracles. In crop insurance, for example, the oracle would be responsible for periodically checking the data and updating it within the blockchain. The smart contract would then be able to read this data and, based on a pre-set list of requirements, issue a payment to customers in the case of, say, storms or terrible weather. A real-life prototype like the cases mentioned above was used in travel insurance. The idea in this scenario was to use flight delay data and automatically issue a refund to customers whose flight was delayed [25].

3.2.2 Claim Processing

Usually, this process requires manual steps where the policyholder inputs their details into a report and makes a call to communicate the claim to the insurance company. Insurers verify the proof of the claimed event and might also assess the damage. Then the insurer pays out the claim. The product conditions are inputted using blockchain into a piece of code, which is the smart contract. The claim is automatically paid out once it receives the correct requirements. The trigger event of the claim can be verified from publicly available data but is also dependable—for example, flight delays or natural catastrophes [43]. Claims are documented on the blockchain to get audited and eliminate numerous claims reported for the same insured event. This defends against insurance fraud. Implementing blockchain also improves a customer's experience.

3.2.3 Reinsurance and Swaps

Reinsurance is insurance for insurers. This happens when insurers transfer a part of their risk portfolios to other parties in some form of agreement to decrease the chances of having to pay a large sum that results from a claim. This is typically settled in 2 to 3 months, from the time the insurer pays out a claim to the time recovery is received from the reinsurer. This results from the time taken to compute the data for claims, determine reinsured claims and premiums, coordinate claims and premiums and resolve conflicts, to name a few. Reinsurance treaties/swap terms can be written into smart contracts using blockchain that automatically execute payments (premiums and claims) to and from reinsurers when pre-established agreements are satisfied. 'Experience data is recorded on the blockchain, tamper-resistant and immediately auditable' [43]. All people involved in reinsurance, like insurers, reinsurers, third-party data providers and asset managers, record data on the blockchain so anyone can access it without needing to ask someone for it. Cleansed data can also be stored on the blockchain and cannot be tampered with, and everyone can see what alterations have been made. 'All transactions (premiums and claims) are recorded on the blockchain for visibility to future transacting parties' that may need the information [43].

3.2.4 Peer-To-Peer Insurance

This has been around for some time. It refers to individuals purchasing premiums and pooling them to insure against a risk. Blockchain technology brings new opportunities due to the decentralised autonomous organisation (DAO) principle. Smart contracts are a representation of the first level of the decentralised application. They often involve human input, especially when the contract is to be signed by several different parties. DAOs allow P2P insurance to be available on a large scale. This is due to their capacity to manage complex rules among many stakeholders. Both established insurers and new players can position themselves more efficiently in the P2P market [44]. The peer-to-peer models currently being used within the industry aren't technically peer-to-peer models but rather modified models. The implementation of smart contracts could reap serious benefits with the implementation and creation of DAOs. A prototype example of this is DYNAMIS (Dynamic InsurTech) [31], based on the currently used Ethereum blockchain. DYNAMIS uses Ethereum to provide a peer-to-peer insurance model for unemployment insurance and similar niche markets. DYNAMIS's main aim is to create a DAO to restore trust between clients and insurance providers. This can be done by using consensus-based infrastructures that circulate the costs and processes traditionally applied by insurance companies. It is important to note that this level of implementation of peer-to-peer insurance is still in its initial stages and is not readily available to the public. Moreover, a recent study concluded that customers are more willing to pay extra costs to interact with intermediaries rather than benefit from lower costs and do it themselves [25].

3.2.5 Micro-Insurance/Pay-Per-Use

Blockchain implementation in the insurance industry also can open income opportunities for insurance firms, irrespective of their size. Pay-Per-Use insurance policies were hard to finance and implement in the past due to high administrative costs and high levels of human intervention needed. However, with the intervention of smart contracts, the possibility of quick and cheap policies is enabled. Car renting is a use case of this in action, and GPS data may automatically calculate the premium while driving the car [25]. Companies could potentially rely on the public blockchain. The smart contract would collect funds from customers and keep them until a specified period has passed, then pass them on to the insurance company. From the point of view of the insurance firm, employing this type of new technology could further differentiate them from their competition and attract a new type of young market.

3.2.6 Other Examples of Blockchain Use in Insurance

The InsurTech Company, Etherisc, started building blockchain-enabled products and testing them publicly in October 2017. Etherisc has designed a decentralised

insurance protocol to build insurance products collectively [15]. The Co-founder of Etherisc, Stephan Karpische, explained how blockchain technology could ease some of the problems that arise in traditional insurance practices. He also explained how it could enable the production of new insurance products. ‘In traditional insurance practices, you have an inherent conflict of interest’—Stephan explained [14]. Using smart contracts makes processes much easier because they eliminate having to balance out shareholders’ and customers’ interests. This decentralised approach also brings various features, such as fairness and transparency, low conflict of interest, lower cost and faster time to market. Since smart contracts are open source, they can be accessed and verified. Also, they do not need to earn money as decentralised insurance is not run by multiple intermediaries that take a cut of the premium. The decentralised approach to insurance products means lower operating costs. Moreover, with blockchain technology and smart contracts, new insurance products can be developed in weeks and sometimes even days.

Etherisc is also developing insurance solutions such as weather-based crop insurance, multi-signature wallet insurance and collateral protection for crypto-backed lending. [14]. Another example of how Etherisc utilises blockchain is through their ‘cryptocurrency-based flight delay program’ that lets passengers buy flight insurance ‘using either cryptocurrency or fiat money’ [6]. Then they receive pay-outs automatically after a qualifying event [6]. The smart contracts that Etherisc uses can autonomously verify claims by using various ‘oracles’ of data sources. An example of this would be when dealing with a crop insurance claim. Etherisc can examine satellite images, data from weather stations and videos and pictures captured by drones provided by the insured person. This automated inspection can uncover fraudulent claims before they are submitted for human review, allowing insurers more time for urgent and complex tasks.

3.3 Limitations of Blockchain Use in Insurance

Previously, the advantages and benefits of implementing blockchain technology within the insurance industry were highlighted. However, as with any other technology, blockchain has drawbacks and limitations. The primary limitations blockchain has when being adopted within the insurance industry include security, scalability and standardisation and regulations.

3.3.1 The Security

Although blockchain is very secure, there are still some unlikely instances in which security is an issue. This instance is referred to as a ‘51% Attack’. A 51% attack refers to an attack on the proof-of-work blockchain, where an attacker, usually a group of miners, takes control of 51% or more of the computing power, putting them in control [48]. Also, another security concern is smart contracts. Smart contracts often

rely on external data from sources known as ‘oracles’. Oracles are centralised data point which contains large amounts of data. Since oracles are found in a centralised manner, they are prone and vulnerable to attacks. Also, oracles cannot understand their processing data; thus, shared data may be invalid, outdated and even unusable. This situation brings up what is known as the ‘oracle problem’. This problem refers to the fact that the process of smart contracts could be heavily impacted by unreliable or invalid data given by the oracles [43].

3.3.2 The Scalability

Since for a transaction to be verified, it needs to go through a consensus-based approach and continuous replications, the scalability of the blockchain quickly becomes an issue. Moreover, stored data is continuously becoming an issue as data within a blockchain is said to be immutable. Thus, if a record needs to be updated—a new one must be uploaded. The scalability issue can be seen when comparing a transaction done with Bitcoin through blockchain and a transaction done through VISA. Visa has the potential to process well over 65,000 transactions per second. However, Bitcoins’ maximum is only seven transactions per second—thus, it can be seen that although it may be more secure, it isn’t as efficient as it could be [43].

3.3.3 Regulations and Standardisation

Many finance and insurance industry regulators are starting to understand the potential blockchain could have on their industries and thus is forcing these regulators to reconsider regulations and laws. Although EU regulators are aware of these changes, they must also realise that technology is overtaking current regulations. EU regulators are aware of the high potential blockchain has and thus are monitoring its growth within the sector. While observing blockchain’s features and elements, queries were raised about data ownership, different applications, transparency and solvency. For this reason, sandbox testing and monitoring need to pursue this technology. As a result of this, newer regulations and legislations will likely be published to protect better both the consumer and insurance agency [47].

4 Literature Review and Bibliometric Analysis

Figures 1 and 2 show the Google Trends for the search term ‘Blockchain Insurance’ since 2004 and for the last 5 years, respectively. These show that the interest in searching for Blockchain Insurance increased steadily from 2015 to 2017 and remained almost consistent in the previous 5 years.

While searching Blockchain and Insurance in Scopus, we got 619 articles. When carefully analysing these articles, we excluded 294 articles discussing general

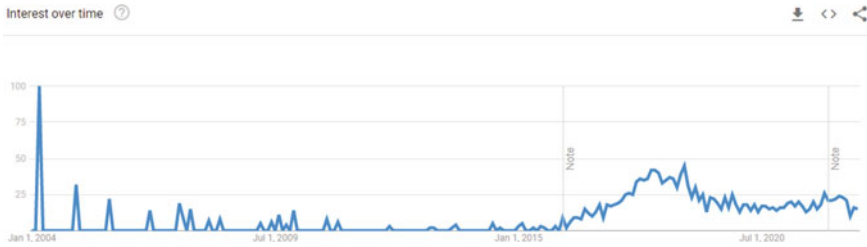


Fig. 1 Google Trends for the search term ‘Blockchain Insurance’ since 2004

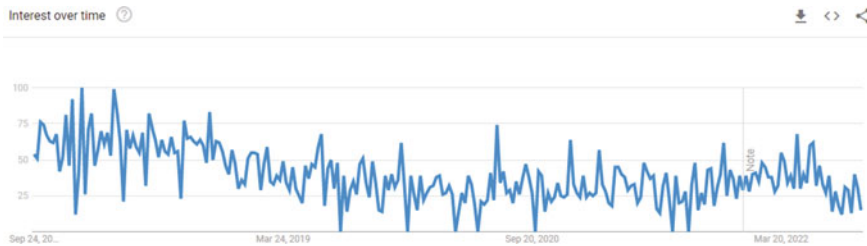


Fig. 2 The last 5 years’ Google Trends for the search term ‘Blockchain Insurance’

blockchain applications, including insurance. The remaining 325 articles presented blockchain applications in insurance. Figures 3, 4 and 5 show the distribution of these articles based on their publication year, country and article type, respectively. We can see steady growth in publication numbers during the last 6 years. The maximum number of articles were published by Indian authors, followed by Chinese and US authors. More than half (50.5%) of articles are conference papers, 38.2% are Journal articles and 7.7% are book chapters.

When analysed the articles based on the insurance type, the maximum proportion was health insurance (147), followed by general insurance or applying blockchain in the insurance industry as a whole (80), then auto insurance (50) and agri-insurance (18). Also, there were other insurance types, including Cyber Insurance (10), Parametric Insurance (7), Financial Insurance (6), Travel insurance (3), Liability Insurance (2) and property and casualty (P & C) insurance (2) as shown in Table 1.

5 Conclusion and Future Perspectives

As previously mentioned, blockchain is still well in its early stages of implementation within the insurance industry. However, the abundance of products it can offer us in the future is already apparent, and blockchain can heavily impact the macro- and micro-levels of the insurance industry [4]. At the macro-level, researchers believe that blockchain has the potential to overcome the high prices of data acquisition

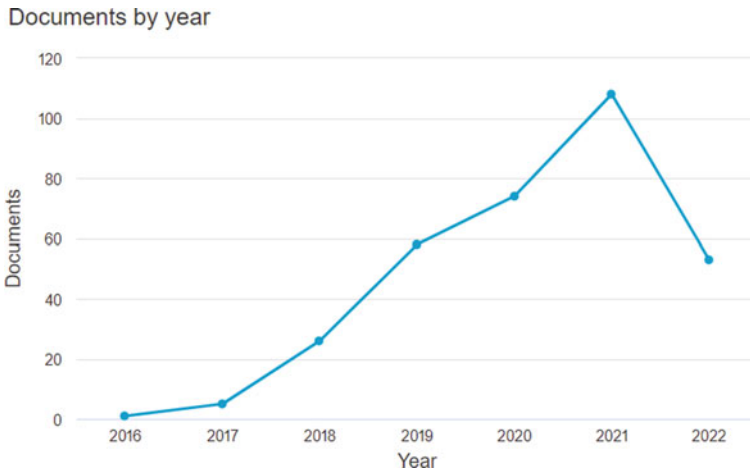


Fig. 3 Scopus results for blockchain and insurance by year

Documents by country or territory

Compare the document counts for up to 15 countries/territories.

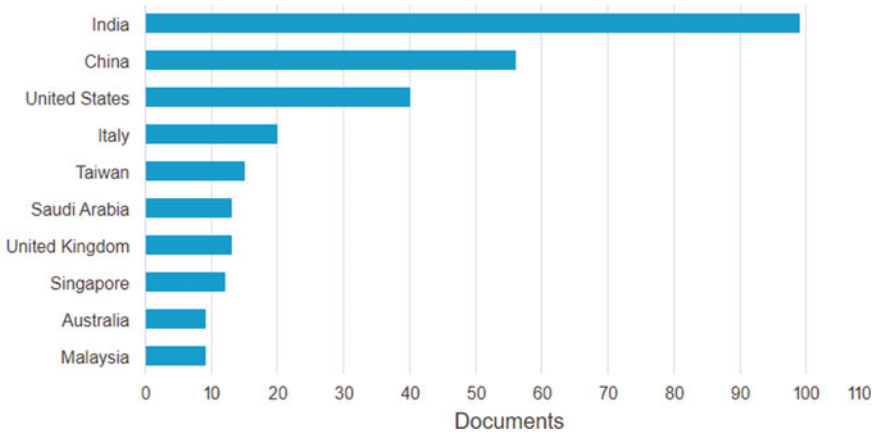


Fig. 4 Scopus results by the country or territory for blockchain and insurance

and modernise the ways data is shared, exchanged and stored. This would benefit small-to-medium-sized insurance agencies, enabling them to access high-quality, reliable, complete information. Moreover, this data would open newer and more diverse opportunities for smaller enterprises using better and more accurate pricing and improved niche market targeting with better product designs. Blockchain may be used in the future through short-term insurance or what is better known as micro-insurance. This type of insurance could utilise blockchain when used in car-sharing or renting accommodation. Currently, these insurance packages are often pre-purchased

Documents by type

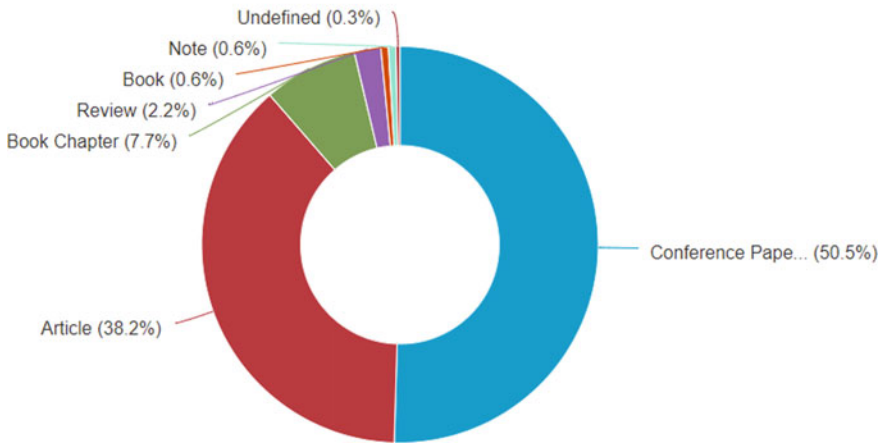


Fig. 5 Scopus results by article type for blockchain and insurance

Table 1 Number of articles by insurance type

S. No.	Insurance type	Number of articles
1	Health insurance	147
2	General insurance	80
3	Auto insurance	50
4	Agri-Insurance	18
5	Cyber insurance	10
6	Parametric insurance	7
7	Financial insurance	6
8	Travel insurance	3
9	Liability insurance	2
10	Property and casualty (P & C) insurance	2

at higher rates by the service provider and then passed on to the end-users. However, with blockchain, clients can purchase insurance at any time. With this, clients can buy insurance for the time they need, based on their usage and expiring time and date. In the case of any claims, records of events will be far more accurate, and potential disputes will surely be avoided. Moving onto the micro-level of Insurance, blockchain can heavily impact product design, pricing and claims services. Blockchain may be implemented in various types of insurance, including parametric insurance. Firstly, parametric insurance covers the probability of a predefined event from happening [52]. As a result, large amounts of real-time analytics are needed and exchanged among third parties. Although this is proven to be an efficient risk

management method, there is still room for improvement. Two examples of parametric insurance include agricultural insurance and flight delay insurance, as a lot of human intervention is required. With the use of blockchain, the data exchange process is greatly enhanced. Also, with smart contracts, human intervention error is significantly reduced; thus, the claim settlement and payment processes are significantly improved. This will lead to lower operational costs, increase efficiency and improve customer satisfaction. Health insurance providers can also exploit the advancement in health data analytics, [1, 17–22, 24, 26, 27, 38–40] to understand and forecast the healthcare resource requirements better, demand and estimate insurance premium. Furthermore, mobile health insurance systems [9] can facilitate better accessibility and broader reach.

References

1. Barton M, McClean S, Garg L, Fullerton K (2010) Modelling costs of bed occupancy and delayed discharge of post-stroke patients. In: 2010 IEEE workshop on health care management (WHCM), February. IEEE, pp 1–6
2. Blockchainhub (2019) Tokenized networks: what is a DAO? <https://blockchainhub.net/dao-decentralized-autonomous-organization/>. Accessed Nov 2022
3. Blockchainhub (2019) Tokens, cryptocurrencies & other cryptoassets. <https://blockchainhub.net/tokens/#:~:text=These%20native%20tokens%20%E2%80%93%20also%20referred,inc%20entive%20scheme%20of%20blockchain%20infrastructure.&text=These%20relatively%20imple%20smart%20contracts,features%20of%20a%20fungible%20commodity>. Accessed Nov 2022
4. Brophy R (2019) Blockchain and insurance: a review for operations and regulation. *J Financ Regul Compliance* 28(2)
5. Built In (2018) Blockchain. <https://builtin.com/blockchain>. Accessed Nov 2022
6. CBinsights (2019) How blockchain could disrupt insurance. <https://www.cbinsights.com/research/blockchain-insurance-disruption/>. Accessed Nov 2022
7. Christidis K (2016) Blockchains and smart contracts for the Internet of Things. *IEEE* 4:2292–2303
8. Chukwu E, Ekong I, Garg L (2022) Scaling up a decentralized offline patient ID generation and matching algorithm to accelerate universal health coverage: insights from a literature review and health facility survey in Nigeria. *Front Digital Health* 4
9. Chukwu E, Garg L, Eze G (2016) Mobile health insurance system and associated costs: a cross-sectional survey of primary health centers in Abuja Nigeria. *JMIR mHealth uHealth* 4(2):e4342
10. Chukwu E, Garg L (2020) A systematic review of Blockchain in healthcare: frameworks, prototypes, and implementations. *IEEE Access* 8:21196–21214
11. Consensus (2020) Blockchain in insurance. <https://consensus.net/blockchain-use-cases/finance/insurance/>. Accessed 12 Nov 2020
12. Crosby M (2016) Blockchain technology: beyond Bitcoin. *AIR—Applied Innov Rev* 1(2):6–18
13. Dominique (2017) Public blockchain versus private blockchain, s.l.: s.n.
14. Etherisc (2018) Democratizing insurance using blockchain. <https://blog.etherisc.com/democratizing-insurance-using-blockchain-2cdac647e957>. Accessed Nov 2022
15. Etherisc (2020) Make insurance fair and accessible. <https://etherisc.com/>. Accessed Nov 2022
16. Euromoney (2019) What is blockchain? <https://www.euromoney.com/learning/blockchain-explained/what-is-blockchain#:~:text=Blockchain%20is%20a%20system%20of,computer%20systems%20on%20the%20blockchain>. Accessed Nov 2022

17. Garg L, McClean S, Meenan B, Millard P (2009a) Non-homogeneous Markov models for sequential pattern mining of healthcare data. *IMA J Manag Math* 20(4):327–344
18. Garg L, McClean S, Meenan B, El-Darzi E, Millard P (2009b) Clustering patient length of stay using mixtures of Gaussian models and phase type distributions. In: 2009 22nd IEEE international symposium on computer-based medical systems, August. IEEE, pp 1–7
19. Garg L, McClean S, Barton M, Meenan B, Fullerton K (2010a) Forecasting hospital bed requirements and cost of care using phase type survival trees. In: 2010 5th IEEE international conference intelligent systems, July. IEEE, pp 185–190
20. Garg L, McClean S, Meenan B, Millard P (2010b) A non-homogeneous discrete time Markov model for admission scheduling and resource planning in a cost or capacity constrained healthcare system. *Health Care Manag Sci* 13(2):155–169
21. Garg L, McClean S, Meenan BJ, Millard P (2011) Phase-type survival trees and mixed distribution survival trees for clustering patients' hospital length of stay. *Informatica* 22(1):57–72
22. Garg L, McClean SI, Barton M, Meenan BJ, Fullerton K (2012) Intelligent patient management and resource planning for complex, heterogeneous, and stochastic healthcare systems. *IEEE Trans Syst, Man, Cybern-Part A: Syst Humans* 42(6):1332–1345
23. Garg L, Chukwu E, Nasser N, Chakraborty C, Garg G (2020) Anonymity preserving IoT-based COVID-19 and other infectious disease contact tracing model. *IEEE Access* 8:159402–159414
24. Garg L, McClean SI, Meenan BJ, Barton M, Fullerton K, Buttigieg SC, Micallef A (2022) Phase-type survival trees to model a delayed discharge and its effect in a stroke care unit. *Algorithms* 15(11):414. <https://doi.org/10.3390/a15110414>
25. Gatteschi V, Lamberti F, Demartini C, Pranteda C, Santamaria V (2017) Blockchain and smart contracts for insurance: is the technology mature enough? *Future Internet* 10(2)
26. Gillespie J, McClean S, Scotney B, Garg L, Barton M, Fullerton K (2011) Costing hospital resources for stroke patients using phase-type models. *Health Care Manag Sci* 14(3):279–291
27. Gillespie J, McClean S, Garg L, Barton M, Scotney B, Fullerton K (2016) A multi-phase DES modelling framework for patient-centred care. *J Oper Res Soc* 67(10):1239–1249
28. Chen G, Xu B, Lu M (2018) Exploring blockchain technology and its potential applications for education. *Smart Learn Environ* 5
29. Gupta V (2017). A brief history to blockchain. <https://hbr.org/2017/02/a-brief-history-of-blockchain>. Accessed Nov 2022
30. Harrington J (2017) The present use and promise of blockchain in insurance. Wells Media Group Inc., San Diego
31. Hugh T (2017) DYNAMIS—Ethereum-based DAO for distributed P2P insurance. <https://www.the-digital-insurer.com/dia/dynamis-ethereum-based-dao-for-distributed-p2p-insurance/>. Accessed Dec 2020
32. Iredale G (2018) 6 key blockchain features you need to know now. <https://101blockchains.com/introduction-to-blockchain-features/>. Accessed Nov 2022
33. Jain A (2019) What are the three pillars of blockchain technology? <https://medium.com/@anijain/what-are-the-three-pillars-of-blockchain-technology-9ed9ca3bd754>. Accessed Nov 2022
34. Khekade A (2018) If you thought blockchain was amazing, wait till you read about hybrid blockchain. <https://entrepreneur.com/article/307794>. Accessed Nov 2022
35. Kosba A (2016) Hawk: The blockchain model of cryptography and privacy-preserving smart contracts. *IEEE Symp Secur Privacy* 55
36. Kot I (2020) Blockchain in insurance: 3 use cases. <https://www.insurancethoughtleadership.com/3-big-use-cases-of-blockchain-in-insurance/>. Accessed Nov 2022
37. Manley G (2019) The history of blockchain. <https://www.section.io/engineering-education/history-of-blockchain/>. Accessed Nov 2022
38. McClean S, Garg L, Meenan B, Millard P (2007). Using Markov models to find interesting patient pathways. In: Twentieth IEEE international symposium on computer-based medical systems (CBMS'07), June. IEEE, pp 713–718
39. McClean S, Barton M, Garg L, Fullerton K (2011) A modeling framework that combines markov models and discrete-event simulation for stroke patient care. *ACM Trans Model Comput Simul (TOMACS)* 21(4):1–26

40. McClean S, Gillespie J, Garg L, Barton M, Scotney B, Kullerton K (2014) Using phase-type models to cost stroke patient care across health, social and community services. *Eur J Oper Res* 236(1):190–199
41. Nam S (2018) How much are insurance consumers willing to pay for blockchain and smart contracts? A contingent valuation study. *Sustainability* 10(11):4332
42. Nofer M (2017) Blockchain. *Bus Inf Syst Eng* 59
43. Popovic (2020) Understanding blockchain for insurance use cases, s.l.: Institute and Faculty of Actuaries
44. PWC (2020) Blockchain, a catalyst for new approaches in insurance, s.l.: s.n.
45. Rawlings P (2017) Insurance fraud and the role of the civil law. *Mod Law Rev* 80(3):525–539
46. Rosic A (2019) What is blockchain technology? A step-by-step guide for beginners. <https://blockgeeks.com/guides/what-is-blockchain-technology/>. Accessed Nov 2022
47. Salmon J (2019) Blockchain and associated legal issues. *Emcompass* 63
48. Sayeed S (2019) Assessing blockchain consensus and security mechanisms against the 51% attack. *Adv Blockchain Technol Appl* 9(9)
49. Sultan K (2018) Conceptualizing blockchain: characteristics and applications. *Int Conf Inf Syst* 11:49–57
50. Swan M (2015) Blockchain—Blueprint for a new economy. 1st ed. O’Reilly Media, Sebastopol
51. Swanson T (2015) Consensus-as-a-service: a brief report on the emergence of permissioned, distributed ledger systems, s.l.: s.n.
52. Swiss Re (2018) What is parametric insurance. https://corporatesolutions.swissre.com/insights/knowledge/what_is_parametric_insurance.html. Accessed Nov 2022
53. Tarr JA (2018) Distributed ledger technology, blockchain and insurance: opportunities risks and challenges. *Insurance Law J* 29(3):254–268

CNN-LSTM: A Deep Learning Model to Detect Botnet Attacks in Internet of Things



Chetanya Kunndra, Arjun Choudhary, Prashant Mathur, Kapil Pareek, and Gaurav Choudhary

1 Introduction

With the increased penetration of technology in our day to day lives, we usually come across the term ‘*Internet Of Things*’ or *IoT*. To explain this concept in simple terms, IoTs are nothing but physical world objects that are connected to the internet with the help of sensors [1]. An internet connected camera or even an internet connected thermostat can all be considered as an IoT. Connecting physical world objects to the internet is not a new concept. As a matter of fact, the first internet connected device, a toaster, that could be operated over the internet, was showcased in the Internet Conference in 1990 [2]. With humanity’s innate need to reduce the amount of work that goes into doing anything. IoT since then has found itself being used for automating daily recurring tasks that initially required human intervention, such as controlling temperatures, operating machineries or industry lines. Soon with the onset of the concept of smart devices, IoT was transitioned into smart IoT. Such smart IoTs were then used to create smart systems, which were not only used by individual consumers but also found their use in the industry.

C. Kunndra (✉) · A. Choudhary · P. Mathur · K. Pareek
Sardar Patel University of Police Security and Criminal Justice, Jodhpur, India
e-mail: mtcs20ck@policeuniversity.ac.in

A. Choudhary
e-mail: a.choudhary@policeuniversity.ac.in

P. Mathur
e-mail: mtcs20pm@policeuniversity.ac.in

K. Pareek
e-mail: spu19cskp@policeuniversity.ac.in

G. Choudhary
Technical University of Denmark, Lyngby, Denmark

IoT's can be classified into two broad categories. These classifications are done on the basis of the sector employing the IoT device and the smartness of the device.

On the basis of the area of use, we can classify IoT's into the following two categories:

- **Consumer IoT (CIoT)**—These are IoT devices that are used by consumers, normal people for their everyday use, smart devices such as smart ACs, smart thermostats come under this category [3], based on the device, they can perform a multitude of tasks.
- **Industrial IoT (IIoT)**—Industrial IoT's are manufactured for the sole purpose of being used in the industry. Unlike CIoT, IIoT's are more designed to be durable, they are ruggedized to counter the harsh working conditions of industrial processes and programmed to perform a specific task with utmost precision and accuracy [4].

One can also classify IoT's on the basis of the intelligence of the device, based on this classification scheme, there can be two major categories of IoT's:

- **Smart IoT**—This category of IoT devices are paired with sensors and actuators and are programmed to perform smart actions based on the physical inputs, these actions can vary from dynamically controlling the physical environment to performing automated mechanical tasks [5].
- **Non smart IoT**—Unlike smart IoT devices non-smart IoT devices do not possess any abilities to dynamically interact with the physical environment based upon the inputs or perform any smart actions all together. They can only perform a set of restrictive actions that are programmed into them.

Figure 1 illustrates various components that are present in an IoT ecosystem and how an IoT device interacts with the physical environment.

IoT's since its conception have found their use in almost every niche of the society. IoT's have become an integral part of modern society. Even though a typical person might not use them directly in their daily lives, IoT's are affecting our lives indirectly.

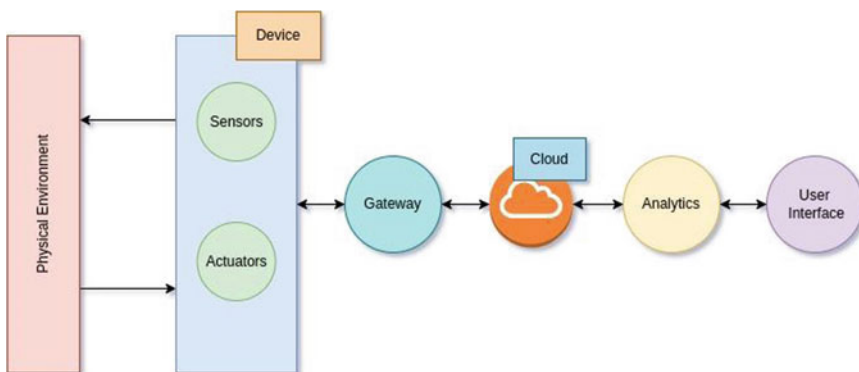


Fig. 1 Basic representation of various components in IoT

To elaborate on this statement, below are some of the few critical sectors that employ and benefit from the use IoTs:

- **Agriculture**—The agriculture sector is a human-intensive sector. The use of IoTs has proven to be beneficial and a boon for it. IoT has been integrated into traditional machinery to increase the efficiency of agriculture practices [6]. IoT has been used for crop management, livestock management [7]. Smart IoTs have given rise to smart farming, with drones that are built to monitor soil, air and water quality. Overall the state of farming has enhanced with the introduction of IoTs.
- **Healthcare**—Healthcare is one of the most important sectors in which IoTs are being used. With smart wearable technology, they can help in early detection and thereby timely prevention of certain conditions, such as issues related to the human heart [8]. Other than that, IoTs are also being used to maintain a unique digital identity of each patient, relatively easing up the process of data collection and assessment [9].
- **Transport**—Self-driving automobiles and smart transportation have been increasing in popularity. IoT devices are readily and extensively being used in this sector. One can see their applications in real-time vehicle tracking [10], automated toll tax collection and even in smart cars [11].
- **Automation**—IoT is now being readily adopted for automating daily recurring tasks not only in our homes but at our workplaces as well. The amalgamation of IoT, artificial intelligence and regular home devices has given birth to smart home devices such as Amazon Echo, Google Assistant which are being used to control and automate daily tasks within our homes [12].
- **Manufacturing/Industrial Processes**—The industrial sector is the one that has benefited the most due to use of IoT devices, IoTs are used to control and monitor industrial processes, which initially required human intervention, in doing so they provide increased precision and accuracy, extremely low error rate and the ability to remotely control the process over the internet. Moreover, their use in the critical industries has yielded beneficial results. The industry using IoT is also dubbed as ‘**Industry 4.0**’ [13].
- **Energy**—Smart IoT devices have become a boon to the energy sector, while providing the energy sector with remote monitoring capabilities [14], IoTs have also proven useful in providing a way to minimize energy wastage thereby making it more sustainable.

Overall, IoTs are used in numerous sectors, these are just a few examples, they can also be found in waste management and disposal sectors, sectors related to water management and even the aviation industry.

The recent developments in 5G communication technologies have brought us faster communication speeds, low latency connections and secure communication links, not just to personal smartphones, computers but also to the entire domain of Internet of Things [15]. 5G ensures better-connected networks for IoT and is considered a game changer for every form of IoT, ranging from your smart assistant that controls your AC, to the sensors that help a self-driving car change lanes [16]. The criticality of IoT in our day to day lives also brings out the urgent need to protect

them from attacks from adversaries, these attacks are not just limited to cyberattacks but can also be physical attacks. As opposed to physical attacks, cyberattacks pose a much larger threat to IoTs. The anti-virus giant Kaspersky in a report stated that from January to June 2021, there had been nearly 1.15 billion breaches in IoT, which is a substantial increase from a mere 639 million breaches in 2020 [17].

One of the most notable cyber threats that comes to mind when IoT security is discussed is the Mirai malware. Mirai is capable of controlling any IoT device, running an ARC processor and has its default credentials set [18], turning the infected IoT device into a zombie and part of its botnet. The exact size of the Mirai botnet is unclear but it is known that Mirai worm infected at least 65,000 devices within the first 20 h and between 2016 and 2017, boasted of at least 600,000 devices in its botnet [19]. Mirai botnet can also be linked to some of the most massive Distributed Denial of Service (DDoS) attacks the cyberspace has ever seen. Some of its notable targets include Dyn, a major DNS service provider [20] (1 Terabit/second traffic from Mirai) and an attempted take-down of Lonestar Cell, a major Liberian telecom provider. Since the release of Mirai source code, the cyberspace has seen many new variants of the original Mirai malware, the most recent one being the ‘wormable’ log4j variant of mirai [21] can be considered a highly critical threat, primarily due to the underlying critical nature of the log4j bugs [22–24].

Due to their critical nature and an ever increasing threat landscape, IoT devices should be protected. It is estimated that there will be more than 27 billion live IoT devices connected to the internet installed by 2025 [25]. While traditional security mechanisms do work to a certain extent, they however do not suffice, newer emerging technologies relying on the fields of machine learning and deep learning to intelligently tackle the problems of protecting IoT devices are being devised and are also the need of the hour. This paper proposes a novel hybrid deep learning model to detect compromised IoT devices by analyzing the network traffic emerging from the IoT.

Section 2 of this paper discusses other machine learning and deep learning models that have been proposed to tackle the problem of IoT security. Section 3 gives a brief overview of botnets with respect to IoT and how a particular IoT interacts with a C2 server. Section 4 explains our proposed model, Sect. 5 briefly explains the experiments that were conducted to test the model, Sect. 6 explains the evaluation metrics employed to test the effectiveness of the model, Sect. 7 elaborates on the results of the experiments performed in Sects. 5 and 7 is followed by the conclusion.

2 Related Work

IoT is being used predominantly everywhere, with the advent of 5G, which supports IoT. This huge number proposes itself that IoT security is one of the top concerns in the cyber security domain. There have been numerous research in the field of machine learning and protecting IoT using it.

Table 1 Comparative table of papers on malware detection using machine learning

Paper	Dataset	Accuracy (%)
Bhatt et al. [26]	CTU-13	81.80
Parraa et al. [27]	N-BaloT	94.80
Alzahrani et al. [28]	N-BaloT	88.62
Yin et al. [29]	ISCX 2012, ISOT	74.04
Our model	CTU-13	98

Bhatt et al. [26] propose a two-step progressive algorithm that uses an ensemble-based classifier, graph structure-based anomaly detection and KNN algorithm to determine botnet from network flow. The ensemble-based classifier efficiently creates instances of the network traffic that is then used in their graph-based anomaly detection. Their proposed framework is built using the CTU-13 dataset and has an average accuracy of 81.8%.

Parraa et al. [27] propose a deep learning-based security add-on for IoT devices, their research proposes two different deep learning add-ons, one for detecting DDoS attacks that work on the application layer and phishing and another module for detecting Botnet attacks. Their add-on for botnet attack detection uses Long-Short Term Memory network model, while their phishing attack module uses Distributed CNN model. Their botnet add-on is built using the N BaIoT dataset and gives an accuracy of 94.80%.

Alzahrani et al. [28] combined two deep learning algorithms, namely, CNN and LSTM to propose a model that can detect BASHLITE and Mirai botnet attacks on IoT. Their research uses a dataset curated from various connected cameras, collected in real time. Their proposed model shows an average accuracy of 88.62%.

Yin et al. [29] propose a novel botnet detection framework, built using Generative Adversarial Networks (GAN) called Bot-GAN, their proposed framework is essentially a standard GAN discriminator, replaced with a botnet detector, with the use of the softmax function they are able to convert the binary classifier into a ternary classifier. They test their proposed framework on ISCX 2012 and ISOT datasets, their proposed framework gives a peak accuracy of 74.04%.

Table 1 gives a comparative analysis of various researches done in the field.

3 Overview

In order to understand the issue of botnets in IoT, we need to get acquainted with certain terms:

- **Bots**—Bots are devices, which once infected by a botnet malware, execute commands given to them by the attacker. They become **robots** of the attacker [30].

- **Botnets**—A botnet can be considered as a network of multiple bots, typically in a botnet network, two bots do not interact with each other, but rather only interact with the attacker. Typically a small botnet can consist of a few hundred bots, whereas a large botnet can have more than 50,000 bots, it is difficult to estimate the exact number of bots in a botnet [31].
- **C2 Server**—Command and Control Server or C2 server or C&C server is the master server of all the bots [32], all bots send and receive data from the C2 server, which is under direct control of the attacker.

Figure 2 illustrates a typical botnet network architecture and how an attacker controls bots using C2 servers.

IoT, just like any other computer resource, is vulnerable due to software vulnerabilities that have unintentionally crept up in the system. Malwares such as Mirai and BASHLITE take advantage of such vulnerabilities to hijack an IoT device and make it part of their botnets. Once part of the botnet, the device is at the mercy of the attacker. Attackers use this botnet to commit nefarious actions, whilst monetary

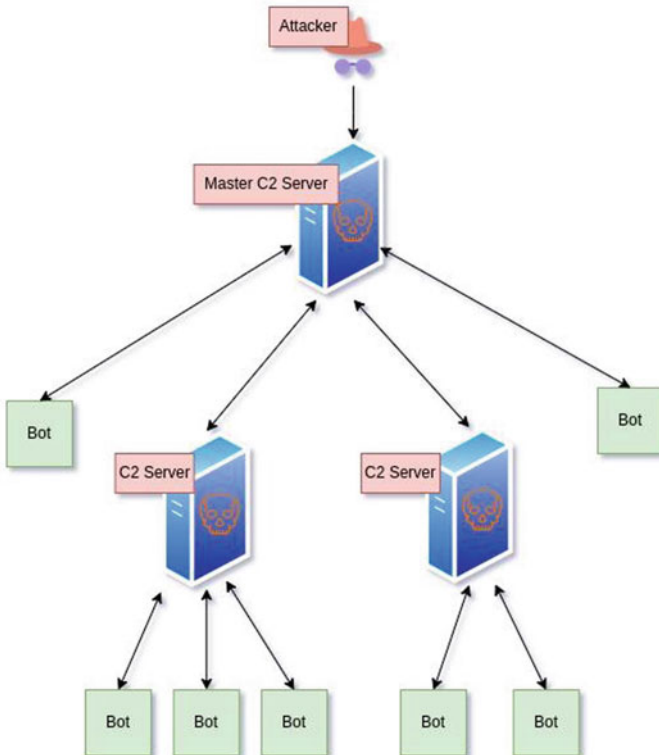


Fig. 2 A typical botnet network

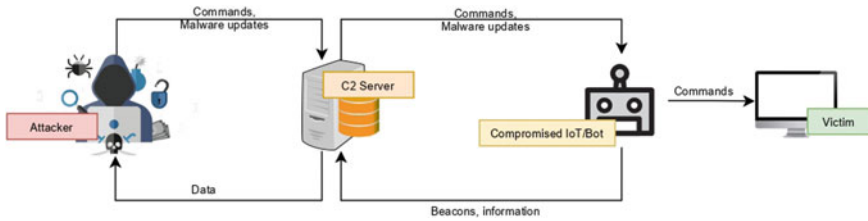


Fig. 3 Network interactions of a compromised bot with the attacker

gains being the primary objective. The attacker can set up botnet-as-a-service and render services such as Distributed Denial of Service Attacks (DDoS) [33] to clients in exchange for monetary gains.

Once an IoT device is part of the botnet, it remains in constant touch with the attacker, via the C2 server. Once a device is infected the network traffic of an infected IoT device will contain beacons, to elaborate beacons can be thought of as pings to the C2 server, essentially telling the C2 server that the particular IoT device is active and still part of the botnet [34]. Beacons are one way to identify a plausible botnet, apart from beacons there are other network traffic indicators that can help identify a botnet, these can include commands sent by the attacker or even updates to the malware sent by the C2 server. Figure 3 illustrates various forms of network interactions between a compromised bot and the attacker. Such network interactions are crucial in identification of a compromised device.

Depending on the area of use, the amount of network traffic flowing to and from a device is tremendous. Manually viewing the network traffic for anomalies is a hassle and a humanly taxing task. Artificial intelligence can be leveraged to ease up this process of detection. With automated systems and deep learning models in place, identification of botnet traffic from the network traffic can become relatively easy and a hassle-free task. Our proposed model, which is built using Convolution Neural Network (CNN) and Long-Short Term Machine (LSTM) algorithms, is trained, validated and tested on the CTU-13 dataset [35]. The proposed model aims at identifying compromised IoT devices by analyzing the network traffic.

4 Proposed Model

CNNs are a form of feed-forward neural network that are primarily used for processing images and natural language due to their high efficiency. They can even be used to successfully forecast time series [36]. The use of CNN’s weight sharing and local perception can greatly reduce the number of parameters used to train the model, making the training phase have less space and computational complexity, thereby increasing overall efficiency. The two essential components of CNN are the convolution layer and the pooling layer. The number of convolution kernels in each

convolution layer may vary depending upon the problem that is to be solved, the number of kernels to be used in each layer can be calculated using this mathematical equation -

$$l_t = \tanh(x_t * k_t + b_t) \tag{1}$$

The features extracted after the convolution operation have very large dimensionality, thus a pooling layer is used after the convolution layer to reduce the feature dimension and thereby reducing the cost of training the model.

LSTM is a deep learning model proposed by Schmidhuber et al. in 1997 [37]. Unlike CNN, LSTM are majorly used to tackle problems of speech recognition, text analysis, analysis of emotions, their use in such areas are largely attributed to their design, LSTM have their own memory and thus can make relatively precise predictions and forecasting. The LSTM memory cell comprises three modules, namely, the input, forget and output gates, the remaining structure of LSTM is similar to that of RNN.

A classification model can be created for a complex problem using an amalgamation of CNN and LSTM algorithms. In this amalgamation, features are extracted with the help of CNN, the output of the CNN layer is then fed into the LSTM layer as an input. The amalgamation of CNN and LSTM yields a deep learning model that has highly optimized space and computational complexity. These combinations have multidimensional uses in both research and industrial areas.

Figure 4. figuratively depicts the logical structure of the proposed model and the various layers that work in sequence to tackle the problem of botnets in the IoT.

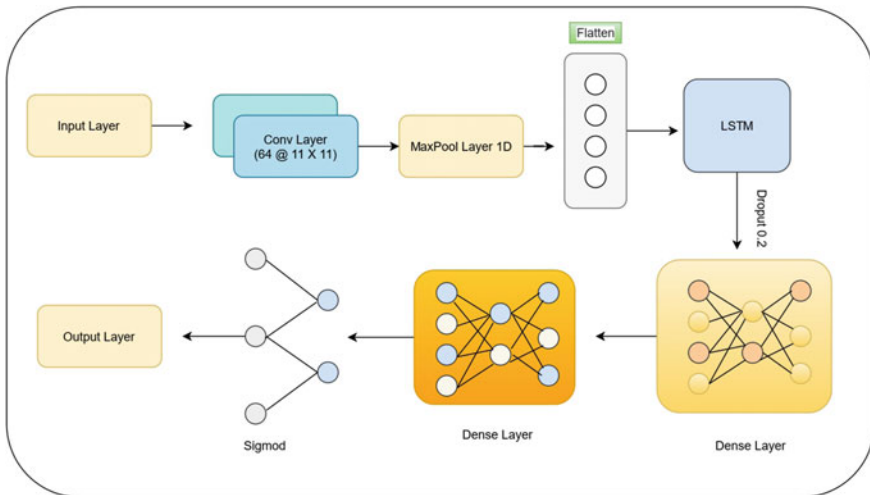


Fig. 4 Proposed CNN + LSTM model

The proposed model comprises numerous hidden layers, namely: convolution layer, maxpool layer, flatten layer, LSTM layer, dense layer and softmax function. Data features are extracted from the network traffic using the convolution layer, a data preprocessing step is involved that data features without losing any of the attributes. The CNN layer comprises 64 filters. The CNN layer is followed by the Max Pool layer, which outputs the maximum value out of the data vector matrices for the next layer. Post ‘Max Pool’, ‘flattening’ is performed, which merges all visible layers into the background layer to reduce output size. The output layer obtained from flatten is inserted into the LSTM layer in the form of input layer. The dropout rate in LSTM determines what value of the neuron should be discarded from it. A Dense Layer is just a simple layer of neurons in which each neuron receives input from all the neurons of the previous layer. Softmax function extends the binary class world into a multi-class world, to elaborate, Softmax attributes decimal probabilities to each class in a multi-class problem. These decimal probabilities must add up to 1.0.

5 Experiments

To compute our CNN-LSTM model’s effectiveness, we compared it to CNN-RNN, CNN, LSTM, RNN and MLP models, using the same dataset and in the same computational environment. Figure 5 illustrates the logical flow of the experiments performed. All of the tests were performed on a 2.6 GHz Intel i7-8850H CPU with 16 GB RAM, 500 GB of hard disk space and Ubuntu 20.04 LTS operating system. As a result of our experiments, we found out that our proposed deep learning model showed the best results and accuracy for botnet detection in the IoT as compared to other models.

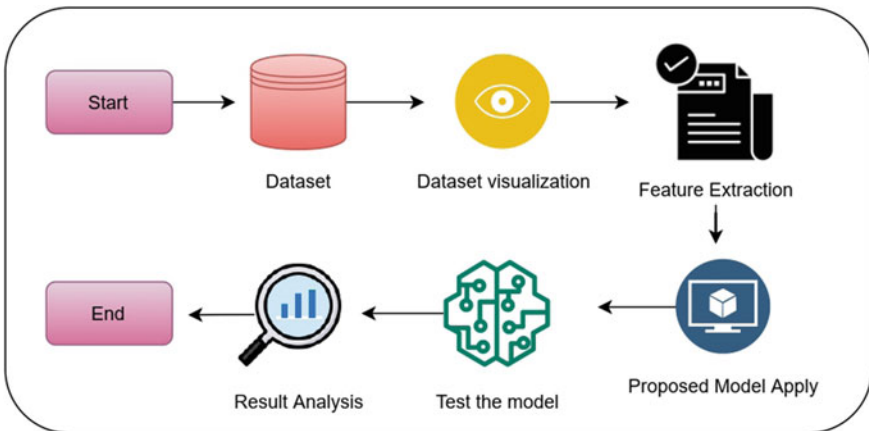


Fig. 5 Logic flow of experiments

6 Evaluation Metrics

In order to evaluate the proposed model, we need to understand the following terminology:

- **True Positives**—True positives (T_P) refers to the predictions that are labeled as ‘True’ and are also ‘True’.
- **True Negatives**—True negatives (T_N) refers to the predictions that are labeled as ‘False’ and are also ‘False’.
- **False Positives**—False positives (F_P) refers to the predictions that are labeled as ‘True’ but are actually ‘False’.
- **False Negatives**—False negatives (F_N) refers to the predictions that are labeled as ‘False’ but are actually ‘True’.

Keeping the above terms in mind, the proposed model is evaluated by calculating the following metrics:

- **Accuracy**—Accuracy is equivalent to the ratio of the total number of correct predictions, to the total number of predictions made by the model. Therefore, accuracy of a model can be defined as

$$A_{cc} = (T_P + T_N) / (T_P + T_N + F_P + F_N). \quad (2)$$

- **Error Rate**—Error rate of a model is equivalent as the ratio of the total number of incorrect predictions, to the total number of predictions made by a model. Thus its mathematical representation is defined as

$$E_{rr} = (F_P + F_N) / (T_P + T_N + F_P + F_N). \quad (3)$$

7 Results

The proposed model performed better as compared to other models that could be used in the problem of botnet detection. Figure 6 depicts the comparisons between the performance of various machine and deep learning algorithms and our proposed model. This model has been trained and tested on 7588 samples and has an accuracy of 98%. One of the advantages of supervised learning is that we can use testing sets to get objective measurements of the learning process. Accuracy is one metric for evaluating classification models. Table 2 shows the results of our experiment run on the CTU-13 dataset using our proposed model.

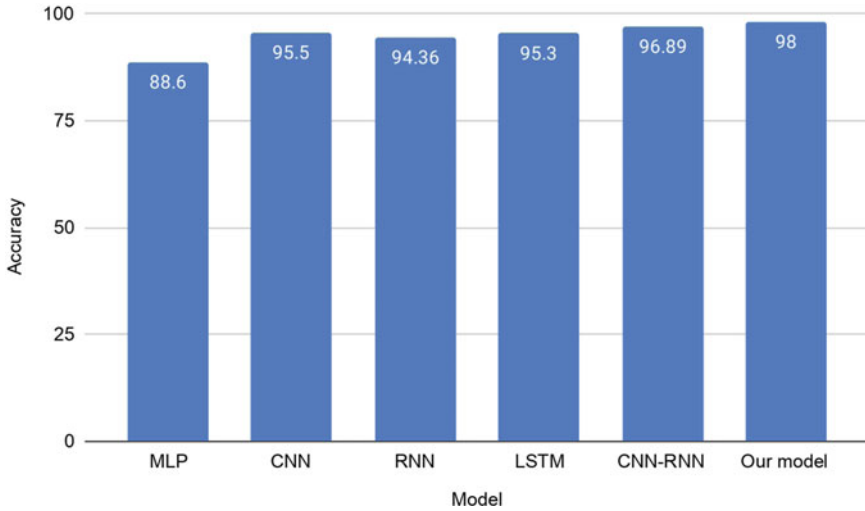


Fig. 6 Comparison of accuracy of various models on the CTU-13 dataset

Table 2 Performance of the proposed model on CTU-13 dataset

Model	Dataset	Accuracy	Error rate
Proposed CNN-LSTM model	CTU-13	0.98	0.23

8 Conclusion

This paper presents a CNN-LSTM-based deep learning model for botnet detection in IoT which relies on network flow to actively perform detection of attacks. This proposed model can be used to protect highly vulnerable IoT devices that are being used everywhere. Features extracted from the network flow of the IoT devices are used to identify anomalous behavior. This proposed model has shown an accuracy of 98% and shows better performance as compared to other proposed models that are built using the CTU-13 dataset.

References

1. Rose K, Eldridge S, Chapin L (2015) The internet of things: an overview. Internet Soc (ISOC) 80:1–50
2. Albishi S, Soh B, Ullah A, Algarni F (2017) Challenges and solutions for applications and technologies in the Internet of Things. *Procedia Comput Sci* 124:608–614
3. Doshi R, Apthorpe N, Feamster N (2018) Machine learning DDoS detection for consumer internet of things devices. In 2018 IEEE security and privacy workshops (SPW). IEEE, pp 29–35

4. Boyes H, Hallaq B, Cunningham J, Watson T (2018) The industrial internet of things (IIoT): An analysis framework. *Comput Ind* 101:1–12
5. Zheng X, Cai Z, Li Y (2018) Data linkage in smart internet of things systems: a consideration from a privacy perspective. *IEEE Commun Mag* 56(9):55–61
6. Stočes M, Vaněk J, Masner J, Pavlík J (2016) Internet of things (IoT) in agriculture—selected aspects. *Agris On-Line Pap Econ Inform* 8(665–2016–45107):83–88
7. Alonso RS, Sittón-Candanedo I, García Ó, Prieto J, Rodríguez-González S (2020) An intelligent Edge-IoT platform for monitoring livestock and crops in a dairy farming scenario. *Ad Hoc Netw* 98:102047
8. Li C, Hu X, Zhang L (2017) The IoT-based heart disease monitoring system for pervasive healthcare service. *Procedia Comput Sci* 112:2328–2334
9. Kodali RK, Swamy G, Lakshmi B (2015) An implementation of IoT for healthcare. In: 2015 IEEE recent advances in intelligent computational systems (RAICS). IEEE, pp 411–416
10. Sridevi K, Jeevitha A, Kavitha K, Sathya K, Narmadha K (2017) Smart bus tracking and management system using IoT. *Asian J Appl Sci Technol (AJAST)* 1
11. Krasniqi X, Hajrizi E (2016) Use of IoT technology to drive the automotive industry from connected to full autonomous vehicles. *IFAC-PapersOnLine* 49(29):269–274
12. Stojkoska BLR, Trivodaliev KV (2017) A review of Internet of Things for smart home: challenges and solutions. *J Clean Prod* 140:1454–1464
13. Wan J, Chen B, Imran M, Tao F, Li D, Liu C, Ahmad S (2018) Toward dynamic resources management for IoT-based manufacturing. *IEEE Commun Mag* 56(2):52–59
14. Hossein Motlagh N, Mohammadrezaei M, Hunt J, Zakeri B (2020) Internet of Things (IoT) and the energy sector. *Energies* 13(2):494
15. Akpakwu GA, Silva BJ, Hancke GP, Abu-Mahfouz AM (2017) A survey on 5G networks for the Internet of Things: communication technologies and challenges. *IEEE Access* 6:3619–3647
16. Dangi R, Lalwani P, Choudhary G, You I, Pau G (2021) Study and investigation on 5G technology: a systematic review. *Sensors* 22(1):26
17. NEC (2022) 5 of the biggest threats to Cyber Security in 2022 - NEC New Zealand. NEC. <https://www.nec.co.nz/market-leadership/publications-media/5-of-the-biggest-threats-to-cyber-security/>. Accessed 28 May 2022
18. Xu Y, Koide H, Vargas DV, Sakurai K (2018) Tracing Mirai malware in networked system. In: 2018 sixth international symposium on computing and networking workshops (CANDARW). IEEE, pp 534–538
19. Antonakakis M, April T, Bailey M, Bernhard M, Bursztein E, Cochran J, Zhou Y (2017) Understanding the Mirai botnet. In: The 26th USENIX security symposium (USENIX Security 17), pp 1093–1110
20. Scott Sr J, Summit W (2016) Rise of the machines: the dyn attack was just a practice run December 2016. Institute for Critical Infrastructure Technology, Washington, DC, USA
21. Hiesgen R, Nawrocki M, Schmidt TC, Wählisch M (2022) The race to the vulnerable: measuring the log4j shell incident. [arXiv:2205.02544](https://arxiv.org/abs/2205.02544)
22. National Vulnerability Database (2021) CVE-2021-44228 Detail. NVD - CVE-2021-44228. <https://nvd.nist.gov/vuln/detail/CVE-2021-44228>. Accessed 27 May 2022
23. National Vulnerability Database (2021) CVE-2021-45046 Detail. NVD - CVE-2021-45046. <https://nvd.nist.gov/vuln/detail/CVE-2021-45046>. Accessed 27 May 2022
24. National Vulnerability Database (2021) CVE-2021-45105 Detail. NVD - CVE-2021-45105. <https://nvd.nist.gov/vuln/detail/CVE-2021-45105>. Accessed 27 May 2022
25. Gamal I, Abdel-Galil H, Ghalwash A (2022) Osmotic message-oriented middleware for Internet of Things. *Computers* 11(4):56
26. Bhatt P, Thakker B (2021) A novel forecastive anomaly based botnet revelation framework for competing concerns in Internet of Things. *J Appl Secur Res* 16(2):258–278
27. Parra GDLT, Rad P, Choo KKR, Beebe N (2020) Detecting Internet of Things attacks using distributed deep learning. *J Netw Comput Appl* 163:102662
28. Alzahrani MY, Bamhdi AM (2021) Hybrid deep-learning model to detect botnet attacks over Internet of Things environments. *Soft Comput* 2022:1–15

29. Yin C, Zhu Y, Liu S, Fei J, Zhang H (2018) An enhancing framework for botnet detection using generative adversarial networks. In 2018 international conference on artificial intelligence and big data (ICAIBD). IEEE, pp 228–234
30. Feily M, Shahrestani A, Ramadass S (2009) A survey of botnet and botnet detection. In: 2009 third international conference on emerging security information, systems and technologies. IEEE, pp 268–273
31. Freiling FC, Holz T, Wicherski G (2005) Botnet tracking: exploring a root-cause methodology to prevent distributed denial-of-service attacks. In: European symposium on research in computer security. Springer, Berlin, pp 319–335
32. Alberts DS, Hayes RE (2006) Understanding command and control. Assistant secretary of defense (C3I/Command Control Research Program) Washington DC
33. Kebande VR, Venter HS (2014) A cloud forensic readiness model using a Botnet as a Service. In: The international conference on digital security and forensics (DigitalSec2014). The Society of Digital Information and Wireless Communication, pp 23–32
34. Richer TJ (2017) Entropy-based detection of botnet command and control. In: Proceedings of the Australasian computer science week multiconference, pp 1–4
35. Garcia S, Grill M, Stiborek J, Zunino A (2014) An empirical comparison of botnet detection methods. *Comput Secur* 45:100–123
36. Khan A, Sohail A, Zahoora U, Qureshi AS (2020) A survey of the recent architectures of deep convolutional neural networks. *Artif Intell Rev* 53(8):5455–5516
37. Hochreiter S, Schmidhuber J (1997) Long short-term memory. *Neural Comput* 9(8):1735–1780

Prediction of Covid-19 Using Artificial Intelligence [AI] Applications



**R. Kishore Kanna, Mohammed Ishaque, Bhawani Sankar Panigrahi,
and Chimaya Ranjan Pattnaik**

1 Introduction

Covid is identified as a respiratory illness. Older people and people with underlying medical conditions are most likely to be seriously ill and in need of critical care. The primary instance of Covid-19 in China was accounted for on December 1, 2019. It has been identified as COVID-19. Researchers have initially concluded that the cause of the illness could have been originated from animals and later mutated to transmit to humans [1]. It is believed that the COVID-19 outbreak was caused by a zoonotic origin. It can be easily spread by people. It began at the fish market in Huanan, Wuhan. As of March 12, 2020, over 81,000 human infections with SARS-CoV-2 and a minimum of 3100 deaths related to COVID-19 are confirmed in China alone [2]. The COVID-19 might prompt some genuine respiratory intricacies like intense respiratory pain condition (ARDS). ARDS requires ICU confirmation and oxygen treatment [3]. A few side effects of COVID-19 are fever, cough, chest pain, shortness of breath, diarrhea, fatigue, etc. The asymptomatic cases are highly contagious, and they progress rapidly, leading to a high fatality rate. For the rapid development of a

R. K. Kanna (✉)

Department of Biomedical Engineering, Jerusalem College of Engineering, Chennai, India
e-mail: kishorekanna007@gmail.com

M. Ishaque

Department of Computer Science and Information Technology, Jeddah International College,
Jeddah, Saudi Arabia
e-mail: m.ishaq@jiccollege.edu.sa

B. S. Panigrahi

Department of Computer Science & Engineering (AI&ML), Vardhaman College of Engineering
(Autonomous), Hyderabad, India

C. R. Pattnaik

Department of Computer Science & Engineering, Ajay Binay Institute of Technology, Cuttack,
India

diagnostic test of COVID-19 gene sequencing of the disease has been employed [3]. The high amount of false-negative results of the RT-PCR test will potentially increase the problem in managing the outbreak; the misdiagnosed patients might miss the most effective timing for correct treatment and cause the spread of the disease [4]. So, to avoid the misdiagnosed issues faced by clinicians due to the low viral content in the test sample or uneven distribution of the samples, some imaging techniques have been employed to rectify or improvise such issues. Chest computed axial tomography is recommended for early identification and prediction of asymptomatic individuals [5].

Due to its similarity with pneumonia cases artificial intelligence has been used to assist the CT images of a COVID-19 patient to differentiate it from other clinical features. The AI-based deep learning technique can localize different entities and classify them from COVID-19 images, this method has achieved an accuracy of 90% [7].

2 Methodology

CT Imaging

Computed Tomography (CT) is an imaging strategy used to examine the internal organs of the body. It is otherwise called modernized tomography or electronic axial tomography (CAT). CT machines take constant pictures in a helical way as opposed to taking a progression of pictures of individual cuts of the body [6]. Helical CT is used because it is faster, produces better quality 3D images of the internal organs, and may also predict small abnormalities better. Computed tomography imaging is highly sensitive in predicting early disease, assessing the abnormalities, progression of the disease, etc. [7]. The CT image of COVID-19 is described by glass opacity, patchy lesions, crazy-paving patterns, and some areas of consolidation [10]. They have disseminated alongside the lung bronchovesicular groups and subpleural space. COVID-19 is predicted using chest CT in four different stages.

1. Early-stage: Single and numerous dispersed fixed ground-glass opacity are conveyed in the peripheral and subpleural spaces of the lung. The interlobular and intralobular septal thickening prompts a crazy-paving pattern at this stage [8].
2. Advanced stage: Expanded degree and thickness of two-sided lung parenchymal haziness can be seen. There are areas of ground-glass opacification and area of consolidation in both lungs, which co-exist with variation in size and presence of air bronchogram [9].
3. Severe stage: The CT scans of diffuse consolidation of the lungs at this stage reveal variations in density due to fibrous exudate into the alveolar cavity, air bronchograms, and bronchial dilation [10]. A portion of the lung that appears as patchy ground-glass opacity is required to prove. When all its components are integrated, the lungs seem to be “whited out.”
4. Dissipation stage: Ground glass opacities and solidification show resolution, leaving some remaining curvilinear spaces of thickness (Fig. 1).

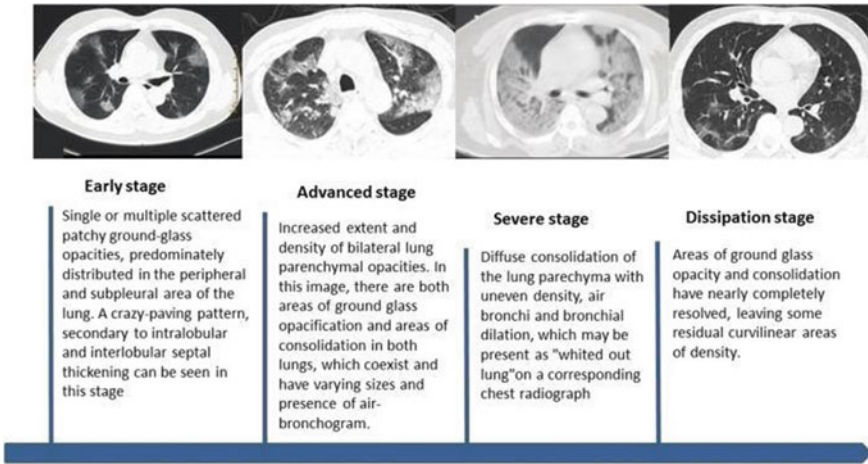


Fig. 1 CT image of different stages of infection [9]

3 Artificial Intelligence in Covid-19

Artificial intelligence (AI) has been suggested highly in clinical practice for the identification of diseases, the diagnosis of COVID-19, monitoring of cases, tracking and maintaining records, the prediction of a sudden future outbreak, mortality risk, disease management, and pattern recognition for studying the future outbreak of diseases [11]. During this pandemic, the AI algorithm has immensely developed in diagnosing diseases. Researchers and scientists have developed and employed numerous techniques using the advancement of AI technology and its use has been constantly increasing since the COVID-19 outbreak in 2020. AI has contributed as a medium for contactless delivery of clinical services during the outbreak of COVID-19 in hospitals and clinics [12]. This system has been adopted to avoid the asymptomatic and misdiagnosis of the patient.

Clinically, AI technology is used for rapid diagnosis and medical analysis of a variety of diseases, including COVID-19. AI technology has adopted two methods for the identification of disease: machine learning (ML) and deep learning (DL). Therefore, the application of machine learning and deep learning has provided a correct and accurate diagnosis of COVID-19 using CT images and X-rays.

3.1 Artificial Intelligence for COVID-19

4 Machine Learning-Based Diagnosis

Machine learning was the recommended technology used to support frontliners for the early prediction and diagnosis of diseases and infections. Recent studies have shown that computationally trained models on a large clinical database can provide more precise diagnostics. The COVID-19 patient could be differentiated based on serum levels of immune cells, gender, and symptoms documented. Computational models were used to differentiate between COVID-19 patients and influenza, which gave a sensitivity rate of 92.5% and a specificity rate of 97.90%.

Integrated advanced AI models are highly sensitive and have been compared to clinical professionals because of their sensitivity. AI-assisted screening for chest images is in high demand for radiologists with less expertise. An artificial intelligence algorithm combined with chest CT has been successful in predicting COVID-19 in some tests set up by the expert.

5 Deep Learning-Based Diagnosis

A DL-based screening framework for COVID-19 has been developed through multi-view chest CT imaging.

Multitask deep learning technology can be used to identify and observe lesions on CT scans caused by the infection of COVID-19. A multitask deep learning model is used, characterized by segmentation, classification, and reconstruction techniques. This shows an accuracy of 87%. For early prediction of coronavirus based on X-ray imaging, a transfer learning method has been applied. Some techniques are being employed in the identification and regulation of acute respiratory disorders such as Lung Ultrasonography. Recent findings suggest that the deep learning technology could assist the clinician in analyzing the images of the COVID-19 patient with Lung Ultrasonography (LUS). Lung ultrasonography has been shown to have a correct prediction and prediction of image biomarkers in a COVID-19 patient.

6 Applications of Artificial Intelligence

1. AI may be used in forecasting to construct early warning systems by gathering information from news sources, social media, and telephone systems to anticipate morbidity and death. Using existing data, machine learning may identify a cluster to anticipate an outbreak's geographical location.
2. AI-powered mobile apps are used to trace COVID-19 contacts. Smartphones and wearables with mobile health apps can diagnose and monitor diseases.

3. AI can screen COVID-19 patients and predict therapy success. The AI-based approach might give clinical resource allocation and decision-making information from clinical parameter data. They can predict recovery and mortality rates.
4. AI predicts, analyzes, quantifies, and segments COVID-19 instances from chest CT and X-ray pictures. Machine learning and deep learning can diagnose the illness. AI-based approaches employing chest CT and X-rays have led to accurate diagnosis and prediction. AI to minimize the workload for medical practitioners and healthcare staff: The AI-based system can reduce the workload for medical practitioners and health workers using the pattern recognition technique to analyze the clinical applications. This would lead to the automatization of several clinical procedures such as training the health practitioners and determining the mode of treatment to minimize contact with the patients.
5. AI could also minimize the frequent visit to the hospital by incorporating telemedicine for distant monitoring of the patient. AI-based robotic models could also be used for delivering essentials and other services.
6. AI in protein structure prediction: AI has also been used in protein studies. It is important to know the useful insight of protein structure crucial for virus entry and replication for the development of a drug within a short period. To determine the protein complex structure of Coronavirus, a deep convolutional neural network was trained using amino acid sequences and high-resolution cryoelectron microscopy density [11].
7. AI in the development of vaccines: Artificial intelligence innovation has changed medical revelation lately. By combining the AI-based model for drug discovery with machine learning's characteristic feature pattern recognition, machine learning makes this possible. Deep learning has given a program included in extricating the information. This deep learning feature has been shown to be promising and superior in performance than the computer-aided models.
8. AI in curbing the spread of misinformation: The pandemic has showered us with lots of information, awareness, and practices. Artificial Intelligence-based machine learning techniques have been used to separate misleading information from rumors and interpret the information to determine the source of information. Additionally, these methods can be used to provide accurate information about recovery rates, accessibility, and the availability of healthcare.

7 Result

As a result of machine learning and deep learning methods, it is possible to predict COVID-19 disease at an early stage and manage outbreak-related issues. Researchers have used different analytical methods to assess the clinical image of a COVID-19 patient and a non-COVID patient. Numerous methods have been developed to observe the best-performing computer-aided system to distinguish between a COVID-19-positive patient and other viral infections. These two patients can be

differentiated by the location of the lung, the number of lesions on the image, ground-glass opacity, and some crazy-paving patterns. The deep learning convolutional neural network has also been used to evaluate chest CT and X-rays to predict COVID-19 patients [12–15].

8 Discussion

COVID-19 can be diagnosed using high-dimensional features of medical images to discriminate from one another. AI-based deep learning and machine learning have been employed. Apart from the assessment of the chest CT image, an application of X-ray has shown promising results and it's easily accessible and low cost. X-ray is cheaper than CT scans. Most of the studies have incorporated both methods. The CNN-based method has achieved an accuracy rate of 99% in the identification of COVID-19 patients from other viral diseases. Deep learning required large data sets to give the final result, whereas machine learning needs a very less amount of data provided by the user. Deep learning requires good-performance hardware, whereas machine learning needs precise user input. One of the most promising methods applied to the study is transfer learning, it requires knowledge on a huge database to transfer it to another new set of problems. Transfer learning is highly useful in medical imaging. The data obtained after analyzing the CT image show a high accuracy rate which will be helpful in future studies and research [8, 10].

9 Conclusion

Coronavirus which started in the seafood market of China has spread throughout the country and beyond. Predicting the virus is essential for the isolation of the patient, treatment, development of drugs and vaccines, etc. Some nucleic acid tests have shown false-negative errors, and such issues could alarm the health professionals. A chest CT imaging of COVID-19 is highly crucial due to its recognition feature, which enables clinicians to make a primary diagnosis within the first few minutes of contact with a suspected patient. AI (artificial intelligence) has been employed to improve healthcare systems, its algorithm is used for diagnostic purposes, decision-making, pattern recognition, protein sequencing, and other health-related issues. These techniques could be used to develop contactless services for frontline workers and to improve public health.

References

1. Yang W, Sirajuddin A, Zhang X, Liu G, Teng Z, Zhao S, Lu M The role of imaging in 2019 novel coronavirus pneumonia (COVID-19)
2. Huang C, Wang Y, Li X, et al (2020) Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet*
3. Riou J, Althaus CL (2020) Pattern of early human-to-human transmission of Wuhan 2019 novel coronavirus (2019-nCoV), December 2019 to January 2020
4. Kanna RK, Ansari AA, Kripa N, Jyothi G, Mutheeswaran U, Hema LK (2022) Automated defective ECG signal detection using MATLAB applications. In 2022 IEEE International conference on current development in engineering and technology (CCET) (pp. 1–7). IEEE, December 2022
5. Harmon SA, Sanford TH, Turkbey B Artificial intelligence for the detection of COVID-19 pneumonia on chest CT using multinational datasets
6. Kanna RK, Vasuki R (2019) Advanced study of ICA in EEG and signal acquisition using mydaq and lab view application. *Int J Innov Technol Explor Eng (IJITEE)* ISSN:2278-3075
7. Mohammad-Rahimi H, Nadimi M, Ghalyanchi-Langeroudi A, Taheri M, Ghafouri-Fard S Application of machine learning in diagnosis of COVID-19 through X-ray and CT images: a scoping review
8. Kripa N, Vasuki R, Kanna RK (2019) Realtime neural interface controlled au-pair BIMA bot. *Int J Recent Technol Eng* 8(1):992–994
9. Ravikumar KK, Ishaque M, Panigrahi BS, Pattnaik CR (2023) Detection of Covid-19 using AI application. *EAI endorsed transactions on pervasive health and technology*, 9
10. Kanna RK et al (2022) Nursing assist module compact patient monitoring system using Iot application. *J Pharm Negat Results* 236–239
11. Kanna RK, Chandrasekaran R, Khafel AA, Brayyich M, Jabbar KA, Al-Chlidi H (2023) “Study on diabetic conditions monitoring using deep learning application,” 3rd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE) 363–366, Greater Noida, India
12. Shome D, Kar T, Mohanty SN, Tiwari P, Muhammad K, Altameem A, Zhang Y, Saudagar AKJ (2021) COVID-transformer: interpretable COVID-19 detection using vision transformer for healthcare. *Int J Environ Res Public Health* 18(21):1–14. <https://doi.org/10.3390/ijerph182111086>, ISSN: 1660-4601
13. Prasath Alias Surendhar S, Kanna RK, Indumathi R (2023) Ensemble feature extraction with classification integrated with mask RCNN architecture in breast cancer detection based on deep learning techniques. *SN Comput. Sci.* 4:618
14. Shankar K, Mohanty SN, Yadav K, Gopalakrishnan T (2021) Automated COVID-19 diagnosis and classification using convolutional neural network with fusion based feature extraction model. *Cogn Neurodynamics* 16(1). <https://doi.org/10.1007/s11571-021-09712-y>. ISSN: 1871-4099
15. Dash S, Chakravati S, Mohanty SN, Patnaik CR, Jain S (2021) A deep learning method to forecast Covid-19 outbreak. *New Gener Comput* 39(2):437–461. <https://doi.org/10.1007/s00354-021-00129-z>. ISSN: 02883635

A New Authentication Protocol for RFID and Applications in E-Passport



Vandani Verma and Garima Jain

1 Introduction

RFID (Radio frequency identification) is used to identify objects or person automatically. It uses electromagnetic fields to recognize, follow the location and transmit the data. It consists of three kinds of roles: RFID Tags, RFID Readers, and Antenna. First, the tag sends the coding information to the reader with the help of the antenna. For this, tag generates the current. Now, the reader gathers all the data and hence decodes it. After collecting and decoding, the reader will send all the information to the computer system to proceed further. Hence, now all the information will be analyzed and can be secured safely too. At last, the receiver conveys the frequency signal with the help of an antenna to the tag.

Cryptography plays a very important role in RFID transactions and authentication. There are different types of cryptographic techniques that we can use in RFID some of them are as follows: (a) Public key techniques: There are numerous types of public key technique that can be used for authentication, confidentiality, non-repudiation, or data integrity provided by symmetric cryptography. We can take an example of 'Authentication' which can be achieved based on various primitives and schemes by different cryptographic protocols. One more example we can take that of entity authentication. To achieve entity authentication, zero-knowledge proof of knowledge and digital signature schemes can be used. Further, public key techniques are divided into four levels, i.e., services, protocols, schemes, and primitives based on the type of requirement. (b) Authentication Techniques: It aims to protect the data from various attacks. There are two types of attacks: active and passive. In active attacks, the verifier communicates and extracts the information by active participation. But in a passive attack, the verifier works passively, and information extraction is done just by

V. Verma (✉) · G. Jain

Department of Mathematics, Amity Institute of Applied Sciences, Amity University Uttar Pradesh, Noida, India

e-mail: vandaniverma@yahoo.com

monitoring the executions of the protocol. These authentication protocols protect the whole information from these passive attacks. Elliptic curve cryptography algorithm designed for encryption, signature, and identification can be used to protect RFID tags from various kinds of attacks.

Many researchers [1–6, 8] have proposed RFID authentication protocols in literature and their applications to health care systems, e-passports, and others. The rest of the paper is organized as follows: Sect. 2 discusses the background concepts; Sect. 3 proposes Authentication Protocol for RFID; Sect. 4 presents the security analysis and comparative analysis of the proposed scheme; Sect. 5 discusses the Application of RFID in Electronic Passport along with security issues and enhancement. Finally, conclude in Sect. 6.

2 Background Concepts

This section discusses the background concepts of elliptic curves [7], how the addition of points is carried out on elliptic curves and RFID as follows.

2.1 Elliptic Curve

In this section we will discuss the elliptic curve cryptography. The elliptic curves are those curves that are typically described as Eq (a, b), where a and b are constrained to mod q and q is a prime number. It can be shown by the equation.

$n^2 = m^3 + am + b$, where m gives the positive and negative value. Arithmetic operations on the elliptic curve over field Z_p are as follows:

- **Addition:** Adding two points $X (m_x, n_x)$ and $Y (m_y, n_y)$ gives $Z (m_z, n_z)$. For this, we need to find the slope λ , using the following formulas

$$\lambda = \frac{(n_y - n_x)}{(m_y - m_x)} \text{ if } X \neq Y$$

$$\lambda = \frac{(3m_x^2 + a)}{2n_x} \text{ if } X = Y$$

where a is obtained from Eq(a, b). Next step is to find the sum $Z(m_z, n_z) = X + Y$ where $m_z = \lambda^2 - m_x - m_y$ and $n_z = \lambda(m_x - m_z) - n_x$

- **Negation of a point:** If $Y = (m_y, n_y)$ is a point on the elliptic curve then its negation is $-Y = (m_y, -n_y)$
- **Subtraction of a point:** To subtract Y from X, we need to write $X - Y$ as $X + (-Y) = X - Y$

$$\begin{aligned}
 &= (m_x, n_x) - (m_y, n_y) \pmod q \\
 &= (m_x, n_x) + (m_y, -n_y) \pmod q
 \end{aligned}$$

Now proceed same as addition

- **Multiplication:** Only scalar multiplication can be done on elliptical curves. Multiplication between two points on the elliptical curves is not possible and it is performed by repeated addition like $2X = X + X$, $3X = X + X + X$, and so on. For multiplication, use the formula for $X = Y$
- **Division:** On elliptic curve only scalar division can be done, i.e., $\frac{1}{[a(m_x, n_x)]} = a - 1$ (m_x, n_x) . Now, further proceed it by multiplication process.

2.2 RFID and Its Types

RFID (Radio Frequency Identification) is a wireless system that consists of two components: tags and readers. The reader is a device with one or more antennas that broadcast radio waves and receive signals from RFID tags. Tags can be passive or active, communicating their identity and other information to nearby readers via radio waves. Passive RFID tags do not have batteries and are powered by the reader. Passive RFID uses 13.56 MHz for near field communication. It operates at a low frequency of (125–134) KHz. It draws power from the RFID reader and has a very high frequency of (865–960) MHz Batteries are used to power active RFID tags. Active RFID operates at 433 MHz or 915 MHz frequencies. It lasts for three to five years. Transponders are classified as having a higher battery efficiency than beacons, whose batteries degrade more quickly.

The use of RFID as a barcode substitute is expanding. Although inventory tracking with RFID and barcode technologies is comparable, there are some significant differences between the two discussed as follows (Table 1).

Table 1 Comparison between RFID Tag and Barcode

RFID Tag	Barcode
It can recognize distinct things without a clear line of sight	Scanning needs a clear line of sight
Depending on the type of tag and reader, items can be scanned from inches to feet away	Closer proximity is necessary for scanning
Real-time data updates are possible	The data is read-only and cannot be altered
A power source is required	No external power supply is required
The read time per tag is less than 100 ms	Each barcode takes at least a half-second to read

3 Proposed Authentication Protocol for RFID

In this section, we propose a new authentication protocol between RFID Server and the tag using a finite field and elliptic curve. The following are the pre-requisites for the proposed scheme:

$F(q)$: Finite field where q is the size of the field.

a, b : An elliptic curve E has two parameters.

E : Elliptic curve represents by the equation $y^2 = x^3 + ax + b$ over the finite field $F(q)$.

P : Generator point.

y : Server's private key.

x : Secret key for the tag.

U, R', R_1, R_3 : Message forward from server to tag.

C_2, R_2 : Message forward from tag to server.

r_1, r_2 : Randomized number taken by server and tag resp.

Communication Round 1: SERVER \longrightarrow TAG

Server chooses a random number $r_1 \in Z_q$ and computes.

$$U = r_1 P.$$

$$R' = r_1^{-1} P + r_1 P - y X.$$

$$R_1 = R' - r_1 P + y X.$$

Now, server sends U, R' and R_1 to the tag for further communication.

Communication Round 2: TAG \longrightarrow SERVER

Tag on receiving U, R' and R_1 from server, now chooses a random number $r_2 \in Z_q$ and computes

$$R'' = r_2^{-1} P + r_2 P + x Y$$

$$R_2 = R'' - r_2 P - x Y$$

$$C_1 = (R' + R'') - (R_1 + R_2)$$

$$C_2 = r_2^{-1} C_1$$

Now, Tag sends C_2 and R_2 to the server.

Communication Round 3: SERVER \longrightarrow TAG

Server on receiving the C_2 and R_2 verifies the tag's identity through the following series of calculation:

Table 2 Protocol Summary

Server	Communication	Tag
$r_1 \in Z_q$ $U = r_1 P$ $R' = r_1^{-1} P + r_1 P - y X$ $R_1 = R' - r_1 P + y X$	$\xrightarrow{U, R', R_1}$ $\xleftarrow{C_2, R_2}$	$r_2 \in Z_q$ $R'' = r_2^{-1} P + r_2 P + x Y$ $R_2 = R'' - r_2 P - x Y$ $C_1 = (R' + R'') - (R_1 + R_2)$ $C_2 = r_2^{-1} C_1$
$C_3 = r_1^{-1} C_2$ $R_1 = C_3 - R_2$ Checks if $R_3 = C_2 + r_1 P$	$\xrightarrow{R_3}$	Checks if $R_4 = r_2 R_3$ $= C_1 + r_2 U$

$$C_3 = r_1^{-1} C_2 \text{ and checks if } R_1 = C_3 - R_2 \tag{1}$$

If Eq. (1) is true, then Server calculates.

$R_3 = C_2 + r_1 P$ and sends it to Tag.

Verification Phase is completed when the tag computes

$$R_4 = r_2 R_3 = C_1 + r_2 U \tag{2}$$

And checks if Eq. (2) is true otherwise Tag is invalid (Table 2).

4 Security Analysis and Comparative analysis

In this section, we test our proposed scheme against various parameters like availability, mutual authentication, anonymity, cloning attack, and impersonation attack. We also compare the operations required for generating our proposed algorithm with the similar kind of algorithms existing in the literature.

4.1 Availability

The protocol proposed by us can be available easily. There is no specific requirement for updates of the private key to execute the proposed protocol. So, the execution process can be done without having any problems. Hence, our proposed protocol provides the availability.

4.2 *Mutual Authentication*

In our proposed protocol, we cannot generate the message R_1 without knowing the values of r_1 , i.e., the random number chosen by the server and also the value of y , i.e., server's private key. These values were not forwarded to the tag and also these values are known by the server only. Hence, both these values can maintain their secrecy. Therefore, values can be stored only on the server. Similarly, the values of C_2 cannot be calculated without knowing the values of r_2 , i.e., a randomized number chosen by the tag and x , i.e., the secret key of the tag. Therefore, our proposed scheme provides mutual authentication.

4.3 *Anonymity*

Both tag and server have their different secret keys which will never forward in the whole process. The tag has a secret key x and the private key of the server is y . They both can never be fetched. Therefore, our proposed protocol provides anonymity.

4.4 *Cloning Attack*

Our proposed protocol consists of both tag and server, having their own individual private keys. If any hacker wants to hack these secret keys, then they will not be successful in this work because the secret keys have no correlation. So, the hacker fails to do so. Therefore, our proposed protocol can overcome cloning attack.

4.5 *Impersonation Attack*

If we observe our second proposed protocol, C_1 cannot be generated by the hacker without knowing R' , R_1 , R'' , R_2 because the hacker doesn't have any idea of x , r_2 , y , and r_1 . Therefore, our proposed protocol can overcome impersonation attack.

4.6 *Comparative Analysis*

Here we compare the proposed scheme with Dukyil et al. [1], Nikitin et al. [10], and Liu et al. [9] scheme on the following parameters: addition over the elliptical curve, random no. requirements, number of communications, and total computation

Table 3 Comparison table

Scheme operations		[1]	[10]	[9]	Proposed
Addition over elliptical curve	S	3	3	2	3
	T	3	2	2	2
Random no. requirements	S	2	1	3	1
	T	2	1	2	1
No. of communications	S + T	3	5	4	3
Total computation	S + T	13	12	13	10

required for the construction of protocol. From Table 3, we can see that our proposed scheme requires least number of computations as compared to Dukyil et al. [1], Nikitin et al. [10], and Liu et al. [9] scheme. Also, a number of communications are less than Nikitin et al. [10] and Liu et al. [9] scheme. Also, the random number requirement is less in our scheme as compared to Dukyil et al. [1] and Liu et al. [9] protocols. In the following table, consider S = Server and T = Tag.

5 Application of RFID in Electronic Passport

The e-passports develop the technologies named biometric and radio frequency identification (RFID). Many national governments decide to develop the ID cards which includes biometrics and RFID. Initiatives were taken by the national government to integrate the technologies biometrics and RFID with the existing identity cards. By combining these technologies with ID cards, we will be able to improve security, minimize fraud, make identification verification easier, and introduce new hazards. RFID can play a crucial role in e-passports because of the following reasons:

- *Security*: Use of RFID in e-passports, improves the security of the tag data. It prevents the data from steal, copy, on official travel document, and from frauds [11–13].
- *Speed*: RFID helps the e-passports to work quickly. Now, there is no need for humans to queue, and thanks to the digitization process, we can simply handle information. We can even convey data in a fraction of the time and maintain the security of the tag’s information while sitting at home.
- *Automation*: RFID helps to increase the automation. The reason for automation is that the e-passport shortens the travel process and allows for work to be completed entirely online. The human interaction will also decrease. Because these e-passports work with automated border control gates. These barriers are only open to circumstances that demand an immediate response. So, these gates get resources free. That is how e-passports increase automation.

- *Influence Other Travel Documents:* For travelers, who travel from one place to another through the border crossing, they are in requirement to show their driving license, Aadhaar card, and their ID, etc. But instead of these, they can show a soft copy of e-passports. And now other IDs, Aadhaar card, and driving license are also getting electronic.

5.1 Security Issues in E-Passports

Following are some issues or threats related to privacy and security in e-passports:

- **Clandestine Scanning:** As we all know that RFID system is important for clandestine scanning. According to the guidelines given in the ICAO, the well-encrypted communications between the readers and their passports will not require. Thus, an e-passport chip that is not secured will do short-range clandestine scanning which leads to the personal information to be unauthenticated like place of birth and birth date, etc. thus, it is a privacy issue in e-passports.
- **Clandestine Tracking:** Tracking can be possible even while using RFID chips in checking the authentication of e-passports. Although we can't read the data on the chip. If the ID of the RFID chip is distinctive for every passport, then the tracking of all the movements can be done by the parties which are unauthorized of the passport holder.
- **Skimming and Cloning:** According to the regulations of ICAO, we need digital signatures for the data on the e-passport. This depicts that the information on the passport is authenticated, or the data is of the correct person and that is checked by the passport issuing authority. So, if passport cloning happens then digital signatures have no role to play to defend against it.
- **Eavesdropping:** A Faraday cage would prevent the entry of RFID signals because of having its countermeasure ability to clandestine RFID scanning as a Faraday cage would be like metallic material on its cover. Faraday cages would do the scanning only. So, it does not prevent eavesdropping on the transmission between reader and passport that happened in the airport also.
- **Biometric Data Leakage:** If the physical environment is in controlled condition, then the biometric images which include in the e-passport would not require a secret for identification and authentication. Earlier thumbprints were used in the e-passport of Malaysia, and these were all digitalized. Hence, automation will be done for e-passports and human oversight will get weakened. Hence, biometrics data leakage is also an issue [14–16].
- **Cryptographic Leakage:** According to the guidelines given by ICAO, they used a modern technique for authentication of the e-passport. In this, firstly reader derives a cryptographic key k . For this, he needs to make some documents to be scanned like his name, date of birth, and passport number [17–20]. By marking these optical contacts with a passport, he cannot get the key k . The functions of this key k are used for the encryption of the data of the interaction between the reader and the passport and to allow them to interact with the reader to the

passport. And before making them aware of their RFID tag information, this work will be completed. Hence, after knowing about his key k , the person will travel everywhere by showing this key and scanning his or her passport. Later, it was discovered that this method was not very effective. It has a few problems.

5.2 Suggested Improvements in E-Passports

- **Faraday Cages:** To make our e-passport more powerful and full of strength, then we need to add the RF blocking material made up of aluminum fiber, which we can easily use to create a Faraday cage to an e-passport's cover. Now, this created Faraday cage helps in preventing the reading of RFID devices inside the e-passport. We must open it physically, only then we can read it, otherwise is no option. Now, the US state department used this method for better authentication. This method helps to reducing the situation where unauthorized reading can be possible of the passport. Further, the research department also proposed some tools to enhance the security of RFID like "Antenna energy analysis" and "Blocker tags." By doing this, Faraday cages would be more powerful and authenticated. It protects the RFID from leakage of the data.
- **Larger Secrets for Basic Access Control:** Larger secrets for larger authentication. Every passport must have a 128-bit secret added to it so that it may be combined with the other passport data in the algorithm. It will show in the form of an ID number which would be a very big one. But as we all know that all we have is 52 bits of entropy which is very low for the attack of brute force. So, we just need to make it larger.
- **Private Collision Avoidance:** A method to enhance the strength of the e-passport is that to select a random identifier that is random on each tag read and we should take care that UID should be dissimilar and un-linkable across sessions. We should employ all of them because, even though UID is an element of ISO 4443's collision avoidance protocol, we should use a larger passport secret as part of key derivation. So, we conclude that e-passports and various IDs should use private collision protocols.

6 Conclusion and Future Work

The proposed authentication protocol between server and tag can combat many drawbacks and provides mutual authentication at a greater level as compared to the other schemes [1, 9, 10] at low computational cost. We also discuss various approaches to RFID tag which can be applied to the security and privacy of electronic passport. In the future, the security analysis of the proposed protocol can be done using various Linux-based software.

References

1. Dukyil A, Mohammed A, Darwish M (2018) Design and optimization of an RFID-enabled passport tracking system. *J Comput Des Eng* 5(1):94–103. <https://doi.org/10.1016/j.jcde.2017.06.002>
2. Hutter MRFID authentication protocols based on elliptic curves - a top-down evaluation survey. In: *Proceedings of the international conference on security and cryptography*, pp 101–110. <https://doi.org/10.5220/0002186201010110>
3. Batina L, Guajardo J, Kerins T, Mentens N, Tuyls P, Verbauwhede I (2006) An elliptic curve processor suitable for RFID-tags. <https://eprint.iacr.org/2006/227.pdf>
4. Moosavi SR, Nigussie E, Virtanen S, Isoaho J (2014) An elliptic curve based mutual authentication scheme for RFID implant systems. In: *Procedia computer science*, 5th international conference on ambient systems, networks and technologies, vol 32, pp 198–206
5. Noori D, Shakeri H, Niazi Torshiz M (2020) Scalable, efficient, and secure RFID with elliptic curve cryptosystem for Internet of Things in healthcare environment. *EURASIP J Info Secur* 2020:13. <https://doi.org/10.1186/s13635-020-00114-x>
6. Lee C-I, Chien H-Y (2015) An elliptic curve cryptography based RFID authentication securing e-health system. *Hindawi Publ Coop Int J Distrib Sens Netw Article ID 642425*, 7 p
7. Chandraul R, Paliwal R, Jain A (2015) Efficient security of RFID devices using HECC algorithms and performance analysis by simulation. *Innov Syst Des Eng* 6(12). ISSN 2222-1727 (paper) ISSN 2222-2871
8. Koblitz N (1994, 1987). *A course in number theory and cryptography*, 2nd edn. Springer, New York
9. Liu Y-L, Qin X-L, Wang C, Li B-H (2013) A lightweight RFID authentication protocol based on elliptic curve cryptography. *J Comput* 8(11)
10. Nikitin PV, Ramamurthy S, Martinez R, Rao KS (2012) Passive tag-to-tag communication. In: *2012 IEEE international conference on RFID (RFID)*. IEEE, pp. 177–184
11. Shukla V, Chaturvedi A, Srivastava N (2019) Nanotechnology and cryptographic protocols: issues and possible solutions. *Nanomater Energy* 8(1):1–6. <https://doi.org/10.1680/jnaen.18.00006>
12. Misra MK, Chaturvedi A, Tripathi SP, Shukla V (2019) A unique key sharing protocol among three users using non-commutative group for electronic health record system. *J Discret Math Sci Cryptogr* 22(8):1435–1451. <https://doi.org/10.1080/09720529.2019.1692450>
13. Shukla V, Chaturvedi A, Misra MK (2021) On authentication schemes using polynomials over non commutative rings. *Wireless Pers Commun* 118(1):1–9. <https://doi.org/10.1007/s11277-020-08008-4>
14. Chaturvedi A, Shukla V, Misra MK (2018) Three party key sharing protocol using polynomial rings. In: *5th IEEE Uttar Pradesh section international conference on electrical, electronics and computer engineering (UPCON)*, pp 1–5. <https://doi.org/10.1109/UPCON.2018.8596905>
15. Shukla V, Misra MK, Chaturvedi A (2022) Journey of cryptocurrency in India in view of financial budget 2022–23, Cornell university 1–6. <https://doi.org/10.48550/arXiv.2203.12606>
16. Shukla V, Srivastava N, Chaturvedi A (2016) A bit commitment signcryption protocol for wireless transport layer security (WTLS). In: *IEEE international conference on electrical, computer and electronics engineering*, pp 83–86. <https://doi.org/10.1109/UPCON.2016.7894629>
17. Shukla V, Chaturvedi A, Srivastava N (2019) A secure stop and wait communication protocol for disturbed networks. *Wireless Pers Commun* 110:861–872. <https://doi.org/10.1007/s11277-019-06760-w>
18. Chaturvedi A, Shukla V, Misra MK (2021) A random encoding method for secure data communication: an extension of sequential coding. *J Discret Math Sci Cryptogr* 24(5):1189–1204. <https://doi.org/10.1080/09720529.2021.1932902>

19. Shukla V, Chaturvedi A, Srivastava N (2019) A new one time password mechanism for client-server applications. *J Discret Math Sci Cryptogr* 22:1393–1406. <https://doi.org/10.1080/09720529.2019.1692447>
20. Shukla V, Misra MK, Chaturvedi A (2021) A new authentication procedure for client-server applications using HMAC. *J Discret Math Sci Cryptogr* 24(5):1241–1256. <https://doi.org/10.1080/09720529.2021.1932908>

Comprehensive Survey on AQI Prediction Using Machine Learning Algorithms



Imran Khan and Rashi Agarwal

1 Introduction

With the rapid rise of businesses and human civilisation, air pollution is steadily expanding. It has an impact on the lives of all living organisms on the earth. The World Health Organization estimates that about 7 million people die each year from air pollution. Polluted air is a major worry for humans since it causes health problems such as lung cancer, stroke, heart disease, and many more. Furthermore, it has a major environmental impact, producing global warming, ozone layer depletion, and contamination of water and soil. As a result, it is critical to investigate and monitor the air quality patterns induced by various pollutants.

Air pollution is not only harmful to humans but it also has a significant impact on the planet's general health. The increase in air pollution has also been slowly raising the global temperature. This phenomenon is known as global warming, and if not controlled, it can have disastrous consequences. Global warming has the potential to bring a plethora of disastrous impacts, such as the melting of the ice caps at the earth's poles, which have already been shown to be shrinking in size. As a result, there is an imperative need to control and monitor air pollution to guarantee that global warming does not reach dangerous levels for our planet. There are many air quality monitoring methods around the world that have proven to be highly effective in detecting changing air quality in real time. The importance of air quality has prompted the WHO (World Health Organization) to warn that air quality is one of the most important elements of healthcare and must be maintained at all costs.

Air Quality Index: The most fundamental requirement for regular life on Earth is the presence of air. The recent increase in the number of industries, private vehicles,

I. Khan (✉) · R. Agarwal
Harcourt Butler Technical University, Kanpur, India
e-mail: imran.k@hbtu.ac.in

R. Agarwal
e-mail: rashi@hbtu.ac.in

and other activities that require the burning of fuel has led to a significant decline in the quality of the air. NO_2 , SO_2 , NO , CO_2 , NO_x , CO , PM_{10} , and $\text{PM}_{2.5}$ are all examples of pollutants that can be found in the environment. Pollution in the air has an effect not only on the ability of living things to continue existing but also on the communities in which they live. The amount of pollution in the air for a specific amount of time, as estimated by an air monitoring model, is required for the AQI calculation. The product of the concentration at a certain time and the amount of time since the pollution was released is the amount of pollution in the air. Disease research is performed to determine the effects that a particular dose has on a person's health. Both the amount of air pollution and the process that is used to turn air pollutants into an AQI reading are subject to change. There is a significant range of readings for the air quality index. Glossaries, colour codes, and basic information are provided to students at each grade level. Advisory opinions on matters pertaining to public health The Air Quality Index (AQI) might go up if there is less air pollution or if the pressure in the atmosphere is higher. Dry air, which is frequently caused by anticyclones, temperature variations, or low air speed, allows air pollution to persist, resulting in high concentrations, chemical interactions between air pollutants, and dangerous circumstances. This can be avoided by maintaining a constant level of air movement. Numerous academics have embarked on extensive research projects concerning the prediction and forecasting of pollution. Particles and gases that are released into the environments in which people live combine to form what is known as air pollution. There are two different kinds of air pollutants: those that come from natural sources and those that come from anthropogenic, or man-made, sources. Natural sources include sulphate, sulphur dioxide, nitrogen dioxide, and carbon dioxide. Combustion of fossil fuels, emissions from transportation, discharge of waste from industrial processes, global warming, climate change, and acid rain are all examples of pollution caused by human activity. Table 1 presents the Air Quality Index along with its Ranges and Description for your perusal.

Table 1 Air quality index, ranges, and description

AQI value range	Level of health concern	Description
0–50	Good/low	No impact/normal activities
51–100	Satisfactory	Cause minor breathing problem
101–300	Moderate	Breathing distress
201–300	Poor	Breathing and discomfort in heart patients
301–400	Very poor/high	Breathing problems and effects on heart and lung disease
401–500	Harmful/very high	Effects on healthy persons, serious health problems, and immediate effects on persons with pulmonary/heart disease

2 Various Techniques of AQI Prediction

It is possible to evaluate and forecast pollution concentrations in metropolitan areas using a variety of air quality prediction models. Common types of statistical models used for forecasting include chemical transfer and air dispersion models. Air quality forecasting models have increasingly relied on machine learning approaches as the primary technique.

Statistical model: The statistical model is predicated on a method that looks to the past in order to understand how to predict the future behaviour of the variable of interest. The precision of these models is exceptional. Multiple linear regression and autoregressive moving average (ARMA) are two important statistical models used for assessing the accuracy of aerial weather forecasts [1, 2]. However, they cannot provide reliable estimates of exposure levels because they fail to account for the dynamic behaviour of climatic data.

Machine Learning Models: Thanks to advancements in technology, algorithms based on artificial intelligence are increasingly being utilised for forecasting purposes, such as predicting air quality. Unlike a purely statistical model, an auto-learning method incorporates a variety of parameters into its predictive calculations. It appears that Artificial Neural Networks (ANN) are the most widely utilised technique for predicting air quality [3, 4]. A neural network-based model has been demonstrated to be effective for prediction in other research, while hybrid or mixed models have also been shown to be effective.

1. Support Vector Machines (SVM): Support Vector Machines were introduced [5] (SVM). Hyperplanes split support vector data into classes. If data can't be linearly separated, kernel functions project them to higher dimensions. Non-linearly separable data is linearised [6]. SVR (SVR). This allowed SVM regression with a new loss function. SVR has been used in time series forecasting [7–9]. SVR models offer faster training and better forecasting with fewer parameters.

Let us represent the training set with m data points as

$$D = \{(a_1, b_1), (a_2, b_2), (a_3, b_3), (a_4, b_4) \dots (a_m, b_m)\}$$

where, $a \in \mathbb{A} \subset \mathbb{R}^n$ are the inputs in the training set and $b \in \mathbb{B} \subset \mathbb{R}$ are the corresponding expected outputs in the training set.

A nonlinear kernel function is represented as

$$f(x) = \omega' \Phi(a_i) + c$$

2. Linear Regression: Using linear regression, we can find a formula that faithfully represents the data. The model can now be used to create predictions thanks to the known regression equation. Regression analysis can be carried out in a number of ways, including the more common linear form. You can use simple linear regression to get the predictive function if the correlation coefficient suggests the

data are useful for making predictions and the scatterplot of the data looks like a straight line. You might recall from elementary algebra that the equation for a straight line is $y = mx + b$. Data collection, linear regression computation, and the solution of $y' = a + bx$ are all covered in detail. This statement can also be written as $b_0 + b_1x$, which is equivalent to the original (simply substitute $a + b$ for $b_0 + b_1$), and is commonly found in AP statistics.

3. **Artificial Neural Networks (ANNs):** Neural networks are a type of machine learning. Their names and shapes mimic the way organic neurons interact. An artificial neural network (ANN) has input, hidden, and output node layers. Each node or artificial neuron contains weights and thresholds. When a node's output surpasses a threshold, it wakes up and transfers data to the next layer. The data won't reach the next network layer otherwise.

The neural network equations are:

$$W = \text{bias} + w_1x_1 + w_2x_2 + \dots + w_nx_n$$

where,

W : is the output function

w_i , : the weights of coefficients

x_i : are the independent variables or the inputs, and bias : intercept

4. A decision tree is one of the most commonly used Machine Learning algorithms for tackling regression and classification issues. As the name implies, the technique employs a decision tree-like model to either forecast the target value (regression) or predict the target class (classification). Decision trees with a target variable that can have continuous values are known as regression trees.
5. A supervised learning technique for regression called random forest regression makes use of ensemble learning techniques. To provide more accurate predictions than a single model, ensemble learning techniques aggregate predictions from many machine learning algorithms. The Bootstrap Random Forest methodology creates a large number of randomly generated decision trees from data and averages the outputs to get a new result. It blends ensemble learning techniques with a decision tree architecture. This frequently results in accurate forecasts (Fig. 1).

3 Methodology

Figure 2 depicts the five-step approach for assessing air quality. The entire procedure is as follows:

Here's how it works:

- A. *Data Gathering*
- B. *Data Preparation*
- C. *Feature Choice*

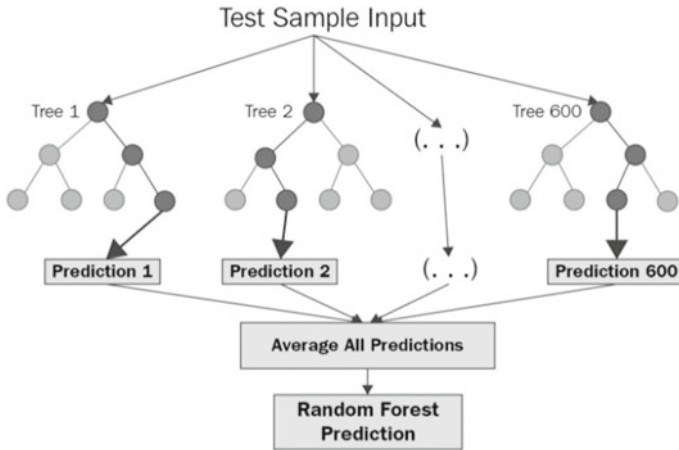


Fig. 1 Radom forest [10]



Fig. 2 Overall process of AQI prediction

D. Model Training

E. Result Evaluation

Data Collection: Manual ambient air quality monitoring stations are used to monitor air pollution (CAAQM). According to the Central Pollution Control Board’s (CPCB) National Air Quality Monitoring Program (NAMP) [11] Delhi’s Sarojini Nagar, Chandni Chowk, Mayapuri Industrial Zone, Pitampura, Shahadra, Shahzada Bagh, Nizamuddin, Janakpuri, Fort Siri, and ITO all have manual air pollution monitoring programmes in place. Data Pre-processing: Missing values in input parameters were removed from the data to be studied. An imputation function is used to estimate missing data at the target object, i.e. the pollutant.

Interpolate. The average is the estimation approach utilised here. All parameters are modified before normalising the dataset to facilitate calculations. As a result, the wind direction, which is stated in degrees, has been converted to wind direction Index as the input parameter (dimensionless). If the input includes several qualities with different units, it is necessary to scale these attributes to a certain region so that all available attributes have the same weight. This assures that there is a minor a meaningful account that may have a broader scope removed potentially more important qualities [12].

Feature Selection: The process of picking a subset of beginning characteristics having important information that predicts the output data is known as feature selection. Function extraction is utilised when there is redundant data. The selection of suitable input parameters for the specified input dataset is part of feature extraction. Analysis is applied to the resulting reduced data set. Because there are a maximum of six entries accessible for analysis, all inputs are chosen for calculations.

Training the Model: Once the data have been prepared, the next step is to develop the model for prediction. In the domain of AQI prediction, various training models are available, including statistical, regression, and deep learning techniques.

Result Analysis: Following model building, the model's performance must be analysed in order to validate the model's prediction capabilities. A model that predicts well on an unknown dataset is regarded good and can be utilised for deployment.

Experiments and Results

In order to analyse the results of various algorithms discussed in the previous section, we have collected climate data for AQI prediction from [13]. We collected the AQI data of last 3 years (from 2018 to 2021) of Indian city Bangalore. Following are the attributes/characteristics of data that we have collected.

1. NO₂
2. SO₂
3. NO
4. CO₂
5. NO_x
6. CO
7. PM 2.5

We applied the above-mentioned steps like data collection, data pre-processing, feature selection etc. and then we applied different machine learning models onto that pre-processed data. For experimentation we have Anaconda Jupyter Notebook and Python as our programming language. We experimented on the following algorithms and compared their performances. For performance evaluation metrics, mean squared error (MSE), mean absolute error (MAE), and root mean square error were employed (RMSE). Table 2 shows the overall results of the experiment, and Fig. 3. shows the overall comparison of various algorithms with respect to MAE, MSE, and RMSE.

The result shows that the performance of the algorithms ANN and Random Forest is almost same. Also, the Linear Regression algorithm is giving the poorest results for AQI prediction.

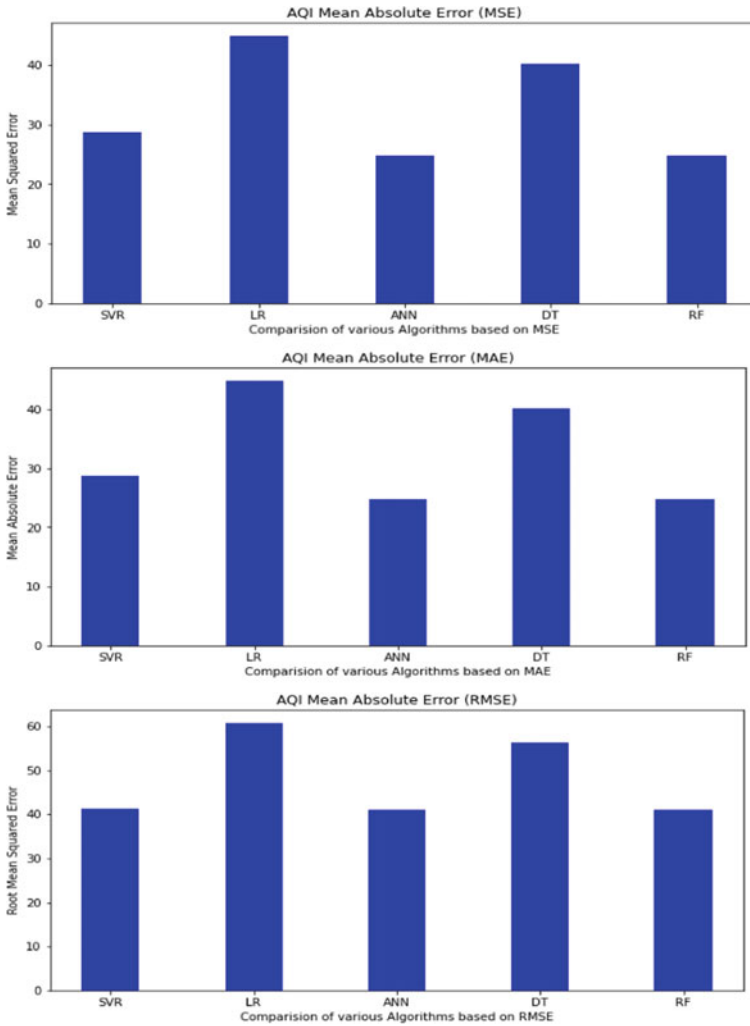


Fig. 3 Comparison of various algorithms with respect to MAE, MSE, and RMSE

4 Challenges and Future Plans

There are several problems involved in forecasting and predicting air pollution. This section includes some examples of forecasting research challenges and future directions. Some of the major challenges [14–19] to consider in pollution forecasting are as follows:

1. Temporal (Time series): Forecasting pollution, timing-wise updating, and concurrent updating are the main issues that need to be resolved for prediction. Prior research on pollutant forecasting based on time has been presented by a few of

Table 2 MAE, MSE, and RMSE in AQI prediction using different algorithms

Algorithm	Mean absolute error (MAE)	Mean squared error (MSE)	Root mean squared error (RMSE)
Support vector machine regression	28.75	1708.21	41.33
Linear regression	44.84	3687.54	60.72
Artificial neural network	24.66	1677.07	40.95
Decision tree	40.14	3171.81	56.32
Random forest	24.66	1677.08	40.95

the researchers. The continuous updating of pollution forecasts, however, is one of the most challenging problems.

2. Spatial (Location series) forecasting is another major difficulty in the dynamic changes of places. One of the most difficult difficulties in dynamic mobility is location-based updation and forecasting.
3. Statistical analysis: The only thing that the statistical analysis models take into consideration are the input data; they do not take into account any changes or processes that are social, chemical, or biological in nature.
4. One of the key research directions is improving accuracy. The earlier techniques only took a few parameters into account. The precision should be maintained when utilising several parameters.
5. Spatial–temporal updating: Researchers have developed various sorts of models for forecasting pollution. However, only a few articles [20, 21] employed location and timing-based updates. As a result, different parameters must be considered for spatiotemporal updating.
6. Model Selection: Mathematical models are effective and can be developed in a shorter amount of time, but they call for a larger quantity of historical data and place a significant amount of reliance on time-series processes within the data. AI techniques such as NN methods perform very well and are capable of solving nonlinear data; however, the models are less stable and are more data dependent.

5 Conclusion

Air pollution and forecasting is an essential scientific field since pollution affects the overall living environment. Pollution levels are rising as a result of the advancement of technologies, industry, and deforestation. In this article, we explore the pollution forecasting methodologies and algorithms used by different nations, as well as the air quality index ranges for those countries. Also, we have performed an experimental study on popular algorithms and found out ANN and Random Forest algorithms were showing the best performance. Although, the results are not very exhaustive and require further considerations like applying Lasso and Ridge regularisation and

other hyperparameter tuning. At last, we also discussed about the open challenges, which may help the new researchers for giving a path in this domain.

References

1. Li NH, Tsay S (2011) A study on the potential applications of satellite data in air quality monitoring and forecasting. *Atmos Environ* 45(22):3663–3675
2. Box G, Jenkins G (1970) *Time series analysis: forecasting and control*. Wiley S. Pro., Hoboken
3. Baawain M (2014) Systematic approach for the prediction of ground-level air pollution (around an Industrial Port) using an artificial neural network. *Aerosol Air Qual Res*
4. Huang M, Zhang T, Wang J, Zhu L A new air quality forecasting model using data mining and artificial neural network. In: 6th IEEE
5. Cortes C, Vapnik V (1995) Support-vector networks. *Mach Learn* 20(3):273–297
6. Smola AJ et al (1996) Regression estimation with support vector learning machines. Ph.D. dissertation, Master's thesis, Technische Universität München, 1996
7. Cao L, Tay FE (2001) Financial forecasting using support vector machines. *Neural Comput Appl* 10(2):184–192
8. Drucker H, Burges CJ, Kaufman L, Smola A, Vapnik V et al (1997) Support vector regression machines. *Adv Neural Inf Process Syst* 9:155–161
9. Müller K-R, Smola AJ, Rätsch G, Schölkopf B, Kohlmorgen J, Vapnik V (1997) Predicting time series with support vector machines. In: *International conference on artificial neural networks*. Springer, pp 999–1004
10. <https://levelup.gitconnected.com/random-forest-regression-209c0f354c84>
11. Xiao C, Chen N, Hu C, Wang K, Gong J, Chen Z (2019) Short and mid-term sea surface temperature prediction using time-series satellite data and LSTM-AdaBoost combination approach. *Remote Sens Environ* 233:111358
12. Guerreiro CB, Foltescu V, De Leeuw F (2014) Air quality status and trends in Europe. *Atmos Environ* 98:376–384
13. Data Source: <https://en.tutiempo.net/climate/ws-422600.html>
14. Bai L, Wang J, Ma X, Lu H (2018) Air pollut ion forecasts: an overview. *Int J Environ Res Public Health* 15(4):780
15. Chen S, Wang JQ, Zhang HY (2019) A hybrid PSOSVM model based on clustering algorithm for short -term atmospheric pollutant concentration forecasting. *Technol Forecast Soc Chang* 146:41–54
16. Baklanov A, Zhang Y (2020) Advances in air quality modeling and forecasting. *Glob Transit* 2:261–270
17. Li L, Dai S, Cao Z, Hong J, Jiang S, Yang K (2020) Using improved gradient-boosted decision tree algorithm based on Kalman filter (GBDT-KF) in time series prediction. *J Supercomput* 1–14
18. Zhou Q, Jiang H, Wang J, Zhou J (2014) A hybrid model for PM_{2.5} forecasting based on ensemble empirical mode decomposition and a general regression neural network. *Sci Total Environ* 496:264–274
19. Sinnott RO, Guan Z (2018) Prediction of air pollution through machine learning approaches on the cloud. In: *2018 IEEE/ACM 5th international conference on big data computing applications and technologies (BDCAT)*, IEEE, 2018, pp 51–60
20. Seng D et al (2021) Spatiotemporal prediction of air quality based on LSTM neural network. *Alex Eng J* 60(2)
21. Baca-López K et al (2020) Spatio-temporal representativeness of air quality monitoring stations in Mexico City: implications for public health. *Front Public Health* 8

Ransomware 3.0—A Weapon for Next-Generation Information Warfare



Mohiuddin Ahmed, A. N. M. Bazlur Rashid, and Al-Sakib Khan Pathan

1 Introduction

Ransomware is a kind of malicious software (malware) that prohibits victims from accessing their computers and the information they save. The victims are usually asked to pay the ransom using cryptocurrency such as Bitcoin to regain access. Ransomware attacks pose a significant threat to national security. There has been a substantial surge of such attacks in the post-Covid era. In the last couple of years, the victims, primarily large enterprises, have started practicing good cyber hygiene; for instance, deploying data loss prevention mechanisms, improved backup strategies as possible countermeasures to avoid ransomware attacks, etc. However, cybercriminals have devised a hybrid variant called Ransomware 2.0. The sensitive data is first stolen before encryption, allowing the evil entities to release the information publicly if the ransom is not paid. Further, cybercriminals exploit the advantage of cryptocurrencies being anonymous and untraceable. Figure 1 shows the most notorious ransomware family members [1].

Recently, Russia invaded Ukraine, and several countries retaliated against Russia. A ransomware group threatened cyber-attacks on those countries' critical infrastructure (Fig. 2). Experts have warned that this could be the most widespread ransomware gang globally and is linked to a trend of Russian hackers who support the Kremlin's ideology [3]. Figure 3 illustrates the tracked ransom amount paid by victims as ran-

M. Ahmed (✉) · A. N. M. B. Rashid
School of Science, Edith Cowan University, Perth, WA, Australia
e-mail: mohiuddin.ahmed@ecu.edu.au

A. N. M. B. Rashid
e-mail: a.rashid@ecu.edu.au

A.-S. K. Pathan
Department of Computer Science and Engineering, United International University, Dhaka,
Bangladesh
e-mail: sakib.pathan@gmail.com

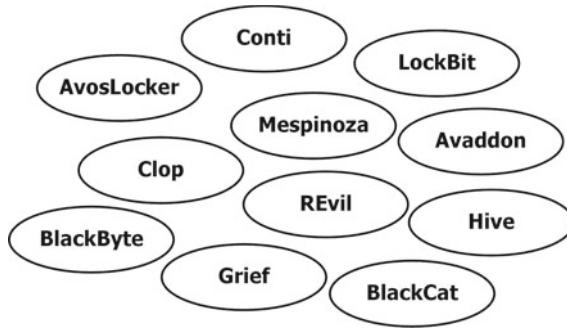


Fig. 1 Top members of the ransomware family

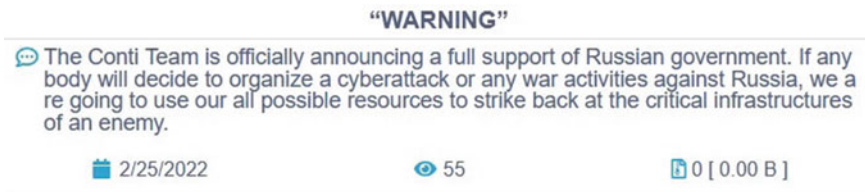


Fig. 2 Conti siding with Russia over the invasion of Ukraine

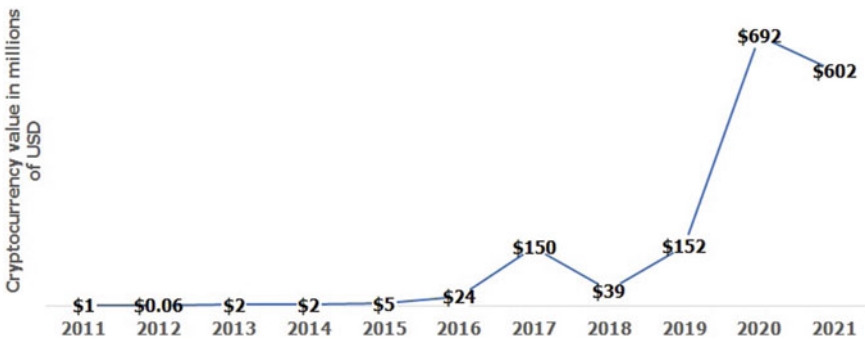


Fig. 3 Total cryptocurrency value received by ransomware addresses

som money by Chainalysis from 2011 to 2021 [2]. In 2020, the ransom amount was roughly US \$692 million, which is about four times higher than that of the previous year. An incomplete picture of 2021 shows a near ransom amount to that of 2020. It is expected that the accurate portrayal of 2021 will undoubtedly exceed the amount of 2020.

1.1 Key Contributions

The key contributions of the paper are as follows:

- We introduce a new variant of ransomware attack called Ransomware 3.0.
- We propose a framework to address Ransomware 3.0 using the combination of Mobility Markov Chain (MMC) and Differential Privacy.
- We showcase preliminary results to reflect the effectiveness of our proposed approach.

1.2 Paper Roadmap

The rest of the paper is organized as follows. Section 2 showcases the underlying architecture of Ransomware 3.0. Section 3 summarises the existing countermeasures for ransomware attacks. Section 4 discusses mobility Markov chain and differential privacy to address ransomware 3.0 attacks. Section 5 highlights preliminary experimental results and the paper is concluded in Sect. 6.

2 Ransomware 3.0

Figure 4 showcases the infection chain of Ransomware 3.0. It is shown that Ransomware 3.0 is unlike other ransomware attacks where cyber criminals are primarily concerned with encrypting the victims' devices to demand ransom. Instead, the cyber criminals are more interested in the sensitive data stored on the victims' devices and they attempt to replace the original data with misinformation. Then, finally, they encrypt the devices. Also, instead of ransom in the form of cryptocurrency to ensure anonymity, the sophisticated state-based cyber attackers will demand classified intelligence which would pose serious threats to national security.

3 Existing Ransomware Countermeasures

Ransomware is such an attack that requires defense in depth. Following is a brief discussion on the most generic approaches. Even if there is a plethora of techniques available, it is clear that the existing techniques are not suitable to address the Ransomware 3.0 attacks [10].

- EldeRan [11] uses a sandbox for performing the static and dynamic analysis of applications, such as directory operations, application programming interface (API) calls, strings of executables, registry key additions, and modifications and

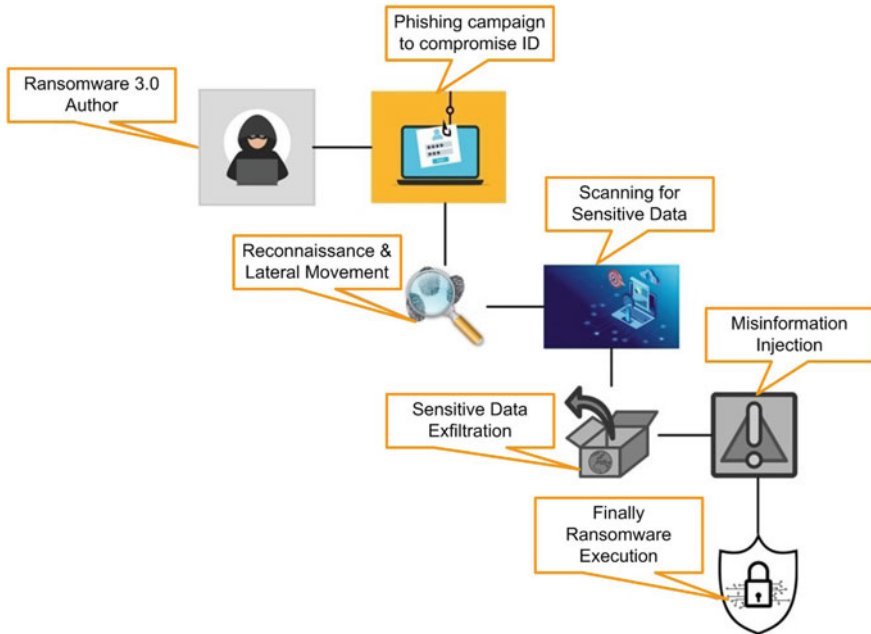


Fig. 4 Infection chain of Ransomware 3.0

dropped files analysis. The EldeRan sandbox performing these operations is an isolated system that can test the ransomware behavior. This sandbox can presuppose that ransomware executes behaviors and processes, which are significantly different compared to harmless software.

- RansomWall [12] was designed and developed for detecting ransomware attacks in real time. It can also conduct behavioral analysis by making use of a sandbox. Moreover, it was designed for the Windows operating systems. The system uses five different layers for analysis: (1) static analysis layer, (2) a trap layer, (3) dynamic analysis layer, (4) backup layer, and (5) machine learning layer.
- RansHunt [9] was designed and developed for detecting prevalent characteristics during a ransomware attack. RansHunt can deploy static and dynamic features, and they were built using the analysis of 21 ransomware families.

Regardless of the quality and usability of state-of-the-art ransomware detection techniques, the threat of Ransomware 3.0 is still present. Rather, it is more dangerous and has the capability to disrupt a nation’s regular business operations.

4 Effectiveness of Mobility Markov Chain and Differential Privacy

Algorithm 1 showcases the proposed countermeasure for Ransomware 3.0. The basic principle to address Ransomware 3.0 is to create mobility Markov chains ($MMC - SD_t$) for the sensitive data (SD) mobility, i.e., how the sensitive data travels in the regular and legitimate business operations. Depending on the data and computing environment, at regular intervals, the initial MMC will be compared with an MMC ($MMC - SD_{t+1}$) from a different time. If there is a significant difference, it is expected to come from a data exfiltration attempt. It should be mentioned here that exfiltration occurs when a malware or any malicious actor carries out an unauthorized data transfer from a device. The threshold, T , will be based on the context of the critical infrastructure and the data type itself.

Algorithm 1 Mitigation of Ransomware 3.0 Using Differential Privacy

```

Create an MMC for SD;
if  $d(MMC - SD_t, MMC - SD_{t+1}) \geq T$  then
     $SD_{DP} \leftarrow \text{Apply DP to SD}$ 
else if  $SD_{DP} \leftarrow \emptyset$  then
end if

```

4.1 Mobility Markov Chain

Mobility Markov Chain (MMC) has been widely used in the literature for solving many problems. As such, MMC has been effectively used in the de-anonymization attack on geolocation data that is accurate and resilient to sanitization techniques [6]. MMC is a probabilistic automation. A probabilistic transition between the two states, i.e., POIs (Points of Interest), is denoted by an edge. The states or POIs are the data movements inside the computing facility in this context. Each state is assigned a semantic term or a unique identifier, usually a string of letters and numbers, that serves as an address. An MMC can be built from the sensitive data movements observed during the training phase. Different observations can be used during the testing phase to run the test. Following that, numerous distance metrics (or similarity matrices) can be generated to determine how dissimilar two MMCs are. The process involves splitting the extensive collection of sensitive data movements/transactions into training and testing data. The MMC model is then applied to the training data (with identities) to build a trained model. A distance matrix or similarity matrix is formed that links MMCs with identities using the trained model and testing data. The similarity matrix is used to identify the suspicious data movements, i.e., data exfiltration attacks. An MMC describes a sensitive data movement (i.e., mobility behavior) as a discrete stochastic process [7]. The probability of the movement is

solely determined by the previous location visited and the probability distribution of the locations' transitions. In general, an MMC is divided into two stages:

- Set of addresses, $P = p_1, \dots, p_n$, is a collection of addresses in which each address represents a frequent POI (in descending order of importance), except for the last address p_n that represents the set consisting of the union of all infrequent POIs.
- Transitions, such as $t_{(i,j)}$, show the chance of moving from location p_i to location p_j .

An MMC can be represented as a graph or as a $(n \times m)$ -dimensional transition matrix. When the MMC is a transition matrix, the rows and columns represent the addresses, while the value of a cell represents the related transition between the addresses. Distances between MMCs can be computed using the following distance measures:

- Stationary Distance: The total of the distances between the nearest addresses of two MMCs can be used to calculate the stationary distance between them. To minimize the distance, the MMCs' addresses must be paired, resulting in an address of the first MMC that can be paired with several addresses of the second MMC.
- Proximity Distance: When two MMCs share "important" addresses between them, then a proximity distance can be calculated. How many times an address is visited corresponds to its importance.
- Matching Distance: This is similar to the stationary distance in the sense that it considers the sum of the distances between the addresses of the two MMCs. Unlike the stationary distance measure, however, one address of the first MMC can only be coupled with one address of the second MMC.
- Density-Based Distance: The density-based distance can be calculated using the sum of distances between the two MMCs' addresses, just as the stationary and matching distance measures. The main difference between the density-based distance and the other two is that MMC addresses are matched based on their rank after being sorted according to their associated probability vectors.

4.2 Differential Privacy

Differential privacy (DP) [5] is a technique that injects random noise into the dataset at the time of data analysis and ensures the privacy of an individual. Identifying individual information based on the outcome of an analysis will be impossible if noises are introduced. The analysis shows an approximation instead of accurate result or insight; hence, this will be effective when cybercriminals are trying to steal sensitive data. The differential privacy will be triggered when the MMC dissimilarity threshold is reached. Even in the case of a false alarm, the actual data owner will have no issues reproducing or retrieving the data as the differential privacy will be designed by the in-house team instead of the outsiders.

Epsilon (ϵ) is a parameter for the privacy loss for determining the amount of noise that needs to be injected into the dataset. Laplace distribution is a probability distribution function and can be used to derive the Epsilon that can determine the deviation in computation when an attribute is removed from the dataset. When a user's data is removed from the dataset, the Epsilon will be smaller, resulting in a smaller deviation in the computations. In other words, a higher Epsilon value can depict more accurate and lower private results. In comparison, a lower Epsilon value can provide a highly randomized result, and attackers will not be able to learn significantly. Therefore, a smaller Epsilon value can lead to significantly robust data preservation even when the computations are barely accurate. Accordingly, the differential privacy technique is suitable for incorporation into the data analysis process. However, this adoption depends on the trade-off between the accuracy and privacy of the data.

5 Preliminary Experimental Analysis

Figure 5 shows the effectiveness of differential privacy on four different datasets [4]. The code used to run the experiment is adapted from [8]. Unlike the original intent of differential privacy [5], we tried to see whether the added random noise to the sensitive data helps obfuscate the data for cyber criminals. If the data is made private as such the analysis on the data will provide no valuable intelligence, then it would be best to find the appropriate parameters for making the data private. In this case, it is seen that for four different datasets, the accuracy changes with the different values of ϵ , which is a crucial parameter. However, this is the second stage of the Ransomware 3.0 mitigation strategy. If the data exfiltration is identified successfully with the help of MMC, then all the associated data will be the input for differential privacy approach.

6 Conclusions and Future Research Directions

In this paper, we have incorporated Ransomware 3.0, which is deadlier than the existing ransomware variants. In this age of information warfare, state-based cyber actors are more interested in disrupting Internet-enabled day-to-day operations and misleading them by injecting false data after stealing sensitive data. Hence, it is crucial to identify the intrusion and data exfiltration activities before the Ransomware is executed. A potential solution is to combine the mobility Markov chain with differential privacy to mitigate the impacts of attempts taken by cyber criminals. Based on the preliminary experimental and theoretical analysis, it is evident that the cyber actors will find it challenging to make sense of and even harder to steal sensitive data from critical infrastructures. As part of our future research, we will endeavor to develop robust automated detection and prevention algorithms to address Ransomware 3.0.

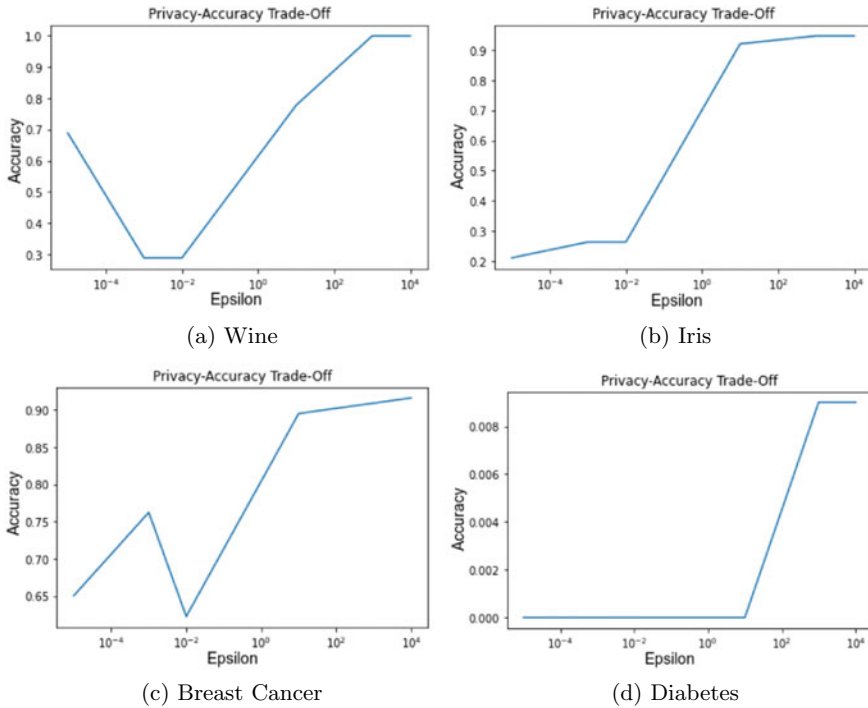


Fig. 5 Effectiveness of differential privacy to defend Ransomware 3.0

References

1. The 2022 threatlabz state of ransomware report (2022). <https://www.zscaler.com/blogs/security-research/2022-threatlabz-state-ransomware-report>. Last accessed 25 Oct 2022
2. Chainalysis, the 2022 crypto crime report (2022). <https://go.chainalysis.com/2022-Crypto-Crime-Report.html>. Last accessed 25 Oct 2022
3. Abrams L (2022) Conti ransomware finally shuts down data leak, negotiation sites. <https://www.bleepingcomputer.com/news/security/conti-ransomware-finally-shuts-down-data-leak-negotiation-sites/>. Last accessed 25 Oct 2022
4. Dua D, Graff C (2017) UCI machine learning repository. <http://archive.ics.uci.edu/ml>. Last accessed 25 Oct 2022
5. Dwork C (2008) Differential privacy: a survey of results. In: Agrawal M, Du D, Duan Z, Li A (eds) Theory and applications of models of computation. Springer, Berlin, pp 1–19
6. Gambs S, Killijian MO, Núñez del Prado Cortez M (2014) De-anonymization attack on geolocated data. *J Comput Syst Sci, Spec Issue Theory Appl Parallel Distrib Comput Syst* **80**(8), 1597–1614 (2014). <https://doi.org/10.1016/j.jcss.2014.04.024>, <https://www.sciencedirect.com/science/article/pii/S0022000014000683>. Last accessed 25 Oct 2022
7. Gambs S, Killijian MO, del Prado Cortez MNN (2010) Show me how you move and i will tell you who you are. In: Proceedings of the 3rd ACM sigspatial international workshop on security and privacy in GIS and LBS. SPRINGL '10, Association for Computing Machinery, New York, NY, USA, pp 34–41. <https://doi.org/10.1145/1868470.1868479>
8. Harrod J (2017) Privacy in machine learning—demo code | jordan harrod.ipynb. <https://github.com/harrodjordan>. Last accessed 25 Oct 2022

9. Hasan MM, Rahman MM (2017) Ranshunt: a support vector machines based ransomware analysis framework with integrated feature set. In: 2017 20th international conference of computer and information technology (ICCIT), pp 1–7. <https://doi.org/10.1109/ICCITECHN.2017.8281835>
10. Ahmed M, Bachmann SD, Ullah AB, Barnett S (2022) Ransomware 2.0: an emerging threat to national security. *Aust J Def Strat Stud* 4(1), 125–132 (2022). <https://doi.org/10.51174/AJDSS.0401/EMQH2521>, <https://defence.gov.au/ADC/publications/AJDSS/volume4-number1/ransomware2-0.asp>. Last accessed 25 Oct 2022
11. Sgandurra D, Muñoz-González L, Mohsen R, Lupu EC (2016) Automated dynamic analysis of ransomware: benefits, limitations and use for detection. <https://doi.org/10.48550/ARXIV.1609.03020>, <https://arxiv.org/abs/1609.03020>
12. Shaukat SK, Ribeiro VJ (2018) Ransomwall: a layered defense system against cryptographic ransomware attacks using machine learning. In: 2018 10th international conference on communication systems and networks (COMSNETS), pp 356–363. <https://doi.org/10.1109/COMSNETS.2018.8328219>

Securing Transmission of Medical Images Using Cryptography Steganography and Watermarking Technique



Satish Kumar, Pawan Kumar Chaurasia, and Raees Ahmad Khan

1 Introduction

Healthcare data systems are increasingly deployed in the recent healthcare domain. In actuality, radiology information systems (RIS), Please be aware that your name and affiliation and if applicable those of your co-author(s) will be published as presented in this proof. If you want to make any changes, please correct the details now. Please note that after publication corrections won't be possible. Due to data protection we standardly publish professional email addresses, but not private ones, even if they have been provided in the manuscript and are visible in this proof. If you or your co-author(s) have a different preference regarding the publication of your mail address(s) please indicate this clearly. picture archiving and communication systems (PACS), hospital information systems (HIS), and much other information and communications methods are essential to the operation of several hospitals and healthcare facilities across the world. These systems made it easier for doctors, radiologists, and patients to share medical data and EPR (Electronic Patient Records) data for telemedicine applications like tele-surgery, tele-consulting, and tele-diagnosis. Even with these inventive developments, when using public networks, then a malicious attacker can easily tamper and intercept transferred medical data. It is crucial to create protected medical transfer systems in telemedicine applications. To protect the transmission of healthcare data between healthcare organizations and telemedicine systems [1, 24, 25], it provides three pivotal security factors: Authentication, Confidentiality, and Integrity. Authentication verifies that the medical data obtained are from the right source and belong to the right persons or right patients. In confidentiality, only authorized persons are allowed access to the transmission of the medical data, however, integrity verifies that no unauthorized person has modified the received

S. Kumar (✉) · P. K. Chaurasia · R. A. Khan
Babasaheb Bhimrao Ambedkar University, ('A Central University') Vidya Vihar, Raebareli Road,
Lucknow, India
e-mail: satish993596@gmail.com

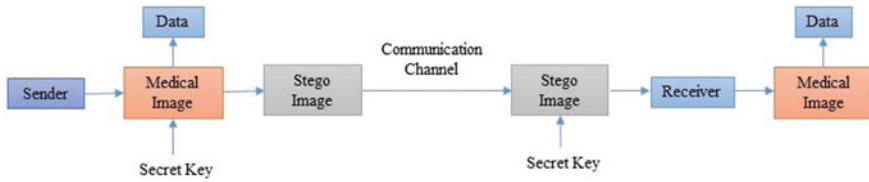


Fig. 1 General transmission of medical data through cryptography, watermarking, and steganography techniques

medical data. These securities are provided for secured telemedicine applications and are reliable [1, 2].

Nowadays, these security services are provided by watermarking, steganography, and cryptography technologies [3]. The cryptography technique is used in the DICOM standard. It provides authenticity and integrity by utilizing hashing functions, symmetric encryption, and digital signatures, but a main drawback of cryptography is that if the digital signature associated with the medical image is deleted or lost, the medical image becomes unreliable and thus its integrity and authenticity are difficult to verify. This shows that cryptography is only effective as an a priori security measure. In watermarking techniques, the use of both fragile and robust watermarks to secure telemedicine. Characterize robust watermarks that are resistant to malicious and signal processing attacks; so, these are suitable for authenticating identities and verifying ownership. Fragile watermarks, it is not surviving signal processing attacks, making them suitable for tamper detection and data integrity control [3–5]. In steganography techniques, it provides secure data confidentiality because it communicates imperceptibly inside digital entities, especially medical images [3, 6]. A combination of three transmission technique: cryptography, watermarking, and steganography. These hybrid techniques generally transmission of medical data are shown in Fig. 1.

The embedding process of data is hidden within the cover image by securely transmission algorithms. Commonly, to improve embedding algorithms, the sender and receiver of the data utilize a secret key and securities method. Stego-images are medical data with data embedding inside them because embedding algorithms assure that embedded data are invisible and, stego-image transfer generally takes place across a public communication channel. This embedding procedure is shown in Fig. 1. The receiver receives the embedded message from the stego-image using the data-decoded technique. The data have been received in original form by extract algorithms with the secret key and may contain any medical data distortion embedding process across communication channels when transmission a stego-images. On the receiver side, Fig. 1 also shows this, the original medical data can be received. The ability to fully recover the original medical data after extracting the embedded data [6]. Individual details of embedding and extracting procedures are cryptography, watermarking, and steganography algorithms in review papers [2, 7, 8]. The security concerns of telemedicine applications have led to the literature's proposal of cryptography, watermarking, and steganography algorithms.

The remaining portions of the paper are structured as follows: Section 2, introduction of transmission technique: cryptography, watermarking, and steganography. Section 3 describes the medical image analysis. Section 4, providers of healthcare data breaches. Sections 5 and 6, the discussion section addresses and identifies current issues and challenges. The last section is the conclusion of the paper.

2 Material and Methods

2.1 *Cryptography in Transform Domain*

The transmission domain-based image encryption techniques are regarded as one of the fundamental image encryptions. The transformed model is used to convert the given material from the spatial to the transform domain during the encryption process. These methods include the discrete cosine transform (DCT), fractional Fourier transform (FrFT), and gyrator transform (GT). Two complex matrices, one real and one imaginary, are used to evaluate the complex function for medical imaging security utilizing the gyrator transform and phase-truncated Fourier transform. Furthermore, one of the triplet functions applies a random distribution function. Attack evaluation and simulation results demonstrated the technique's usefulness and robustness against a variety of attacks. A fractional (DCT) system with a lot of freedom is used to secure medical data. After that, for FrDCT coefficients, a chaotic map is used. An examination of the proposed method's efficiency followed by a comparison with cutting-edge methods revealed that exceed is more effective than other algorithms [9]. A discrete cosine transform (DCT) paired with chaotic-based encryption for medical imaging. Once employing DCT, the medical image is compressed twice utilizing the mathematical encoding method once more. The output of the compressed image is then encrypted using a chaotic sequence. A medical image's host is secured using an encryption method to secure sensitive data. Bandelet Transform is used to retrieve the features from the medical image after it has been encrypted using a Logistic Chaotic Map [3, 10].

2.2 *Watermarking in the Transform Domain*

The decomposition of any digital data file into frequency coefficients occurs during the transform technique before the process of embedding important data. The robustness of this method against many attacks has several benefits. It can defend against numerous types of attacks on data that can be modified invisibly, as well as provide hidden data that can withstand changes in stego-file. The transform domain technique also has disadvantages, including a lesser payload than the spatial domain technique

and extremely complex computing [11, 26–28]. Furthermore, Among the techniques employed are DCT, DWT, IWT, and DFT [3, 12, 13, 27, 28].

DCT is utilized in signal processing applications and for image compression. According to a secret image quality criterion, DCT divides images into high-frequency sub-bands, medium-frequency sub-bands, and low-frequency sub-bands. DFT is used in various image processing methods. The DFT transform method is frequently employed in signal processing. The components of sine and cosine are separated from a signal. These elements can be changed to serve as a crucial tool for hiding data. The original image and the changing coefficients are both referred to as the stego-image [3].

2.3 Steganography in Transform Domain

The host media coefficients are regenerated in frequency domain mechanisms following the embedding procedure. The method uses DFT, DCT, RDWT, SVD, DWT, and other techniques. The computation for spatial-domain approaches is simpler than that for frequency domain, but they are less resistant to geometrical attacks. The image in DCT is divided into three parts: low (L), medium (M), and high (H). The low-frequency band is where the majority of the image energy is located. The medical image was divided into four LL, HL, LH, and HH segments in DWT. “L” denotes low while “H” denotes high. Additionally, the LL sub-band was constantly split to accomplish additional levels according to the quantity necessary by the application. In DFT, the image’s sine and cosine forms were resolved, and direct or template-based hiding was used to embed a watermark. DCT techniques are used for high robustness capacity hiding against signal processing and geometric attacks. The chosen medical image sub-bands contain an embedded electronic patient record [3].

2.4 Characteristics of Cryptography, Watermarking, and Steganography Techniques

The major characteristics of the cryptography, steganography, and watermarking scheme are shown in Table 1.

Medical image security, ownership authentication, digital right management, copyright protection, media forensics, tamper detection, localization, and many other uses are among the several applications for the transmission of medical images [3, 5, 14–16].

Table 1 Characteristics of cryptography, watermarking, and steganography techniques

Cryptography [3]	Watermarking [3, 6, 7]	Steganography [3, 6]
The cryptography method can be sensitive to any attacks such as filter, cropping, scaling, etc	In robustness, the watermarking method can be very sensitive data to any geometric attacks, including scaling, translating, compressing, rotating, copying, and cropping attacks	In robustness, the steganography method can be sensitive to any malicious attacks before the embedded secret message convert into a stego-image such as scaling, rotation, and loose compression
Cryptography security is embedding hidden data with a secret key in plain text to cipher text of extraction procedure and any unauthorized persons not access any data	In security, the watermark security is not to detect unauthorized persons and to protect against malicious attacks	In security, the main purpose of hidden data transmission with a secret key, and not to detect any unauthorized person and to protect against malicious attacks
In capacity, it provides the number of bits that a cryptography method can plain text to cipher text without any losses to original data quality	In capacity, it provides the number of bits that a watermarking method can encode of the data without degrading the original medical data's quality	In capacity, it provides the number of bits that a steganography method can embed without any losses to the quality of the data
In imperceptibility, it means the original image and extracted image are similar	In imperceptibility, it means similarity between the host image and the watermarked image. The process of watermarking affects the host image's perceptual quality since the watermark is embedded into it. It is always preferred that the watermark be integrated into the cover image to prevent a significant loss in the image's visual quality	In imperceptibility, it means that the cover image and the extract steganography image are similar
In cost, it is generally computational cost that involves the whole cryptography procedure	In cost, it is a typically computational cost that included the entire watermarking procedure, i.e. watermarking procedure is embedding and extraction watermark	In cost, it is a generally computational cost that involves the whole steganography process
Parties interested in secure communication must each have a secret key according to cryptography	Fragile refers to sensitivity to even the smallest changes. The most important characteristic of a fragile watermark is that it becomes undetected if it is subjected to unauthorized changes	Invisible is based on the human visual system (HVS) or human audio system (HAS)

3 Medical Image Analysis

A medical image is closely related to the area of diagnostic imaging, however, it is dependent on the computer interpretation of the images, not their collection. These techniques are classified into three main categories, including methodological tasks, clinical tasks, and anatomical applications [17], as presented in Fig. 2.

A technique is used for analysis in medical image registration. These techniques are also known as image mapping, fusion, or warping. It is a technique for combining various medical datasets into a single matched coordinate system with identical imaging information, and it is highly significant from a medical perspective, which has substantial medical significance, and a crucial step in the processing of images that allows for the transmission of meaningful information across many images. Medical image registration, image fusion, learning-based image registration, and image reconstruction are all utilized in various clinical applications. It can be put together from the perspective of the region of interest according to anatomical locations such as the brain, liver, lung, and so on. Enrolment strategies may be divided into three-dimensional to two-dimensional, three-dimensional to three-dimensional, and two-dimensional to two-dimensional from the perspective of the image pair measurement. Radiologists might find localization to be a simple procedure. In classification and detection, assign an absence, presence, normal, or abnormal illness to a diagnostic exam image, and a lesion or another intriguing object an image within an image. In segmentation, determine the organ's curvature or the area of its interior that is of interest, allowing the volume to be quantitatively analyzed concerning shape and form, as in the case of the heart or brain. In a wide range of computer-aided diagnostic system applications, clinical image segmentation is essential [17, 18].

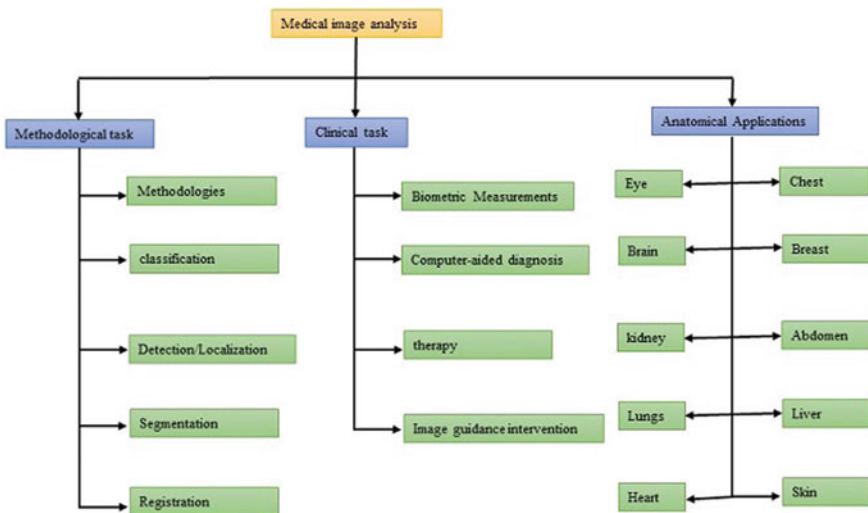


Fig. 2 Medical image analysis

Medical image processing methods are being used through X-ray, Ultrasound, CT scan, Microscopy, MRI, Elastography, Echocardiography, Nuclear Medicine, and PET [17, 19], as shown in Fig. 3. Medical imaging scans are commonly used to determine the many abnormalities that they reflect in the body’s many organs. The medical community is increasingly using these scans to make diagnoses, staging, and therapy of many diseases as a result of the intrinsic advancement in low-cost imaging and computational technologies. The two main categories of radiation are electromagnetic radiation and particle radiation. Infrared, ultraviolet, X-rays, and gamma radiation are the four main categories of electromagnetic radiation, whereas electrons, protons, positrons, and neutrons are particulate radiation. The general uses of several imaging techniques are provided [19].

X-rays are the most widely used type of medical imaging technique for identifying hard tissues and bone structures. The CT scan was able to record the minor change in tissues with more than 4% but less than 5% of contrast variation. It is clearly seen in the subject contrast, while the spatial resolution of items with a millimeter or less is better represented with an X-ray. Ultrasound imaging, which is particularly helpful in obstetrics, allows for the visualization of internal body structures such as tendons, joints, uterus, ovaries, muscles, and tendons. The intricate architecture of the brain, nerves, spinal cord, tendons, and ligaments may all be imaged using MRIs since they produce a far sharper image than CT. There are many applications for PET and SPECT in clinical oncology. Detecting cancer, revealing the spread of cancer, determining if the tumor is responding to treatment, and identifying the faulty chemical functioning of numerous other human organs are just a few examples. The most common cancers and lesions seen by PET and SPECT on the body are lymphomas, esophageal, melanoma, cervical, prostate, pancreatic, brain tumors, and Alzheimer’s disease [18, 19].

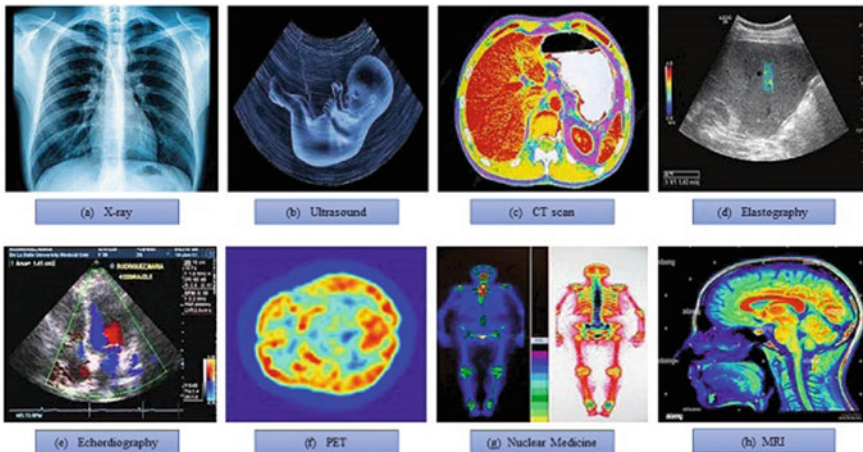


Fig. 3 Common types of medical images

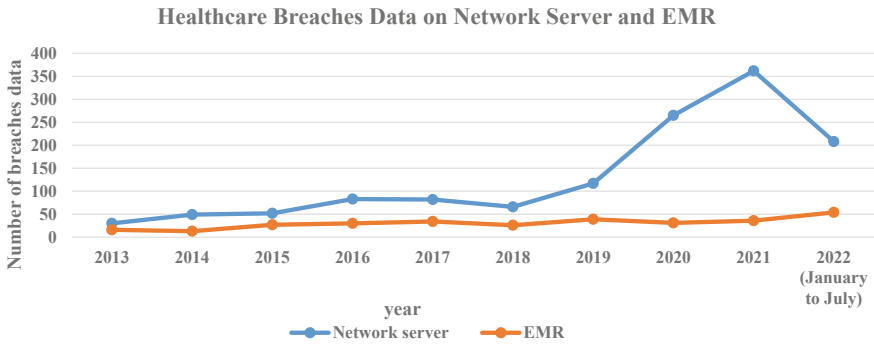


Fig. 4 Graph of healthcare breaches data on network server and EMR

4 Healthcare Data Breaches on Network Server and EMR

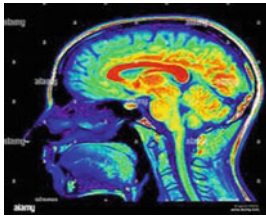
Electromechanical storage and paper system are both used to store secured healthcare data. This section describes the network server and electronic medical report (EMR) that allow access secure healthcare data. Annual data on the locations of data breach cases are presented in Table 2. The information shown in this table was gathered from HIPAA reports (Fig. 4). The HIPAA journal’s publication of healthcare data breaches between 2013 and 2022 is a 10-year healthcare data breach in two locations of network server and electronic medical record (EMR) [20–23].

In Table 3, two locations, i.e. Network Server and Electronic Medical Records (EMR) are two locations from where secured medical data was breached according to the 10 years’ data breach analysis, out of the 2019 to 2022 breaches, healthcare data are the maximum network server and EMR breaches. A comparison of these locations is shown in Table 2 based on the annual average number of breaches that occur at each location. The reader will better understand the findings of this analysis if they are presented in a graphical format, which will also make it easier to map the variation in healthcare information breach instances that occurred in various regions during a 10-year (2013–2022) period.

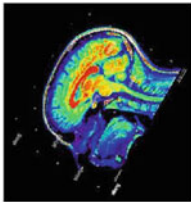
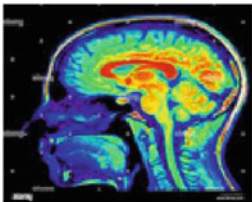
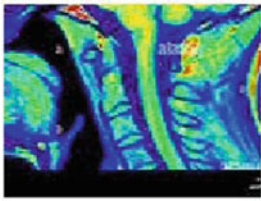
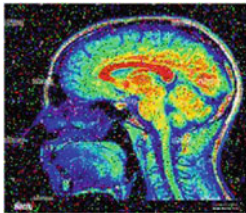
5 Discussion

Transmission of the medical images is done so that data are not lost through any networks and no attackers can do attacks. To secure these medical data, we have told use of three combined transmission techniques (cryptography, steganography, and watermarking). So that if we transmission the data from any network to the medical image. Then our data remain secure and data are not redundant so that we do not have any problems in getting treatment from doctors. We have told about some medical image attacks, what types of medical attacks affect the medical image and

Table 2 Medical image attacks

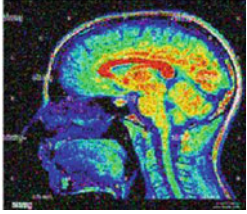
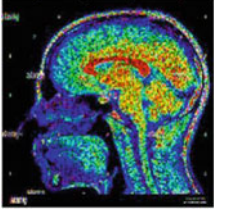
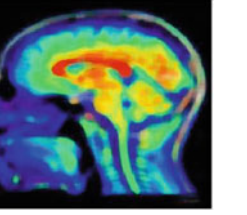
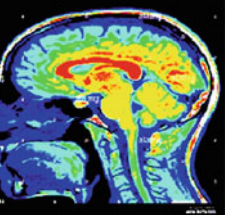


Original image

Category	Attacks	Description	Effect image
Geometric [4] attacks	Rotation	You can rotate an image in either a clockwise or counterclockwise direction using this feature	
	Scaling	It is the process of resizing a medical image	
	Cropping	To improve framing or composition, bring attention to the topic of the image, or change the size, or aspect ratio of an image, it may be necessary to crop or eliminate the image's outer edges	
Noise attacks [1, 4]	Salt and peppers	Images are added with some white and black pixels as a form of noise	

(continued)

Table 2 (continued)

	Gaussian noise	Gaussian noise is a probability density function (PDF) with statistical noise equal to the normal distribution	
	Speckle noise	It is a multiplicative noise that affects pixels in a medical image	
De-noising attacks	Median filter	It is a method of non-linear digital filtering for eliminating the noise of medical images or signals	
Image processing attacks	Histogram equalization	It is a technique for adjusting image contrast in image processing	

it can be difficult for the doctor to treat, which are Medical Attacks like Geometrics Attacks, Noise Attacks, De-noising Attacks, and Image Processing Attacks as shown in Table 2. We have told about the medical image analyses are X-ray, Ultrasound, CT scan, Electrography, Echocardiography, and MRI. Which image is used for what and its quality has been told. Medical data breaches and their location have been told. How much data is breached through server networks and EMR? Which we have mentioned in Table 3. The breaches data year-wise (2013–2022) have been shown so that we can know how much data is lost through server networks and EMR. To secure these data, we have described the use of hybrid technology, which is CSW (cryptology, steganography, and watermarking) transmission domain techniques.

Table 3 Healthcare data breaches on network servers and EMR

Year	Network server	EMR
2013	30	16
2014	49	13
2015	52	27
2016	83	30
2017	82	34
2018	66	26
2019	117	39
2020	265	31
2021	362	36
2022 (January to July)	208	54

By which medical data can be reduced and secured. After discussing the entire research paper, I have identified several issues and challenges.

6 Issues and Challenges

The security of transmission medical data faces increasing challenges as a consequence of growing cryptography, steganography, and watermarking (CSW)-based multimedia applications. After discussing the many research paper [1–19], we have come out with some research issues and challenges, which are given below.

- Major issues for media data security and privacy protection in CSW include communication and computation costs. Although certain encryption methods provide greater security, their computational costs rise as data volumes rise. The energy consumption and computing power of CSW methods are, thus, furthermore challenged through increased security and privacy requirements. Lightweight security and privacy solutions are necessary to manage situations involving scarce resources (such as memory, processing power, and energy) and insufficient cyber security measures.
- Data exchange. Healthcare professionals are given access to data, but it is unclear how that data are handled and how many people are engaged before it is delivered to the doctors. Patients may receive the incorrect medication or perhaps no prescription from the doctors if a data hacker tampers with the information. It can make their treatment take longer or be more dangerous to their health. So, it is necessary to regulate data sharing or implement a mechanism to safeguard data transfer in the healthcare system.
- The lack of openness and governance surrounding the usage of multimedia data makes it difficult to secure patient privacy while utilizing CSW technique advantages. It is challenging to inform people that their medical information is being

collected. Users may not be aware that their medical data are being collected automatically in healthcare. So, the CSW technique needs to be transparent and well-governed to safeguard the use of patients' medical information and maintain their right to privacy.

- Various types of medical data require additional protection and privacy. Additionally, because the selection of storage is arbitrary, miners are unable to identify health data repositories.
- The processing of plain text data by miner nodes compromises patients' privacy.
- The transfer of medical data to support eHealth applications may be secured by additional security techniques proposed by researchers.
- To strike a balance between data security and utility while lowering costs through appropriate classification.
- Most image CSW methods use a grayscale image as covert information and more research is needed to hide the medical image.
- Several techniques have considered imperceptibility, robustness, security, and data-hiding capacity as performance measures. However, when medical data transfers happen through untrusted channels, there are possibilities of man-in-the-middle attacks. Medical images can also be tampered with during transfer. The effectiveness of the designed method against all these attacks as well as these attacks themselves might be taken into consideration for evaluating metrics.

7 Conclusion

In this paper, the transmission of the medical data using secret key and hybrid cryptography, steganography and watermarking (CSW) techniques and converting medical image into stego-image so that medical image is secure and data is not reduction. The hybrid CSW transmission domain technique is suggested for secure data transmission, data reduction, and quality of medical images. This technique is mainly focused on medical image security and quality. Medical image analysis (i.e. X-ray, Ultrasound, CT scan, Elastography, Echocardiography, Nuclear medicine, PET, and MRI) for medical image quality that is suitable for diagnosis of disease. These suggest that schemes are largely based on CSW technology to enable communication security. There are many methods specially designed for medical image security and privacy protection. Further, last 10 years (2013–2022) healthcare data breaches of network servers and electronic medical records (EMR) were the maximum number of data breaches on network servers in the last 4 years (2018–2022) than EMR.

References

1. Al-Haj A, Mohammad A, Amer A (2017) Crypto-watermarking of transmitted medical images. *J Digit Imaging* 30(1):26–38. <https://doi.org/10.1007/s10278-016-9901-1>
2. Kasim Ö (2022) Secure medical image encryption with Walsh—Hadamard transform and lightweight cryptography algorithm. *Med Biol Eng Comput* 1585–1594. <https://doi.org/10.1007/s11517-022-02565-5>
3. Magdy M, Hosny KM, Ghali NI, Ghoniemy S (2022) Security of medical images for telemedicine: a systematic review. *Multimedia Tools and Applications*
4. Benrhouma O (2022) Cryptanalysis and improvement of a semi-fragile watermarking technique for tamper detection and recovery
5. Borra S, Thanki R (2020) Crypto-watermarking scheme for tamper detection of medical images. *Comput Methods Biomech Biomed Eng Imaging Vis* 8(4):345–355. <https://doi.org/10.1080/21681163.2019.1595730>
6. Evsutin O, Melman A, Meshcheryakov R (2020) Digital steganography and watermarking for digital images: a review of current research directions. *IEEE Access* 8:166589–166611. <https://doi.org/10.1109/ACCESS.2020.3022779>
7. Kadian P, Arora SM, Arora N (2021) Robust digital watermarking techniques for copyright protection of digital data: a survey. *Wirel Pers Commun* 118(4):3225–3249. <https://doi.org/10.1007/s11277-021-08177-w>
8. Bhalerao S, Ahmad I, Kumar A (2022) “Reversible ECG Watermarking for Ownership Detection, Tamper Localization, and Recovery”, *Circuits. Syst Signal Process.* <https://doi.org/10.1007/s00034-022-02024-4>
9. Fang Y, Liu J, Li J (2022) Robust zero-watermarking algorithm for medical images based on SIFT and Bandelet-DCT, pp 16863–16879
10. Kumar S, Panna B, Kumar R (2019) Medical image encryption using fractional discrete cosine transform with chaotic function, pp 2517–2533
11. Kaur J (2017) An adaptive quad tree based transform domain steganography for textual data. In: 2017 international conference on energy, communication, data analytics and soft computing, pp 3194–3199
12. Thanki R, Borra S, Dwivedi V, Borisagar K (2017) A steganographic approach for secure communication of medical images based on the DCT-SVD and the compressed sensing (CS) theory. *Imaging Sci J* 00:1–11. <https://doi.org/10.1080/13682199.2017.1367129>
13. Arunkumar S, Subramaniaswamy V, Vijayakumar V, Chilamkurti N, Logesh R (2019) SVD-based robust image steganographic scheme using RIWT and DCT for secure transmission of medical images. *Measurement.* <https://doi.org/10.1016/j.measurement.2019.02.069>
14. Thabit R (2021) Review of medical image authentication techniques and their recent trends. *Multimed. Tools Appl.* 80(9):13439–13473. <https://doi.org/10.1007/s11042-020-10421-7>
15. Kaur S, Singh S, Kaur M, Lee HN (2022) A systematic review of computational image steganography approaches. *Arch Comput Methods Eng* 0123456789 (2022). <https://doi.org/10.1007/s11831-022-09749-0>
16. Dhawan S, Chakraborty C, Frnda J, Gupta R, Rana AK, Pani SK (2021) SSII: secured and high-quality steganography using intelligent hybrid optimization algorithms for IoT. *IEEE Access* 9:87563–87578. <https://doi.org/10.1109/ACCESS.2021.3089357>
17. Balasamy SSVSK (2022) A review on deep learning in medical image analysis. *Int J Multimed Inf Retr* 11(1):19–38. <https://doi.org/10.1007/s13735-021-00218-1>
18. Muhammad S, Muhammad A, Adnan M, Muhammad Q, Majdi A, Khan MK (2018) Medical image analysis using convolutional neural networks: a review, pp 1–13
19. Journal P (2018) Noise issues prevailing in various types of medical images, vol 11, p. 1227–1237
20. Seh AH et al (2020) Healthcare data breaches: insights and implications. *Healthc* 8(2):1–18. <https://doi.org/10.3390/healthcare8020133>
21. January to December 2020 each months of Healthcare Data Breach Report. <https://www.hipaajournal.com/january-2020-healthcare-data-breach-report/>. Accessed 07 Sept 2022

22. January to December 2021 each months of Healthcare Data Breach Report. <https://www.hipaajournal.com/january-2021-healthcare-data-breach-report/>. Accessed 07 Sept 2022
23. January to December 2022 each months of Healthcare Data Breach Report. <https://www.hipaajournal.com/january-2022-healthcare-data-breach-report/>. Accessed 07 Sept 2022
24. Saraswat D, Chaurasia BK (2013) AHP based trust model in VANETs. In: The IEEE 5th international conference on computational intelligence and communication networks (CICN2013), Mathura, India, pp 391–393. ISBN No: 978-0-7695-5069-5. <https://doi.org/10.1109/CICN.2013.86>
25. Sharma K, Soni S, Chaurasia BK (2014) Reputation and trust computation in VANETs. In: International conference on electronics engineering and computer science (IEMCON2014) organized by Elsevier, Kolkata, India, ISBN No: 9789351072485
26. Singh MP, Rai A, Chaurasia PK (2020) Study and analysis of digital watermarking for medical images. *Invertis J Sci Technol* 13(1):1–7
27. Chaurasia PK, Tiwari SK, Ansar SA, Yadav N, Soni N, Singh S (2022) The security of transforming digital medical image using RDWT, HT, and SVD techniques. *Harbin Gongye Daxue Xuebao/J Harbin Inst Technol* 54(10):270–276
28. Kumar S, Srivastava A, Chaurasiya PK, Kushawaha A, Vishal V (2022) DCT and SVD-based watermarking technique for imperceptibility and robustness of medical images. In: 2022 4th international conference on advances in computing, communication control and networking (ICAC3N), pp 2335–2339. IEEE

New Commitment-Based Client–Server Key Establishment Protocol



Varun Shukla , Surendra Talari , Shishir Kumar , P. Vinooth ,
and Harikesh Singh 

1 Introduction

In the modern scenario, the importance of information security is increasing day by day. Cryptography bears the responsibility to provide necessary information security. Cryptography provides secrecy (or confidentiality), authentication, data integrity and non-repudiation. They are known as goals of information security or cryptographic goals. These goals can be studied from any suitable source for better understanding [1–4]. Key agreement is a process where participating entities contribute in such a way that a session key can be deduced securely. This session key is used in various symmetric key encryption algorithms. The concept of commitment is taken from game theory, and it shows binding and revealing properties, which can be very useful for the development of new key agreement protocols in order to surprise intruders [5, 6]. The binding property binds a participating entity to a specific value that can't be changed later on. Revealing property gives the freedom to show the selected value at a selected point of time or a definite stage of a protocol. The remaining paper is organized as follows. Section 2 talks about the proposed method. Security analysis and advantages are discussed in Sect. 3. Conclusion and future scope are given in Sect. 4.

V. Shukla (✉)

Department of ECE, Pranveer Singh Institute of Technology, Kanpur, India
e-mail: varun.shukla@gmail.com

S. Talari

Department of Mathematics, GIS, GITAM Deemed to be University, Visakhapatnam, India

S. Kumar

BBAU (A Central University of India), Lucknow, India

P. Vinooth · H. Singh

JSS Academy of Technical Education, Noida, India

2 Proposed Method

The proposed method is a session key establishment protocol between client and server. The following abbreviations are used in the proposed protocol, and they are defined in Table 1 for easy understanding of the readers of this paper.

The proposed protocol is based on core commitment theory so it is important here to understand it before the implementation results.

- **Core commitment theory:** The core commitment concept acts like a game between two players where a player writes his choice in a piece of paper and puts it into a sealed box [7, 8]. This step is called binding step as a particular user can't change his choice now but has a freedom to reveal it at any point of time. Based on mutual understanding or appropriate reply by the second player, the first player shows his choice and it is called as revealing step. Here, two parties are client and server and both perform commitment steps. It is assumed that the mutual authentication is already done between client and server and now it is the turn to deduce the shared session key. Initially, as a first step, client sends b_c to server and binds to a particular value. In reply, server takes the similar action and sends b_s to client. At this stage, both the entities are bounded and can't alter their choices further. Now client sends r_c to server and as a reply server sends r_s to client. The core illustration is shown in Fig. 1.
- **Implementation and test results:** Now the test readings with the required steps are shown here.

Table 1 Showing the abbreviations used in the proposed method

Serial number	Symbol	Meaning
1	b_c	Binding step initiated by client
2	r_c	Revealing step initiated by client
3	b_s	Binding step initiated by server
4	r_s	Revealing step initiated by server
5	C	Client
6	S	Server
7	I_{c1}	First image component of client
8	I_{c2}	Second image component of client
9	I_{s1}	First image component of server
10	I_{s2}	Second image component of server
11	I_c	$I_c = I_{c1} + I_{c2}$ Performed at server's side
12	I_s	$I_s = I_{s1} + I_{s2}$ Performed at client's side
13	K_c	Key component of client
14	K_s	Key component of server
15	K	Shared session key, i.e. $K = K_c K_s$

Fig. 1 Showing the illustration of core commitment concept

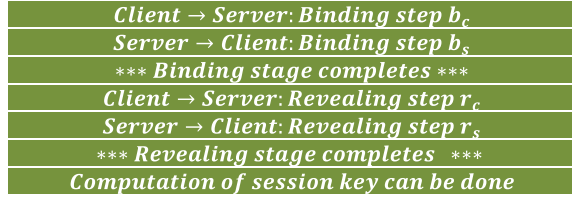


Table 2 Showing the key components in various formats

Key	Decimal	Binary	Hexadecimal
K_c	3543350479	11110011001101110011110011000011	B335BCC1
K_s	2599950188	10111101101100001101011111011101	B98FD8E3

Step 1: b_c : C sends I_{c1} to S

Step 2: b_s : S sends I_{s1} to C

Step 3: r_c : C sends I_{c2} to S and hence S performs $I_c = I_{c1} + I_{c2}$ & deduce K_c

Step 4: r_s : S sends I_{s2} to C and hence C performs $I_s = I_{s1} + I_{s2}$ & deduce K_s

Session key establishment: Both the parties will compute the session key as $K = K_c || K_s$. The following readings (for different key sets) are taken as shown in Table 2.

The corresponding test results for I_c , I_{c1} and I_{c2} are shown in Table 3.

Similarly, the corresponding test results for I_s , I_{s1} and I_{s2} are shown in Table 4.

The deduced session keys in all the above three cases are shown in Table 5.

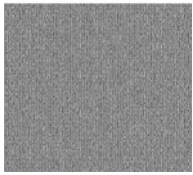

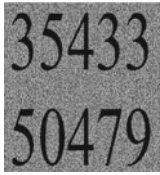
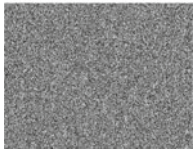
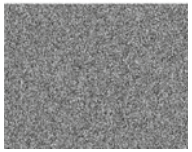
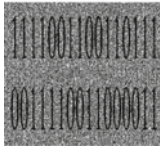
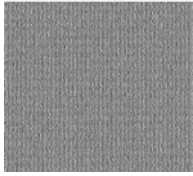
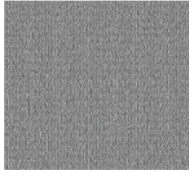
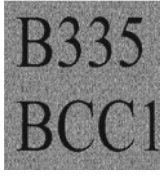
The logic behind is that I_{c1} and I_{s1} are unique to I_{c2} and I_{s2} , respectively. That means I_{c1} and I_{s1} can't express anything alone but bind both the entities because I_c and I_s can only be calculated when the second components, i.e. I_{c2} and I_{s2} are provided such that $I_c = I_{c1} + I_{c2}$ and $I_s = I_{s1} + I_{s2}$. The logic utilized for white and black pixels processing is given in Fig. 2.

Protocol execution representation: The overall run of the proposed key establishment protocol is shown in Fig. 3.

3 Security Analysis and Advantages

- Core commitment-based key agreement:** The proposed protocol is a commitment-based key establishment protocol that utilizes image components as binding and revealing parameters. In the binding phase, the occurrence of b_c and b_s , i.e. I_{c1} and I_{s1} between C and S makes sure that both the entities now have to stick on a particular value. I_{c1} and I_{s1} individually do not express anything, and it shows the way to revealing step as C and S have the freedom to reveal I_{c2} and I_{s2} based on mutual exchange strategy. The deduced key $K = K_c || K_s$ is the resultant of the process and C and S both have contributed equally in the process. If any



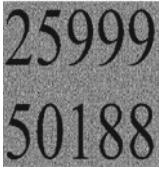


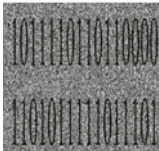
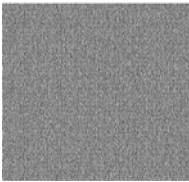

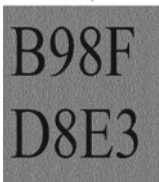
Table 3 Showing the test results for I_c , I_{c1} and I_{c2}

Serial number	I_c	I_{c1}	I_{c2}	$I_c = I_{c1} + I_{c2}$ Calculated by S
1	35433 50479			
2	1111001100110111 0011110011000011			
3	B335 BCC1			

one of the entities is not satisfied in the binding steps, then the protocol run can be terminated because r_c and r_s are also required to calculate I_c and I_s and hence K .

- Role of TTP:** In the proposed key establishment protocol, Trusted Third Party (TTP) is not involved in any manner. It has been observed plenty of times that inclusion of TTP creates many complications and difficulties [9, 10]. The involvement of TTP always turns the security into conditional one because if the TTP is not reliable, there is no question of any security. The proposed method overcomes all these difficulties as it is a direct commitment-based exchange of image components between C and S . The key establishment is a result of successful occurrences of b_c , b_s , r_c and r_s which require no involvement of any TTP.
- Computational overheads:** The key establishment occurs with very minimal computational expenses in the proposed protocol. The binding and revealing steps contain images I_{c1} , I_{s1} , I_{c2} and I_{s2} , respectively. The size of these images is in KBs and hence saves memory. It is also important to mention that the key is computed by C and S if and only if binding and revealing steps occur successfully. If any of the step is missing, the participating entities will not go for session key establishment and hence saves computational resources (in failure cases). It is an interesting feature as it is noticed that clients/servers are required to do computations even in failure cases, and it is rigorously enjoyed by intruders.

Table 4 Showing the test results for I_s , I_{s1} and I_{s2}

Serial number	I_s	I_{s1}	I_{s2}	$I_s = I_{s1} + I_{s2}$ Calculated by C
1	25999 50188			
2	1011110110110000 110101111011101			
3	B98F D8E3			

Intruders keep the servers busy by launching DoS (Denial of Service) attack and consume the available bandwidth in such a way that legitimate users will not get it which in turn reduces QoS (Quality of Service) [11, 12].

- Key length variations:** It is very interesting that the proposed key establishment protocol provides flexible key length. C and S can make the key variable at any point of time by changing the corresponding images. The key can also be represented in any format, i.e. binary, decimal or hexadecimal etc. These size and format variations pose difficulties for intruders and make the proposed method customized and suitable for high-security applications [13, 14].
- Incorporation of timer:** Incorporation of timer can be applied in the proposed method for increasing the level of security [15, 16]. After the binding step, a timer can be started that means both C and S will have limited time to perform r_c and r_s , respectively. It will add another layer of security because in case of any intruding efforts, the intruders have a limited time (equal to timer duration) to launch the attack as session will be terminated automatically after the timer duration. So, the timer bounds an intruder in a very short time span to launch any attack and needless to say if this timer is made variable (in every protocol run), it will create another obstacle for intruders.

Table 5 Showing shared session key between C and S

Serial number	Representation	$K = K_c \parallel K_s$
1	Decimal	35433504792599950188
2	Binary	11110011001101111001110011000011101110110110101111011101
3	Hexadecimal	B335BCC1B98FD8E3

```

%For White Pixels
%White Pixel shares
disp('White Pixel Processing...');
sla=[1 0];
s1b=[1 0];
[x y] = find(inImg == 1);
len = length(x);

for i=1:len
    a=x(i);b=y(i);
    pixShare=generateShare(sla,s1b);
    share1((a),(2*b-1):(2*b))=pixShare(1,1:2);
    share2((a),(2*b-1):(2*b))=pixShare(2,1:2);
end
        
```

```

%For Black Pixels
%Black Pixel shares
disp('Black Pixel Processing...');
s0a=[1 0];
s0b=[0 1];
[x y] = find(inImg == 0);
len = length(x);

for i=1:len
    a=x(i);b=y(i);
    pixShare=generateShare(s0a,s0b);
    share1((a),(2*b-1):(2*b))=pixShare(1,1:2);
    share2((a),(2*b-1):(2*b))=pixShare(2,1:2);
end
        
```

Fig. 2 Representing the logic for the processing of white pixels (left side) and black pixels (right side)

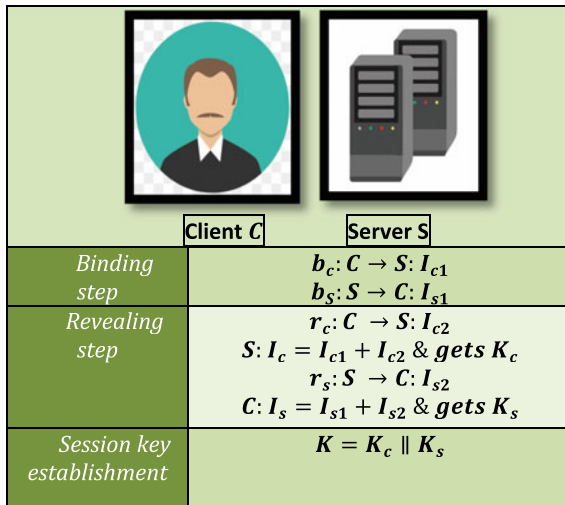


Fig. 3 Illustrating the overall run of the proposed protocol

- Machine Learning (ML)-based procedure:** The incorporation of timer can convert the proposed key establishment protocol into a ML-driven protocol [17]. Here, server plays a very important role. Based on the any malicious activity observed at client’s side, the server can reduce the timer duration and puts the timer into a variable mode. The reduced timer will create hurdles for intruders because they have to launch the security attacks in a reduced time interval now. So, server as a machine learns the usage pattern of client and based on the learning, server decides the timer value. So, it can be said that the proposed method is easily convertible to a ML-based key establishment protocol depending on the security requirements.

- **Kerckhoff's law:** Famous cryptographer Auguste Kerckhoff had mentioned a very important rule for the security of cryptosystems. He said that a given cryptosystem should remain secure even if intruders know everything but not the key [18]. The proposed key establishment is a resultant of binding and revealing steps (with no involvement of TTP) and the key prediction is infeasible to intruders. Since the key is secured, the further encrypted communication between C and S will remain secure even if other things are known to intruders and it clearly indicated that the proposed system satisfies Kerckhoff's law perfectly.
- **Lightweight protocol for IoT applications:** The proposed key establishment protocol is a lightweight, computationally optimized protocol based on image components and commitment concept. The binding and revealing steps require only the exchange of images, i.e. I_{c1} , I_{s1} , I_{c2} and I_{s2} sequentially. No complicated calculation is required for the session key establishment and it makes the protocol very suitable for IoT (Internet of Things) devices. As IoT devices are low power and low memory devices and can't bear computationally expensive procedures. The security of IoT devices is in current trend and excessive research is going on in it [19, 20]. So we believe that the proposed method can set a new path for the security of IoT devices along with other applications as well [21–38].

4 Conclusion and Future Scope

A commitment-based key establishment protocol between client and server is presented in this paper. The protocol uses image components in binding and revealing steps. The shared session key is the outcome of binding and revealing steps and requires an equal contribution of client and server. The proposed protocol saves computational overheads and provides unconditional security because of no involvement of any TTP. The shared key length can be made variable very easily and can be represented in any format. The incorporation of timer can convert the proposed protocol into a ML-driven key establishment protocol, which can be customized in plenty of applications. The proposed method satisfies Kerckhoff's law of secrecy and suitable for IoT devices as well. Many variations of the proposed protocol can also be presented in future.

Acknowledgements The authors thank the editor and the anonymous reviewers for reviewing this article and providing valuable and kind suggestions.

Conflict of Interest The authors declare no competing interests.

References




1. Shukla V, Chaturvedi A, Misra MK (2021) On authentication schemes using polynomials over non commutative rings. *Wirel Pers Commun* 118(1):1–9. <https://doi.org/10.1007/s11277-020-08008-4>
2. Misra MK, Chaturvedi A, Tripathi SP, Shukla V (2019) A unique key sharing protocol among three users using non-commutative group for electronic health record system. *J Discrete Math Sci Cryptogr* 22(8):1435–1451. <https://doi.org/10.1080/09720529.2019.1692450>
3. Shukla V, Chaturvedi A, Srivastava N (2019) A secure stop and wait communication protocol for disturbed networks. *Wirel Pers Commun* 110:861–872. <https://doi.org/10.1007/s11277-019-06760-w>
4. Chaturvedi A, Shukla V, Misra MK (2018) Three party key sharing protocol using polynomial rings. In: 5th IEEE Uttar Pradesh section international conference on electrical, electronics and computer engineering (UPCON), pp 1–5. <https://doi.org/10.1109/UPCON.2018.8596905>
5. Antikainen M, Sethi M, Matetic S, Aura T (2015) Commitment-based device-pairing protocol with synchronized drawings and comparison metrics. *Pervasive Mob Comput* 16(part B):205–219. <https://doi.org/10.1016/j.pmcj.2014.10.006>
6. Jiang Q, Chen Z, Ma J, Ma X, Shen J, Wu D (2021) Optimized fuzzy commitment based key agreement protocol for wireless body area network. *IEEE Trans Emerg Top Comput* 9(2):839–853. <https://doi.org/10.1109/TETC.2019.2949137>
7. Shukla V, Srivastava N, Chaturvedi A (2016) A bit commitment signcryption protocol for wireless transport layer security (wtls). In: IEEE international conference on electrical, computer and electronics engineering, pp 83–86. <https://doi.org/10.1109/UPCON.2016.7894629>
8. Shukla V, Misra MK, Chaturvedi A (2021) A new authentication procedure for client-server applications using HMAC. *J Discrete Math Sci Cryptogr* 24(5):1241–1256. <https://doi.org/10.1080/09720529.2021.1932908>
9. Alshanty A, Ersan I (2016) Trusted third party authentication protocol development for clustered wireless sensor networks. *Int J Commun Netw Syst Sci* 9(11):451–470. <https://doi.org/10.4236/ijcns.2016.911037>
10. Trivedi HS, Patel SJ (2021) Privacy preserving scalable authentication protocol with partially trusted third party for distributed internet-of-things. In: 18th international conference on security and cryptography, pp 812–818. ISBN: 978-989-758-524-1
11. Eliyan LF, Pietro RD (2021) DoS and DDoS attacks in software defined networks: a survey of existing solutions and research challenges. *Futur Gener Comput Syst* 122:149–171. <https://doi.org/10.1016/j.future.2021.03.011>
12. Kaur G, Saxena V, Gupta JP (2020) Detection of TCP targeted high bandwidth attacks using self-similarity. *J King Saud Univ—Comput Inf Sci* 32(1):35–49. <https://doi.org/10.1016/j.jksuci.2017.05.004>
13. Amalarethnam IG, Leena HM (2017) Enhanced RSA algorithm with varying key sizes for data security in cloud. In: World congress on computing and communication technologies, pp 172–175. <https://doi.org/10.1109/WCCCT.2016.50>
14. Bibak K, Kapron BM, Srinivasan V (2022) Authentication of variable length messages in quantum key distribution. *EPJ Quantum Technol* 9: 1–20. Article number 8. <https://doi.org/10.1140/epjqt/s40507-022-00127-0>
15. Afifi MH, Zhou L, Chakrabarty S, Ren J (2018) Dynamic authentication protocol using self-powered timers for passive internet of things. *IEEE Internet Things J* 5(4):2927–2935. <https://doi.org/10.1109/JIOT.2017.2757918>
16. Alezabi KA, Hashim F, Hashim SJ, Ali BM, Jamalipour A (2020) Efficient authentication and re-authentication protocols for 4G/5G heterogeneous networks. *EURASIP J Wirel Commun Netw* 105:1–34. <https://doi.org/10.1186/s13638-020-01702-8>
17. Ma Z, Liu Y, Wang Z, Ge H, Zhao M (2020) A machine learning-based scheme for the security analysis of authentication and key agreement protocols. *Neural Comput Appl* 32:16819–16831. <https://doi.org/10.1007/s00521-018-3929-8>

18. Kerckhoffs A (1883) LA cryptographie militaire. *Journal des sciences militaires* 9:161–191. https://www.petitcolas.net/kerckhoffs/crypto_militaire_2.pdf
19. Rana S, Obaidat MS, Mishra D, Mishra A, Rao YS (2022) Efficient design of an authenticated key agreement protocol for dew-assisted IoT systems. *J Supercomput* 78:3696–3714. <https://doi.org/10.1007/s11227-021-04003-z>
20. Vinoth R, Deborah LJ (2021) An efficient key agreement and authentication protocol for secure communication in industrial IoT applications. *J Ambient Intell Human Comput* 1–13. <https://doi.org/10.1007/s12652-021-03167-z>
21. Shukla V, Chaturvedi A, Srivastava N (2019) Nanotechnology and cryptographic protocols: issues and possible solutions. *Nanomater Energy* 8(1):1–6. <https://doi.org/10.1680/jnaen.18.00006>
22. Chaturvedi A, Srivastava N, Shukla V (2015) A secure wireless communication protocol using Diffie-Hellman key exchange. *Int J Comput Appl* 126(5):35–38. <https://doi.org/10.5120/ijca2015906060>
23. Shukla V, Mishra A, Agarwal S (2020) A new one time password generation method for financial transactions with randomness analysis. *Innovations in electrical and electronic engineering (Part of the lecture notes in electrical engineering book series (LNEE, volume 661))*, 713–720. https://doi.org/10.1007/978-981-15-4692-1_54
24. Shukla V, Mishra A, Yadav A (2019) An authenticated and secure electronic health record system. In: *IEEE international conference on information and communication technology*, pp 1–5. <https://doi.org/10.1109/CICT48419.2019.9066168>
25. Shukla V, Chaturvedi A, Srivastava N (2017) Secure wireless communication protocol: to avoid vulnerabilities in shared authentication. *Commun Appl Electron* 7(6):4–7. <https://doi.org/10.5120/cae2017652680>
26. Shukla V, Misra MK, Chaturvedi A (2022) Journey of cryptocurrency in India in view of financial budget 2022–23. *Cornell University arxiv*, 1–6. <https://doi.org/10.48550/arXiv.2203.12606>
27. Chaturvedi A, Shukla V (2020) Miracle of number theory. *Everyman's Science* 50(3–4):131–134. http://www.sciencecongress.nic.in/pdf/e-book/august_nov_2020.pdf
28. Shukla V, Chaturvedi A (2018) Cryptocurrency: characteristics and future perspectives 53(2):77–80. <http://164.100.161.164/pdf/e-book/june-july-18.pdf#page=14>
29. Shukla V, Chaturvedi A, Srivastava N (2015) A new secure authenticated key agreement scheme for wireless (mobile) communication in an EHR system using cryptography. *Commun Appl Electron* 3(3):16–21. <https://doi.org/10.5120/cae2015651903>
30. Chaturvedi A, Srivastava N, Shukla V, Tripathi SP, Misra MK (2015) A secure zero knowledge authentication protocol for wireless (mobile) ad-hoc networks. *Int J Comput Appl* 128(2):36–39. <https://doi.org/10.5120/ijca2015906437>
31. Shukla V, Chaturvedi A, Srivastava N (2019) Authentication aspects of dynamic routing protocols: associated problem & proposed solution. *Int J Recent Technol Eng* 8(2):412–419. <https://doi.org/10.35940/ijrte.B1503.078219>
32. Shukla V, Kushwaha A, Parihar SS, Srivastava S, Singh VP (2016) Authenticated wireless information display system using GSM module. *Commun Appl Electron* 5(3):7–11. <https://doi.org/10.5120/cae2016652251>
33. Shukla V, Chaturvedi A, Srivastava N (2017) Double layer cryptographic protocol for mobile ad-hoc networks (MANETs) by commitment scheme. *Commun Appl Electron* 7(9):32–36. <https://doi.org/10.5120/cae2017652716>
34. Shukla V, Dixit S, Dixit P (2022) An IoT based user authenticated soil monitoring system. *Adhoc Sensor Wirel Netw* 53(3–4):269–283. <https://doi.org/10.32908/ahsw.n.v53.9453>
35. Chaturvedi A, Shukla V, Srivastava N (2017) A secure wireless peer to peer authentication protocol using triple decomposition problem. *Asian J Math Comput Res* 22(2):63–69. <https://archives.biciconference.co.in/index.php/AJOMCOR/article/view/1167>
36. Shukla V, Mishra A (2020) A new sequential coding method for secure data communication. In: *IEEE international conference on computing, power and communication technologies*, pp 529–533. <https://doi.org/10.1109/GUCON48875.2020.9231252>

37. Chaturvedi A, Shukla V, Misra MK (2021) A random encoding method for secure data communication: an extension of sequential coding. *J Discrete Math Sci Cryptogr* 24(5):1189–1204. <https://doi.org/10.1080/09720529.2021.1932902>
38. Shukla V, Chaturvedi A, Srivastava N (2019) A new one time password mechanism for client-server applications. *J Discrete Math Sci Cryptogr* 22:1393–1406. <https://doi.org/10.1080/09720529.2019.1692447>

Role of Virtualization in Secure Network



Anju Shukla , Shishir Kumar , and Varun Shukla 

1 Introduction

Virtualization is the mechanism to run multiple operating systems simultaneously on the same hardware by creating virtual machines in secure cloud network [1, 2]. The virtual machine monitor (VMM) or hypervisor, which sits in between one or more running operating systems and the hardware and provides each one the appearance that it controls the computer, is the essential component of virtual machines. On the other hand, the VMM is in control of the hardware and is required to run multiple operating systems over the physical resources of the machine. Transparency is thus one of the most crucial objectives in whole scenario.

Virtual machines have become increasingly popular in cloud paradigm for a variety of reasons. One of these reasons is server consolidation. It also serves as a security feature, as well as making configuration easier. Full virtualization, para-virtualization, and hardware-assisted virtualization are three different types of virtualization techniques. Full virtualization exactly provides the replica of underlying hardware. The virtualized guest OS is totally ignorant that it is being used.

Direct execution and binary translation are both used in full virtualization. Non-sensitive CPU instructions can now be executed directly, whereas sensitive CPU instructions must first be translated. The hypervisor doesn't simulate the underlying hardware in para-virtualization. Hypercalls are instead made available. Hypercalls are used by the guest OS to carry out delicate CPU commands. However, because the guest OS is aware that it is being virtualized, it performs better. System calls in

A. Shukla (✉)
VIT Bhopal University, Bhopal 466114, India
e-mail: anjushukla.iitb@gmail.com

S. Kumar
Babasaheb Bhimrao Ambedkar University Lucknow, Lucknow 226025, India

V. Shukla
Pranveer Singh Institute of Technology, Kanpur, Uttar Pradesh, India

the kernel are comparable to hypercalls. They provide communication between the guest OS and the hypervisor. AMDV and Intel VT technologies enable hardware virtualization, which integrates virtualization into the x86 CPU architecture itself. Virtualization hardware assistance has made it unnecessary to para-virtualized guest operating systems.

2 Related Works

There have been numerous studies comparing and evaluating the performance of various VMM using various benchmarks [3–7]. Chierici et al. [8] used kernel compilation tests and I/O heavy tests to compare Xen and KVM in terms of overall performance, performance isolation, scalability, and fault tolerant [9, 10].

3 Hypervisor or Virtual Machine Monitor (VMM)

VMM functions as an operating system at the low level for operating systems. In order to give each operating system the idea that it is in charge of the computer, by virtually sharing its resources, a hypervisor enables a single host computer to support numerous guest VMs. Description of VirtualBox and VMware's virtual machine monitors is presented here.

VMware

VMware creates the VMware desktop virtualization software platform. We can run different operating systems on a single server thanks to hypervisor software that is installed on physical servers. It provides virtualization, software storage, and networking tool products.

It operates an operating system that is installed in a virtual environment using the resources of the host. It uses resources in the designated way to perform properly. All servers use the same resources that the actual server does.

X86 assumes that entire hardware is under their ownership because it is bare metal architecture. Bare metal means it is more advanced and secure than other hosted architecture. It has four privileged levels: Ring 0, Ring 1, Ring 2, and Ring 3. Ring 3 is basically used for user requests since it requires direct access to the memory. Ring 0 has most highly privileged than other rings. For any action or user request, ring 0 will firstly executed. Ring 3 must have the request permission to ring 0 to execute those requests. Virtual layer in x86 placed under operating system with expectation of being in ring 0, so that it will get more privileged. Instructions having distinct semantics are not executed if not in ring 0 or if they are in ring 3, they cannot be virtualized. It has to be trapped and translated for privileged instruction requests. The challenges of capturing and translating these private and privileged instruction requests initially made it seem impossible to virtualize the x86 architecture. VMware

uses binary translation mechanism to resolve the issue. In this, operating system executes in user level greater than ring 3. But VMM has the highest priority than other rings. VMware allows the VMM to run in Ring 0 for isolation and performance.

VirtualBox

Oracle offers VirtualBox, a program that allows you to set up virtual machines on your computer. It was first presented by Innotek GmbH in 2007 and afterwards developed by Oracle. A software virtualization package with the ability to load various operating systems is what it is also known as. There are several different operating systems available, including Windows XP, Linux, Ubuntu, and macOS. It is a very complete, strong, and high-performance product that is available for both home and business use. In the form of Open Source Software, it is employed by professionals. Virtualization is made possible by open-source software that can operate on practically all host operating systems.

VirtualBox is a hypervisor that is “hosted.” VirtualBox is functionally identical on all host platforms to a considerable extent, and the same file and image formats are used. On the host hardware, the guest OS code that was previously operating in ring 0 is re-configured to run in ring 1. This switches out the instruction for a hypervisor memory hop to a piece of compiled code that is VM-safe. Nested paging is a relatively new technology that manages memory in hardware. Because the virtualization software no longer needs to handle these processes, this can significantly speed up hardware virtualization.

4 Experiments and Performance Study

Here, two benchmarks, LMBench and IOzone, are used to perform the simulation. LMBench is an open micro-benchmark used to evaluate the performance of operating systems and hardware. IOzone is a free file system benchmark that has a wide range of characteristics. Its ANSI C-based source code is 64-bit compatible. The virtual machines with same configuration details are created on both hypervisors (VMware and VirtualBox):-

- Memory: 1 GB
- Number of CPUs: 1
- Number of threads: 1
- CPU frequency: 1.344 GHz
- Number of cores: 1

There were four virtual machines created. Each with this arrangement, two on VMware and two on VirtualBox. One virtual machine on VMware runs Windows XP Dark Edition as a guest OS, while the other runs Ubuntu 10.10 as a guest OS. One of the two virtual machines on VirtualBox runs Windows XP Dark Edition as the guest OS, while the other runs Ubuntu 10.10. Ubuntu 10.10 was used as the host operating system for each virtual machine.

In order to research memory virtualization, we benchmarked virtualized versions of Ubuntu in VirtualBox and VMware using Lmbench. First, the two virtual machines' read and write bandwidths in memory were examined. Memory write is the amount of time needed to save data in memory, whereas memory read is the amount of time needed to read data into the processor. Figure 1 shows the outcomes. The size specifications "k" and "m" in the following figures denote kilobytes and megabytes, respectively. Figure 1 demonstrates that VMware's memory read is higher than VirtualBox's. Except for a few tiny block sizes, write operations in VMware outperform those in VirtualBox. Table 1 shows comparisons of other system operations. Interprocess communication over TCP/IP is assessed by transmitting a token back and forth between two processes (lat tcp), as well as context switching. In comparison to VirtualBox, VMware is performed better. We assessed interprocess connection latencies, which is the time it takes to create and connect an AF INET socket to a remote server, and found that VirtualBox outperformed VMware in the following tests. The time it takes to open and close a file, also known as syscall. We used two processes interacting through a UNIX pipe and passing a token back and forth to measure interprocess latency.

Forking a process into two identical copies with a single exit is known as forking, and executing a process is known as creating a new process and having that process run a new program. Context Switch times were then computed utilizing the size of the process and the quantity of processes as inputs. The results are summarized in Table 2.

The results of file system create/delete performance are shown in Tables 3 and 4, and we can infer that VMware outperforms VirtualBox in terms of creations and deletions per second for file sizes of 0 k, 1 k, 4 k, and 10 k. The time it takes for data to pass across pipes is measured by connecting two processes with a UNIX pipe.

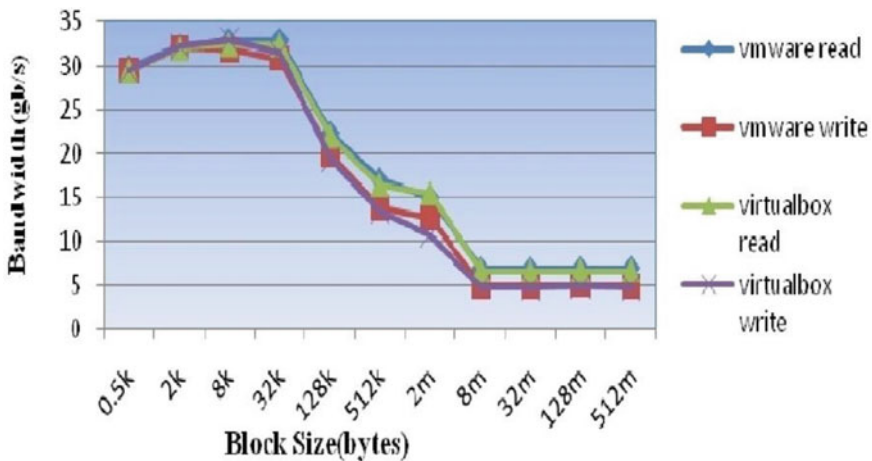


Fig. 1 Memory bandwidth for read and write in Lmbench

Table 1 System operation times

	VMware(μ s)	VirtualBox(μ s)
lat_tcp	1.7625	0.6127
connect	25.78575	91.436
syscall	3.45395	2.217
lat_pipe	24.59705	177.7406
fork	497.4091	2948.6757
exec	524.0909	3225.7659

Table 2 Context switch times

	VMware(μ s)	VirtualBox(μ s)
2 p 0 k	9.46	66.86
2 p 16 k	11.61	75.52
2 p 64 k	11.76	75.05
8 p 16 k	12.12	83.93
8 p 64 k	11.98	71.90
16 p 16 k	12.25	79.36
16 p 64 k	12.65	76.27

The message is divided into 64 KB parts and sent in that order. Table 5 shows the results, which show that VMware outperforms VirtualBox.

Except for tiny file sizes, VMware outperforms VirtualBox, as seen in Fig. 2. Because VMware’s reread and rewrite speed is better, this suggests that the buffer utilization is efficient (Fig. 3).

Table 3 VMware file system

Size of the file	Number of files created	Creations per second	Removal per second
0 k	339	62,639	106,269
1 k	194	34,282	82,499
4 k	130	22,891	82,569
10 k	71	12,147	66,508

Table 4 VirtualBox file system

Size of the file	Number of files created	Creations per second	Removal per second
0 k	82,148	58,896	103,235
1 k	54,659	13,454	48,254
4 k	45,129	19,493	50,692
10 k	31,206	12,497	46,094

Table 5 Data movement through pipes

Total bytes	VMware (MB/s)	VirtualBox (MB/s)
30 M	1528.41	152.16
100 M	1060.59	149.59

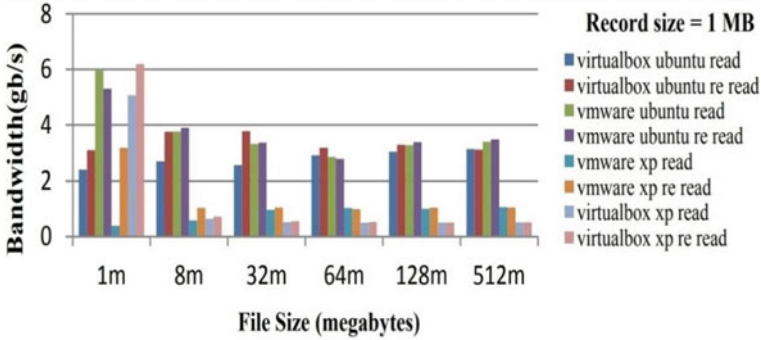


Fig. 2 Read bandwidth comparison in IOzone

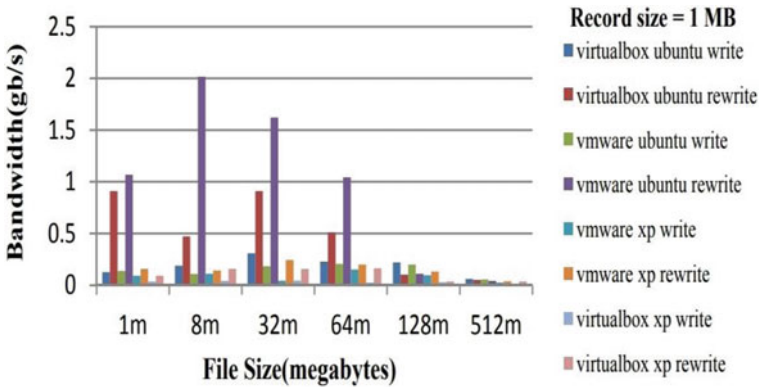


Fig. 3 Write bandwidth comparison in IOzone

5 Conclusion and Future Works

Virtualization is the mechanism to create the virtual environments that can be shown in testing, platform deployment, or access to different operating systems. There are several options for building virtual machines on a system, and each has advantages and disadvantages. Here, various virtualization techniques are presented and also analyzed the two hypervisors: VMware and VirtualBox. Two benchmarks, Lmbench and IOzone, are used to evaluate the performance of virtualized Ubuntu and Windows XP on VMware and VirtualBox. Overhead of CPU, memory, and I/O

operations are considered and outcomes are summarized. In future, other hypervisors and benchmarks can be used to test the network performance.

References

1. Shukla A, Kumar S, Singh H (2021) MLP-ANN-Based execution time prediction model and assessment of input parameters through structural modeling. *Proc Natl Acad Sci, India, Sect A* 91(3):577–585
2. Shukla A, Kumar S, Singh H (2019) An improved resource allocation model for grid computing environment. *Int J Intell Eng Syst* 12(1):104–113
3. Apparao P, Iyer R, Zhang X, Newell D, Adelmeyer T (2008) Characterization & analysis of a server consolidation benchmark. In: *Proceedings of the fourth ACM SIGPLAN/SIGOPS international conference on Virtual execution environments*, March, pp 21–30
4. Menon A, Santos JR, Turner Y, Janakiraman G, Zwaenepoel W (2005) Diagnosing performance overheads in the xen virtual machine environment. In: *Proceedings of the 1st ACM/USENIX international conference on Virtual execution environments*, June, pp 13–23
5. Barham P, Dragovic B, Fraser K, Hand S, Harris T, Ho A, Neugebauer R, Prat I, Warfield A (2003) Xen and the art of virtualization. *ACM SIGOPS Oper Syst Rev* 37(5):164–177
6. Vojnak DT, Đorđević BS, Timčenko VV, Štrbac SM (2019) Performance comparison of the type-2 hypervisor VirtualBox and VMWare workstation. In: *2019 27th telecommunications forum (TELFOR)*, November. IEEE, pp 1–4
7. Che J, He Q, Gao Q, Huang D (2008) Performance measuring and comparing of virtual machine monitors. In: *2008 IEEE/IFIP international conference on embedded and ubiquitous computing*, December, vol 2. IEEE, pp 381–386
8. Chierici A, Veraldi R (2010) A quantitative comparison between Xen and KVM. *J Phys: Conf Series* 219(4):042005. IOP Publishing
9. Shukla A, Kumar S, Singh H (2019) Fault tolerance based load balancing approach for web resources. *J Chin Inst Eng* 42(7):583–592
10. Shukla A, Kumar S, Singh H (2020) Fault tolerance based load balancing approach for web resources in cloud environment. *Int Arab J Inf Technol* 17(2):225–232

Blockchain-Based NFT for Evidence System



Aditya Kumar Sharma and Brijesh Kumar Chaurasia

1 Introduction

A Blockchain is a decentralized, immutable, and transparent distributed ledger [1]. It is essentially an append-only data structure maintained by assets of nodes that do not fully trust each other [2]. Blockchain is the technology that underpins Bitcoin, the cryptocurrency. However, Bitcoin is just one application of the Blockchain. A Blockchain is essentially a “chain of blocks,” where each block represents a set of records. Each such record could, in turn, represent a cryptocurrency, a land plot, insurance, gaming, or even an identity and an evidence system [3]. This technology has the potential to revolutionize the way value is stored and exchanged with immutability and fast access. Even though there are doubts about the feasibility of cryptocurrencies, the unanimous opinion is that the basic technology of Blockchain will revolutionize the way value is stored and transferred. Blockchain may be categorized into public Blockchain, private Blockchain, and hybrid Blockchain. In this work, we used a permissionless Blockchain. Ethereum [4] is a permissionless Blockchain and permits extending its practicality with the assistance of smart contracts. The proposed work is the application of NFT using Ethereum. Ethereum powers its own cryptocurrency ETH. The cryptocurrency ether is also known as fungible currency. Because of its unique characteristics such as unique identity, non-divisibility, and non-mergeability, the work focuses on NFT in the Ethereum Blockchain [5].

Organization of the paper: The remainder of the paper is organized as follows: Sect. 2 presents the state of art of NFT, Sect. 3 discusses about NFT architecture with its standard and properties, and issues and challenges with NFT Usage are

A. K. Sharma (✉) · B. K. Chaurasia
Department of Computer Science and Engineering, Pranveer Singh Institute of Technology,
Kanpur, India
e-mail: ad.sharma171@gmail.com

B. K. Chaurasia
e-mail: brijesh.chaurasia@psit.ac.in

Table 1 Applications of NFTs in various domains

Domain	Applications
Art and Music [10]	Music files and paintings are stored on blockchain that keeps them safe from piracy and ensure proper compensation to the artists
Gaming [10]	NFTs allow games to have in-game items, land and collectibles. Also allows them to trade them on secondary markets
Ownerships [10]	People performing property deeds with NFT
Insurance [11]	Insurance documents in the form of NFT can keep track of all the health records

discussed in Sect. 4. The proposed NFT structure under the Ethereum Blockchain along with the working process of the proposed system is presented by Sect. 5. Results and discussion are depicted in Sect. 6. Finally, conclusion and future work are presented in Sect. 7.

2 State of Art

Evidence is one of the most controversial and litigious subjects in the Indian judiciary system. Individual criminal records are stored on Blockchain to store the data transactions in terms of logs alongside encrypting the data. As a result, the records cannot be altered in order to assist the court [6]. A Blockchain-enabled law enforcement chain of evidence is presented in [7]. This evidence system is considered a type of supply chain for evidence collection. In this system, a police or evidence agency is considered a distributed node. This distributed node plays the role of miner and is responsible for mining the evidence along with performing the authentication process and distributing the IDs and passwords. According to the report [8], the use of Blockchain to issue warrants and the impact of Blockchain on the justice system are discussed (Table 1).

3 NFT Architecture

A NFT is a digital asset that can be stored on a Blockchain and tracked using a unique hash value to determine its identity [9]. Generally, NFTs are tokens that we can use to represent ownership of unique items. Each NFT has some distinguishing characteristics, such as a distinct identity, the inability to be divided, and the inability to merge. NFT may be designed for digital art, gaming, real state, and metaverses. However, to the best of my knowledge in the evidence system, we are proposing for the first time using Ethereum. Smart contracts in Ethereum allow for the attachment of unique identifiers to each NFT—no one can modify the record of ownership or

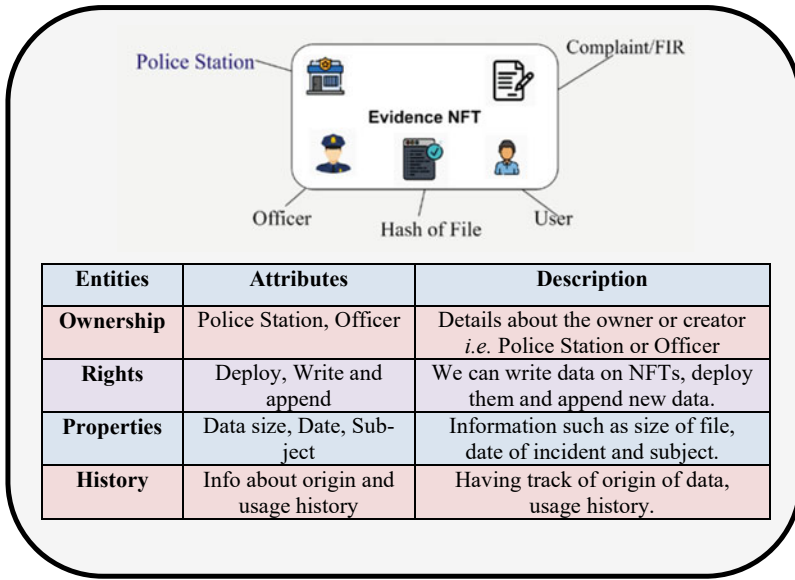


Fig. 1 Proposed architecture of NFT

copy/paste a new NFT into existence. They have real-world market value and can be traded through Blockchain as well. The ownership of these is changed and registered in their smart contracts. NFTs can be wholly or fractionally held depending upon the token value being traded [10]. The architecture of NFT is illustrated in Fig. 1.

Every NFT has at least four entities, which are as follows: **properties** to define uniqueness; **rights** to define utility; **ownership** for record; and **historical data** of change made. In Fig. 1, proposed NFT, historical data are assumed as the previous record available at any police station or logged FIR, etc. Similarly, ownership is considered as a police station, police officer, and police staff, etc., and rights are assumed as permission to create NFT, store NFT, append NFT, and view NFT, etc., and properties may be considered as the size of the data, date, subject, etc.

Moreover, NFTs are simply smart contracts or programs written in Solidity, the programming language for Blockchain. The attributes of our token are actually the attributes defined in the objects created by the program as presented in Fig. 3. Values set for these variables allow an NFT to take several forms, as illustrated by Fig. 2.

3.1 Standards for NFTs

In this sub-section, standard protocols for NFTs are presented. These standard protocols are used for issuing tokens over the Ethereum Blockchain [12].

A. ERC-20

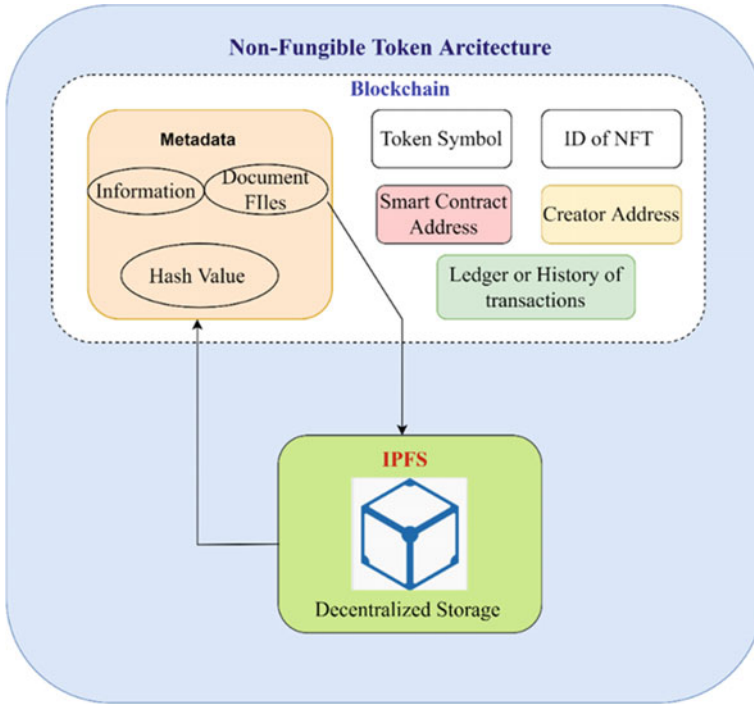


Fig. 2 Key components of an NFT

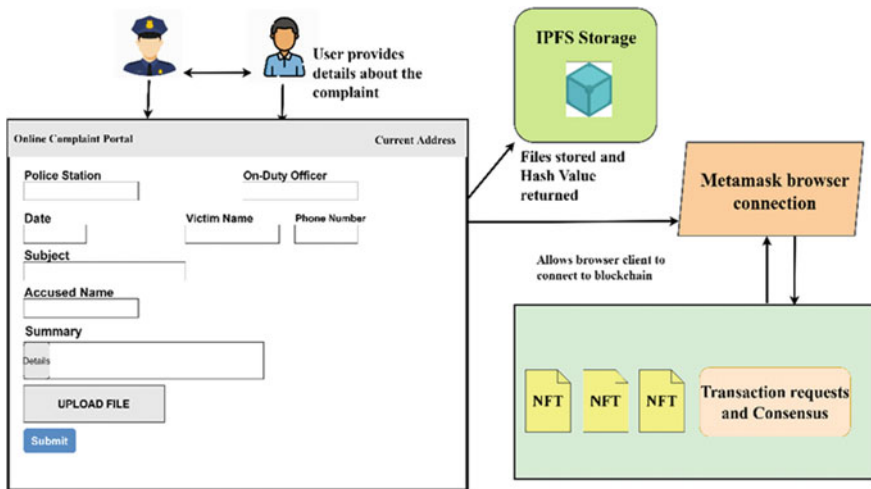


Fig. 3 Workflow of evidence system

Since 2015, Ethereum has been the most important standard. It has emerged as the technological standard for token implementation in blockchain development. It lays out simple guidelines for all Ethereum tokens to adhere to. *ERC-217* came with certain bug fixes in *ERC-20*. A more updated version of this comes out in *ERC-777*, which enables more features that save users from sending tokens to the wrong address [12].

B. *ERC-721 and ERC-1155*

These standards are for NFT specifications [13]. These standards are used by developers all over the world to create smart contracts. NFTs are further used in a variety of places, including music, art, fashion, and IoT.

C. *ERC-725*

It is a digital identity standard. Different smart contracts with multiple keys can be managed with *ERC-725* [12], which defines the proxy smart contract. Machines, objects, and individuals can be easily demarcated through identity smart contracts.

D. *ERC-223*

This was established by European Research. It proposes that this standard will notify users who sent their tokens to the wrong addresses and then let them even cancel these transactions. Although not yet widely accepted, *ERC-223* provides a better UX for the layman user.

E. *ERC-1400 and ERC-1404*

These are security token specifications [12]. These allow you to create tokens with limited features to enhance security.

Due to the novelty of the NFT ecosystem, the breadth of use cases is centered on applications that represent items of scarcity. The idea of using NFTs for this purpose is strengthened by the advantages and features offered by them.

3.2 *Properties of NFTs*

In this sub-section, properties of NFTs are presented as follows:

A. *Managing and protecting critical data*

Somebody who uses public services is rightly concerned that, despite agencies' best efforts to protect their systems, criminals might figure out a way to steal or manipulate their records. With NFTs, the hash values of the records or documents can be used to identify records but not be used to reconstruct them in the file itself. The hash values stored on the Blockchain are distributed over the private network. Whenever a change is made to the file, a new hash value is appended to the chain [14].

B. *Digital property ownership*

The method of owning and transferring assets—whether physical property or financial tools—usually involves multiple interactions and a long trail of paper work. All this is simply done to keep a general track of things that take place during the process. The same can be achieved with even more accuracy and authenticity with NFTs due to their inherent nature of maintaining records of each transaction [15].

C. *Building networked public services*

Governments could create central repositories or enterprise systems for sharing information across agencies with small efforts. A crucial point is security. Authorities can never, in any circumstance, allow indiscriminate access to these records or data. A system in which data can easily be shared across environments but in which individuals and organizations can take back ownership of their data and control the flow of personal information—who sees it, what they see and when.

Blockchain technology may upkeep such a scenario. Information is stored in a dedicated ledger within an encrypted database and each user can access it through the internet. End users could have restricted access using a private cryptographic key [14].

4 Issues and Challenges with NFT Usage

As the usage of NFTs increases, various new aspects of their characteristics will be discovered. Both their benefits and limitations are getting highlighted as we move ahead with them. Some of the challenges currently being faced with the use of NFTs are mentioned further [16].

4.1 *Validation of Ownership for NFTs*

It is significantly important to identify the correct ownership of a certain NFT. It becomes hard to know whether the owner genuinely possesses an NFT before making any transactions, because as of now, when an NFT is purchased, the owner gets the rights to utilize it but not the rights to intellectual property. The metadata may contain information regarding the owner. They face different kinds of problems. The hype created by celebrities who endorse cryptocurrencies affects their valuation. The trading volume of NFTs has surged more than once in just one year [16].

4.2 Privacy Issues with NFTs

For a technological system used by an organization, the concern for security is critical. The data integrity of so many users relies on these security systems. While the privacy offered by NFTs is still being studied, the Ethereum platform used by them provides a level of partial anonymous behavior over the network. The identities of users can be hidden only to an extent.

4.3 A Contributor to Climate Change

The working of a Blockchain is such that to perform any kind of operation, we need to make transactions over the Blockchain. These transactions are in the form of computing power being transferred from one node to another. Computing power directly impacts the electricity consumption of a computer processor. Thus, as the number of concurrent transactions increases, the resources consumed by the nodes also increase. Ethereum is estimated to consume 44.94 terawatt-hours of electricity per year [17].

4.4 Smart Contracts are at Risk

With no industry-wide validation in the case of Solidity, smart contracts become vulnerable to attacks inherently. Although the code and smart contracts themselves are immutable, this appears to be the cause of another problem as well. The developers need to ensure that the code they are deploying is error-free and robust. There have been recent attacks on the Decentralized Finance (DeFi) protocols [22] such as Poly Network and Meerkat Finance [17], which have resulted in huge losses of user data and money. In all, it cost Binance 31 million dollar in the attack on their chain-based lending protocol [17].

5 Proposed NFT Architecture for Evidence System

In our proposed system, the information provided while writing the complaint or an FIR is used to create the NFT. The information may be filled in online or by any police station. On behalf of FIR, a warrant will be issued by the lower court or lawful competent authority. The system must also store any files or documents included with the data. But saving files directly over the Blockchain is an expensive task and not feasible for a stream of data. Thus, we have used IPFS to store the files separately while including the generated hash of the files in our NFT data. It

not only allows NFTs to be lightweight but also allows for quick access if we only need the basic information from the warrant. Our proposed scheme provides non-integrity to evidence, which is a must for any evidence system. IPFS, being the host for decentralized data, helps to secure files from being altered. The working flow of the same is shown in Fig. 3.

The process of filing a complaint and how the system handles this request is outlined in Fig. 3. While designing the system, we kept in mind that we wanted to keep things simple for the users and make it easier for them to adapt with minimal adjustments. The client-side interface is a basic form that requires the user to input details as required, with the option to upload related documents if any.

Upon submitting the details, the information is saved. Documents related to the complaint, such as a copy of the physical FIR, complaint, and evidence photos, are stored directly on IPFS storage. A unique address (hash) of the file is returned by IPFS. This hash value can be used to access the file later. This transfer takes place over a Blockchain and conventional browsers such as Chrome, Firefox, and Edge do not support direct connections to decentralized networks. Hence, a web host such as Metamask is used that allows this kind of connection.

The hash value and details are passed to the server, which invokes the smart contracts to create an NFT. This interaction takes place through Metamask, and upon confirmation of the transaction cost, the NFT gets deployed over the Blockchain.

6 Results and Discussion

In this section, the results and discussion are explained. We have considered the NFT data include the details of the person who lodged a FIR with the details of the accused. After filing for the first time and the transaction being done at IPFS, the warrant gets stored on the Blockchain. It can grow in size as more details are added, but to reduce the size of the transaction, we have stored the data on IPFS and are only using the hash value of the FIR transaction in Blockchain. The Smart Contract will keep track of any changes applied. The simulation ran on two systems with varying hardware. The first being the HP G5 Workstation with a 2.70 GHz CPU with 20 cores, considered a high-performance node, and 32 GB of RAM, and the second being the HP Notebook with a 1.80 GHz CPU with 4 cores and 8 GB of RAM, which is assumed to be a low computation node. On both machines, the same software dependencies were used: Ethereum Blockchain [18] using Ganache [19], Solidity v0.5.0 [20], and Node.js [21] for client development.

The following results demonstrate the efficacy of the proposed NFT on both systems. The results provide us with some key findings. First, in parallel to the nature of blockchain, with an increase in the size of NFT data, the transaction fee increases. Secondly, comparing both the machines, it is also observed that the system with higher computation power used less gas (gwei) while still having a higher gas limit. The data show that for small data documents such as FIRs, data flow is smooth. Therefore, NFTs have a viable candidature in the evidence system, bringing both

the aspects of being indivisible and non-mergeable into one. Figure 4 shows the transaction fees of NFT in terms of ETH. The result shows that size is not increasing continuously as per the data increases. However, NFT fees are increasing with the use of computation power. Similarly, Fig. 5 depicts the consumption of transaction fees in terms of gas used and gas limit. Transaction fees are increasing exponentially at the rate of increasing transaction size, except in the case of the 800-byte size of NFT.

Figure 6 shows the comparison between the transaction fees of NFT with a high-performance node and a low-performance node. It is observed that NFT at the size of 2000 byte transaction fees consumed by NFT in terms of ETH at the high-performance node is around just half. We have also analyzed similar analysis in the case of transaction fees in terms of gas limit and gas used, which is around 1.6 times in both cases. As for the outcome of the proposed work, we can say that NFT transaction fees depend on computation power along with size.

Fig. 4 Transaction cost with different sizes of NFT

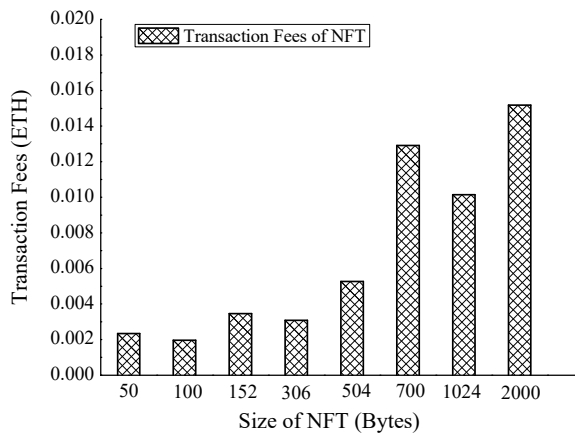


Fig. 5 Variation in gas fee used and gas limit set for different sizes of NFT

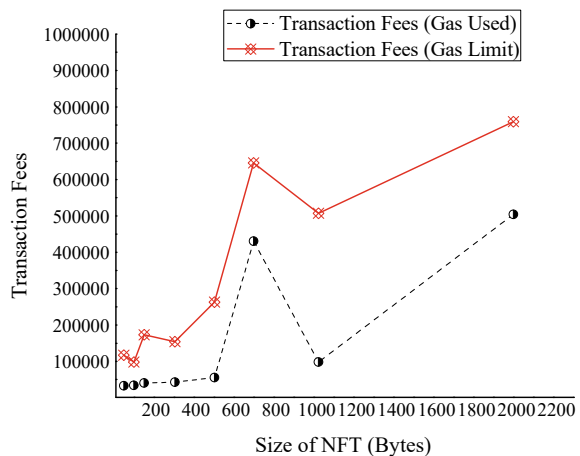
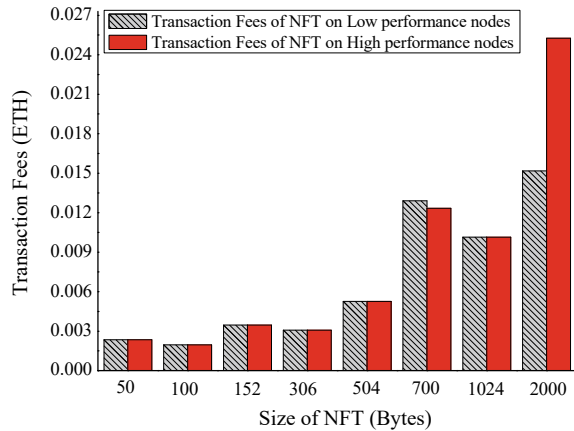


Fig. 6 Comparative analysis of transaction cost consumed by NFT between two systems



7 Conclusion

The digital existence of documents and records constantly faces the threat of being altered unethically. A similar problem affects the evidence records, which are crucial for the judiciary. The technical aspects of using NFTs for these records and the findings observed in the data indicate that NFTs can become a feasible solution. Using appropriate computation power would result in a balanced transaction cost, thus making this use-case of NFTs a viable option for users to file their complaints with the robust security of blockchain. The results show the efficacy of our proposed system. The proposed NFT using public Blockchain is able to achieve lightweight in terms of storage data, fast access using IPFS, and integrity of evidence.

References

1. Hasan HR, Salah K, Battah A, Madine M, Yaqoob I, Jayaraman R, Omar M (2022) Incorporating registration, reputation, and incentivization Into the NFT ecosystem. *IEEE Access* 2:76416–76433
2. Anh DTT, Zhang M, Ooi BC, Chen G (2018) Untangling blockchain: a data processing view of blockchain systems. *IEEE Trans Knowl Data Eng* 30(7):1366–1385
3. Srivastava S, Chaurasia BK, Singh D (2023) Blockchain-based IoT security solutions. In: *Distributed computing to blockchain: architecture, technology, and applications*, Ch 18, section 2. Elsevier
4. www.Ehtereun.org. Last Accessed 10 July 2022
5. Entriken W, Shirley D (2018) Eip 721: Erc-721 non-fungible token standard, Jan. <https://eips.ethereum.org/EIPS/eip-721>. Last Accessed 11 July 2022
6. Tasnim MA, Omar A, Al, Rahman MS, Bhuiya Md ZA (2018) CRAB: Blockchain based criminal record management system, *SpaCCS 2018*, LNCS 11342, 294–303
7. Use of Blockchain system in evidence management system. <https://f.hubspotusercontent10.net/hubfs/5260862/Ebooks%20and%20Whitepapers/Blockchain%20of%20Evidence%20FINAL%20DRAFT-3.pdf>. Last Accessed 12 July 2022

8. When might blockchain appear in your court. [blockchaininthecourts.pdf](#). Last Accessed 12 July 2022
9. Casale-Brunet S, Ribeca P, Doyle P, Mattavelli M (2022) Networks of Ethereum non-fungible tokens: a graph-based analysis of the ERC-721 ecosystem. In: IEEE international conference on blockchain (Blockchain), pp 188–195
10. NFT use cases & applications examples—Data analytics. <https://vitalflux.com/nft-use-cases-applications-examples>. Last Accessed 16 July 2022
11. The impact of metaverse on insurance industry. www.pwc.com/jp/en/knowledge/column/metaverse-impact-on-the-insurance-industry.html. Last Accessed 25 July 2022
12. Quick guide on Ethereum ERC standards. <https://blockchain.oodles.io/blog/quick-guide-ethereum-erc-token-standards/>. Last Accessed 30 July 2022
13. ERC-Standards. <https://101blockchains.com/erc-standards/>. Last Accessed 30 July 2022
14. Louw L. Blockchain NFTs—properties, value propositions & applications. <https://bitcoinsv.academy/blog/blockchain-nfts-properties-value-proposition-application>. Last Accessed 07 Aug 2022
15. Insurance in the Metaverse (Part-2)—Its more than a playground. www.vertafore.com/resources/blog/insurance-metaverse. Last Accessed 08 Aug 2022.
16. A new world of assets—the top challenges and risks of NFTs. www.bitcrunch.com/blogs/nft-risks-and-challenges. Last Accessed 08 Aug 2022
17. Biggest DeFi Hack and Heists. <https://decrypt.co/93874/biggest-defi-hacks-heists>. Last Accessed 08 Aug 2022
18. Ethereum 2.0. www.ethereum.org. Last Accessed 08 Aug 2022
19. Ganache by Truffle Framework. <https://trufflesuite.com/ganache/>. Last Accessed 08 July 2022
20. Solidity, programming language. <https://soliditylang.org>. Last Accessed 08 Aug 2022
21. Node.js. <https://nodejs.org>. Last Accessed 08 Aug 2022
22. Blockchain for Decentralized Finance (DeFi). <https://consensys.net/blockchain-use-cases/decentralized-finance/>. Last Accessed 08 Aug 2022

Securing Digital Audio Files Using Rotation and XOR Operations



Abdul Gaffar 

1 Introduction

Every day millions (perhaps billions) of messages in the form of texts, audio, images, and videos, are communicated on the Internet, which is an open (unsecure) network. So, there must be robust technique(s) in order to communicate secretly. In the context of secure communication, encryption is the best choice, which encodes a secret message into a form which is unrecognizable, except by the intended one. Broadly, there are two types of encryption schemes: symmetric-key encryption and asymmetric-key encryption. The symmetric-key encryption, also known as (a.k.a.) *private-key* encryption, uses the same secret key for encoding and decoding a message. The foremost application of the private-key encryption is to provide *confidentiality*. On the other hand, asymmetric-key encryption, a.k.a. *public-key* encryption, uses different keys for encoding and decoding a message. In particular, public key is used for encoding, while private (secret) key is used for decoding a message. The foremost applications of the public-key encryption are authentication and non-repudiation, besides confidentiality.

Since the symmetric-key encryption methods are much faster and more efficient, for attaining confidentiality, as compared to the asymmetric-key encryption methods, therefore, we adopt the symmetric-key encryption method in the proposed technique. Note that the Rotation-XOR (RX) operations, used in the proposed technique, are the primitive operations, which are efficiently and directly supported by most of the computer processors. These operations aid in the possible improvement of speed of the designed technique.

The rest of the paper has been put in the following order: Sect. 2 provides related works; Sect. 3 gives preliminaries; Sect. 4 describes the encryption and decryption algorithms of the proposed technique; Sect. 5 describes the implementation and

A. Gaffar (✉)

Department of Mathematics and Statistics, Integral University, Lucknow 226 026, UP, India
e-mail: abdulgaffar.lu@gmail.com

experimental results; Sect. 6 discusses security analyses of the proposed technique; Sect. 7 gives comparison of the proposed technique with the recent state-of-the-art techniques; and Sect. 8 concludes the paper, followed by the references.

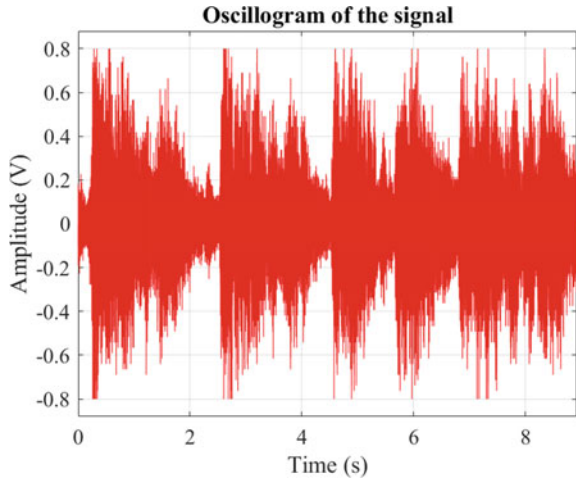
2 Related Works

Abouelkheir and El-Sherbiny [1] in 2022 proposed a technique for the security of digital audio files based on a modified RSA (Rivest, Shamir, and Adleman) algorithm. The authors modified the RSA algorithm via using dynamic keys—for enhancing security of the proposed technique, and five numbers (two primes and three random numbers)—for enhancing speed of the proposed technique. Several metrics have been utilized in order to validate the aims of the designed scheme. Although the scheme performs well in terms of encryption, but in terms of decryption, it is not a good scheme. It performs lossy decryption, i.e., the decrypted audio files are not exactly identical to the original audio files.

Shah et al. [2] in 2021 proposed a technique for the secure communication of digital audio files based on the finite fields. The authors generated a sequence of pseudo-random numbers via an elliptic curve, which is used to scramble the samples of the plain audio files. Further, the scrambled audio samples are substituted via the newly constructed S-boxes, to ensure the *confusion–diffusion* properties [3] required for a secure encryption algorithm. Faragallah and El-Sayed [4] in 2021 proposed an encryption scheme for securing the audio files based on the XOR (eXclusive OR) operation and the Hartley Transform (HT). First of all, a plain audio file is reshaped into a two-dimensional (2D) data block, and then it is XOR-ed with a grayscale image (treated as a secret key). The obtained XOR-ed blocks are then transposed via a chaotic map, followed by an optical encryption using HT. Naskar et al. [5] in 2021 suggested an encryption scheme for audio files based on the distinct key blocks together with the Piece-Wise Linear Chaotic Map (PWLCM) and the Elementary Cellular Automata (ECA). The scheme encrypts a plain audio file in three stages: cyclic shift, substitution, and scrambling. The cyclic shift is used for reducing the correlation between the samples of each audio block. The shifted audio data blocks are substituted (modified) via PWLCM, and finally, modified blocks are scrambled via ECA for better diffusion.

Abdelfatah [6] in 2020 proposed an algorithm for securing audio files in three phases utilizing three secret keys. The first phase is the self-adaptive scrambling of the plain audio files via the first secret key. The second phase is the dynamic DNA (deoxyribonucleic acid) encoding of the scrambled audio data via the second secret key. The last phase is the cipher feedback mode via the third secret key, which aids in achieving better confusion and diffusion properties.

Fig. 1 Oscillogram of the audio file “handel.wav”



3 Preliminaries

3.1 Digital Audio

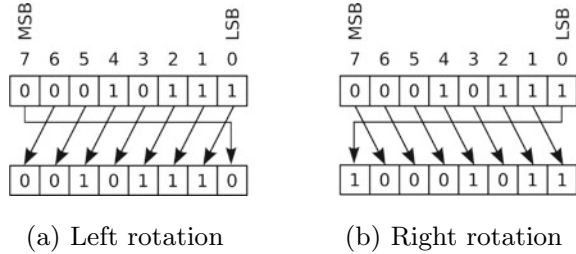
A digital audio, say, P is a l -by- c matrix, consisting of elements called *samples*, where l and c denote the number of samples and the number of channels in P , respectively. If $c = 1$, then P is said to be a *single* (or *mono*) channel audio file, and if $c = 2$, then P is said to be a *dual* (or *stereo*) channel audio file. Note that the samples in P are the floating-point values, i.e., real values. Figure 1 shows the oscillogram (a graph between amplitude and time) of the audio file “handel.wav”, which is of size 73113×1 , i.e., a single-channel audio file containing 73113 samples. For other details of the audio file “handel.wav”, namely, sample rate (in Hz—Hertz), duration (in sec—seconds), bits per sample, bit rate (in kbps—1000 bits per second), and size (in KB—1024 Bytes), see Table 1.

3.2 Rotation operation

By rotation operation, we mean “circular shift” or “bit-wise” rotation. It is of two types:

1. **Left rotation.** It is denoted by “ \lll ”. By $x \lll y$, it is meant that x is *left* rotated by y bits. For example, if $x = 0001\ 0111$ and $y = 1$, then $x \lll y$ gives $0010\ 1110$. Figure 2a demonstrates the concept, wherein MSB is the Most Significant Bit and LSB is the Least Significant Bit.

Fig. 2 **a** Left rotation of $x = 0001\ 0111$ by 1-bit and **b** right rotation of x by 1-bit



2. **Right rotation.** It is denoted by “ \gg ”. By $x \gg y$, it is meant that x is *right* rotated by y bits. For example, if $x = 0001\ 0111$ and $y = 1$, then $x \gg y$ gives $1000\ 1011$. Figure 2b demonstrates the concept.

3.3 XOR Operation

It is one of the simplest operations in a computer’s processor. It is a bit-wise operation that takes two strings of bits of equal length and performs the XOR (denoted by \oplus) operation as: if two bits are same, the result is 0; and if not same, the result is 1. It’s actually addition modulo 2.

For example, if $a = 1010\ 1011$ and $b = 0101\ 1100$, then $a \oplus b = 1111\ 0111$.

4 Description of the Proposed Encryption and Decryption Algorithms

4.1 Preprocessing on the Audio File

Input. An audio file P of size $l \times 1$.

1. Convert the audio samples of P from the floating point values (real values) to binary (matrix) via single-precision floating point (32-bit).¹
2. Convert the binary (matrix) to non-negative integers (bytes) array, i.e., P is of size $1 \times l$. Note that, here samples of P are in bytes ($0-2^8 - 1$).
3. Now, if l is a multiple of 4, then no *padding* is required, else pad $(4 - r)$ elements “post” with zeros to P , where r is a remainder on dividing l by 4.
4. Convert the bytes of P into WORDS, where WORD is a collection of 4 bytes, and rename the audio file P as P_w .

¹ See [7, 8].

Output. The audio file P_w of size $1 \times m$, where m denotes the number of WORDS in P_w .

4.2 Reverse Preprocessing on the Audio File

Input. The audio file P_w of size $1 \times m$, where m being the number of WORDS in P_w .

1. Convert the WORDS of the audio file P_w into bytes ($0-2^8 - 1$), and now, the size of P_w is $1 \times 4m$. Rename P_w as P .
2. Remove “last” zero (padded) bytes, if any, from P , and let the size of P becomes $1 \times l$ bytes.
3. Convert the bytes (non-negative integers— $0-2^8 - 1$) into a binary (matrix).
4. Convert the binary (matrix) into the floating-point values via the single-precision floating point (32-bit).
5. Take the transpose of P so that the size of P becomes $l \times 1$.

Output. The audio file P of size $l \times 1$.

4.3 Preprocessing on Secret Key

Input. Secret key $K = \{k_1, k_2, k_3, k_4, k_5, k_6, k_7, k_8, k_9, k_{10}, k_{11}, k_{12}, k_{13}, k_{14}, k_{15}, k_{16}, k_{17}, k_{18}, k_{19}, k_{20}, k_{21}, k_{22}, k_{23}, k_{24}, k_{25}, k_{26}, k_{27}, k_{28}, k_{29}, k_{30}, k_{31}, k_{32}\}$ of 32 bytes.

1. Split the secret key K into two equal parts, say, K_1 and K_2 as $K_1 = \{k_1, k_2, k_3, k_4, k_5, k_6, k_7, k_8, k_9, k_{10}, k_{11}, k_{12}, k_{13}, k_{14}, k_{15}, k_{16}\}$ and $K_2 = \{k_{17}, k_{18}, k_{19}, k_{20}, k_{21}, k_{22}, k_{23}, k_{24}, k_{25}, k_{26}, k_{27}, k_{28}, k_{29}, k_{30}, k_{31}, k_{32}\}$.
2. Convert the key-bytes of K_1 and K_2 into WORDS as $K_{1w} = \{q_{1w}, q_{2w}, q_{3w}, q_{4w}\}$, and $K_{2w} = \{r_{1w}, r_{2w}, r_{3w}, r_{4w}\}$, where $q_{1w} = k_1k_2k_3k_4$, $q_{2w} = k_5k_6k_7k_8$, $q_{3w} = k_9k_{10}k_{11}k_{12}$, and $q_{4w} = \{k_{13}k_{14}k_{15}k_{16}\}$; $r_{1w} = k_{17}k_{18}k_{19}k_{20}$, $r_{2w} = k_{21}k_{22}k_{23}k_{24}$, $r_{3w} = k_{25}k_{26}k_{27}k_{28}$, and $r_{4w} = \{k_{29}k_{30}k_{31}k_{32}\}$.
3. **Expansion of K_{1w} .**

► Expand K_{1w} to the size m as:

- (a) For $i = 1, 2, 3, 4$; $T_1[i] = K_{1w}[i]$, i.e., $T_1[1] = q_{1w}$, $T_1[2] = q_{2w}$, $T_1[3] = q_{3w}$, and $T_1[4] = q_{4w}$.
- (b) Calculate $T_1[5]$ as

$$T_1[5] = \text{mod}(\lceil \text{mean}(T_1[i]) \rceil, 2^{32}), \quad i = 1, 2, 3, 4.$$

where “mean” denotes the average function, $\lceil \cdot \rceil$ denotes the ceiling function, and “mod” denotes the modulus function.

(c) Calculate $T_1[i]$, for $i = 6, 7, \dots, m$, as

$$T_1[i] = \text{mod}(T_1[i - 1] + T_1[i - 2], 2^{32}), \quad i = 6, 7, \dots, m.$$

4. *Expansion of K_{2w} .*

► Expand K_{2w} to the size m as:

(a) For $i = 1, 2, 3, 4$; $T_2[i] = K_{2w}[i]$, i.e., $T_2[1] = r_{1w}$, $T_2[2] = r_{2w}$, $T_2[3] = r_{3w}$, and $T_2[4] = r_{4w}$.

(b) Calculate $T_2[5]$ as

$$T_2[5] = \text{mod}(\lceil \text{mean}(T_2[i]) \rceil, 2^{32}), \quad i = 1, 2, 3, 4.$$

where symbols have their usual meanings.

(c) Calculate $T_2[i]$, for $i = 6, 7, \dots, m$, as

$$T_2[i] = \text{mod}(T_2[i - 1] + T_2[i - 2], 2^{32}), \quad i = 6, 7, \dots, m.$$

5. *Generation of a third key.*

► Generate a third key K_{3w} from K_{1w} and K_{2w} as:

$$K_{3w} = \text{mod}(K_{1w} \cdot K_{2w}, 2^{32})$$

where “ \cdot ” denotes component-wise multiplication.

Output. The expanded keys T_1 and T_2 of size m , and the generated key K_{3w} of size 4.

4.4 *Encryption Algorithm*

Input. An audio file P of size $l \times 1$ and the secret key K of 32-byte.

1. Apply preprocessing on the audio file P (see Sect. 4.1), and let the obtained file be P_w of size $1 \times m$.
2. Apply preprocessing on the secret key K (see Sect. 4.3) to obtain the expanded keys T_1 & T_2 of size m , and the generated key K_{3w} of size 4 (in WORDS).
3. **Initial round substitution.** XOR P_w with T_1 , i.e.,

$$B[i] = P_w[i] \oplus T_1[i], \quad i = 1, 2, \dots, m.$$

4. *First round substitution.*

(a) Let $B = \{b_1, b_2, \dots, b_m\}$, then do the following:

```

for  $i = 1$  to  $m$ 
   $b_{i-1} = c_{i-1}$ 
   $c_i = [b_i \lll \sigma(b_{i-1})] \oplus b_{i-1}$ 
end for

```

where $b_0 = c_0 = b_m$; “ \lll ” denotes the *left rotation* operator; and “ σ ” in $\sigma(b_{i-1})$ denotes *sum-of-digits* function, and $\sigma(b_{i-1})$ denotes sum-of-digits of b_{i-1} . For instance, if $b_{i-1} = 123$, then $\sigma(123) = 1 + 2 + 3 = 6$.

- (b) Let $C = \{c_1, c_2, \dots, c_m\}$, then do the following:

$$C[i] = C[i] \oplus K_{3w}[i] \quad i = 1, 2, 3, \text{ and}$$

$$C[m] = C[m] \oplus K_{3w}[4].$$

5. *Second round substitution.*

- (a) Do the following:

```

for  $j = 1$  to  $m$ 
   $c_{j-1} = d_{j-1}$ 
   $d_j = [c_j \lll \sigma(c_{j-1})] \oplus c_{j-1}$ 
end for

```

where $d_0 = c_m$ and the rest symbols have their usual meanings.

- (b) Let $D = \{d_1, d_2, \dots, d_m\}$, then do the following:

$$E[j] = D[j] \oplus T_2[j], \quad j = 1, 2, \dots, m.$$

6. Apply the reverse preprocessing on the audio file E of size $1 \times m$ (see Sect. 4.2), and let the obtained audio file be F of size $l \times 1$.

Output. The encrypted audio file F of size $l \times 1$.

4.5 *Decryption Algorithm*

Input. The encrypted audio file F of size $l \times 1$ and the secret key K (32-byte).

1. Apply the preprocessing on the audio file F (see Sect. 4.1) to obtain an audio file E of size $1 \times m$, m being number of WORDS in E .
2. *Second round substitution.*

- (a) XOR the audio file E with T_2 , i.e.,:

$$D[j] = E[j] \oplus T_2[j], \quad j = 1, 2, \dots, m.$$

(b) Let $D = \{d_1, d_2, \dots, d_m\}$, then do the following:

```

for  $j = m$  to 1
   $c_j = [d_j \oplus d_{j-1}] \ggg \sigma(d_{i-1})$ 
end for

```

where “ $j = m$ to 1” means $j = m, m - 1, \dots, 2, 1$; $d_0 = d_m$; and “ \ggg ” denotes the right rotation.

3. *First round substitution.*

(a) Let $C = \{c_1, c_2, \dots, c_m\}$, then do the following:

$$C[i] = C[i] \oplus K_{3w}[i], \quad i = 1, 2, 3, \text{ and}$$

$$C[m] = C[m] \oplus K_{3w}[4].$$

(b) Do the following:

```

for  $i = m$  to 1
   $b_i = [c_i \oplus c_{i-1}] \ggg \sigma(c_{i-1})$ 
end for

```

where $c_0 = C_m$ and the rest symbols have their usual meanings.

4. *Initial round substitution.* Let $B = \{b_1, b_2, \dots, b_m\}$, then do the following:

$$P_w[i] = B[i] \oplus T_1[i], \quad i = 1, 2, \dots, m.$$

5. Apply the reverse preprocessing on the audio file P_w (see Sect. 4.2) of size $1 \times m$, to obtain the audio file P of size $l \times 1$.

Output. The decrypted (original) audio file P of size $l \times 1$.

5 Implementation and Experimental Results

The proposed technique is implemented on *MATLAB (R2021a)* software under the Windows 10 operating system. To evaluate the performance (encryption and decryption qualities) of the proposed technique, two test audio files of different sample lengths are taken from the *MATLAB IPT* (Image Processing Toolbox).² The details of these audio files are provided in Table 1. Also, the oscillograms of the original, encrypted, and the decrypted audio files are shown in Fig. 3.

² Available in, C:\Program Files\Polyspace\R2021a\toolbox\images\imdata.

Table 1 Description of the test audio files

File name (.wav)	Channels	Sample rate (in Hz)	Total samples (length)	Duration (in s)	Bits/Sample	Bit rate (in kbps)	Size (in KB)
Handel	1	8192	73113	8.9249	16	131.0720	142.7988
Splat	1	8192	10001	1.2208	16	131.0720	25.1562

6 Security Analyses

6.1 Key Space Analysis

The space of all potential combinations of a key constitutes a key space of any encryption/decryption algorithm. Keyspace should be very large so that attacks, such as the brute-force [9], known/chosen plaintext [10], etc., could become unsuccessful. Our proposed technique is based on a secret key of 32 bytes (256 bits), which produces a key space of 2^{256} , and as of today, it is believed to be unbreakable.

6.2 Encryption Evaluation Metrics

Since any single metric cannot evaluate any encryption algorithm (or any encrypted audio file) fully, so we employ two important metrics, namely, the oscillogram and the number of sample change rates.

6.2.1 Oscillogram Analysis

The oscillogram is a 2D graph of an audio file (or signal) between the amplitude and the time, generated by the *oscilloscope*, a.k.a. *oscillograph* [11, Chap. 5]. It represents the change in amplitude of an audio file over the time. X-axis represents the time in seconds, while Y-axis represents the amplitude in volts. The oscillograms of the original, encrypted, and the decrypted audio files are shown in Fig. 3.

From Fig. 3, we observe that the oscillograms of the encrypted audio files are uniform, unlike those of the corresponding original audio files. Also, the oscillograms of the decrypted audio files are identical to those of the corresponding original files. Thus, our proposed technique performs a robust encryption. Also, since the audio files are successfully decrypted without any data loss, the designed technique performs lossless decryption.

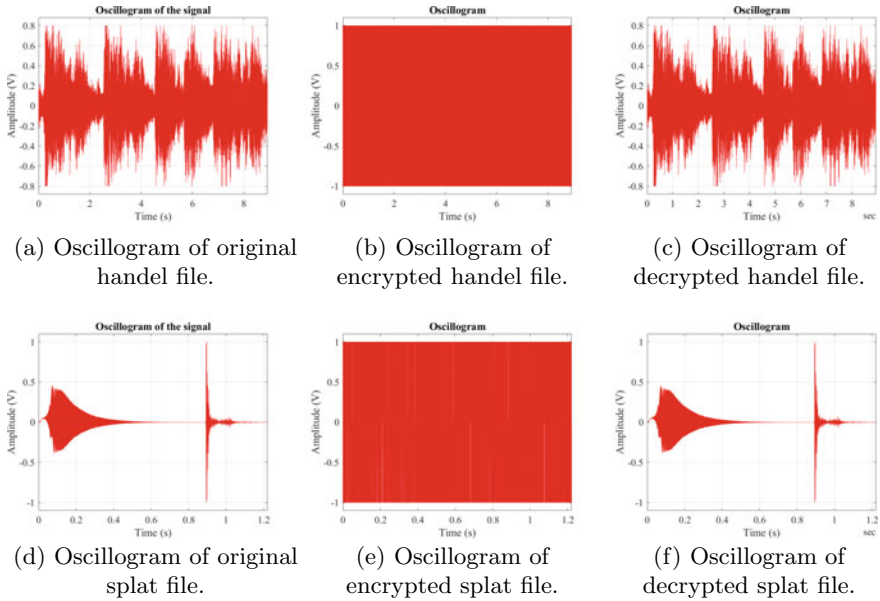


Fig. 3 Experimental results: Figs. **a** and **d** show the oscillograms of the original audio files; Figs. **b** and **e** show the oscillograms of the corresponding encrypted audio files; and Figs. **c** and **f** show the oscillograms of the corresponding decrypted audio files

6.2.2 Number of Sample Change Rate (NSCR) Test

The NSCR [13] is used to test the resistance of the differential attack [14], or judging the Shannon’s *diffusion* property [3]. The NSCR scores between the encrypted audio files E_1 and E_2 can be calculated via Eq. 1:

$$NSCR = \sum_{s=1}^l \frac{\beta(s, 1)}{l} \times 100\% , \tag{1}$$

where $\beta(s, 1)$ is given by Eq. 2:

$$\beta(s, 1) = \begin{cases} 0, & \text{if } E_1(s, 1) = E_2(s, 1) \\ 1, & \text{if } E_1(s, 1) \neq E_2(s, 1) \end{cases} \tag{2}$$

where $E_1(s, 1)$ and $E_2(s, 1)$ are the samples of the encrypted audio files prior to and after alteration of only one sample of the original audio file.

We have calculated the NSCR scores by changing only one sample of the test audio files at different positions (from beginning—(1, 1)th sample as well as from the last—(l, 1)th sample), l being the total number of samples in an audio file. The obtained NSCR scores are shown in Table 2. Note that, if the calculated/reported

Table 2 NSCR scores of the encrypted images

Method	File Name (.wav)	Duration (in s)	Size (in KB)	Position altered	NSCR Score (in %)
Proposed	Handel	8.9249	142.7988	(1, 1) (l, 1)	100 100
	Splat	1.2208	25.1562	(1, 1) (l, 1)	100 100
Shah [2]	Bells sound	–	–	–	99.9884
Faragallah [4]	Alarm	–	–	–	99.7500
Naskar [5]	Audio-4	–	162	–	99.9958
Abdelfatah [6]	Audio-2	1.7600	296.875	–	99.9700

NSCR score is greater than the theoretical NSCR value, which is 99.5527 at 0.01 significance level and 99.5693% at 0.05 level [13], then the NSCR test is *passed*. The proposed technique passes the NSCR test for all the audio files, and thus, ensures the property of diffusion, and also, outperforms the methods listed in Table 2, which are vulnerable to the differential attack.

6.3 Decryption Evaluation Metric

To evaluate the decryption algorithm, i.e., the decrypted audio files, we use an important metric: the mean square error.

6.3.1 Mean Square Error (MSE) Analysis

The MSE [15] is used to judge the decryption quality of any decrypted audio file. The MSE value can be any non-negative integer. *Lower* the MSE, *better* is the decryption quality, in particular, value 0 denotes the perfect decryption, i.e., the original and the decrypted audio files are exactly identical—lossless decryption. The MSE can be calculated via Eq. 3:

$$MSE = \sum_{j=1}^l \frac{(P_j - D_j)^2}{l}, \tag{3}$$

where P_j and D_j denote the j th samples of the original and the decrypted audio files, respectively, while the other symbols have their usual meanings.

The values of the MSE between the original and the decrypted audio files are provided in Table 3. From the table, we observe that the MSE values are 0 (zero), endorsing that the decrypted audio files are perfectly identical to the original audio files.

Table 3 The MSE values between the decrypted and the original audio files

Method	File name (.wav)	Duration (in s)	Size (in KB)	MSE
Proposed	Handel splat	8.9249	142.7988	0
		1.2208	25.1562	0
Abouelkheir [1]	Sen_4	1.1901	41.0156	3.3161×10^{-11}

6.4 Key Sensitivity Analysis

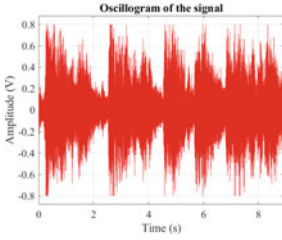
This test is utilized to judge the *confusion* property [3] of any encryption/decryption algorithm. According to Shannon [3], a secure cryptographic algorithm must have the confusion property to thwart statistical attacks. It is the property of confusion that hides the relationship between the encrypted data and the secret key. The key sensitivity test is utilized to judge this confusion property. The sensitivity of the secret key is assessed in two aspects:

1. **Encryption.** It is used to measure the dissimilarity between the two encrypted audio files E_1 and E_2 w.r.t. the same plain audio file P using two different encryption keys λ_1 and λ_2 , where λ_1 and λ_2 are obtained from the original secret key K by altering merely the LSB corresponding to the last and the first bytes of K , respectively.
2. **Decryption.** It is used to measure the dissimilarity between the two decrypted audio files D_1 and D_2 w.r.t. the same encrypted audio file E , encrypted via secret key K , using the decryption keys λ_1 and λ_2 , respectively. Note that both the encryption/decryption keys λ_1 and λ_2 differ from each other as well as from the secret key K merely by 1-bit.

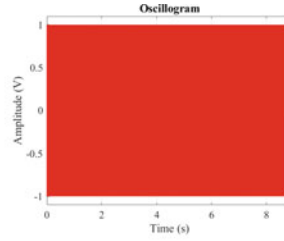
The results of the key sensitivity analysis w.r.t. the encryption (*enc*—in short) and decryption (*dec*—in short) aspects are shown in Figs. 4 and 5, respectively, whence we infer that the proposed technique has a very high bit-level sensitivity, and thus, ensures the property of confusion.

7 Comparison with the Existing Techniques

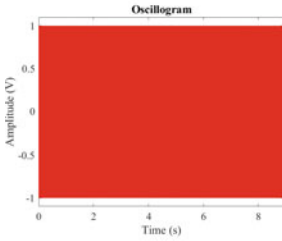
The proposed technique is compared with the recent state-of-the-art techniques based on the commonly available metrics, namely, the NSCR and the MSE. The comparisons of the proposed approach with the recent approaches based on the NSCR and the MSE metrics are provided in Tables 2 and 3, respectively. From these tables, we infer that our proposed technique performs well in terms of the respective compared metrics.



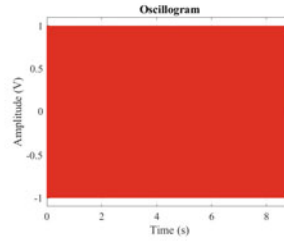
(a) Oscilloscope of the original handel file (P).



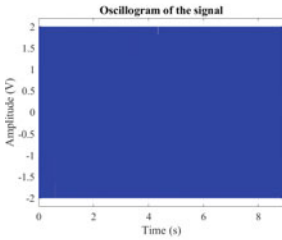
(b) Oscilloscope of the encrypted handel file E , where $E = enc(P, K)$.



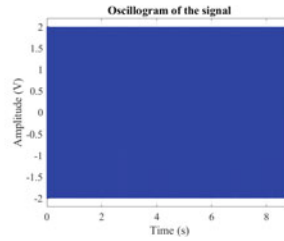
(c) Oscilloscope of E_1 , where $E_1 = enc(P, \lambda_1)$.



(d) Oscilloscope of E_2 , where $E_2 = enc(P, \lambda_2)$.



(e) Oscilloscope of $|E_1 - E|$, where $|\cdot|$ denotes the absolute difference.

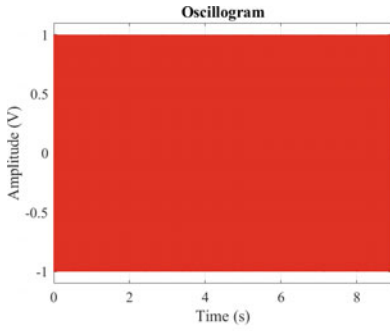


(f) Oscilloscope of $|E_2 - E|$.

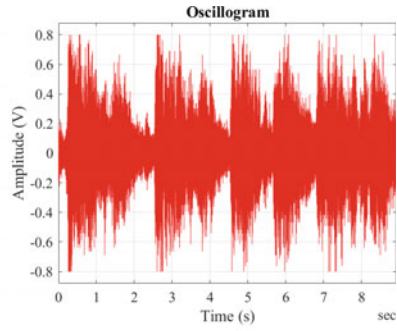
Fig. 4 Key sensitivity analysis w.r.t. the encryption

8 Conclusion

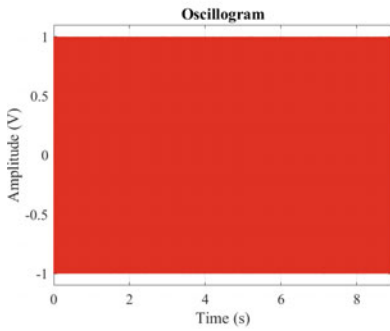
In this paper, we proposed a technique for securing digital audio files based on the WORD-oriented RX operations. Several performance evaluation metrics, i.e., encryption and decryption evaluation metrics, have been used on the audio files of varying sizes from the standard database, in order to empirically assess the efficiency and robustness of the designed approach. The results of these performance evaluation metrics validate the goals of the proposed approach. Moreover, a thorough comparison with the recent state-of-the-art techniques, based on several metrics, have also been made.



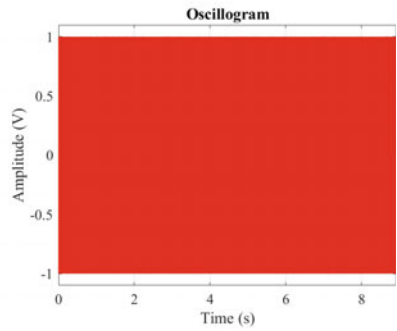
(a) Oscillogram of the encrypted handel file E , where $E = enc(P, K)$.



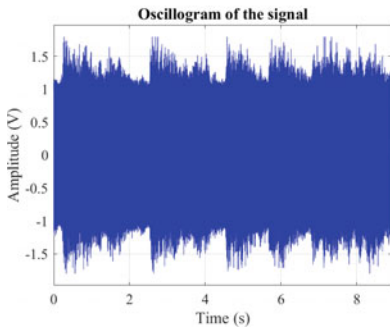
(b) Oscillogram of the decrypted handel file (D) using correct secret key K .



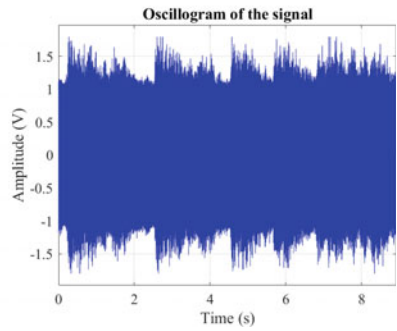
(c) Oscillogram of D_1 , where $D_1 = dec(E, \lambda_1)$.



(d) Oscillogram of D_2 , where $D_2 = dec(E, \lambda_2)$.



(e) Oscillogram of $|D_1 - D|$.



(f) Oscillogram of $|D_2 - D|$.

Fig. 5 Key sensitivity analysis w.r.t. the decryption

References

1. Abouelkheir E, Sherbiny SE (2022) Enhancement of speech encryption/decryption process using RSA algorithm variants. *Hum-Centric Comput Inf Sci* 12(6). <https://doi.org/10.22967/HCCIS.2022.12.006>
2. Shah D, Shah T, Hazzazi MM, Haider MI, Aljaedia, Hussain I (2021) An efficient audio encryption scheme based on finite fields. *IEEE Access* 9:144385–144394. <https://doi.org/10.1109/ACCESS.2021.3119515>
3. Shannon CE (1949) Communication theory of secrecy systems. *Bell Syst Tech J* 28(4):656–715. <https://doi.org/10.1002/j.1538-7305.1949.tb00928.x>
4. Faragallah OS, El-Sayed HS (2021) Secure opto-audio cryptosystem using XOR-ing mask and Hartley transform. *IEEE Access* 9:25437–25449. <https://doi.org/10.1109/ACCESS.2021.3055738>
5. Naskar PK, Bhattacharyya S, Chaudhuri A (2021) An audio encryption based on distinct key blocks along with PWLCM and ECA. *Nonlinear Dyn* 103:2019–2042. <https://doi.org/10.1007/s11071-020-06164-7>
6. Abdelfatah RI (2020) Audio encryption scheme using self-adaptive bit scrambling and two multi chaotic-based dynamic DNA computations. *IEEE Access* 8:69894–69907. <https://doi.org/10.1109/ACCESS.2020.2987197>
7. Available at https://in.mathworks.com/help/matlab/matlab_prog/floating-point-numbers.html. Accessed 05 Nov 2022
8. Available at https://en.wikipedia.org/wiki/Single-precision_floating-point_format. Accessed 05 Nov 2022
9. ECRYPT II yearly report on algorithms, key sizes NS (eds) (BRIS) 2011–2012. <https://www.ecrypt.eu.org/ecrypt2/documents/D.SPA.20.pdf>. Accessed 05 Nov 2022
10. Stinson DR (2006) *Cryptography: theory and practice*. Chapman and Hall CRC, UK
11. Kularatna N (2002) *Digital and analogue instrumentation: testing and measurement*. IET, UK. <https://doi.org/10.1049/PBEL011E>
12. Belmeguenai A, Ahmida Z, Ouchtati S, and Dejmi R (2017) A novel approach based on stream cipher for selective speech encryption. *Int J Speech Technol* 20:685–698. <https://doi.org/10.1007/s10772-017-9439-8>
13. Wu Y, Noonan JP, Aghaian S (2011) NPCR and UACI randomness tests for image encryption. *J Sel Areas Telecommun* 31–38
14. Biham E, Shamir A (1993) *differential cryptanalysis of the data encryption standard (DES)*. Springer, US
15. Hossein PN (2014) *Introduction to probability, statistics, and random processes*. Kappa Research LLC, USA

Author Index

A

Abdul Gaffar, 453
Abhishek Bajpai, 265
Abhishek Singh Rathore, 1
Adhiraj Gupta, 53
Aditya Kumar Sharma, 441
Ajay Pratap, 77
Al-Sakib Khan Pathan, 397
Amit Verma, 279
Anand B. Joshi, 241
Anita Yadav, 265
Anju Shukla, 433
Arjun Choudhary, 29, 353
Arsh, 109

B

Bazlur Rashid, A. N. M., 397
Bhawani Sankar Panigrahi, 367
Bhogeswar Borah, 279
Bodhi Chakraborty, 179
Brijesh Kumar Chaurasia, 441

C

Chetanya Kundra, 29, 353
Chimaya Ranjan Pattnaik, 367
Corinne Marie Formosa, 335

D

Debanjan Sadhya, 179
Dhananjay Dey, 71
Dheerendra Mishra, 147

G

Garima Jain, 375
Garima Thakran, 15
Gaurav Choudhary, 29, 353
Girish Kumar, B. C., 157
Girraj Kumar Verma, 147
Gopal, P. V. S. S. N., 41
Gowri, T., 41
Guna Shekar, P., 179

H

Harikesh Singh, 421
Himanshu Dhumras, 125, 203

I

Imran Khan, 387

H

Harsha, K. G., 157

K

Kapil Pareek, 29, 353
Kishore Kanna, R., 367
Krishan Kumar Goyal, 225

L

Lalit Garg, 335
Luca Bugeja, 335

M

Mahesh, G., 157

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2024

B. K. Roy et al. (eds.), *Cryptology and Network Security with Machine Learning, Algorithms for Intelligent Systems*, <https://doi.org/10.1007/978-981-99-2229-1>

Mahesh Kumar, 295
 Manish Kumar, 309
 Manoj Kumar Misra, 203
 Manoj Srivastava, 309
 Md Tarique Jamal Ansari, 91
 Meenakshi Srivastava, 109
 Mohammed Ishaque, 367
 Mohammed Kamran Siddiqui, 225
 Mohiuddin Ahmed, 397

N

Nand Kishore Sharma, 1
 Naseem Ahmad Khan, 91
 Neeraj Kumar, 147

P

Paras Sharma, 53
 Pawan Kumar Chaurasia, 407
 Pradeep Malik, 63
 Prashant Mathur, 29, 353
 Prateek Singh, 91
 Prateek Thakral, 53
 Pratyush Shukla, 295

R

Radiant Ambesh, 91
 Raees Ahmad Khan, 407
 Raghuraj Singh Suryavanshi, 77, 135
 Raghwendra Singh, 179
 Rahul Deo Shukla, 77
 Rakesh Kumar Bajaj, 53, 125
 Ramakant Kumar, 99

Rashi Agarwal, 387
 Ruchi Agarwal, 279

S

Sahadeo Padhye, 71, 99, 325
 Saru Kumari, 63
 Satish Kumar, 407
 Satyam Omar, 71
 Shahnaz Fatima, 135
 Shishir Kumar, 421, 433
 Shivani Dixit, 203
 Smita Bedekar, 193
 Sonali Singh, 241
 Sonam Yadav, 63
 Sonika Singh, 325
 Sreenivasulu Reddy, L., 213
 Sujit Beborra, 169
 Sumanta Kumar Singh, 169
 Sumit Chandra, 135
 Surendra Rahamatkar, 1
 Surendra Talari, 157, 421

U

Udai Bhan Trivedi, 309

V

Vanashree Gupta, 193
 Vandani Verma, 15, 375
 Varun Shukla, 125, 157, 203, 335, 421, 433
 Vasudeva Reddy, P., 41
 Vinooth, P., 421
 Vivek Dabra, 63