

Sandeep Joshi
Amit Kumar Bairwa
Milena Radenkovic
Anton Pljonkin (Eds.)

Communications in Computer and Information Science 2195

Cyber Warfare, Security and Space Computing

Second International Conference on Cyber Warfare
Security and Space Computing, SpacSec 2024
Jaipur, India, February 22–23, 2024, Proceedings

Editorial Board Members

Joaquim Filipe , *Polytechnic Institute of Setúbal, Setúbal, Portugal*

Ashish Ghosh , *Indian Statistical Institute, Kolkata, India*

Lizhu Zhou, *Tsinghua University, Beijing, China*

Rationale

The CCIS series is devoted to the publication of proceedings of computer science conferences. Its aim is to efficiently disseminate original research results in informatics in printed and electronic form. While the focus is on publication of peer-reviewed full papers presenting mature work, inclusion of reviewed short papers reporting on work in progress is welcome, too. Besides globally relevant meetings with internationally representative program committees guaranteeing a strict peer-reviewing and paper selection process, conferences run by societies or of high regional or national relevance are also considered for publication.

Topics

The topical scope of CCIS spans the entire spectrum of informatics ranging from foundational topics in the theory of computing to information and communications science and technology and a broad variety of interdisciplinary application fields.

Information for Volume Editors and Authors

Publication in CCIS is free of charge. No royalties are paid, however, we offer registered conference participants temporary free access to the online version of the conference proceedings on SpringerLink (<http://link.springer.com>) by means of an http referrer from the conference website and/or a number of complimentary printed copies, as specified in the official acceptance email of the event.

CCIS proceedings can be published in time for distribution at conferences or as post-proceedings, and delivered in the form of printed books and/or electronically as USBs and/or e-content licenses for accessing proceedings at SpringerLink. Furthermore, CCIS proceedings are included in the CCIS electronic book series hosted in the SpringerLink digital library at <http://link.springer.com/bookseries/7899>. Conferences publishing in CCIS are allowed to use Online Conference Service (OCS) for managing the whole proceedings lifecycle (from submission and reviewing to preparing for publication) free of charge.

Publication process

The language of publication is exclusively English. Authors publishing in CCIS have to sign the Springer CCIS copyright transfer form, however, they are free to use their material published in CCIS for substantially changed, more elaborate subsequent publications elsewhere. For the preparation of the camera-ready papers/files, authors have to strictly adhere to the Springer CCIS Authors' Instructions and are strongly encouraged to use the CCIS LaTeX style files or templates.

Abstracting/Indexing

CCIS is abstracted/indexed in DBLP, Google Scholar, EI-Compendex, Mathematical Reviews, SCImago, Scopus. CCIS volumes are also submitted for the inclusion in ISI Proceedings.

How to start

To start the evaluation of your proposal for inclusion in the CCIS series, please send an e-mail to ccis@springer.com.

Sandeep Joshi · Amit Kumar Bairwa ·
Milena Radenkovic · Anton Pljonkin
Editors

Cyber Warfare, Security and Space Computing

Second International Conference on Cyber Warfare
Security and Space Computing, SpacSec 2024
Jaipur, India, February 22–23, 2024
Proceedings



Springer

Editors

Sandeep Joshi  Manipal University Jaipur
Jaipur, Rajasthan, India

Amit Kumar Bairwa  Manipal University Jaipur
Jaipur, Rajasthan, India

Milena Radenkovic  University of Nottingham
Nottingham, UK

Anton Pljonkin  Southern Federal University
Rostov-on-Don, Russia

ISSN 1865-0929

ISSN 1865-0937 (electronic)

Communications in Computer and Information Science

ISBN 978-3-031-73493-9

ISBN 978-3-031-73494-6 (eBook)

<https://doi.org/10.1007/978-3-031-73494-6>

© The Editor(s) (if applicable) and The Author(s), under exclusive license
to Springer Nature Switzerland AG 2025, corrected publication 2025

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 Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

If disposing of this product, please recycle the paper.

Preface

In our increasingly interconnected world, the necessity for secure and reliable communication systems has never been more critical. The advent of cyber warfare and the rapid advancements in space computing have introduced new challenges that threaten the integrity of our data and communications. These domains demand comprehensive exploration to ensure the security and resilience of our communication systems.

The 2nd International Conference on Cyber Warfare, Security & Space Computing (SpacSec 2024), hosted by Manipal University Jaipur on 22–23 February 2024, aimed to address these pressing issues. This conference provided a platform for researchers, industry professionals, policymakers, and other stakeholders to delve into the multi-faceted aspects of cyber warfare and space computing, fostering the exchange of ideas and best practices.

SpacSec 2024 was designed to be an engaging and informative event, featuring a dynamic format that included keynote speeches from leading experts, plenary sessions, panel discussions, and interactive workshops. Attendees also benefited from poster presentations and demonstrations, offering insights into the latest research and innovations. Social events facilitated networking, enabling participants to forge valuable connections and collaborations.

Keynote speeches provided strategic perspectives on the evolving landscape of cyber warfare and space computing, while plenary sessions explored specific challenges and emerging trends. Panel discussions encouraged interactive dialogue, promoting the sharing of diverse viewpoints and collaborative problem-solving. Workshops offered hands-on experiences with the latest tools and techniques, enhancing practical knowledge and skills.

Beyond the technical discussions, the conference addressed the legal and international implications of cyber warfare and space computing. By examining these broader contexts, SpacSec 2024 aimed to foster a holistic understanding of the issues and promote the development of comprehensive strategies for effective security measures.

We believe that SpacSec 2024 served as a vital platform for advancing our collective knowledge and capabilities. The insights gained and collaborations forged at this conference contributed to the development of innovative methods and techniques for securing communication systems against emerging threats. By bringing together a diverse community of experts and stakeholders, we aimed to inspire new approaches and drive state-of-the-art cyber and space security forward.

We are pleased to present the proceedings of this year's conference, which showcase a diverse range of research contributions from across the globe. The number of submissions sent for peer review was impressive, with a total of 111 papers submitted. This high volume of submissions reflects the growing interest and active engagement of the research community in this field.

After a rigorous and thorough peer-review process, 20 full papers have been accepted for inclusion in these proceedings. These papers represent some of the most innovative and impactful research currently being conducted, offering new insights and advancements in their respective areas.

In addition to the full papers, we are also pleased to include 7 short papers. These shorter contributions were selected based on their quality and the relevance of their findings. The inclusion of short papers was discussed and agreed upon with the Springer Editor, ensuring that these concise yet significant pieces of research are also given a platform in our proceedings.

We extend our warmest welcome to all participants and express our gratitude for your contributions to this critical dialogue. Your expertise, experience, and insights are invaluable as we work together to navigate the complexities of cyber warfare and space computing, building a safer and more secure future for all.

July 2024

Sandeep Joshi
Amit Kumar Bairwa
Milena Radenkovic
Anton Pljonkin

Organization

Chief Patron

S. Vaitheswaran

Manipal University Jaipur, India

Patron

Gopalkrishna K. Prabhu

Manipal University Jaipur, India

Co-Patrons

Jawahar M. Jangir
Nitu Bhatnagar

Manipal University Jaipur, India
Manipal University Jaipur, India

Conference Chairs

Arun Shanbhag
Sandeep Chaurasia

Manipal University Jaipur, India
Manipal University Jaipur, India

General Chair

Sandeep Joshi

Manipal University Jaipur, India

Program Chairs

Milena Radenkovic
Anton Plionkin

University of Nottingham, UK
Southern Federal University, Russia

Organizing Secretary

Amit Kumar Bairwa

Manipal University Jaipur, India

Technical Program Committee

Puneet Mittal	Manipal University Jaipur, India
Vivek Bhardwaj	Manipal University Jaipur, India
Saurabh Srivastava	Manipal University Jaipur, India
Surbhi Sharma	Manipal University Jaipur, India
Preeti Narooka	Manipal University Jaipur, India
Shikha Maheshwari	Manipal University Jaipur, India

Finance Chair

Deepika Shekhawat	Manipal University Jaipur, India
-------------------	----------------------------------

Sponsorship Chairs

Shishir Singh Chauhan	Manipal University Jaipur, India
Shilpi Birla	Manipal University Jaipur, India

Programme Committee

Anita Shotriya	Manipal University Jaipur, India
Sayar Singh Shekhawat	Manipal University Jaipur, India
Neelam Chaplot	Manipal University Jaipur, India
Shubh Lakshmi Agrwal	Manipal University Jaipur, India
Manish Rai	Manipal University Jaipur, India
Upendra Singh	Manipal University Jaipur, India
Lokesh Malviya	Manipal University Jaipur, India
Yadvendra Pratap Singh	Manipal University Jaipur, India

Publication Chairs

Gautam Kumar	Manipal University Jaipur, India
Ajay Kumar	Manipal University Jaipur, India
Surendra Solanki	Manipal University Jaipur, India
Siddharth Kumar	Manipal University Jaipur, India
Jayesh Gangrade	Manipal University Jaipur, India

Programme Advisory Committee

Dinesh Kumar Saini	Manipal University Jaipur, India
Amit Soni	Manipal University Jaipur, India
Roheet Bhatnagar	Manipal University Jaipur, India
Sumit Srivastava	Manipal University Jaipur, India
Vijaypal Singh Dhaka	Manipal University Jaipur, India
Prakash Ramani	Manipal University Jaipur, India
Akhilesh Kumar Sharma	Manipal University Jaipur, India
Shakti Kundu	Manipal University Jaipur, India
Sunil Kumar Vasistha	Manipal University Jaipur, India
Pankaj Vyas	Manipal University Jaipur, India

Technical Session (Online)

Satpal Singh Kushwaha	Manipal University Jaipur, India
Lav Upadhyay	Manipal University Jaipur, India
Hemlata Parmar	Manipal University Jaipur, India
Neetu Gupta	Manipal University Jaipur, India
Preeti Narooka	Manipal University Jaipur, India

Technical Session (Offline)

Deepak Panwar	Manipal University Jaipur, India
Mahesh Jangid	Manipal University Jaipur, India
Jaydeep Kishore	Manipal University Jaipur, India

Registration Committee

Babita Tiwari	Manipal University Jaipur, India
Sushama Tanwar	Manipal University Jaipur, India
Bali Devi	Manipal University Jaipur, India
Vaishali Chauhan	Manipal University Jaipur, India
Pallavi	Manipal University Jaipur, India

Food Committee

Prakash Chandra Sharma
Venkatesh Gauri Shankar

Manipal University Jaipur, India
Manipal University Jaipur, India

Accommodation Committee

Puneet Mittal
Dinesh Swami

Manipal University Jaipur, India
Manipal University Jaipur, India

Transport Committee

Ajit Noonia
Ajay Kumar
Dinesh Swami

Manipal University Jaipur, India
Manipal University Jaipur, India
Manipal University Jaipur, India

Cultural Committee

Harish Sharma

Manipal University Jaipur, India

Liaison Committee

Satyabrata Roy
Deepjyoti Choudhury

Manipal University Jaipur, India
Manipal University Jaipur, India

Poster Presentation

Juhi Singh

Manipal University Jaipur, India

Media and Photography

Surbhi Sharma
Rishika Singh

Manipal University Jaipur, India
Manipal University Jaipur, India

Website Coordinator

Vineeta Soni

Manipal University Jaipur, India

Additional Reviewers

Adel Al-Jumaily	University of Technology Sydney, Australia
Cihangir Tezcan	Middle East Technical University, Turkey
Ekaterina Maro	Southern Federal University, Taganrog, Russia
Sedat Akleylek	Middle East Technical University, Turkey
Selwyn Piramuthu	Warrington College of Business, USA
Dmitry Namiot	Lomonosov Moscow State University, Russia
Victor Govindaswamy	Concordia University Chicago, USA
Muthu Kumar S.	National Institute of Technology, Puducherry, India
Nisheeth Joshi	Banasthali University, India
Rasib Khan	Northern Kentucky University, USA
Ekaterina Maro	Southern Federal University, Rostov-on-Don, Russia
Mohiuddin Ahmed	University of New South Wales, Australia
Muhammad Mostafa Monowar	King Abdul-Aziz University, Saudi Arabia
Sanjay Bansal	Acropolis Inst. of Tech. & Research, India
Jagannathan Sarangapani	Missouri University of Science and Tech., USA
B.K. Verma	Shridhar University, India
Muhammad Mostafa Monowar	King Abdulaziz University, Saudi Arabia
Nurilla Avazov	Inha University, South Korea
Abhineet Anand	Galgotias University, India
Sukanta Ganguly	Entryless, USA
P. K. Gupta	University of Pretoria, South Africa
M. M. Hafizur Rahman	International Islamic University, Malaysia
Xianglin Wei	NTTRI, Shanghai, China
Elena Basan	Southern Federal University, Rostov-on-Don, Russia
Maxim Anikeev	Southern Federal University, Rostov-on-Don, Russia
Parag Narkhede	Symbiosis Institute of Technology, India
Snehanshu Shekhar	BIT Mesra, India
Ekaterina Pakulova	Southern Federal University, Rostov-on-Don, Russia
G. S. Mahapatra	National Institute of Technology, Puducherry, India

Mohiuddin Ahmed	University of New South Wales, Australia
Wilfred Lin	PuraPharm, China
Syed Hassan Ahmed	Kyungpook National University, South Korea
Eduard Babulak	Fort Hays State University, USA
Carlos Eduardo de Barros Paes	PUC-SP, Brazil
Jalel Ben Othman	Institut Galilée, France
Adel Al-Jumaily	University of Technology Sydney, Australia
Krithi Ramamritham	IIT, Bombay, India
Lei Zhang	East China Normal University, China
Edmond C. Prakash	University of Westminster, UK
D. M. Akbar Hussain	Aalborg University, Denmark
Prabhaker Mateti	Russ Engineering Center, USA
Vijay Laxmi	Malaviya National Institute of Tech., India
Olariu Stephan	Old Dominion University, USA
Atilla Elci Aksaray	Aksaray University, Turkey
Mouhamed Abdulla	Chalmers University of Technology, Sweden
Syed Nisar Bukhari	NIELIT Srinagar, MeitY, India
Bert-Jan van Beijnum	University of Twente, Netherland

Contents

Visvesvarayya Space Lab: Preliminary Space Lab Demonstrator	1
<i>Sanjay Lakshminarayana, Vinod Singh Yadav, Vinit Soni, Anil Kumar Jogi, Devendra Kumar, Yashwanth Singh Chauhan, Harsh Vardhan, Jitendra Singh, Raj Kumar, Sachin Kumar, Surendra Kumar, Kuldeep, and Ajay Kumar Vaishnav</i>	
Hybrid Deep Driven Cross Industry Sentiment Analysis Model for Netizen's Behavioral Characterization	26
<i>Santhosh Priya and R. Kalaiarasu</i>	
Fake Product Detection Using Blockchain with Encryption and AI	57
<i>Shyam Pandit, Abhay Jakhere, Siddharth Tated, Omkar Bhakare, and Chandan Prasad</i>	
S-Defender: A Smishing Detection Approach in Mobile Environment	68
<i>Ankit Kumar Jain, Ankur Panday, and Diksha Goel</i>	
Mathematical Socio Analysis of Cybercrimes Preparedness a Simulation Odessey with R	79
<i>S. Dheva Rajan</i>	
Revolutionizing Weather Forecasting: Harnessing Machine Learning and Big Data in Upcoming Technologies	91
<i>Basetty Mallikarjuna and Varun Tiwari</i>	
Trends in Drowsiness Detection & Analysis of the Different Technologies Engaged	101
<i>Sachin B. Honrao and U. D. Shiurkar</i>	
Detecting Local Software Issues Using NSGA Multi-optimization	112
<i>K. R. Jothi, Gireesh Kambala, Chetan Khemraj Lanjewar, Leena Jain, Janjhyam Venkata Naga Ramesh, and Pavitar Parkash Singh</i>	
Deep Learning and IoT Based Robotics to Monitor the Traffic	125
<i>V. Vishwa Priya, Soumitra S. Pande, Md Ilyas, R. Jayasudha, Janjhyam Venkata Naga Ramesh, and D. Suganthi</i>	

Development of Lightweight and Cheaper 5G Mobile Communication System to Analyze the Performance of Espresso Ciphers and Grain Family	140
<i>C. Kamalanathan, J. Balamurugan, Neelam Sharma, A. Basi Reddy, and R. Senthamil Selvan</i>	
Development of Light Weight Authentication Protocol Based on Cryptography to Access the IoT Device	154
<i>Sameer Yadav, Surepalli Venkataratnam, P. Balaji Srikanth, Jetti Madhavi, A. Basi Reddy, and R. Senthamil Selvan</i>	
Development of Elliptical Cryptography Technique to Watermark Embedded and Extrusion for Healthcare Records	167
<i>N. Prajwal Hegde, R. Sivaraman, G. Dharmamoorthy, Ankit Kumar Dubey, Pramoda Patro, and S. Suma Christal Mary</i>	
Step-by-Step Image Encryption Using UACI and PixAdapt	182
<i>J. Balamurugan, Mali Yadav, Jetti Madhavi, A. Basi Reddy, and R. Senthamil Selvan</i>	
Investigation of Post-Quantum Cryptography to Secure the Functionality of Vehicle Hardware Architecture	194
<i>K. R. Jothi, Chetan Khemraj Lanjewar, R. Sivaraman, Bramah Hazela, P. R. Sivaraman, and A. Azhagu Jaisudhan Pazhani</i>	
Secure Data Management Using BlockChain	207
<i>Akshay Tupetewar, Shivprasad Chaudhari, Shashikant Deshmukh, and Sonali. V. Shinkar</i>	
Navigating Through Digital Realm: Role of Cyberpsychology in Fostering Mental Well-Being and Digital Empathy	220
<i>Anadi Trikha, Antima Sharma, Arpita Agarwal, Preeti Nagar, and Ritu Singh</i>	
An Empirical Analysis of Neighborhood-Based Approaches for Trustworthy Recommendations with Apache Mahout	229
<i>Vijay Verma</i>	
Multilingual Sentiment Analysis over Real-Time Voice	238
<i>Samikshya Rath, Ojasvi Nagayach, Asritha Boddu, Raguru Jaya Krishna, and B. Vamshi Krishna</i>	
Identity Verification: A Decentralized KYC Approach Using Blockchain	252
<i>Akshay Chouke, Vikas Kumar Jain, Jitendra Parmar, Shiv Shankar Prasad Shukla, and Atul Kumar Verma</i>	

A Neural Network-Based Facial Expressions Detection Technique Using CK+ Dataset	265
<i>Subhash Chandra Jat, Sadaf Naaz, and Shikha Chaudhary</i>	
Cyber Security Challenges in Industrial Settings with the Internet of Things ...	281
<i>Shailaja Salagrama, Amit Garg, J. Logeshwaran, Satpal Singh Kushwaha, and Rajan Kumar</i>	
Designing Secure Software-Defined Network, Resistant to DDoS Attack Using Non-linear Routing Rule Installation	291
<i>Anoop Kumar Patel and Prince Raj</i>	
NPQuant: A Robust Quantum Inspired Computation Algorithms as an Efficient Solution to NP-Complete Problems	302
<i>Bali Devi, Mehil Bimal Shah, Venkatesh Gauri Shankar, and Gauri Sharma</i>	
Impact of Sentiment Analysis in E-Commerce and Cybersecurity	314
<i>Sonakshi Arora, P. Harika, and Sakshi Shringi</i>	
Cultivating Cyber Vigilance: Shaping Employee Behavior for Security Success	325
<i>Antima Sharma, Anadi Trikha, Preeti Nagar, Arpita Agarwal, and Akeke Niyi Israel</i>	
A Deep Learning Approach to PDF Malware Detection Enhanced with XAI ...	337
<i>Ganapathiappan Kirubavathi and Fathima Noorudheen</i>	
Optimized Deep Learning Technique for the Effective Detection of Windows PE Malware	359
<i>Kirubavathi Ganapathiappan and Abhishek Yadav</i>	
Correction to: A Neural Network-Based Facial Expressions Detection Technique Using CK+ Dataset	C1
<i>Subhash Chandra Jat, Sadaf Naaz, and Shikha Chaudhary</i>	
Correction to: Impact of Sentiment Analysis in E-Commerce and Cybersecurity	C2
<i>Sonakshi Arora, P. Harika, and Sakshi Shringi</i>	
Author Index	371



Visvesvarayya Space Lab: Preliminary Space Lab Demonstrator

Sanjay Lakshminarayana¹ , Vinod Singh Yadav² , Vinit Soni¹ , Anil Kumar Jogi¹ , Devendra Kumar¹ , Yashwanth Singh Chauhan¹ , Harsh Vardhan¹ , Jitendra Singh¹ , Raj Kumar¹ , Sachin Kumar¹ , Surendra Kumar¹ , Kuldeep¹, and Ajay Kumar Vaishnav¹

¹ Rajasthan Institute of Engineering and Technology, Jaipur, Rajasthan 302026, India
sanjaylakshminarayana@gmail.com

² National Institute of Technology (NIT), Uttarakhand 246174, India

Abstract. In this paper preliminary design of a space lab module is realised with bottom-up approach. Initial design iterations are performed on hexagonal cross-section module design assembled together in T shape. Design consists of research and human habitable sections integrated. Structural finite element analysis is performed on the design, further iterated to obtain suitable design configuration for both orbital and ground cases. Modal analysis is performed on the orbital design to study suitability of launch vehicle adaption. Thus obtained preliminary design concept along with theme of sub-system design is implemented in real-world, life-sized demonstrator at the institution. The demonstrator served as test bench to apply, test, observe functionalities, proving capability and engineering principles. Demonstrator performance was optimised for its respective operating condition i.e., orbital and ground. The work serves as preliminary design towards realising an in-space operating space lab.

Keywords: space station demonstrator · space station design · prototype

1 Introduction

Idea to create a habitable environment in the low earth orbit is age old. There exists hundreds or more engineering concepts for a space station like habitat. While designs pre-exist, each new generation of design is profoundly new. It is diverse in terms of application and uses for space missions that it renders to. Designs are reflection of the technology readiness level of that time and with evolution of technology, designs tend follow the trend of readiness level. First of its type to be implemented, skylab consisted of minimal components to study human habitation in space. Possibilities for docking the crew module, establish ground-space transportation ports, components for research that could be performed by crew on-board etc., astronauts stayed for about 90 days. In 1979, Skylab re-entered Earth atmosphere and the mission ended its presence in the orbit. It was replaced with International Space Station (ISS). ISS proved to be technically superior, flexible in terms of operations and far more advanced, with lessons learned

from Skylab. It was launched at 400 km altitude, roughly size of football field and it hosted multiple crews over two decades including from United States, Russia, Japan and European Union. In Skylab and also ISS, mechanical structure of modules among many play key role in deciding the longevity of the mission. The structure is primary defence and also safety element to the crew from micro-meteors [1–4, 18].

Design and concept, operation and research on-board ISS is well known and documented. Third party and third country involvement is severely restricted with ISS missions. By end of 2030, ISS may be decommissioned or operated with new arrangements with private players due to various reasons listed in [4]. As of 2023, every space enabled country is eyeing their own space station with vision to commercialize it. For instance, “Tiangong” space station from China is made up of three large modules connected in a T-shaped configuration launched in April 2021. With further plans on Tiangong to expand the crew and research activities in near future. The costs of such ambitions restrict the possibility to realise the plans. One among those challenges is economic access to low earth orbit (LEO). Launching modules from Earth requires launch vehicles with payload capacity to lift off at least 25–30 tonnes at a time into the prescribed orbit. The cost to run such launch vehicles are still expensive and also availability of such launchers, capable for the job are few in the world. Since 2014, private companies have steadily increased presence and market share overall in comparison to previous decades. Development of reusable micro-launchers are challenging established companies to develop large-payload capacity launchers in order to remain in competition. Price per kilogram of payload insertion to LEO is already as low as 2000 US\$ and possibly when promises of new space era is met, goal is to reach 10 US\$ per kg or similar scale. Such developments will result with ease of access, enable effective launch of large payloads, multiple times to LEO. It also enables low cost human missions to LEO for tourism etc. Collectively, it is a leap towards a sustainable model towards colonising space [4–6].

Engineering solutions and its development are important towards implementation of new ways of living and human life enhancement. Among many important media of communication, especially in democracies that connects government programs and development of space for peaceful purpose is “public interest”. Access to “space-experience” through government setups such as museums etc to provide an informative know-how of space-tech and its influence on everyday life is essential. Such human and technology interfaces are necessary to be established at every foot and corner of country to foster positive informed opinion among public at large and also to motivate involvement, especially students.

In this paper, design for human habitat module with research module is envisioned as preliminary step. Study for its implementation in terms of readiness, capability and maturity to realise flight ready design is investigated. Finite element model is prepared to study the structural and modal behaviour. Based on the results, design iterations were performed to realise suitable design for both ground and orbital implementation. The performance of the design is realised by building real-life sized demonstrator named here after as Visvesvarayya Space Lab (VSL). The demonstrator which was built at academic institution campus served engineering interests in testing, developing the concept, challenges associated with implementation and also was utilised for public outreach. Here, a

small portion of the investigation is showcased as proof of the study. Experimental data and findings, know-how, detailed design of sub-systems are withheld intentionally.

2 Design and Definitions

A hollow tube with square cross-section is welded into hexagonal shape with 'n' nodes extended by length 'L' along-Z axis shown in Figure 1. The shape provides flat surface area on each side of the hexagon. Even though the second moment of inertia is slightly lower compared to the circular cross-section, hexagonal cross-section foresee relatively better resistance to impact loads. Also, it is easier to replace a single panel of hexagonal module rather than to replace entire circular section in case of damage. Stress concentration along the edges of the hexagonal cross-section is overcome by strategic welding and design to an extent. The shape maximises storage space and can be stacked easily on launch vehicle adapter platform for launch. Energy method is applied assuming, work done by external forces equals the energy stored in the structure under applied load. It is safe to assume that deformation and corresponding stress is small. Therefore, the relation between stress (σ) and strain (ϵ) can be considered within the elastic limits of the material. Approximate deformation function can be represented in matrix form using variable 's' as shown in Eq. 1.

$$\mathbf{u} = [u(s, z), v(s, z), w(s, z)]^T \quad (1)$$

Corresponding stress and strain relation can be represented in matrix form as shown in Eq. 2,

$$\sigma = E\epsilon = \begin{bmatrix} E/1 - v^2 & E/1 - v^2 & 0 \\ E/1 - v^2 & E/1 - v^2 & 0 \\ 0 & 0 & E/(2(1 + v)) \end{bmatrix} \epsilon \quad (2)$$

where v is Poisson's ratio, E is Young's modulus of material. Further, Heat flow into the module through thickness is calculated using Fourier's law along each direction and through the thickness. By neglecting convective effects,

$$Q_{heat} = -k_{cond} * A_{CRS} * \frac{\delta(T)}{b} \quad (3)$$

Heat balance is obtained by considering main heat sources,

$$dQ_{total} = dQ_{sun} + dQ_{backgnd-rad} + dQ_{on-board} + dQ_{albedo} \quad (4)$$

where in Eq. 2, A_{CRS} area of cross-section, δ (T) is temperature difference between inner wall and outer wall of the section, b is total thickness of the layer, k_{cond} material equivalent thermal conductivity of multi-layered insulation taken to be 0.85 W/m-K, Q_{heat} heat conducted through the layers of insulation in each direction of thickness of the wall. In Eq. 4, dQ_{total} is the total heat on the module, dQ_{sun} is the heat from sun in orbit which is constant at 1 AU at 1360 W/m² approximately, $dQ_{on-board}$ heat from equipment on-board, dQ_{albedo} , earth albedo in orbit, $dQ_{backgnd-rad}$ background radiation. For each human on-board the space station, heat dissipation is considered equivalent to 100 watt when resting and while performing activity 250 watt is assumed [21].

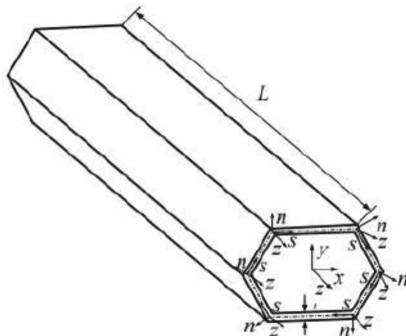


Fig. 1. Global (x, y, z), local coordinate system (s, n, z) of the cross-sections

2.1 Design

Design and development of complex structure such as a space station is not a one step process but a continuous evolution with time as per mission needs. The work presented in this paper serves as a preliminary study with a bottom up approach. Therefore, in this section preliminary orbital design concept referred to as “Module 2014” and ground implemented design concept referred to “Module 1729” or VSL is presented. Both represent the same concept but former is adopted for orbital environmental conditions and later for ground conditions.

Purpose of developing the design with bottom-up approach and then building the demonstrator is:-

1. Project initially drew inspiration from “There are some who question the relevance of space activities in a developing nation. To us, there is no ambiguity of purpose. We do not have the fantasy of competing with the economically advanced nations in the exploration of the moon or the planets or manned space-flight. But we are convinced that if we are to play a meaningful role nationally, and in the community of nations, we must be second to none in the application of advanced technologies to the real problems of man and society.” as quoted by Sarabhai [19].
1. Develop module with state of art facility to enable scientific research on-board and to identify functional challenges.
2. Develop accessibility to on-board data recorded from various studies, ability to live and work in space
3. Provide test-bed facility to develop technology at higher Technology Readiness Level (TRL)
4. Provide a collaborative environment towards easy access to low earth orbit including reduced travel cost and for tourism purpose/space taxi implementation.
5. Build a prototype to apply engineering principles, test it, to work around and optimise the solution. Prove engineering and personal capability, capacity and vigour to make the same possible for space. Finally, to open the prototype for public outreach and experiences.

Table 1. Design principles

Domain	Problems	Remedy
Environment	Vacuum corrosion Solar Radiation, Earth albedo, Asteroids./meteors/space debris Station keeping	Use paints and Multi layered insulation to guard against radiative effects and corrosion. Design multi layer collision barriers to reduce the damage due to asteroids strikes. Use cold gas thrusters for station keeping if required
Launch challenges	Limitation of fairing size Limitation of payload mass cost per kg	Promising developments of heavy launchers i.e., star-ship from SpaceX may enable one time launch of whole concept. While improving the affordability to LEO
Sub systems	Limit Mass, Volume Limit Power consumption, plan for dissipation Location without compromising the attitude dynamics Dimension reduction	Efficient design, material choice
Structural Parts	Reduce stress, improve stability weld-ability and reduce corrosion to environmental and operating conditions limit Thermal expansion and thermal inertia, limit warping over the operation period, Static, dynamic behaviour and match the needs of the launch vehicle, Homogeneity, affordability	Efficient design, material choice, suitable launch vehicle, extensive testing of the design using FEA software to ensure static and dynamic behaviour s in accordance of the mission needs
Operation coverage	Communication range at all times with ground station, Energy consumption, Human waste management, Emergency situation preparedness levels,	efficient design and mission planning
Maintenance	Cleanliness, safe handling of research equipment's, safe disposal of bi-products from experiments, Accessibility, limit EVA technical downtime etc	astronaut training on ground, use autonomous-AI-enabled probe, efficient and detailed design of ventilation system [17]

Design features of laboratory module shall be capable of fulfilling the requirements such as data collection through sensors etc. of various activities on-board. Validation in comparison with mathematical model or simulation results. Repeatability of on-board activities and readability of thus recorded data. Generic equipment to facilitate variety of experiments on-board. Risk tolerance through optimised design and mitigation through

training and precautions. In the case of human habitat, study of observability, training to work in space. Develop ability to adapt for “change of plan” for on-board activity, testing and so on. To design the habitat module with capacity to accommodate adequate number of crew numbers on-board for given mission task at all times.

The module design process is as represented in Fig. 2, follows initial design problem statement of bottom up approach. Based on which an preliminary design was sketched using first principles. CAD model was created and then simulation was performed with different factor of safety (FOS) for both flight and ground conditions. FOS for ground was taken to be at minimal 1.6 and for flight it was about 1.2. In parallel, detailed design calculation were made with material properties. The findings were compared with an optimised solution thus obtained from Finite Element Analysis (FEA) simulation. Design underwent iterations until satisfactory criterion was met thus, constituting the delta phase. When the two models both space and ground design agree to the mechanical constrains, the design is implemented.

2.2 Module 2014

Orbital module – 2014 (first designed in year 2014) is launched to an altitude of 450 kilometer circular orbit with 0 degree inclination. This altitude accounts for long orbital decay time, good connectivity with ground communications thus relatively non-critical conditions for station keeping requirements. The module is initially made of two sections which will be launched separately. The first section of the module is of 15 meter

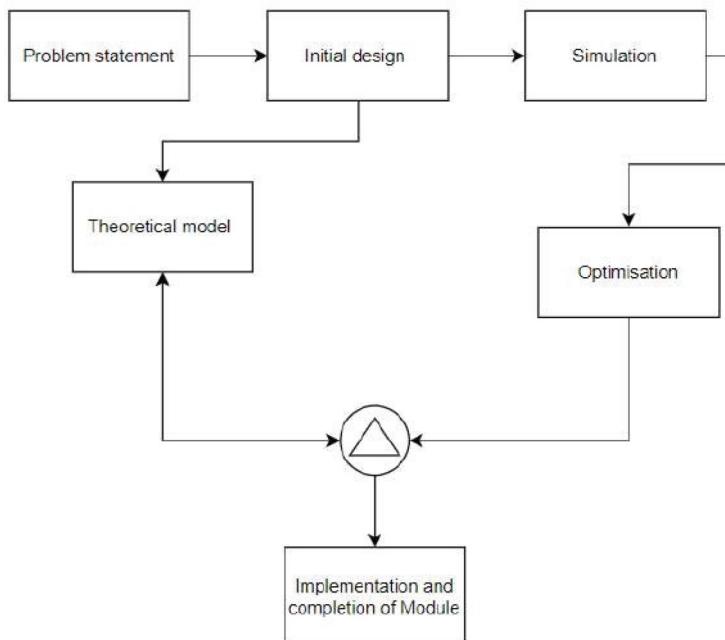


Fig. 2. Design iterative process of module

in length about the horizontal, ~7 m height from base along the vertical direction. The second module, 10.2 m in length and 7.7m in height with same cross-section as of first section. The design size of the modules are by far limited by cost, weight and fairing size of launch vehicle of the current day with possibilities for modifications. Early design concept is representation of the Module 2014 and same holds good for ground module – VSL and is presented in Figs. 3, 4, and 5 represents the individual design cross-sections of both orbital and ground modules. Key points considered while developing the design principles while formulating the solution is presented in Table 1 [8].

Both ground and orbital modules are realised by assembling two sections in T-shape configuration. Geometry of the assembled sections of module is presented here, with -x direction of co-ordinate is parallel to Earth surface shown in Fig. 6. The same reference holds good for both orbital and ground design. The geometry for the primary structure design is chosen to be similar to regular hexagonal prism with varying length of each module. The definition of the geometry provides additional flat-faces both internally and externally. Design in Figs. 3 and 6 is combined with planned functional logic which is presented in Fig. 7. In Fig. 7, functional logic with EVA ports for maintenance support, space walks and docking ports, for cargo/re-supply to the space lab for both research and crew is shown with the associated module.

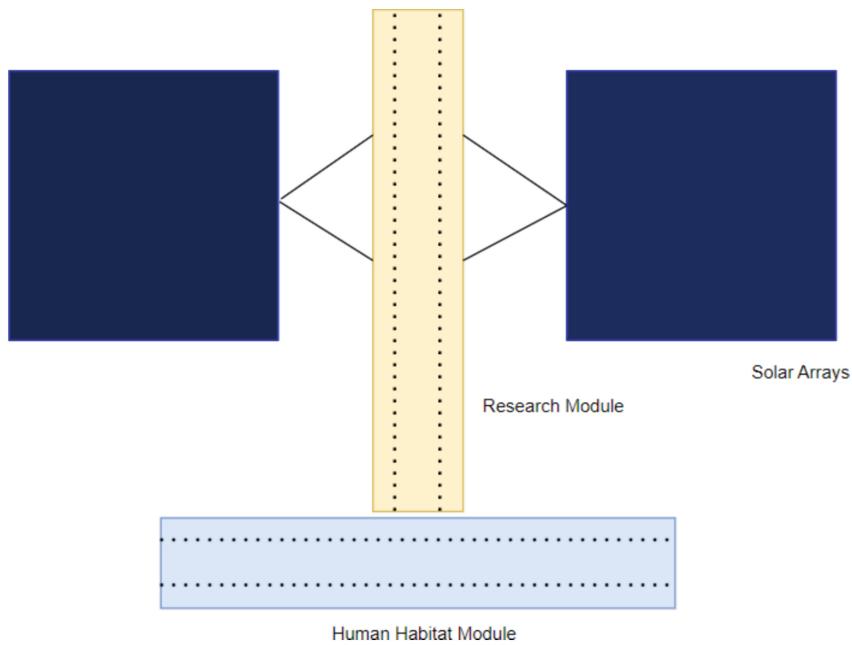


Fig. 3. Early design concept of Space Module of Module 2014/Visvesvarayya Space Lab

2.3 Sub-System Definition – Research Payloads

In this section, preliminary identification of sub-systems that are minimal to satisfy objective is listed. Detailed design of each sub-system is not included in this paper however, function of the same is proven by implementation in the ground module 1729/VSL. A total of five minimum payloads are part of the research module design definition. They satisfy basics of research mission,

1. Inert environment:- During previous missions, samples from asteroids etc. are returned to Earth through capsules and mostly 10 g or similar sample size. Towards end of mission, samples endure critical mission design sequences such as re-entry into the atmosphere. The risks associated with such events can be mitigated to large extent with sample collection at an orbital station. The sample can be studied and tested on-board. This option also prompt to increase the mass of sample to be collected because of lack of need of thermal shielding necessary during re-entry. The inert nature of the chamber enables on-board researcher to perform research activities safely. Able to study the specimen with respect to its fundamental physical properties. Able to derive test specimens from extra-terrestrial samples and also for biological experiments without contamination. The design is a cuboid with volume of approximately 2 m^3 and fitted with high density pump able to operate up to 6 bar. Hand interface is provided to the crew to modify the sample. Within the chamber volume, slots are available to store the samples etc. With availability of gases such as nitrogen, carbon di-oxide, Argon, Xenon and helium for test-chamber environment support. The chamber is operated at wide range of pressure from near vacuum few milli-bar to 5 bar pressure with unique choices for chamber gas, used to create inert environment for testing. The resulting design is represented Fig. 18 [19].
2. Storage bins:- It is well known that anything unattached in space tends to move without resistance, all the secondary equipment such as computer hardware, human needs such as food, water etc., samples used for experiments and on-board solid and liquid waste require a multiple sections of storage bin to keep them in place and also to avoid contamination with local environment inside the module. The same is represented by Fig. 21.

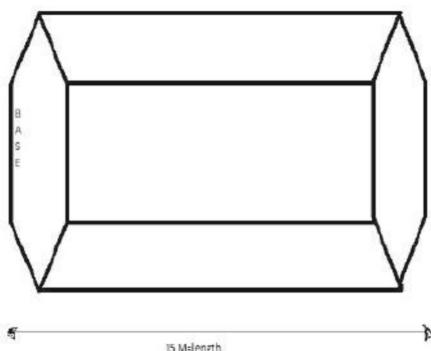


Fig. 4. 15 m hexagonal section for Research activity of T shaped module

3. Waste disposal system/micro gravity-toilet:- For human waste disposal as part of the human habitat. Liquid waste can be recycled with passive filtration system while all waste samples are returned to Earth for further study. The design employs principle of suction creation operating from the line drawn from ventilation pressure line with manipulation of valves. Suction is able to actively clear the solid wastes while liquid wastes are chemically neutralised and recycled. Design is emphasised to be familiar and simple with consideration of human comfort. The design familiar to ground based system are employed. The representation of which is made in Fig. 23.
4. Communication:- Vital to be able to communicate to the ground mission control center and also for tracking the space station while in orbit for station- keeping. The availability of continuous high-bandwidth communications to ground enables quick rescue in an emergency. The bi-directional nature of the existing communications enables an space station module to close iterative investigation and service loops, allow software reconfiguration, and support multiple scientists while one uses the on-board facility. Further, the availability of ever increasing communication features will enable real-time video teleconferencing options as part of daily operations to better create a virtual presence of scientists aboard the space station. Several antennas are placed on the space station for real-time data communication. The space station also provides a platform for Earth observation with synthetic aperture radar (SAR) systems in L and S band frequencies to study various land, sea and ice applications. The design of the same is presented in Fig. 20. [10]
5. On-board habitat:- apart from human habitat, several experiments can be conducted to grow and reap produce from plants while using recycled liquid waste on-board. Sleeping ca-coons, tread-mills for exercise to benefit health of crew on-board. Space suits are provided on-board suitable for extra-vehicular activity (EVA). EVA is a generic term which include both scheduled and unscheduled activity including critical tasks performed by on- board crew during emergencies. The configuration of the same is represented in Fig. 21. However, this particular area of development is best reserved for later stages of mission.

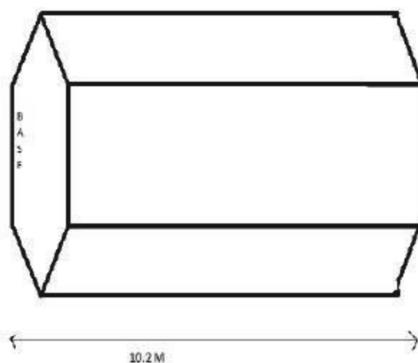


Fig. 5. 10.2 m hexagonal section for human habitat of T shaped module

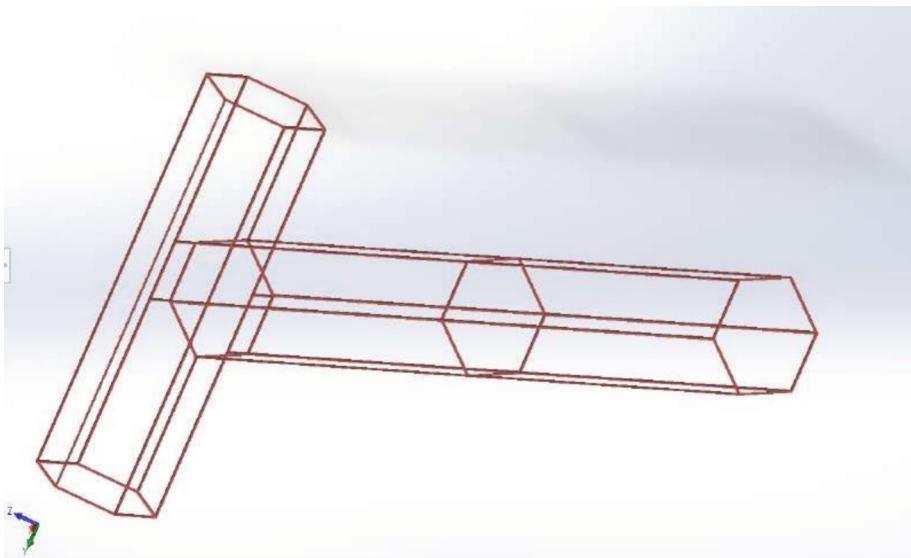


Fig. 6. Primary structure of assembled T shaped module to scale

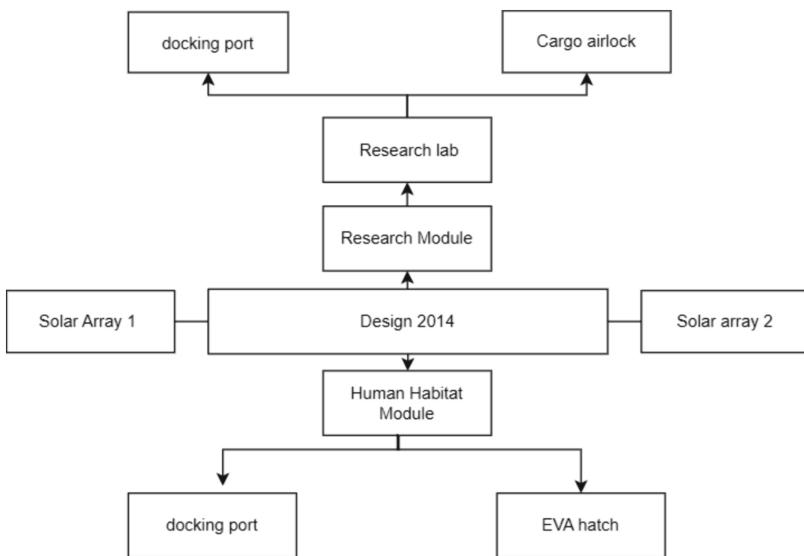


Fig. 7. Module 2014 functional design layout arrangement

Module 2014 In this section material properties defined for orbital module is presented along with its mass budget. The material properties utilised for simulating the primary structure assembly is presented in Table 2. The structure is made of aluminum 7075 alloy suitable for enduring mechanical loads at relatively low mass density. The panels are made of Carbon Fiber Reinforced Plastics (CFRP) for both structural strength and

thermal insulation properties. However, additional surface finishing is necessary in-order to avoid degradation of material in vacuum conditions. Multi-layered insulation (MLI) concept is adopted to insulate the module from surroundings and is applicable for orbital module only. Goal is to restrict the thermal conductivity through walls to 0.8 W/m-K or lower [21].

Table 2. Orbital Module 2014 material reference

Reference	Material	Young's Modulus	Optical Coating
Primary structure	Al-7075	70 GPa	None
Secondary Structure	Al-7075	70 GPa	None
Panels	CFRP + Metallic		
sheet/finishing	-	Thermal	
barrier Coating			
Insulation	MLI	-	-

In Table 2, thermal barrier coating refers to white paint with very low absorptivity and for corrosion protection. The position of the center of mass presented in Table 3 is obtained from CAD model presented in Fig. 8.

Table 3. Center of mass position for both module

Center of Mass position		
X = .61	Y = 3.75	Z = -2.66

Mass budget of preliminary mission for operation in low earth orbit is given in Table 4. The budget is dependent on the duration of the mission, research objectives and launcher capability. Power source for this design is solely solar power from the solar panels. Projected power budget of 7.5 kilo watt is allocated for the operation of two modules and then an additional kilo-watt for sub-systems.

Finite Element Analysis of Module 2014. For Module 2014, loads are applied to full scaled CAD assembly and is presented in Table 5. The study is performed on CAD assembly with mass blocks representing each subsystem of module 2014. Each subsystem is represented by point mass definition in AN- SYS student software. Maximum loads are experienced during flight. Loads are presented in Table 5 and is applied assuming acceleration due to gravity = 9.81 m/s at all times during the flight. Arbitrary force of 4.5 kilo-newton is also applied along the axis to simulate maximum loading conditions with higher margin to better understand the mechanical behaviour. Acceleration experienced during liftoff is also applied and is presented in Table 5. Additionally, modal analysis is performed using the same constrains applied previously for static structure

analysis. In Fig. 9 mesh generated consists of, total nodes = 11147, 5568 total elements for primary structure shown in Fig. 6 and 10930 nodes for assembled module shown in Fig. 10.

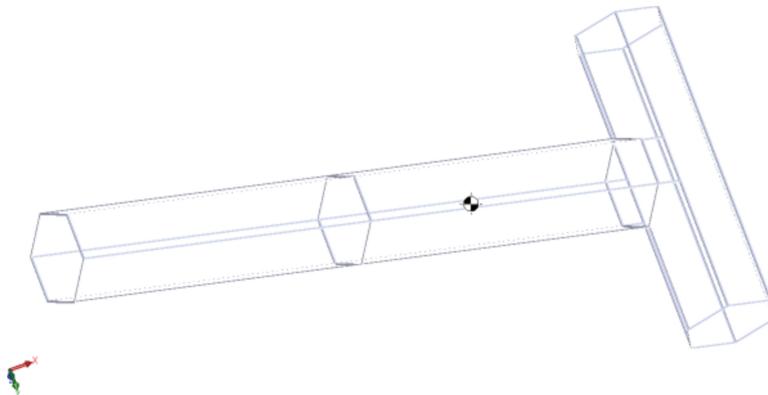


Fig. 8. Center of mass reference in CAD

Table 4. Preliminary mission mass budget in kilogram (approx) for module

Primary structure	3300
Secondary Structure, Walls	3750
Asteroid Impact protection Structure	2100
Batteries	500
Storage bin	210
Per Space suit	90
food etc	250
Research Payload	1850
Communication	750
Misc	1000

Table 5. Loading conditions for finite element analysis used on Module 2014

Load	Magnitude
Lateral Acceleration	$\pm 3.5 \text{ g}$
Axial Acceleration	$\pm 7 \text{ g}$
Axial force	4500 N

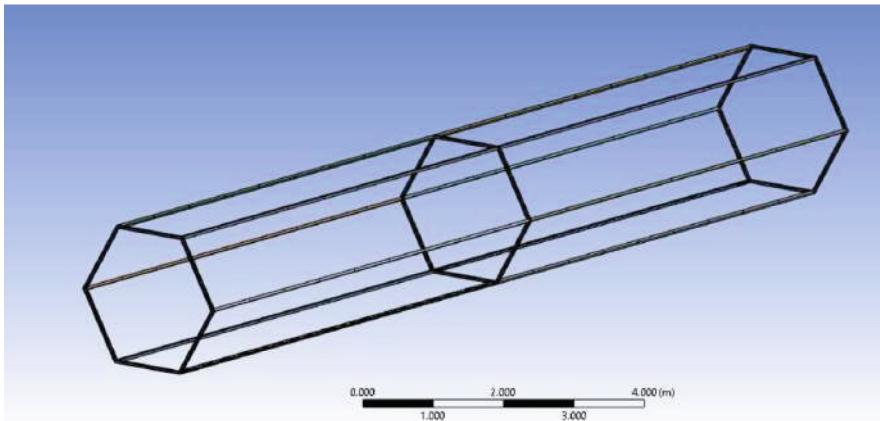


Fig. 9. Meshing of the primary structure of module 2014

At the preliminary design development stage, analysis is performed on the primary structure and fully assembled CAD model of the module. The mesh is presented in Fig. 10. Fixed support is applied on the base of the module. Fixed support represents the interface between the payload and the payload adapter in the launch vehicle.

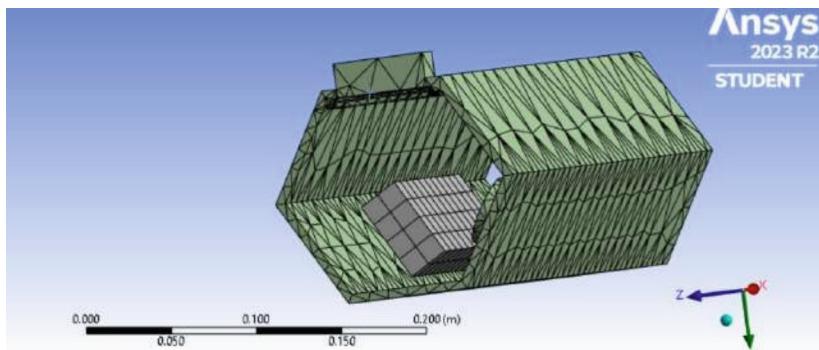


Fig. 10. Meshing of the assembled research section of Module 2014

FEA Results for Module 2014 As expected, maximum deformation is endured by the primary structure along its axis due to the load definition as shown in Fig. 11. Under the same loading constraints, static loading analysis performed on the assembly Module 2014, deformation is negligible while stress and strain values are greatly reduced as shown in Fig. 12. Modal analysis is performed on the assembly module 2014, results indicate that the primary mode is 446.3 Hz as shown in Fig. 13. The maximum deformation of 0.7 meter approximately. The preliminary results satisfy requirements for most launch vehicle currently capable of launching module 2014.

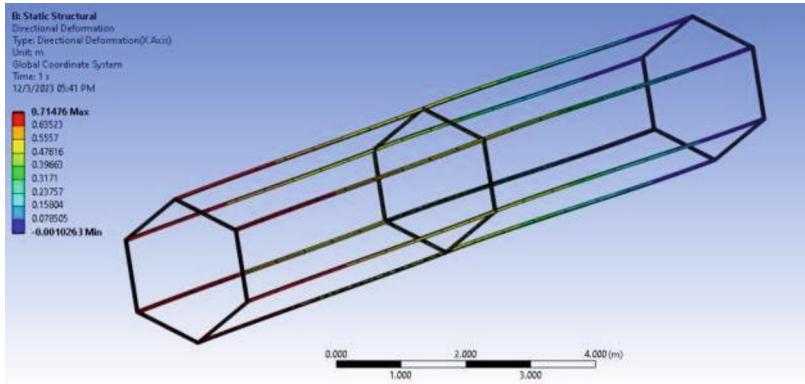


Fig. 11. Static structural result for axial deformation of the primary structure – module 2014

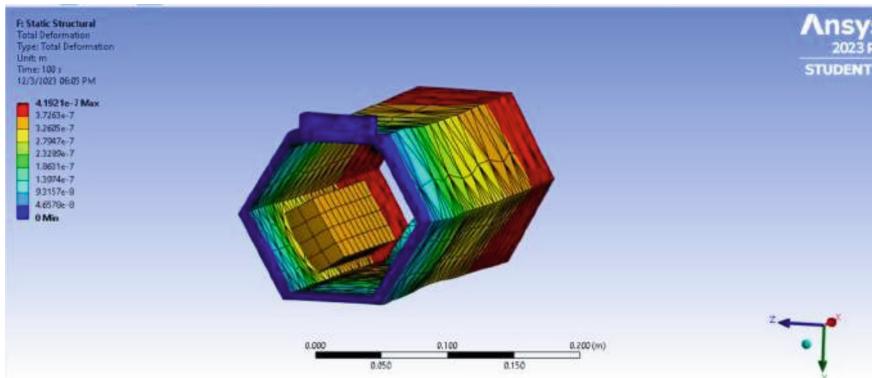


Fig. 12. Static structural result for axial deformation of the assembly- module 2014

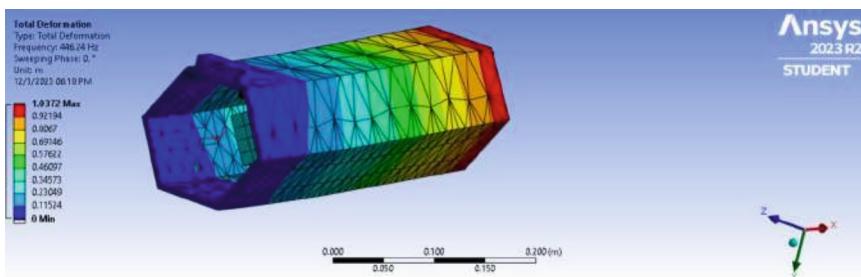


Fig. 13. Modal analysis result for first mode of deformation – module 2014

2.4 Module 1729/VSL

Life sized prototype based on the design definition and study conducted on Module 2014 was produced. In this case, the prototype is adapted for ground environment in Jaipur

city of Rajasthan, India with location at 26.8779° North, 75.6892 East. The module 1729 was realised and named to be known as visves-varayya space lab (VSL). Approximately 65 m of area is occupied by the prototype, satellite image of which is shown in Fig. 14. Motivation to build life sized model began during academic year 2014–15 with following goals :-



Fig. 14. Satellite image of Visvesvarayya Space Lab. Courtesy:- Google earth taken on 24-April-2023

1. To demonstrate exploration, human presence in space..
2. Conducting research to benefit technological progress on Earth.
3. Enable the experience of space station to serve as a construction platform for Lunar and Mars missions and beyond.
4. Learning how to construct large structures for low earth orbit, testing conditions for the same on Earth.
5. Learning how to operate in space, in-situ resource utilisation and acting as an engineering test-bed.
6. Learning how to operate and survive in space for long term, Conducting research to support future long-duration space missions.

Table 6. Module 1729 material reference for ground environment

Reference	Material	Young's Modulus	Optical Coating
Primary structure	Steel	210 GPa	None
Secondary Structure	Steel	210 GPa	None
Panels	Wood, steel sheet		

(continued)

Table 6. (*continued*)

Reference	Material	Young's Modulus	Optical Coating
Insulation	Fermacell + paint and white paint	-	Water repellent
	Cotton, open-cell insulation material		
Insulation	-	-	-

VSL was realised using the same primary structure design as shown in Fig. 6 and the same sub-system design listed in Sect. 2.3 with minor adaptations. Material used is listed in Table 6, steel is used instead of aluminum 7075 because of cost, environmental condition at the location to which the structure will be subjected to. In this case, there is relatively less stringent restrictions on maximum allowable design mass.



Fig. 15. Primary structure, welded with meshed-steel sheet – module 1729

Steel square beams are welded together to form hexagonal cross-section and two sections of which is welded in T-shape. The secondary structure comprises of sheets of steel square mesh which are welded to the primary structure. T section module with research section and human habitat module is distinctly assembled as shown in Figs. 15 and 16. Design of subsystems is based on the theme described in Sect. 2.3. Figure 17 shows the wooden casing (material adopted for ground conditions) of the vacuum chamber/inert environment, every side of the inner wall is provided with vinyl plastic sheet finishing which was recycled in the process, one of the insulation materials used as part of ground MLI is also clearly visible in the Fig. 17. The ground MLI is based on concept of MLI used for space application however in this case, open cell foam, poly vinyl sheets, steel mesh and wood is used in layers. The same is mentioned in Table 6.



Fig. 16. Primary structure and secondary support structure of research section under construction – module 1729

The wooden casing of inert chamber was then provided with a layer of vinyl plastic sheet finishing. Acrylic sheet was cut to meet the dimensions of the casing and was fixed in place with place for hand-glove holdings. Care was taken to ensure the chamber was air tight. One of the hand gloves could be removed in order to place or replace chemical species inside the inert chamber. Piping was connected to ventilation from the compressor situated outside the module and on the steel platform made for the module. The same circuit would be present in Module 2014 only that the pressurised gas would be inert gas or depressurised to create near vacuum conditions in the chamber as shown in Fig. 18. Communication system is setup using sound system, microphone for input and speaker as output via the dish between the ground i.e., outside the module and the same is done inside the human habitat section of the module represented by Fig. 19.

The dish fitted on top of the module acts as an exchange device between the ground and the astronaut on-board the human habitat. The dish is mounted on custom welded tripod joint with ability to rotate 360° parallel to ground, at 35° about the vertical as shown in the Fig. 20.

Further, space suit was installed with modifications to clean-room suits to represent and explain the need for extra vehicular activity or commonly known as spacewalk. Then, the complexity of designing, working and testing of the space suits are represented. Therefore, entry view of the research section of the module 1729 is represented in Fig. 21.

One of the hulls in the hexagonal cross-section of human habitat section was converted into viewing deck with 3D earth picture as shown in Fig. 22. This particular section was installed to promote philosophical thought of oneness among diversity. Waste collection platform (ground reference) or Zero-g toilet was converted to pressure operated waste disposal system which can work in micro-gravity environment and in any orientation. The pneumatic circuit were drawn from parallel line leading from the compressor system, shown in Fig. 23.



Fig. 17. Primary structure and secondary support structure with sub-system casing, insulation, finishing material during construction – module 1729



Fig. 18. Inert chamber demonstrator – module 1729

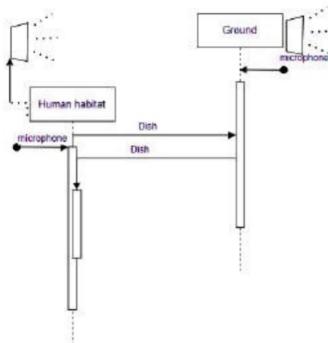


Fig. 19. Communication acoustic logic – module 1729



Fig. 20. Communication acoustic dish with custom welded tripod demonstrator – module 1729

Functionality of the above described sub-systems were performed twice and on special occasions until year 2019. The primary structure designed, assembly welded in year 2014–15, showed no warping until 2018–19. Gas line from the air pump situated outside the module directed fresh air through the pipes along the length of the module, allowing for the module to be completely shut during operation. The pressure was maintained slightly higher than atmospheric to keep away the dust entering from outside. Design and assembly of sub-systems and structure were parallelly started in April 2015, while the entire Phase – I was completed by end of July 2015 and the Phase 2 was completed by beginning of 2017, final prototype is shown in Fig. 24 and at night in Fig. 25. During phase - 2, metallic sheets were replaced with Fermacell at non-critical areas identified from results of finite element analysis to reduce weight significantly.



Fig. 21. Research section demonstrator entry view - module 1729



Fig. 22. 3D Earth view demonstrator - module 1729



Fig. 23. Zero-g toilet / waste collection platform demonstrator - module 1729

Module 1729 when finished was able to host 5 fully grown adults inside at a time. Approximate mass (conservative) of the whole module 1729 is presented in Table 7.

Table 7. Conservative mass budget in kilogram (approx) for each operational Module 1729

Primary structure	500
Secondary Structure, Walls	400
Asteroid Impact protection Structure	0.0
Batteries	20.0
Storage bin	110
Per Space suit	80
food etc	2.0
Research Payload	850
Communication	150
Misc	500



Fig. 24. Top view of module design 1729 (Visvesvarayya Space Lab) by end of Phase 2 (taken during April 2018)



Fig. 25. Side view of module 1729 by end of Phase 2 or final version of visvesvarayya space module taken in year 2019

3 Conclusion

Preliminary setup for space station demonstrator was realised through bottom up approach. The paper illustrates the preliminary design, overview of sub-systems, FEM structure and modal analysis and implementation to build a life size demonstrator. Initially, a design concept is assessed based on engineering judgement, next with finite element method for static deformation, stress levels for the chosen material combination at maximum loading conditions. Modal analysis was performed with design iterations to find the correct configuration which could satisfy the launch vehicle payload requirements. The same is then applied to demonstrator utilizing the engineering principles studied during course. Constructed in two phases starting from year 2014/15 to realise a demonstrator like prototype which is operational on ground conditions. Additionally, an experience of designing, assessing, assembling and overcoming the challenges faced in realising an space station. Design of primary and secondary structure, along with thermal insulation concept during phase – 1 was successfully implemented. Then, implementation of all the sub-systems envisioned with advanced self-made mechanical solutions to various engineering problems during installation. During Phase-2, selective replacement of relatively heavier metallic sheets for light weight fermacell material resulting in significant reduction in overall weight with- out any noticeable change in performance of the structure and relative improvement in thermal behaviour. The demonstrator was opened to visitors, at a time it could house 5 fully grown adults, operate sub-systems and a 3D glass for an in-space earth viewing experience. The demonstrator strove towards motivating public interest in realising country own first space station and associated benefits. However, demonstrator requires constant maintenance due to the location which experiences extreme summers, cold winters and monsoon. Also, structure and material is prone to high degradation due to fine sand dust from dessert carried by wind.

Future work include detailed design of each sub-systems with individual parts that are space qualified. Detailed design of impact protection shield, additional sub-systems such as robotic arms, in-space manufacturing & assembly. Update mass and power budget to accurately account for the final design and the materials considered. Also, to expand the knowledge base to other branches of engineering associated with the operation of space station.

Acknowledgments. We thank Chairman BOG RS Tomar (2014–2019) , Rajasthan Institute of Engineering and Technology (RIET), Jaipur, India for initiating support and full funding towards designing and realisation of space station module at RIET Jaipur campus starting from April 2015 – March 2018. We thank Prof (Dr) Amit Kumar Bairwa, Prof (Dr) Vinod Singh Yadav, Prof Raghav S Dhaker, Prof Mahesh Jangid, Late Prof Bhanu Singh for spending hours in enabling us to come up with the space station module and also teaching the team finite element modeling with special sessions and attention. Former Director, Dr Surendra Kumar, Former Dean – Dr Kapil sharma for ideas and support in the launch of VSL. Late Dr Digamber Singh (Former Cabinet Minister in the Government of Rajasthan from Bharatpur, Rajasthan) inaugurated the Module 1729 – Visvesvarayya Space Lab during March 2017. We thank Rakesh, Vimal and team for supporting with machining, highly skilled welding of hexagonal section without fault and other mechanic assistance. Extend our gratitude to Mahaveer and Naresh from Civil Engineering

department for their support and assistance. Mr Madan, G S chaudhary for enabling us to get all the required raw materials on time. Mr Saxena for aiding us in managing the budget. Security staff Daddu and company for taking good care of team during night time shifts with unlimited tea, helped design many solutions. And finally, hearty gratitude for everyone who protected our work on campus in our absence.

I also, thank my teacher Late Raghavendra Panduranga Rao from Department of Biology and Zoology of Vijaya Pre-Univeristy College of Jaynagar 4th BLK in Bengaluru, Karnataka for motivating, inspiring me to include Plants, to encourage biotechnology research pathways in the demonstrator.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article. Article mainly aims towards remembrance of probably, the first funded (private funding included) - student only designed, built space station prototype in India.

References

1. Design of the Space Station Habitable Modules Gary H. Kitmacher 53rd International Astronautical Congress The World Space Congress – 2002 10-19 October 2002/Houston, Texas IAC-02-IAA.8.2.04
2. International Space Station Transition Report <https://www.nasa.gov/wpcontent/uploads/2015/01/2022isstransitionreport-finaltagged.pdf>
3. <https://www.eoportal.org/other-space-activities/iss-transition>
4. Uri, J.J.: NASA utilization of ISS -past, present and future. Micrograv. Sci. Technol. **19**(5–6), 37–41 (2007). <https://doi.org/10.1007/BF02919450>
5. <https://www.cbsnews.com/news/china-launches-3-man-crew-to-tiangong-space-station/#textThe%20Chinese%20space%20station%20is,airlock%20and%20multiple%20docking%20port>
6. Kulu, E.: Small Launchers – 2021 Industry Survey and Market Analysis. 7 (2021).
7. <https://www.nasa.gov/missions/station/iss-research/ad-astra-future-plans-for-the-international-space-station/>
8. Andrews, S., Berthoud, L.: Characterising satellite aerodynamics in Very Low Earth Orbit inclusive of ion thruster plume-thermosphere/ionosphere interactions. Acta Astronautica **170**, 386–396 (2020). <https://doi.org/10.1016/j.actaastro.2019.12.034>
9. <https://www.bauhaus.info/gipsfaserplatten/fermacell-gipsfaserplatte-1-mann-platte/p/13885507?cidSSAGoo80603011028485488778gadsource1gclidCj0KCQiA67CrBhC1ARIsACKAa8TJ28iGtikJdDyjHL9RCy8YGzOrW8zVWcqHazktTYmuV6TZL-FpqsaAg5EALwwcB>
10. Design Principles for the Development of Space Technology Maturation Laboratories Aboard the International Space Station, Alvar Saenz-Otero , <https://www.mit.edu/ alvarso/thesis-phd/ThesisBook.pdf>
11. <https://www.sfu.ca/mbahrami/ENSC%20388/Notes/Staedy%20Conduction%20Heat%20Transfer.pdf>
12. <https://earth-planets-space.springeropen.com/articles/https://doi.org/10.1186/s40623-021-01363xabbreviations>
13. <https://ntrs.nasa.gov/api/citations/20110015359/downloads/JSC-65829-RevALoadsandDynamicsRequirementsforSFHardwarefinalsigned.pdf>
14. <https://athena.ecs.csus.edu/grandajj/ME296M/Paper53AAIAISS1331sq.pdf>
15. <https://apps.dtic.mil/sti/pdfs/ADA589762.pdf>
16. <https://e-archivo.uc3m.es/bitstream/handle/10016/30436/TFGGonzaloMontesinoValle2019.pdf?sequence1>

17. Lakshminarayana, S., Bhaskar Thakare, S., Duddukuru, K.V.: On-orbit, non-destructive surface surveillance and inspection with convolution neural network. In: International Conference on Cyber Warfare, Security and Space Research, pp. 283–293. Springer International Publishing, Cham (2021). https://doi.org/10.1007/978-3-031-15784-4_22
18. <https://www.space.com/19607-skylab.html>
19. <https://www.isro.gov.in/sarabhaformer.html>
20. OSIRIS-REx, Returning the Asteroid Sample. <https://ntrs.nasa.gov/api/citations/2015000809/downloads/20150000809.pdf>
21. Gilmore David, G.: Spacecraft Thermal Control Handbook. Volume I Fundamental Technologies, 2nd ed. El Segundo Calif. Aerospace Press; American Institute of Aeronautics and Astronautics, Reston VA (2002)



Hybrid Deep Driven Cross Industry Sentiment Analysis Model for Netizen's Behavioral Characterization

Santhosh Priya and R. Kalaiarasi^(✉)

Tamil Nadu Open University, Chennai, Tamil Nadu, India
drrkalaiarasi@tnou.ac.in

Abstract. Sentiment analysis has grown into an essential tool for businesses, government agencies, and a variety of other industries to help them make adaptive decisions. This is in line with the growing number of internet users, or “netizens,” and their online activities, which include reviews and comments. Sentiment analysis is a natural language processing (NLP) technology that has been assisting industries in defining, redefining, or innovating their goods and services by using user or customer reactions and comments. Despite these importance, the method of online sentiment analysis or opinion mining is complicated due to the variety of textual presentation, data heterogeneity, and cross-platform content nature. Opinion mining techniques now in use are designed for stand-alone tasks, such as sentiment analysis tailored to a particular sector, product, individual, or even netizen behaviour. It indicates that a model created for one issue cannot be used to another. This work proposes a robust hybrid deep-driven cross-industry sentiment analysis (HD-CISA) model for netizen digital behaviour analysis and associated opinion mining, taking that into consideration as motivation. First, the HD-CISA processes text inputs gathered from four distinct industries: the healthcare, e-Commerce, social media, and hotel sectors. This is done using data adaptive pre-processing. Concatenated and projected as input to the Bi-LSTM deep network for contextual feature extraction, the pre-processed data were used for Word2Vec word-embedding. In order to establish long-term reliance across the input phrases, the Bi-GRU network subsequently learned the extracted characteristics. In order to create a composite feature vector that was learned and categorised at the Softmax layer for sentiment prediction, the local contextual features of the Bi-LSTM and the global features of the Bi-GRU were finally merged. The outcomes of the simulation produced results that were greater than any known method till now accuracy of 97.91%, precision of 97.73%, recall of 97.48%, and F-measure of 97.60%.

Keywords: Sentiment Analysis · cross-industry model · Hybrid Deep Learning · Semantic Word Embedding · Bi-LSTM (Long Short Term Memories) · Bi-GRU (gated recurrent units) · ABSA (Aspect based sentiment analysis) · HD-CISA (Hybrid Deep Model driven cross Industry Sentiment Analysis)

1 Introduction

The high pace rise in software technology, internet and decentralized computation has broadened the horizon for the applications exploiting data for adaptive and situational awareness-based decision making. In this reference, sentiment analysis has emerged as a vital technology that exploits internet user's (say, netizens) online behavior, including reviews, social media responses, feedback etc. to understand their opinion so as to make optimal decisions [1, 2]. The significance of sentiment analysis can be understood by the fact that now days almost all industries, socio-political entities etc. are using it to understand their audience's opinion and sentiment that helps them making optimal decisions [3]. Sentiment analysis or netizen's online behavioral characterization has become a fundamental part of business systems, especially to serve numerous purposes including corporate business decisions, event-based designs and redesign services and/or decisions, product/service recommendation engines, political analytics etc. [1–3]. Despite such significances, the entirety reliance on data makes sentiment analysis a challenging task, especially over uncertain online media where the users can have written their response in the different languages, with the different word-structures, word-composition etc. [4]. The content heterogeneity and non-linearity makes NLP-based sentiment analysis a challenging task [4–6]. In the past, the major sentiment analysis models are designed towards single domain problem, such as opinion mining for e-Commerce products, social media reviews, opinion or review towards healthcare or hospitality sector etc. [2, 3]. In these solutions, feature learning is often done for a single text data corpus and hence the allied latent information remains intact with the same subject matter and therefore can't be optimal over other data type or other subject matter [7]. This is mainly because of diversity of sentence composition, vocabulary, language differences etc. It infers that a sentiment analysis solution designed for a specific problem can be limited for another data or subject matter (say, industry reviews). On the other hand, contemporarily the varied online web-platforms or Mobile applications serving e-Commerce, social media etc. provide different user-friendly custom tools such as different fonts, text alignment, emoji etc. In this case, the content heterogeneity across industries or cross-platform can limit the ability of a solution which is specially designed to address a specific problem on a single platform [1–5, 7]. To address this problem, training a sentiment analysis model over cross-industry datasets can enable a superior solution, where learning over multiple types of data concurrently can strengthen the cognitive behavior of the model to make reliable opinion mining [2, 3, 8]. Unlike classical opinion mining approaches, where a solution is trained over a specific data, training over cross-industry text corpus can enable a scalable and cost-effective opinion mining solution for industry [8]. Noticeably, here we define "cross-industry" as the ability of the sentiment analysis solution to perform opinion mining for the different subject matters (for the different industries) without demanding distinct textual learning, training or processing. In this manner, a standalone sentiment analysis model can be used for different industries like e-Commerce, healthcare, finance, politics, education, etc. to make netizen's behavioral characterization and allied dynamic decision making. It can be considered as the key driving force behind this study.

To cope up with opinion mining demands, varied NLP solutions are proposed in the past where user's reviews have been exploited to yield respective opinion, inclination,

preference etc. [9]. However, their efficiency primarily depends on data, feature and learning ecosystems [3, 10, 11]. Literatures divide sentiment analysis problem based on the data, and classifies it as document-based, sentence-based and word-based methods. In fact, this classification depends on what way (volumetric) textual features are obtained and trained by the models. In document-based method, it learns over the complete set of sentences or document to predict a conclusive opinion [12]. The sentence-based opinion mining methods process each sentence or netizen's review as input and learns opinion for each sentence. This method often undergoes ambiguity, as in a single sentence there can be multiple inferences or polarity words, where the classical dictionary-(or lexicon)-based methods can fail in providing accurate opinion [12, 13]. For instance, "though, the government is doing good; yet, the economic policy seems inferior". Those opinion mining methods applying target labels like "good" as an eventual polarity variable would classify this sentence as "Positive"; however, the later part of the sentence (i.e., the economic policy seems inferior) indicate unsatisfaction. The word-level sentiment analysis method applies latent information per word to make eventual polarity decision; yet, retrieving optimal set of embedding vectors over input corpus with minimum computational cost remains a bottleneck [13]. Moreover, training a model with word-level embedding, especially over a large data can be a mammoth issue and hence can confine the efficiency [14, 15]. In the past, majority of the sentiment analysis models are designed based on machine learning [12]. These methods either use the frequency of word's co-occurrence or lexicon information or sometime embedding matrix for sentiment analysis [12]. The supervised methods perform training over the labelled data, while unsupervised learning methods involve training over the information which is neither labelled nor classified [47]. Despite numerous efforts, most of these methods are applied for in-domain sentiment analysis and hence it doesn't guarantee whether the trained model be effective to identify opinion of out-domain reviews. Though, word-embedding methods like Word2Vec, GloVe, continuous bag of word (CBOW), term frequency and inverse document frequency (TF-IDF) have helped machine learning methods exhibiting better accuracy; yet, the lack of long-term dependency forces them to exhibit poor accuracy [2, 3, 16]. Recently, the authors suggested feature-level ensemble to improve accuracy; however, can be found suitable only for in-domain polarity analysis [16]. Studies suggest applying both local contextual features as well as long-term dependency to achieve higher reliability [16, 17] for out-domain opinion mining task. Learning emotions can give more appropriate sentiment analysis results than merely training over the shallow lexicons or word-level embedded features [15]. To achieve it, deep learning networks have shown better performance [17–19]. To achieve better performance, these methods require high-level features and suitable training set. Additionally, these methods might suffer overfitting and convergence issues, whose severity increases over non-linear data patterns, like cross-industry reviews. To exploit local and global features together different methods are proposed including convolutional neural network (CNN) for local contextual feature, recurrent neural network (RNN) methods like LSTM, Bi-LSTM, gated recurrent unit (GRU) [18, 19]. The hybrid methods have applied CNN for local feature extraction, while Bi-LSTM or allied attention-based methods for long-term dependency. Despite fast convergence, CNN-based methods don't have distinct target's modelling

via context-specific representations. Moreover, it merely extracts local contextual features and hence lacks long-term dependency. Though, a few literatures indicate that the strategic amalgamation of LSTM, Bi-LSTM, and GRU can achieve better performance because of their ability to retain both contextual as well as long-term dependency features. It can make aspect-based sentiment analysis an easier and more reliable solution for netizen's behavioral analysis [20]. In sync with the targeted cross-industry opinion mining or netizen's behavior analysis, ABSA seems to be a viable approach as it provides more fine-grained information by means of aspects level analysis and decision. To improve ABSA performance, especially over cross-corpus learning unlike using classical CNN or LSTM, more rigorous hidden state feature extraction can be applied, for instance Bidirectional LSTM (Bi-LSTM). The ability to extract both forward and backward features can enable Bi-LSTM more efficient for local feature learning. In addition to the aforesaid local feature learning, the retention of long-term dependency can further improve efficiency [18–20].

In this paper, a hybrid deep driven cross-industry sentiment analysis model is developed for netizen's online behavioral analysis. Being a cross-industry solution, the training is done over different text inputs encompassing reviews from the hospitality sector, E-commerce, healthcare sector, etc. To ensure optimality of input features, pre-processing was done with missing value removal, Unicode normalization, removal of emoji, web-link, HTML text, words with numeric values, non-word characters, stop-words. It also performs lower-case conversion and lemmatization, followed by tokenization and padding. Post-tokenization to improve feature's semantic learning ability, Word2Vec word-embedding was performed that provided uniform embedding matrix for further learning and classification. The embedding matrix was passed to the proposed hybrid deep model encompassing Bi-LSTM and Bi-GRU, where the first extracted contextual feature enriched with inter-element associations, while the second was applied for long-term dependency learning. Finally, the extracted features from Bi-LSTM and Bi-GRU were fused together to yield a composite feature vector for learning and classification. The classification over Softmax classifier with ADAM learning and binary cross-entropy loss function enabled HD-CISA achieving sentiment classification accuracy of 97.91%, precision 97.73%, recall 97.48% and F-measure 97.60%, which was higher than any known approach so far.

The other sections are divided as follows. Section 2 presents the related work, Sect. 3 presents overall proposed HD-CISA model, while the simulation results and analysis are given in Sect. 4. Section 5 discusses overall conclusion, which is followed by references at the end of manuscript.

2 Related Work

Opinion mining or netizen's behavior characterization methods can broadly be classified into three key types; machine learning, deep learning and aspect learning based methods. Some of the recent literatures pertaining to the sentiment analysis are discussed in this section.

A. Machine Learning

Kamps et al. [21] used lexicon information WordNet for sentiment analysis in English texts. Later HowNet [22] was designed with improved lexicon information; however, it could not guarantee optimality of the solution for out-domain tasks. Pang and Lee [23] applied maximum entropy information [24] which was trained over different machine learning methods like Naïve Bayes and support vector machine (SVM) for sentence comparison in sentiment analysis. Similarly, Poirier et al. [25] used naive Bayesian classifier for sentence polarity estimation. Naz et al. [26] obtained n-gram features from the input text corpus, which was learnt over SVM to perform Twitter sentiment analysis. The authors [27] have found that the opinion mining can be more significant if trained over sufficiently large review data with higher intrinsic information [3, 5, 12, 28]. In [29], RapidMiner and KNIME tools were applied to understand consumer behavior and their opinion. The authors inferred that the use of Random Forest classifier can yield accuracy up to 97.06%, which was superior than other machine learning methods like Random Tree, k-Nearest neighbor, Naïve Bayes etc. SentiWordNetlexicon was applied in [30] for sentiment analysis by using SVM classifier. To improve efficiency, the authors applied TF-IDF features from the input text to perform opinion mining. In [31], latent Dirichlet allocation (LDA) was applied for feature extraction from the input reviews. Later, Frequency anti-clustering frequency (TF-ICF) method was applied for text similarity estimation over hotel review dataset. LSTM was trained over TF-ICF feature to perform sentiment classification. Kontopoulos et al. [32] inferred that unlike classical glossary-based methods [31], the ontology method with (hidden) latent information can give better sentiment classification result.

B. Deep Learning

Kim [33] used CNN with single Max-pooling per convolution layer to perform sentence-level feature extraction and training over Fully connected layer for sentiment analysis. Yet, it failed in retaining long-term dependency which could have improved overall efficiency. Zhang et al. [34] used CNN for letter-level feature extraction by applying six convolutional layers and three fully connected layers for sentiment analysis. Xiao et al. [35] performed sentiment analysis over hotel review text that resulted 92.58%. Cheng et al. [36] applied textual hierarchical structure for sentiment analysis by applying CNN [33, 37] and attention network over Chinese text. Yet, it could apply merely the contextual information to perform opining mining. Kalchbrenner et al. [38] used a wide convolution model by replacing max-pooling layer by k-Max layer of CNN for feature estimation, which was later applied for sentiment classification. Yin and Schütze [39] applied multi-channel CNN of varied corpus sizes for text sentiment classification. Tang et al. [40] used two LSTMs to retrieve sentiment features, which were later trained over Softmax layer to perform opinion prediction. Ren et al. [41] used LSTM with Bi-LSTM model for twitter short text sentiment analysis. In [42], the aforesaid deep networks were applied for multi-task learning [43–45] for sentiment prediction. Yousif et al. [48] designed a fully-shared multi-task learning model for sentiment classification. Lu et al. [49] used multi-task learning with variational auto-encoder (VAE) for sentiment classification. Despite their (i.e., CNN, RNN, etc.), the state of art methods often undergoes vanishing affect and hence require other tools like attention module, LSTM and GRU for better performance. Tang et al. [50] used CNN/LSTM for local (single sentence) feature

representation. Additionally, it applied gated RNN to encode the internal or semantic relationship between sentences to perform chapter-level sentiment classification [50]. Wang et al. [51] applied Bi-RNN to retrieve word-level contextual information for sentiment classification. Zhou et al. [52] designed C-LSTM by using CNN for local text feature extraction and LSTM as pooling layer to perform sentiment classification. Ruder et al. [54] applied hierarchical Bi-LSTM for sentiment classification. Rao et al. [55] designed a two-layer LSTM to perform document-level sentiment analysis. Sachin et al. [56] used Bi-LSTM to perform sentiment classification over Amazon product reviews. Wang et al. [57] used single layer CNN with one-layer RNN where each models extracted distinct features for final learning and prediction. Zhou et al. [58] combined Bi-LSTM with CNN along with 2-dimensional Max-Pooling to perform sentiment classification. Zhang et al. [59] designed CNN-LSTM for sentiment analysis in Twitter text. Similarly, a multi-channel CNN-LSTM model was designed in [60] for Twitter text sentiment analysis. Sun et al. [61] performed Tibetan micro-blog's sentiment prediction by amalgamating CNN and LSTM networks. Zhang et al. [62] applied Convolution-GRU (CGRU) over Twitter hate comments for sentiment analysis. Abd El-Jawad et al. [63] used CNN-RNN for text sentiment analysis. To alleviate the issue of vanishing affect, Yang et al. [64] amalgamated Bi-RNN with attention method for text sentiment classification. Wang et al. [65] on the other hand proposed a multi-layered attention network with CNN to perform sentiment classification. Wang et al. [66] applied attention network with LSTM for ABSA. Yet, it failed in retaining long-term dependency feature which could have helped achieving superior results [67]. Cheng et al. [68] designed a Hierarchical Attention (HEAT) network for ABSA. The multi-attention network was designed by Han et al. [69] for sentiment classification over input aspect words. Gao et al. [70] designed collaborative extraction hierarchical attention network to retain aspect features for ABSA sentiment classification. A multi-channel convolution and Bi-GRU model was designed in [71] for sentiment analysis. They applied attention model with Bi-GRU to retain and learn features with high impact on the sentiment polarity. Socher et al. [72, 73] used tree-structured LSTM for more resource efficient semantic information towards sentiment classification. Cho et al. [74] found that the use of the GRU can achieve more efficient solution than the classical LSTM due to its fewer parameter demands. It can extract more significant global features to improve learning and hence classification. Rehman et al. [75] designed CNN-LSTM to perform sentiment analysis in IMDB and Amazon movie review datasets. Recently, Wang et al. [76] designed a hybrid deep model encompassing RNN and capsule network for sentiment classification. Bahdanau et al. [77] designed a capsule network model by applying multi-head attention to improve emotion learning over complex language knowledge. Yet, it found that the amalgamation of CNN and Bi-GRU can enable more efficient solution towards sentiment analysis. Once extracting local text feature and global semantic feature, it applied Max-pooling to fuse features. Subsequently, the attention model was combined to generate emotional capsules for further learning and (sentiment) prediction. Jianqiang et al. [78] observed that the concept of feature level ensemble encompassing word embeddings like global vector for word-representation (GloVe) and n-gram features can improve sentiment scores[79]. Xiao and Zhou [80] applied CNN and GRU to perform ABSA prediction, where CNN was applied

for local text feature extraction, while GRU helped obtaining long-term dependency to perform sentiment analysis over a Chinese hotel review dataset.

3 System Model

This research work contributed a robust hybrid deep model driven sentiment analysis model (HD-CISA) for cross-industry opinion mining. As the name indicates, this work emphasizes on exploiting multiple deep learning methods to learn both local as well as global (say, semantic) features to achieve a cross-industry opinion mining solution for business intelligence. Here, the key hypothesis that exploiting semantic features from cross-industry text reviews and training over corresponding local (say, contextual) as well as global features (say, long-term dependency amongst the words within the sentences) can strengthen the opinion mining approach to yield a universal solution for netizen's online behavior characterization [76, 81–91]. Recalling the fact that the text reviews can differ in one industry to another and hence training any sentiment analysis approach on domain specific or specific text reviews can be applicable only for that standalone industry type or related opinion dataset. In other words, training a model over one dataset can't be applicable for another data or review types and hence it would keep industry often engaged in an iterative development and redevelopment phase to accommodate new dataset(s). To alleviate this problem, we designed HD-CISA that can serve multiple industry types, including hospitality, healthcare, E-commerce, political decision system etc. The core behind this efficiency is the emotion-feature learning capability over the different cross-industry input text datasets. Moreover, to achieve higher accuracy and reliability, the proposed HD-CISA model works on both feature optimization as well as learning method. To serve a robust opinion mining solution, training a sentiment analysis tool is required with both local features (i.e., the features extracted over each word or sentence, individually) as well as global features (i.e., the feature obtained towards input's long-term dependency or inter-words relationship-based features). In this reference, HD-CISA applies two well-known deep networks including Bi-LSTM and Bi-GRU which operate for local contextual feature extraction and global feature extraction, correspondingly over each dataset. Functionally, at first different input datasets pertaining to the different industries are collected, which are processed for pre-processing so as to retain a sufficiently large input corpus for further feature extraction and learning. Once performing pre-processing, the input datasets were processed for word-embedding to generate embedded feature representation for each word in the input dataset. Considering the role of semantic features towards sentiment analysis, we applied Word2Vec word embedding method that transformed input words (in the text's sentences) into equivalent unique numeric presentation. This embedded data was obtained for all input samples including Amazon E-commerce product review, Covid-19 lockdown review, Reddit review, Airline service review. Thus, obtaining the embedded matrix for each input dataset, it was horizontally concatenated and was passed to the Bi-LSTM (Fig. 1). Bi-LSTM was designed with an input layer, two CONV layers and one average pooling layer. The extracted contextual features were concatenated at the average pooling layer (APL) for dual tasks, first as feature for learning and second as input to the Bi-GRU for long-term dependency analysis. Unlike major existing [80–91],

where the authors have applied the extracted features either from CNN (local feature-based training) or LSTM or even Bi-LSTM for learning and classification, to improve long-term dependency learning and allied global feature learning we passed the extracted feature at APL layer to the Bi-GRU, which has proven its robustness towards long-term dependency learning ability. Thus, the output features of B-LSTM were passed to the Bi-GRU for long-term dependency learning. Completing Bi-GRU learning and retrieving the final concatenated features, we fused both Bi-LSTM features as well as BI-GRU features at the Global Max Pooling (GMP) layer (Fig. 1). The GMP features were subsequently processed for learning and classification at the Softmax layer by using cross-entropy loss function armored with ADAM learning algorithm with learning rate of 0.0001. Noticeably, in HD-CISA model, the input features were obtained for each sentence distinctly and therefore HD-CISA approach performed sentence level (say, each review sentiment) sentiment classification. The performance analysis was done in terms of the different statistical parameters such as accuracy, precision, recall and F-Measure. The proposed HD-CISA model is depicted in Fig. 1.

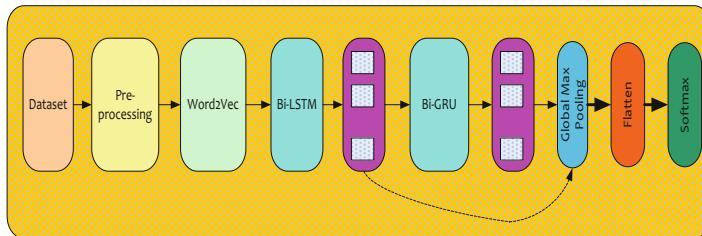


Fig. 1. Proposed Hybrid Deep Model for Sentiment Analysis

The detailed discussion of the proposed HD-CISA model and allied functional components is given in the subsequent sections.

A. Data Acquisition

As HD-CISA intends to develop a robust cross-industry sentiment analysis or opinion mining model, different input datasets were obtained from the different sources, where each dataset represented a specific industry type. Here, the key hypothesis is to train a model with the different vocabularies and allied text-sentiment so that the proposed model can be applied for opinion mining irrespective of the industry or subject matter types. In this reference, a total of five text reviews samples were obtained from Kaggle, representing hospitality sector, healthcare sector, E-commerce, social media etc. More specifically, the datasets were collected from Kaggle for US Airline review data (from AirBNB), Twitter review, Amazon Apple's product review, Reddit social media user's review and Covid-19 lockdown review. Noticeably, these datasets carried the different severity and hence sentence structural contents which makes cumulative learning a mammoth task. A snippet of the different datasets used and their respective sentiment labels is given in Table 1.

Noticeably, the considered datasets (Table 1) can have the high level of content (say, sentence's constituents) heterogeneity and diversity of content can't be ruled out. Hence,

Table 1. Input datasets

Industry Type	Hospitality	Social Media	Social Media	E-Commerce	Healthcare
Dataset	US Airline (AIRBNB)	Twitter	Reddit	Apple Product (Amazon Inc.)	Covid-19
No of reviews	14641	14641	37250	1631	3091
Labels	1-Positive	1	1	1	Sad
	0-Neutral	0	0	0	Anger
	-1 – Negative	-1	-1	-1	Fear

learning over such heterogenous data and components can make learning ambiguous and even unreliable. Considering this fact, to ensure optimal feature learning and classification, a rigorous pre-processing method was applied that retained only those words with meaningful significance towards emotion learning and allied sentiment classification. Here, the pre-processing was performed over each input dataset. The detail of the pre-processing method used is given in the next section.

B. Pre-processing

Realizing the fact that the text reviews on online media often possess differences in emotions, text-compilations, expressions etc., which might even vary over the subject matter discussions. Being a multi-industry fit-to-all sentiment prediction solution, in this research the focus is made on improving input text quality so as to make feature extraction learning and classification effective. To achieve it, HD-CISA applied multiple pre-processing tasks. The HD-CISA model encompassed the following pre-processing tasks:

1. *Missing value removal*
2. *Unicode normalization*
3. *Emoji removal*
4. *Website link removal*
5. *Removal of the words with numeric values*
6. *Non-word character(s) or Punctuations removal*
7. *De-contracting*
8. *Converting to the lower case*
9. *Removing stop-words*
10. *Lemmatization*
11. *Tokenization*

1. Missing Value Removal

In online reviews, Tweets etc., often reviews used to have broken sentences or the content with incomplete communication. In this work, NLTK library was applied to remove such missing values from the words (in the input text or reviews).

2. Unicode Normalization

The social media being an open platform for all invites users with the different languages, personalized writing skills etc. For instance, a user can text review in normal Times New Roman or allied formal Calibri type fonts. On the contrary, there can also be the possibility that the user might have written the content in Scripts or some other languages with italics of special presentations. In this case, the classical word-embedding methods can consider them as non-text information and hence the communication intend and allied contextual significance might be compromised. To alleviate this problem, in this paper Unicode normalization was performed which converted entire text into single font that made input text perceptible and meaningful.

3. *Emoji Removal*

In the last few years, with rise in social media, updated smart phones with a huge emotion-oriented emoji, almost all web or social media platforms provide Emojis to encourage users. On the contrary, such emoji being insignificant in contextual view, might create ambiguity. Therefore, HD-CISA at first removed all kinds of emojis from the input texts.

4. *Hashtag and Website Link Removal*

Generally, in online reviews the different users intend to propagate certain web-links as referral; though, such elements possess no significance towards sentiment analysis. On the other hand, in recent years social medias like Twitter, Instagram, Facebook, LinkedIn etc. provide #HashTag for better content propagation and recommendation across user's network. However, such terms don't have any significance towards sentiment classification. Considering this fact, in this work the aforesaid hashtag and web-URL were removed from the input text by traversing each review or sentence one-by-one. Here, we applied standard expressions and rule-based methods to remove hashtags and URL-links.

5. *Punctuation Removal*

Typically, non-word symbols or the punctuations are the symbols (Ex. " ", "?", "!", ":", etc.), often used in phrases or as remark to improve textual understanding and clarification. However, such symbols can make NLP learning more complex and sometime ambiguous due to reduced learning ability. Thus, applying certain regular expression methods and rule-based approach these punctuations or similar non-word characters were removed from the input texts.

6. *Numeric Word Removal*

In sync with at hand multi-industry, there can be frequent change in textual contents. For instance, a review related to pandemic or healthcare can have the content like temperature 37°C, PSO2, etc. On the contrary, a political review content can have the numeric figures such as 10 lakhs, 10000000 etc. Similar change in text content composition can be in case of hospitality sector where one could have reviewed a restaurant or hotel with rank on 1–5 scales. Despite their perceptual significance, these data elements don't carry much significance for NLP and allied learning, and therefore we have

removed numeric elements from each data sample or dataset which helped in reducing the data (and learning) complexity.

7. Stopping Word Removal

In general, stop-words represent the English words which add no significance or value to the sentence and hence often results a challenge or complexity for the machine learning methods in NLP domains. In practice, such stop-words can be removed from input text content without changing the real-intend of the sentence or phrase. We applied the predefined NLTK corpus stop-words to remove existing stop-words in the text datasets or the samples.

8. Lower Case Conversion

The matter of fact is that the machine learning methods are often case-sensitive and hence the disparity amongst the case often results in complex model training. For instance, in NLP, the term War and WAR are the different words, and hence can create ambiguity in addition to the increased feature space. Considering this fact, we converted each word into the lower case by applying Python's inbuilt lower-case conversion function.

9. Lemmatization

In general, lemmatization transforms an extended word into its root form. This method estimates the real intended component of speech perfectly without losing the sense or intend of a word in a sentence. Typically, stemming and lemmatization differ where the later assesses the context first and then transforms (extended) word into the suitable root word. On the contrary, in stemming only extended characters like "s", "es" etc. are removed at the end of the word. It compromises with the actual or intended meaning of the word in the sentence. For example:

Hobbies → Lemmatization → Hobby

Hobbies → Stemming → Hobbi.

Considering this fact, in this work we performed lemmatization over the input text's sentence over each sample.

10. Tokenization

In this method, the input text corpus or sentences in the dataset were broken down into smaller chunks, often called "tokens". Typically, a token can be a phrase containing a few words, a number, or any symbol that encompasses all the allied relevant information pertaining to the data. We performed sentence-level tokenization in each document by using Python inbuilt function.

C. Word Embedding

In sync with the need of the latent semantic feature information for emotion learning and classification, before executing the proposed contextual (and semantic) feature extraction and learning, the input text was processed for word-embedding. More specifically, in this work, Word2Vec method was applied to convert word-level embedding matrix.

Noticeably, Word2Vec represents a well-known semantic feature extraction method often applied for text-mining and NLP problems. Being developed by Google Inc. Word2Vec embedding method applies neuro-computing method to learn the associations amongst the words and thus retrieves text-features for each word in the input corpus or input reviews datasets. It applies a list of numbers often called “vector” to represent each unique word in the text. In general, the selection of the vector is done in such manner that the cosine similarity can be applied to assess semantic similarity amongst the vectors. The ability to extract latent semantic information from text without human intervention makes it suitable for the at hand text sentiment analysis problem. Before extracting the contextual and semantic (global) features, we executed Word2Vec on each input text data that obtained a set of embedding matrices for the different samples (say, the samples from the different industries’ online reviews). In other words, we obtained embedding matrix for each input samples which was later processed for Bi-LSTM based contextual feature extraction.

D. Hybrid Deep Model Driven (Emotion) Feature Learning

The proposed BI-LSTM and Bi-GRU driven deep learning environment for sentiment analysis is discussed in this section.

1. LSTM Network

The Long- and Short-Term Memory (LSTM) network [14] was at first designed to alleviate the problem of vanishing effect as well as exploding gradient in the recurrent neural network. The fundamental approach behind LSTM is to control the cell-states by means of gates including input gate, forget gate and output gates. As depicted in Fig. 1, in LSTM the forget gate often defined as f_t assesses whether to keep the information of the previous state (c_{t-1}) or forget it by exploring the values of the input (x_t) as well as the hidden state (h_{t-1}). Here, the output value can be either 0 or 1. In the same manner, the input gate (i_t) determines the extent of (volumetric) information pertaining to the input text (x_t) and the hidden layer details (h_{t-1}) to be passed so as to update respective cell-state, so as to achieve output either as 0 or 1. Here, c_t refers the obtained cell state by applying certain mathematical approach on c_{t-1}, f_t and i_t . Here, the information flow between the current cell state to the hidden state is controlled by the output gate (O_t), which often exists as either 0 or 1. At certain time t , let the input to the LSTM model be x_t , and the corresponding previous hidden state be h_{t-1} , with the previous cell-state c_{t-1} . Additionally, let the current output of the hidden state and the current cell state be h_t and c_t , respectively. In reference to these configurations, the different gate elements and their respective outputs are obtained as per (1), (2), (3), (4) and (5).

$$f_t = \text{sigmoid}(W_{fx}x_t + W_{fh}h_{t-1} + b_f) \quad (1)$$

$$i_t = \text{sigmoid}(W_{ix}x_t + W_{ih}h_{t-1} + b_i) \quad (2)$$

$$c_t = c_{t-1} \odot f_t + i_t \odot \tanh(W_{cx}x_t + W_{ch}h_{t-1} + b_c) \quad (3)$$

$$O_t = \text{sigmoid}(W_{ox}x_t + W_{oh}h_{t-1} + b_0) \quad (4)$$

$$h_t = O_t \tanh(c_t) \quad (5)$$

In (1), (2), (3), (4) and (5), $x_t \in R^n$ represents the input vector, $W \in R^{v*n}$, $b \in R^v$. Here, the superscript variables n and v signify the dimensions of the input vector (say, embedded word vector) and the number of words in the corpus, respectively (Fig. 3).

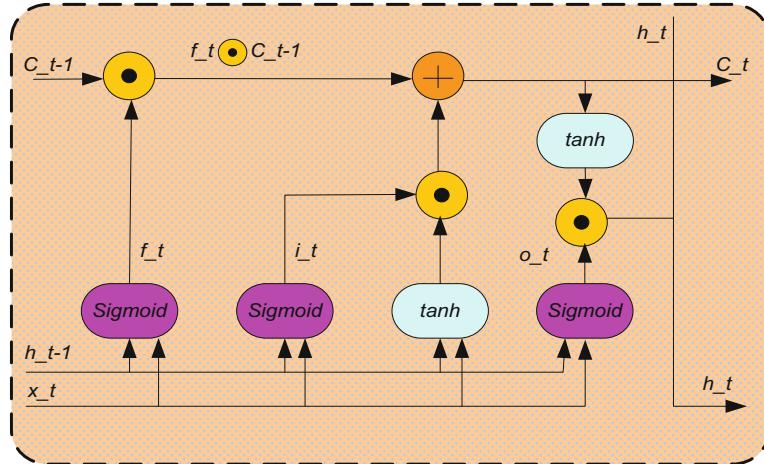


Fig. 2. Functional diagram of LSTM network

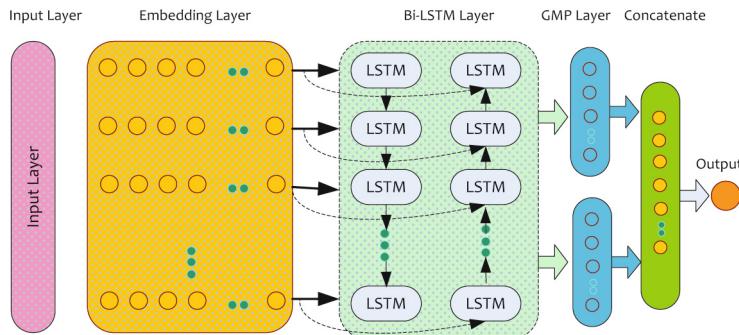


Fig. 3. Bi-LSTM Layer

The detailed discussion of the overall proposed local semantic feature extraction method by using Bi-LSTM is discussed in the subsequent sections.

a. Input Layer

Once embedding the input text datasets, the embedded outputs were passed to the input layer of Bi-LSTM, in the form of input vector x_t . Now, let, $w_1, w_2, w_3, \dots, w_v$ represent the overall number of unique words in each dataset (say, dictionary) $D = d_1, d_2, d_3, \dots, d_m$. In this reference, $i_1, i_2, i_3, \dots, i_v$ belong to the overall unique

indices, signifying the natural numbers. Here, 1 and v be the first and the last index of the data entity or the vocabulary, correspondingly. In the deployed Bi-LSTM model, the input vectors are fed as sequential data with a define length encompassing unique indices. In the subsequent layer, the proposed approach applies an embedding layer in which each word-embedded element index, pertaining to the unique text words in the data corpus, is transferred into the equivalent real-valued feature vector. Thus, the generated real-valued feature vectors are stacked together in the form of a matrix, called embedding matrix (6).

$$R = \begin{matrix} r_{1,1} & r_{1,2} & \cdots & r_{1,n} \\ r_{2,1} & r_{2,2} & \cdots & r_{2,n} \\ r_{3,1} & r_{3,2} & \cdots & r_{3,n} \\ \vdots & \vdots & \ddots & \vdots \\ r_{v-1,1} & r_{v-1,2} & \cdots & r_{v-1,n} \\ r_{v,1} & r_{v,2} & \cdots & r_{v,n} \end{matrix} \quad (6)$$

In above derived embedding matrix, each row signifies a unique index representing a unique word in the corpus or vocabulary. Here, the dimension of the embedding matrix is $v * d$, where v and d be the dataset size and the dimension of the dense vector, correspondingly. In HD-CISA, we assigned $d = 320$ for generate embedding vector.

In order to retain more significant contextual features, in this work, Bi-LSTM was applied. Here, we amalgamated Bi-LSTM layer with the embedding layer, which is discussed as follows:

2. Bi-LSTM Layer

Unlike classical LSTM network, where the information propagates in forward direction, Bi-LSTM accommodates information flow in both forward as well as backwards directions. In other words, in Bi-LSTM network, the state at certain time instant t relies not only on the information before t , but also on other sequential periods. In order to define the complete semantic information over the input text, the use of subsequent information too is can be vital. Therefore, unlike classical CNN and/or LSTM network based local feature extraction, this research applies Bi-LSTM that exploits both the previous states as well as subsequent information to derive more semantically enriched feature (say, contextual information) for further learning and classification. In HD-CISA, Bi-LSTM was designed with two LSTM networks, each capable of processing input vectors in both forward as well as backward direction. Though, we applied each LSTM for feature extraction in single direction (i.e., one for forward direction and another for backward direction). Functionally, the forward LSTM processes input from left-to-right, and thus its hidden layer be represented as (7).

$$\vec{h}_t = LSTM(x_t, \vec{h}_{t-1}) \quad (7)$$

On the contrary, the backward LSTM processes information in right-to-left direction. Thus, the corresponding hidden layer information from backward LSTM is

derived as per (8).

$$\overleftarrow{h}_t = LSTM(x_t, \overleftarrow{h}_{t+1}) \quad (8)$$

Eventually, the extracted outputs from Bi-LSTM (7–8) are concatenated with the final contextual information, (9).

$$h_{t,Bi-LSTM} = [\overrightarrow{h}_t, \overleftarrow{h}_t] \quad (9)$$

In most of the classical approaches, the authors have directly fed the extracted and pooled features for classification. In other words, the state-of-arts LSTM or Bi-LSTM based sentiment analysis models have projected the concatenated contextual feature (9) as the input to the Global Pooling Layer, followed by the output layer for classification. However, such methods often lack “long-term dependency and allied semantic relationship”. On the contrary, to guarantee cross-industry semantic classification retaining both local features (say, contextual features extracted by Bi-LSTM (9)) as well as the global features (i.e., the long-term dependency) is must. With this motivation, in this paper the output of Bi-LSTM is not directly projected for classification rather it is passed to the Bi-GRU for learning and classification.

3. Bi-GRU

The matter of fact is that the major classical machine learning approaches which mainly apply certain limited vocabulary in the form of predefined conditions to model sentiment classification. On the contrary, the RNN networks possesses the capability to employ the preamble vocabulary in the knowledge set. Yet, the likelihood of vanishing effect can't be ruled out in native RNN models. As discussed in the previous sections, LSTM and GRU mainly depend on the gate architecture to permit information to impact the state of each instant or moment which helps in overcoming aforesaid vanishing effect or exploding affect problems. The GRU being a type of LSTM, substitutes both the forget gate as well as the input gate (in the native LSTM architecture) with “update gates”. Figure 2 presents the structure of the GRU network and corresponding mathematical calculations for the different parameter estimation (at the varied layers) is given in following equations.

$$z_t = \sigma(W^z x_t + U^z h_{-1}) \quad (10)$$

$$r_t = \sigma(W^r x_t + U^r h_{-1}) \quad (11)$$

$$\tilde{h}_t = \tanh(W^h x_t + U^h(h_{-1} \odot r_t)) \quad (12)$$

$$h_t = (1 - z_t) \odot \tilde{h}_t + z_t \odot h_{-1} \quad (13)$$

In above Eqs. (10), (11), (12) and (13), W and U represent the GRU's weight matrices, while σ be the logical sigmoid function. Here, \odot states the element multiplication function, while the update gate of GRU be z_t , , which is applied to control the extent or degree of the activation value pertaining to the functional GRU unit. Typically, it remains

abstracted by means of the current input state and the previous layer's state value. Thus, the state of this update layer(s) is obtained cumulatively. In above equation, r_t represents the reset gate that fuses the new input information with the original information. The output at the hidden layer and the candidate hidden layer is given by h_t and \tilde{h}_t , respectively (Fig. 4).

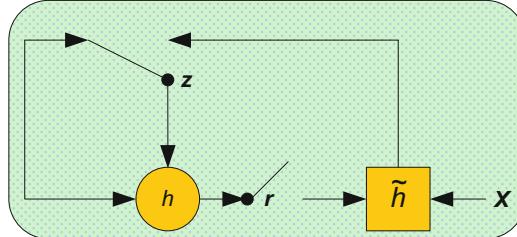


Fig. 4. Structure of the GRU unit.

Summarily, in comparison to the LSTM network, the GRU network minimizes the hyperparameters and their respective complexity that eventually suppresses decisive computational costs. Similar to the classical LSTM network RNN networks (i.e., LSTM), GRU too possesses one-way state transmission. Moreover, there can be certain state, the current output can't be related to the other or previous state, as it can also be related to the next state. On the contrary, predicting the missing words in cross-industry text-corpora or sentences need not merely the previous result, but it also requires the word (embedded matrix) (subsequent) content. To meet this the use of Bi-GRU can be of vital significance. Additionally, the Bi-GRU can help extracting more significant latent semantic information with long-term dependency to make sentiment classification decision better (Fig. 5).

The typical Bi-RNN networks encompass two unidirectional recurrent neural networks. During computation, for each moment (of computation), the input to the two recurrent networks remains in opposite directions concurrently, and the respective output is obtained cumulatively. It helps improving the accuracy of the prediction results. Same as the aforesaid Bi-RNN architecture, we substituted RNN by GRU structure that resulted a Bi-GRU. In our proposed deep model, we applied Bi-GRU network to learn long term-dependency, also called the global semantic information over the Bi-LSTM extracted features (13). In our deployed network, we modelled the network with two GRUs that functions concurrently so as to model emotions towards the two directions (i.e., forward and backward) of the feature sequence (13) for training. Finally, it results the outputs at the hidden layer H_t in the form of (16). The mathematical mechanism applied to train over the forward and backward directions of the feature sequences are derived as per (14), (15), (16) and (17).

$$\overleftarrow{h}_t = GRU(X_t, \overleftarrow{h}_{t+1}), \quad t \in [1, L] \quad (14)$$

$$\overrightarrow{h}_t = GRU(X_t, \overrightarrow{h}_{t-1}), \quad t \in [L, 1] \quad (15)$$

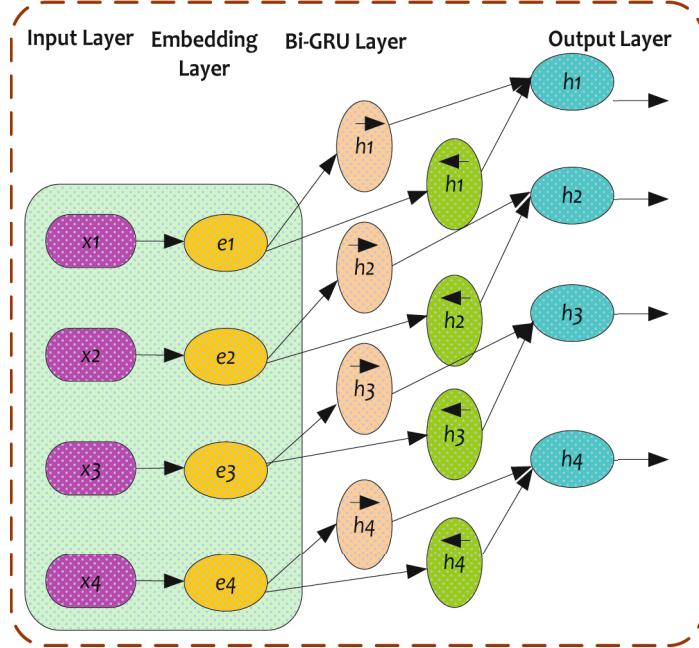


Fig. 5. Structure of the Bi-GRU network

$$H_{t_Bi_GRU} = \left[\overrightarrow{h}_t, \overleftarrow{h}_t \right] \quad (16)$$

$$H_{t_Bi_GRU} = (1 - z_t) \odot \tilde{h}_t + z_t \odot h_{-1} \quad (17)$$

In (14), (15) and (16) h_0 and h_{L+1} are initialized with zero vectors. In Eq. (15), $\vec{h}_t \in R^{L \times d}$ represents the emotional feature representation (EFR) of the trained Bi-LSTM features (13) which is fused with the X_t . Similarly, $\overset{\leftarrow}{h}_t \in R^{L \times d}$ be the EFR of the subsequent fusion. Here, d be the dimension of the output of the GRU unit. In the learnt model, $H_t \in R^{L \times 2d}$ amalgamates two-unit output vector dimension and thus fuses the learnt feature with the contextual emotional information in the form of emotional representation of the input corpuses pertaining to the multiple industry type review texts. Thus, Bi-GRU model retained the global features, signifying the long-term dependency in the input features (of amongst the words) to make learning more efficient. Noticeably, in HD-CISA, the developed hybrid deep model was applied in such manner that it performed feature extraction and learning over each data concurrently so as to reduce computational time. Finally, once extracting the contextual features from the multi-industry review texts inputs and corresponding long-term dependency (say, global features), the feature fusion was performed. A snippet of the feature fusion model applied is given in the subsequent section.

E. Feature Fusion

As discussed above, in HD-CISA, the use of Bi-LSTM model enabled local feature extraction (i.e., contextual information), while the Bi-GRU helped in retrieving global features encompassing the semantic information or long-term dependency information. In HD-CISA, Bi-GRU traverses across the input text sequences to retrieve global semantic features. In this manner, it obtained learnt global features from each input text corpus distinctly, which were then concatenated at the global average pooling layer. In HD-CISA model, we concatenated both Bi-LSTM driven contextual feature as well as Bi-GRU based global (semantic) features to yield text instance feature representation vector V_s . Here, we obtained feature representation vector V_s for each industry related text input. Thus, it yielded the concatenated feature vector as (18).

$$H = concat(h_{t,Bi-LSTM}, H_{t,Bi-GRU}) \quad (18)$$

Let, the number of convolutional kernels in Bi-LSTM be B , then with B convolutional kernels and $2d$ dimensional output vector from Bi-GRU were assigned the same value. In this manner, the generated feature vector by Bi-LSTM and Bi-GRU was connected by concatenation method (18). In (18), $H \in R^{(L \times d) \times 2d}$ be the sliced vector, while the output of the Bi-LSTM be $h_{t,Bi-LSTM} = [C_1, C_2, \dots, C_n], C \in R^{l \times B}$, and $H_{t,Bi-GRU} = [h_1, h_2, \dots, h_L], H_t \in R^{l \times 2d}$ be the output vector of the Bi-GRU. In HD-CISA, we applied global average pooling (GAP) layer that averaged the input vector H (18) to derive feature points or instances (say, vector) $V_s \in R^{2d}$, representing the feature representation of the text instances from the multi-industries reviews. This approach can help avoiding any likelihood of over-fitting problem, as we intend to use multiple corpuses from the different data types, word structure and allied meanings.

$$V_s = GAP(H) \quad (19)$$

Thus, the obtained composite feature vector was projected to the output layer for further sentiment classification. Let, V_{si} be the composite feature encompassing contextual as well as semantic features from the $i - th$ (industry) dataset, and there be N datasets, then the final feature vector obtained was (20), where N be the total number of input (multi-industry) datasets.

$$V_s = concat(V_{s1}, V_{s2}, \dots, V_{sN},) \quad (20)$$

Output Layer

In HD-CISA, the extracted feature vector V_s was projected as input to the output layer, which was later processed for learning and (sentiment) classification by using following loss-function.

$$LossFunction = -\frac{1}{m} \sum_i^m (y_i * log(p(y_i)) + (1 - y_i) * log(1 - p(y_i))) \quad (21)$$

In (19), the parameter m states the number of text-samples (here, four), y_i be the true labels, while $p(y_i)$ be the probability of the true labels. HD-CISA performed classification with ADAM adaptive learning method with learning rate predefined at 0.001.

4 Results and Discussion

This research proposed hybrid deep model driven cross-industry sentiment analysis model(HD-CISA). As the name indicates HD-CISA applied Bi-LSTM and Bi-GRU in cascade design to perform local contextual feature and global (long-term dependency) feature extraction and learning. Realizing the cross-industry review contents and allied composition, this research also focused on pre-processing where suitable set of activities including Missing value removal, Unicode normalization, Emoji removal, Website link removal, Removal of the words with numeric values, non-word character(s) or Punctuations removal, De-contracting, converting to the lower case, removing stop-words, Lemmatization, and Tokenization were performed over input datasets. Noticeably, being a cross-industry sentiment analysis solution at first a total of five input (reviews) samples were collected from Kaggle web-repositories. Amongst the different inputs the US Airline and Twitter reviews (on political activities) comprised a total of 14641 reviews each. Similarly, Reddit reviews encompassed a total of 37250 reviews, while Apple products reviews were collected from Amazon which encompassed 1631 reviews from users. Covid-19 related reviews were counted as 3061 sentences where each sentence represented one individual review. Once obtaining these input datasets, the aforesaid pre-processing methods were applied for which the different rule-based methods and Python's inbuilt libraries including NLTK were taken into consideration. Subsequently, to ensure uniform input vector for deep learning and allied classification, the obtained inputs were processed for Word2Vec word embedding process for which Gensim library was considered. In HD-CISA, Word2Vec method transformed each (text) word into equivalent unique numeric representation. To be noted, being a sentence level sentiment analysis model, we estimated embedding matrix for each sentence or review, and thus a total of 71254 review's embedding matrix were obtained. In other words, in HD-CISA a total of 71254 input reviews from five different industries were processed for learning for sentiment prediction task. Once obtaining the input word-embedding matrix or outputs, these were passed to the input layer of the Bi-LSTM deep model. Here, the inputs were passed to the deep model in moving window approach and thus, Bi-LSTM model retrieved local contextual as well as intrinsic global features. Noticeably, since, in this work Bi-LSTM output is not being applied for eventual classification, we replaced Softmax layer by Bi-GRU deep model or Bi-GRU layer. Here, the key purpose of applying Bi-GRU was to retain long-term dependency and hence inter-element associations to make learning superior towards opinion mining. Unlike existing methods where the authors have applied CNN for local feature extraction followed by Bi-LSTM or GRU for long-term dependency, we horizontally concatenated the extracted features from both Bi-LSTM as well as bi-GRU to form a global feature vector encompassing both local as well as global feature or long-term dependency over complete input texts (i.e., 71254 reviews). Finally, the global feature vector at the fusion layer of flatten layer was passed to the Softmax layer for learning and classification. In this work, cross-entropy loss function was applied for learning and classification. To be noted, considering feature non-linearity over the different input corpuses and allied word-embedding inputs, we applied ADAM adaptive learning method. Here, both Bi-LSTM as well as Bi-GRU applied ReLU activation function at the input layer; though, we assigned the number of neurons in these deep models as 128. The deployed ADAM optimizer was assigned

learning rate of 0.001. Being a batch size of 128, especially over 71254 inputs, to alleviate any likelihood of overfitting and convergence issues, we applied dropout layer post convolution in Bi-LSTM. Here, the dropout layer was assigned filter coefficient of 0.5. We assessed the HD-CISA with the number of epochs assigned as 25; though it indicates the scope of better performance with higher number of epochs.

The proposed HD-CISA model was implemented on a central processing unit armored with Intel Core i5 processor, with 8 GB memory and 3.2 GHz processing elements. We simulated the developed model on Anaconda Notebook, where the algorithms were developed in Python language. More specifically, the Python libraries including Numpy, Keras, Tensorflow, Pandas, NLTK, etc. were taken into consideration.

To assess efficiency of HD-CISA model, statistical performance parameters were obtained by exploiting confusion metrics. To achieve it, the confusion metrics values were obtained in terms of true positive (TP), true negative (TN), false positive (FP) and False negative (FN). Thus, exploiting these metrices the performance parameters were obtained as per Table 2.

Table 2. Performance Parameters

Parameter	Mathematical Expression	Definition
Accuracy	$\frac{(TN+TP)}{(TN+FN+FP+TP)}$	Signifies the proportion of class that were classified correctly over the total number of classes 9say reviews labelled correctly for their original class) inspected out of all modules
Precision	$\frac{TP}{(TP+FP)}$	States the degree to which the repeated measurements under unchanged conditions show the same results
Accuracy	$\frac{(TN+TP)}{(TN+FN+FP+TP)}$	Signifies the proportion of class that were classified correctly over the total number of classes 9say reviews labelled correctly for their original class) inspected out of all modules

Since in this research the focus was made on optimizing both feature as well as learning model, where different datasets were considered and clubbed together to yield a cross-industry sentiment analysis (prediction) solution. To assess whether HD-CISA yielded superior performance over the different datasets individually as well as with the combined datasets, we decoupled cross-industry sentiment data (i.e., training over 71254 input reviews from five different industries) into five different industry specific opinion mining problem. Accordingly, the efficiency of the proposed word-embedding driven hybrid deep networks was examined. Noticeably, to assess the efficiency, we performed same pre-processing followed by word-embedding tasks and feature learning (i.e., Bi-LSTM followed by Bi-GRU) over each data. However, the performance was assessed distinctly over each data. The simulation results are given in Table 3.

Noticeably, to achieve above simulation results (Table 3), the pre-processing, word-embedding and subsequent feature extraction and learning was done for each dataset.

Table 3. The simulation results

SN	Dataset	Data Size	Accuracy (%)	Precision (%)	Recall (%)	F-Measure (%)
1	US Airline	14641	97.89	96.61	96.69	96.69
2	Twitter	14641	96.68	94.99	96.80	95.88
3	Reddit	37250	96.72	94.37	96.57	95.45
4	Apple (Amazon)	1631	93.67	94.49	95.08	94.78
5	Covid-19	3091	92.48	96.01	96.13	96.06
6	Cross Industry (Cumulative 1 to 5)	71254	97.91	97.73	97.48	97.60

Observing the results, it can be found that the accuracy over US Airline data, which comprises a total of 14641 text reviews exhibited sentiment prediction accuracy of almost 97.9%, while the respective precision recall and F-Measure were 96.61%, 96.69% and 96.69%, respectively. Similarly, the performance assessment over Twitter review data about certain political even which embodied a total of 14641 reviews, each from one user showed the sentiment classification accuracy of 96.68%, with precision, recall and F-Measure values as 95%, 96.8% and 95.88%, correspondingly. The simulation results over social media dataset (Reddit) exhibited accuracy of 96.72%, precision of 94.37%, recall 96.57% and F-Measure of 95.45%. On the other hand, Amazon review towards Apple products, which comprised merely 1631 reviews showed the sentiment analysis result with accuracy 93.67%, precision 94.5%, recall 95.08% and F-Measure 94.78%. The Twitter reviews towards Covid-19 pandemic during lockdown period, which comprises almost 3091 reviews or comments from the different users resulted sentiment prediction accuracy of 92.5%, precision 96%, recall 96.13% and F-Measure of 96.6%. Interestingly, when these input reviews were clubbed together, with the cumulative features the proposed hybrid deep driven sentiment analysis model resulted accuracy of 97.91%, precision of 97.73%, recall 97.48%, and F-Measure of 97.60%.

The simulation results infer that the proposed hybrid deep learning model with the cross-industry features can yield superior sentiment analysis or opinion mining solution. Here, the efficiency of HD-CISA seems superior over large inputs spaces or input dataset. However, the efficiency of the proposed Bi-LSTM model followed by Bi-GRU for local as well as global feature learning has resulted optimal performance towards online opinion mining task to characterize online user's (say, netizen's) behavior analysis. The graphical depiction of the results obtained are given in Figs. 6, 7, 8 and 9. Here, to make relative performance analysis clear, the Y-axis scale is taken in the range of 90–100%.

The overall results indicate that the proposed hybrid deep learning model achieves better performance over multiple datasets together, signifying its ability to yield high reliability towards cross-industry sentiment analysis task or netizen's opining mining. Moreover, the performance over cross-industry datasets (Figs. 6, 7, 8 and 9) shows near

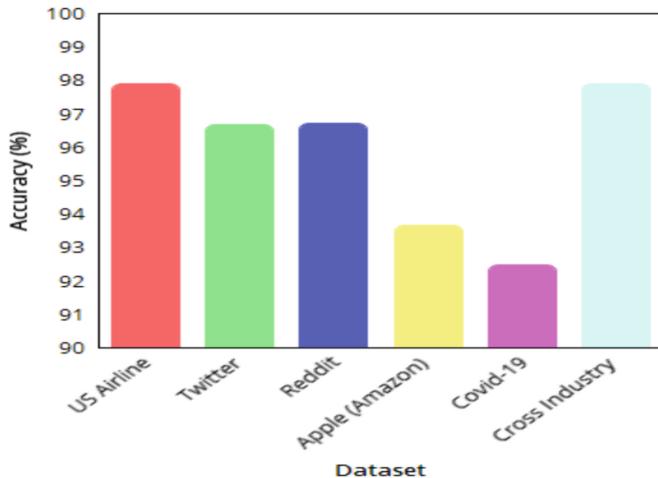


Fig. 6. Accuracy (%) over the different datasets

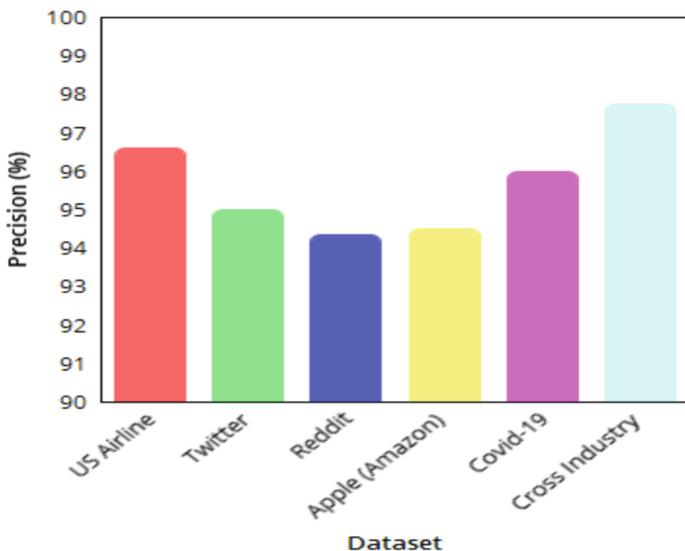


Fig. 7. Precision (%) over the different datasets

performance as that with the standalone input sample or dataset. The results confirm that the proposed model can yield opinion mining irrespective of the industry type and here the contribution of semantic word-embedding followed by hybrid deep feature learning (Bi-LSTM and Bi-GRU) can't be ruled out.

To compare the relative efficiency of HD-CISA, we considered different state-of-arts, especially the one recently published with the similar deep learning driven solutions.

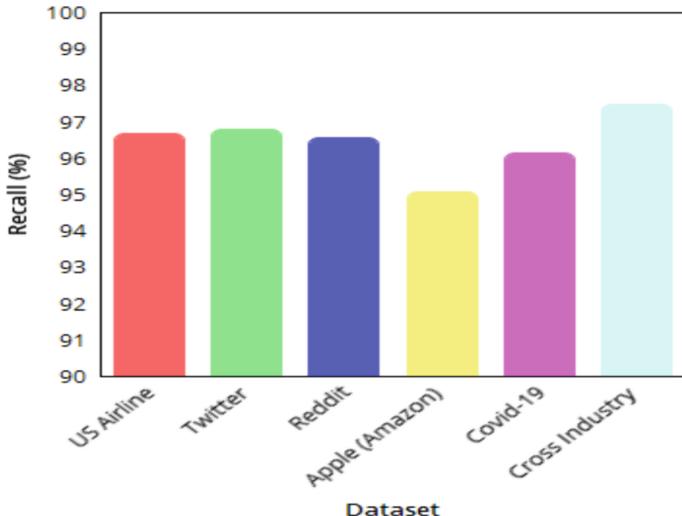


Fig. 8. Recall (%) over the different datasets

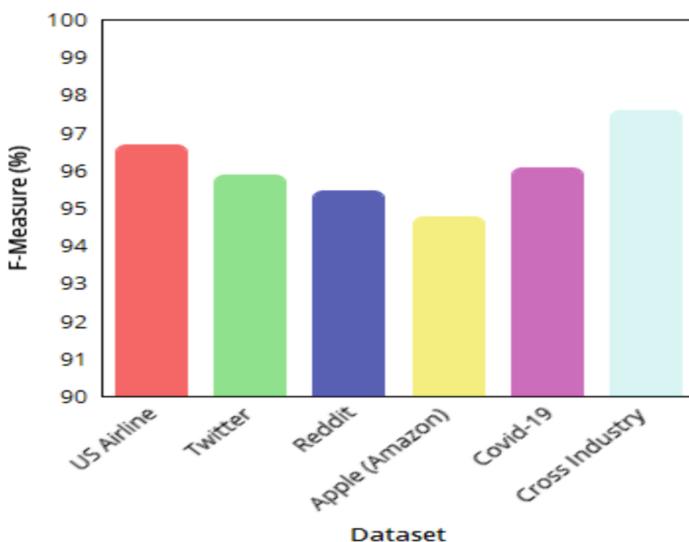


Fig. 9. F-Measure (%) over the different datasets

Xiao et al. [35] in their model could achieve sentiment classification accuracy of 92.58% over hotel review data. Recently, the authors in [81] applied Bi-LSTM deep model for sentiment analysis over the different input datasets including Movie review, Internet Movie Database and Stanford Sentiment Treebank. Noticeably, in their work, post pre-processing a capsule network was designed with Bi-LSTM, especially to extract feature. Moreover, the extracted feature from Bi-LSTM was directly projected to the Softmax layer for classification using Binary cross-entropy loss function. The highest

accuracy obtained by their model was 91.96%, which is almost 5.95% lower than HD-CISA (97.91%). Interestingly, in their work [81], the authors stated that their proposed Capsule-Bi-LSTM performs better than other state-of-art methods including machine learning (SVM) algorithm, LSTM, Tree-LSTM, CNN, ToWE-CBOW etc. It infers that the superiority of the proposed hybrid deep model confirms efficiency over aforesaid secondary approaches. It is worth mentioning that ToWE-CBOW is also a word-embedding method, in which word-embedding takes place not on each word, rather over continuous bag of words so as to reduce computational cost and hence search space. However, their respective performance in [81] revealed that these approaches (i.e., ToWE-CBOW) yields merely 65.10% accuracy, which is very low in comparison to the proposed HD-CISA method. The authors in [82] proposed genetic algorithm (GA) assisted CNN-based aspect level sentiment analysis method. Interestingly, in this work, the authors applied CNN to extract local feature or the contextual feature. To improve efficiency of CNN, the authors applied GA as tuning algorithm. Noticeably, similar to our proposed approach, they too performed Word2Vec word-embedding before executing CNN-based contextual feature extraction. The depth performance analysis revealed that their proposed model could get the accuracy of 95.5%, precision of 94.35, recall 91.1%, and F-Measure of 96.0%. In comparison to this approach [82], HD-CISA exhibited accuracy of 97.91%, precision 97.73%, recall 97.48% and F-Measure of 97.60%, which is higher than their method [82]. It confirms that the use of Bi-LSTM and Bi-GRU enabled HD-CISA retaining both contextual local features as well as long-terms dependency or inter-element association feature (say, global feature). This as a result improved HD-CISA to achieve superior performance. Interestingly, the authors [82] claimed their model exhibiting higher accuracy (95.5%) than the other machine learning methods like decision tree (67%), SVM (92.3%), Linear Dirichlet Analysis (65%), neuro-computing with Levenberg Marquardt (78.5%), random forest ensemble (8.3%), and logistic regression (80.5%). Thus, the superiority of HD-CISA is confirmed over aforesaid approaches. The authors in [83] developed Interactive Gated Convolutional Network (IGRU) driven aspect level sentiment analysis model. They designed their model as extension of CNN-based method, where CNN was at first applied to extract local contextual features, which was later used by IGRU to train over the long-term dependence. The authors simulated their model over Yelp dataset, and performance was obtained in terms of sentiment classification accuracy. Interestingly, they could achieve the highest accuracy of 81.34%, which is almost 16.4% lower than HD-CISA. The authors in [84] applied LSTM driven solution for aspect level sentiment analysis; yet, the highest accuracy obtained was 74.3%. Similarly, in [66] attention-based LSTM was designed for sentiment prediction, where the highest accuracy observed was 77.2%, which is almost 25.5% lower than HD-CISA model. The authors in [85] developed position-aware bidirectional attention network for aspect-level sentiment analysis. Interestingly, despite numerous feature level improvement, it could achieve the highest accuracy of 81.16%, which is almost 16.7% lower than the proposed HD-CISA model. An improved attention-based LSTM model was designed in [86]; however, it could achieve the highest sentiment classification accuracy of 81.2%, which is significantly poor to cope up with contemporary business intelligence demands. Similar to our approach, the authors [87] at first applied CNN for local feature

extraction, which was followed by GRU-based learning to perform aspect level sentiment analysis. Noticeably, unlike our approach, where we fused the features obtained by Bi-GRU and Bi-LSTM together to yield a composite feature, the authors applied GRU output for classification. The performance assessment revealed that the highest accuracy observed in their model was 89.61%, which is almost 8% lower than our proposed HD-CISA model. Here, the use of word-embedding followed by Bi-LSTM and Bi-GRU feature learning with composite feature vector can be found yielding superior results in HD-CISA (accuracy = 97.91%). In reference to [87], HD-CISA can be hypothesized to be yielding superior performance than any known deep learning driven solution. Moreover, the robustness of our proposed multi-industry reviews or data makes it more significant and realistic to cope up with real-time demands. The authors [88] applied BERT-Bi-LSTM network for consumer's opinion mining in power industry. Despite intrinsic value additions in feature learning, the authors [88] could achieve the highest accuracy of 86.20%, and recall 70.8%, which is lower than the HD-CISA. The authors [89] at first applied Bag-of-Word method over input texts, which was later processed with LSTM-GRU deep model. In their approach the feature output from GRU network was classified using the different machine learning methods, where they claimed to have achieved accuracy of 98–99%. Thus, in future, we can examine the efficiency of our fused feature vector with machine learning methods to achieve superior efficiency; yet, generalizing their model seems exaggerated, especially for aspect level-based sentiment analysis. This is because, claiming higher accuracy over merely frequency-based approach (TF-IDF) and claiming it superior over Word2Vec seems questionable. In future, our proposed feature model can be assessed with the other machine learning methods to assess relative performance in comparison to [88]. Though, a complex but significant effort was made in [90] where the authors applied multi-head attention capsule model combining convolutional neural network and Bi-GRU. Their simulation results resulted the highest accuracy of 85.3%, which is lower than HD-CISA model. A single layered LSTM model was applied for sentiment analysis in [91]; however, it resulted sentiment classification accuracy of 80.50%, which is almost 17.2% lower than our proposed method. A multi-channel CNN with Bi-GRU was applied in [53], where the authors claimed to have achieved the highest sentiment classification accuracy of 92.90%. In comparison to this approach as well, our proposed method yields superior accuracy (97.91%). In sync with the depth performance analysis and allied inferences, HD-CISA yields better performance than any other known approaches. Behind such robustness performance efficiency, the contributions of pre-processing, cross-industry text learning, initial word-embedding and followed by Bi-LSTM-Bi-GRU feature learning can be confirmed. Due to space constraints, the comparison of the proposed model could not be done with other existing methods.

5 Conclusion

Unlike major state-of-arts methods for sentiment analysis, this research focused on developing a robust cross-industry sentiment characterization approach. Recalling the fact that the majority of the existing are developed for sentiment analysis for a standalone problem such as E-commerce product review analysis, hotel review analysis, political event or

statement review or reactions analysis or healthcare related reviews or reaction analysis. Interestingly, almost all existing methods have applied topic-specific feature embedding and allied learning methods to perform sentiment analysis; however, training over a single and distinct contextual domain doesn't guarantee applicability of the SA model for other subject matter or domain. For instance, training a model over E-commerce review can be different than the healthcare sector review such as Covid-pandemic. In this case, merely training a model over standalone (single industry) corpus can yield inferior performance over another corpus. Additionally, the use of aspect-based sentiment analysis can yield superior cross-industry feature learning than the traditional word-embedding or lexicon-based approaches. Considering aforesaid challenges and allied scopes, in this paper hybrid deep driven cross-industry sentiment analysis model HD-CISA was proposed. As cross industry solution, the different lengthy corpus including the online reviews from the hospitality sector, E-commerce, healthcare sector and E-commerce were processed altogether to enable a fit-to-all SA solution. Recalling the fact that the different digital platforms can have the different way of expressions such as special fonts, emoji etc., a depth pre-processing mechanism was applied. More specifically, the proposed model performed missing value removal, Unicode normalization, removal of emoji, web-link, HTML text, words with numeric values, non-word characters, stop-words, etc., followed by de-contracting, lower case conversion and lemmatization. This robust pre-processing approaches ensured that the final corpus (pre-embedding) possess optimal intrinsic feature in a common perceptible structure from the cross-platform reviews so as to make optimal feature extraction and learning. Subsequently, the obtained corpus was processed for lemmatization and tokenization that unlike stemming method enabled more contextual information for both local as well as global feature extraction. In the subsequent phase, HD-CISA applied Word2Vec semantic embedding to offer more suitable (in comparison to the linear Dirichlet analysis) vector representation. The obtained embedded features were then passed to the proposed hybrid deep model encompassing Bi-LSTM and Bi-GRU, where the first obtained the local contextual features, while the later trained over the extracted local features to achieve long term dependency learning. Finally, the features obtained from both Bi-LSTM and Bi-GRU were fused on the Softmax layer to perform learning and classification. Here, the use of ADAM learning method with cross-entropy loss function enabled optimal learning and classification that resulted accuracy of 97.91%, precision 97.73%, recall 97.48% and F-measure 97.60%, which was higher than any known approach so far. The depth relative comparison to confirmed that the proposed model exhibits superior over major state-of-arts. Moreover, the training of the proposed model over the different corpus, especially in ABSA paradigm strengthens the proposed HD-CISA method to be applied over the different industrial SA analysis problems. Thus, it can serve as a Fit-to-all SA solution for the different industries and can have broaden scalability to serve masses.

References

1. Liu, B.: *Mining Opinions, Sentiments, and Emotions*. Cambridge Univ. Press, Cambridge, U.K. (2015)
2. Cambria, E.: Affective computing and sentiment analysis. *IEEE Intell. Syst.* **31**(2), 102 (2016)

3. Pang, B., Lee, L.: Opinion mining and sentiment analysis. Found. Trends® Inf. Retr. **2**(1–2), 1–135 (2008). <https://doi.org/10.1561/1500000011>. <https://www.nowpublishers.com/article/Details/INR-011>
4. Liu, B.: Sentiment analysis and opinion mining. Synth. Lectures Hum. Lang. Technol. **5**(1), 1–167 (2012). <https://doi.org/10.2200/S00416ED1V01Y201204HLT016>
5. Hu, M., Liu, B.: ‘Mining and summarizing customer reviews. In: Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery Data Mining (KDD), pp. 168–177 (2004). <https://www.scirp.org/reference/referencespapers?referenceid=3018674>
6. Kim, S.-M., Hovy, E.: ‘Determining the sentiment of opinions. In: Proceeding of the 20th International Conference on Computing Linguistics (COLING), p. 1367 (2004). <https://aclanthology.org/C04-1200/>
7. Rustam, F., Ashraf, I., Mehmood, A., Ullah, S., Choi, G.: Tweets classification on the base of sentiments for US airline companies. Entropy **21**(11), 1078 (2019). https://www.researchgate.net/publication/337050333_Tweets_Classification_on_the_Base_of_Sentiments_for_US_Airline_Companies
8. Ghanbari-Adivi, F., Mosleh, M.: Text emotion detection in social networks using a novel ensemble classifier based on Parzen tree estimator (TPE). Neural Comput. Appl. **31**(12), 8971–8983 (2019). <https://ouci.dnbt.gov.ua/en/works/lo5LqN17/>
9. Yoav, G.: Neural Network Methods for Natural Language Processing (Synthesis Lectures on Human Language Technologies), vol. 10. Morgan & Claypool, San Rafael, CA, USA (2017). <https://www.springer.com/series/16917?srsltid=AfmBOppjN2CvZTF3QGsTvXt8oEDKQ-1XZZbYumBl2UJpmyeEafa9zv>
10. Xu, Y., Ren, J., Wang, G., Zhang, C., Yang, J., Zhang, Y.: A blockchain based nonrepudiation network computing service scheme for industrial IoT. IEEE Trans. Ind. Informat. **15**(6), 3632–3641 (2019)
11. O’Connor, B.: From tweets to polls: Linking text sentiment to public opinion time series. In: Proceedings of the 4th Interanational AAAI Conference on Weblogs social media, pp. 122–129 (2010)
12. Moraes, R., Valiati, J.F., Gavião Neto, W.P.: Document-level sentiment classification: an empirical comparison between SVM and ANN. Expert Syst. Appl. **40**(2), 621–633 (2013)
13. Pontiki, D., et al.: SemEval-2016 task 5: Aspect based sentiment analysis. In: Proceedings of the 10th International Workshop Semantic Evaluation, SemEval. Association for Computational Linguistics, pp. 19–30. San Diego, CA, USA (2016)
14. Yu, W., Zhou, W.N.: Sentiment analysis of commodity reviews based on LSTM. Comput. Sci. Appl. **27**(08), 159–163 (2018)
15. Ren, Y., Wang, R., Ji, D.: A topic-enhanced word embedding for Twitter sentiment classification. Inform. Sci. **369**, 188–198 (2016)
16. Xue, W., Zhou, W., Li, T., Wang, Q.: MTNA: A neural multi-task model for aspect category classification and aspect term extraction on restaurant reviews. In: Proceedings of the 8th International Joint Conference on Natural Language Processing, no. 2, pp. 151–156 (2017)
17. Xu, J.C., Chen, D.L., Qiu, X.P., Huang, X.J.: Cached long short-term memory neural networks for document level sentiment classification. In: Proceedings of the 2016 Conf. Empirical Methods in Natural Language Processing, pp. 1660–1669. Austin, TX, USA (2016)
18. Dong, L., Wei, F., Tan, C., Tang, D., Zhou, M., Xu, K.: Adaptive recursive neural network for target-dependent Twitter sentiment classification. In: Proceedings of the 52nd Annual Meeting Association Computing Linguistics (Short Papers), vol. 2, pp. 49–54 (2014)
19. Liang, B., Quan, L., Jin, X., Qian, Z., Peng, Z.: Aspect-based sentiment analysis based on multi-attention CNN. J. Comput. Res. Develop. **54**(8), 1724 (2017)
20. Zhang, D., Tian, L., Hong, M., Han, F., Ren, Y., Chen, Y.: Combining convolution neural network and bidirectional gated recurrent unit for sentence semantic classification. IEEE Assess **6**, 73750–73759 (2018)

21. Kamps, J., Marx, M., Mokken, R.J., Rijke, M.D.: Words with attitude. In: Proceedings of the Belgian-Netherlands Conference on Artificial Intelligence, pp. 332–341 (2002)
22. Dong, Z., Dong, Q.: HowNet and the Computation of Meaning. World Scientific Publishing Co. Pte. Ltd. (2006)
23. Pang, B., Lee, L.: A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts. In: Proceedings of the 42nd Annual Meeting Association Computing Linguistics-ACL, pp. 271–278 (2004)
24. Lee, H.Y., Renganathan, H.: Chinese sentiment analysis using maximum entropy. In: Proceedings of the Workshop Sentiment Analysis AI Meets Psychol. (SAAIP), pp. 89–93 (2011)
25. Poirier, D., Bothorel, C. Neef, E.G.D., Boullé, M.: Automating opinion analysis in film reviews: the case of statistic versus linguistic approach. In: Proceedings of the Language Resource Evaluation, pp. 12–140 (2011)
26. Naz, S., Sharan, A., Malik, N.: Sentiment classification on Twitter data using support vector machine. In: Proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence (WI), pp. 676–679 (2018)
27. Pantano, E., Giglio, S., Dennis, C.: Making sense of consumers' tweets: Sentiment outcomes for fast fashion retailers through big data analytics. *Int. J. Retail Distrib. Manage.* **47**(9), 915–927 (2019)
28. Vyas, V., Uma, V.: Approaches to sentiment analysis on product reviews. In: Rajput, D.S., Thakur, R.S., Muzamil Basha, S. (eds.) *Sentiment Analysis and Knowledge Discovery in Contemporary Business*, pp. 15–30. IGI Global (2019). <https://doi.org/10.4018/978-1-5225-4999-4.ch002>
29. Alkalbani, A.M., Gadhvi, L., Patel, B., Hussain, F.K., Ghamry, A.M., Hussain, O.K.: Analysing cloud services reviews using opinion mining. In: Proceedings of the IEEE 31st International Conference on Advances Information Network Application (AINA), pp. 1124–1129 (2017)
30. Fikri, M., Sarno, R.: A comparative study of sentiment analysis using SVM and SentiWordNet. *Indonesian J. Electr. Eng. Comput. Sci.* **13**(3), 902–909 (2019)
31. Priyatina, R., Sarno, R.: Sentiment analysis of hotel reviews using latent Dirichlet allocation, semantic similarity and LSTM. *Int. J. Intell. Eng. Syst.* **12**(4), 142–155 (2019)
32. Kontopoulos, E., Berberidis, C., Dergiades, T., Bassiliades, N.: Ontology-based sentiment analysis of Twitter posts. *Expert Syst. Appl.* **40**(10), 4065–4074 (2013)
33. Kim, Y.: Convolutional neural networks for sentence classification. [arXiv:1408.5882](http://arxiv.org/abs/1408.5882) (2014). <http://arxiv.org/abs/1408.5882>
34. Zhang, X., Zhao, J., Lecun, Y.: Character-level convolutional networks for text classification. In: Proceedings of the Advance Neural Information Processing System pp. 645–657 (2015)
35. Xiao, Z., Li, X., Wang, L., Yang, Q., Du, J., Sangaiah, A.K.: Using convolution control block for Chinese sentiment analysis. *J. Parallel and Distributed Comput.* **116**, 18–26 (2018)
36. Cheng, Y., Ye, Z.M., Wang, M.W., Zhang, Q., Zhang, G.H.: Analysis of Chinese text sentiment orientation based on convolutional neural network and hierarchical attention network. *J. Chin. Inf. Process.* **33**(1), 133–142 (2019)
37. Collobert, R., Weston, J., Bottou, L., Karlen, M., Kavukcuoglu, K., Kuksa, P.: Natural language processing (almost) from scratch. *J. Mach. Learn. Res.* **12**, 2493–2537 (2011)
38. Kalchbrenner, N., Grefenstette, E., Blunsom, P.: A convolutional neural network for modelling sentences (2014). [arXiv:1404.2188](http://arxiv.org/abs/1404.2188). <http://arxiv.org/abs/1404.2188>
39. Yin, W., Schütze, H.: Multichannel variable-size convolution for sentence classification (2016). [arXiv:1603.04513](http://arxiv.org/abs/1603.04513). <http://arxiv.org/abs/1603.04513>
40. Tang, D., Qin, B., Feng, X., Liu, T.: Effective LSTMs for target dependent sentiment classification [arXiv:1512.01100](http://arxiv.org/abs/1512.01100) (2015). <http://arxiv.org/abs/1512.01100>

41. Ren, M., Gan, G.: Sentiment analysis of text based on bi-directional long short-term memory model. *Comput. Eng. Design* **39**(379), 272–276 (2018). (in Chinese)
42. Hessel, M., Soyer, H., Espeholt, L., Czarnecki, W., Schmitt, S., van Hasselt, H.: Multi-task deep reinforcement learning with PopArt. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 33, pp. 3796–3803 (2019)
43. Liu, P., Qiu, X., Huang, X.: Deep multi-task learning with shared memory [arXiv:1609.07222](https://arxiv.org/abs/1609.07222) (2016). <http://arxiv.org/abs/1609.07222>
44. Liu, P., Qiu, X., Huang, X.: Recurrent neural network for text classification with multi-task learning. [arXiv:1605.05101](https://arxiv.org/abs/1605.05101) (2016). <http://arxiv.org/abs/1605.05101>
45. Zhang, H., Xiao, L., Wang, Y., Jin, Y.: A generalized recurrent neural architecture for text classification with multi-task learning. [arXiv:1707.02892](https://arxiv.org/abs/1707.02892) (2017). <http://arxiv.org/abs/1707.02892>
46. Liu, P., Qiu, X., Huang, X.: Adversarial multi-task learning for text classification. [arXiv:1704.05742](https://arxiv.org/abs/1704.05742) (2017). <http://arxiv.org/abs/1704.05742>
47. Luong, M.-T., Le, Q.V., Sutskever, I., Vinyals, O., Kaiser, L.: Multitask sequence to sequence learning (2015). [arXiv:1511.06114](https://arxiv.org/abs/1511.06114). <http://arxiv.org/abs/1511.06114>
48. Yousif, A., Niu, Z., Chambua, J., Khan, Z.Y.: Multi-task learning model based on recurrent convolutional neural networks for citation sentiment and purpose classification. *Neurocomputing* **335**, 195–205 (2019)
49. Lu, G., Zhao, X., Yin, J., Yang, W., Li, B.: Multi-task learning using variational auto-encoder for sentiment classification. *Pattern Recognit. Lett.* **132**, 115–122 (2020). <https://doi.org/10.1016/j.patrec.2018.06.027>
50. Tang, D., Qin, B., Liu, T.: Document modeling with gated recurrent neural network for sentiment classification. In: Proceedings of the Conference on Empirical Methods Natural Language Processing, pp. 1422–1432 (2015)
51. Wang, R., Li, Z., Cao, J., Chen, T., Wang, L.: Convolutional recurrent neural networks for text classification. In: Proceedings of the International Joint Conference Neural Networks (IJCNN), pp. 2267–2273 (2019)
52. Zhou, C., Sun, C., Liu, Z., Lau, F.C.M.: A C-LSTM neural network for text classification (2015). [arXiv:1511.08630](https://arxiv.org/abs/1511.08630). <http://arxiv.org/abs/1511.08630>
53. Cheng, Y., Yao, L., Xiang, G., Zhang, G., Tang, T., Zhong, L.: Text sentiment orientation analysis based on multi-channel CNN and bidirectional GRU with attention mechanism. *IEEE Assess* **8**, 134964–134975 (2020)
54. Ruder, S., Ghaffari, P., Breslin, J.G.: A hierarchical model of reviews for aspect-based sentiment analysis. In: Proceedings of the Conference on Empirical Methods Natural Language Processing, pp. 999–1005 (2016)
55. Rao, G., Huang, W., Feng, Z., Cong, Q.: LSTM with sentence representations for document-level sentiment classification. *Neurocomputing* **308**, 49–57 (2018)
56. Sachin, S., Tripathi, A., Mahajan, N., Aggarwal, S., Nagrath, P.: Sentiment analysis using gated recurrent neural networks. *SN Comput. Sci.* **1**(2), 1–13 (2020). <https://doi.org/10.1007/s42979-020-0076-y>
57. Wang, X., Jiang, W., Luo, Z.: Combination of convolutional and recurrent neural network for sentiment analysis of short texts. In: Proceedings of the 26th International Conference on Computing Linguistics, Technical Papers (COLING), pp. 2428–2437 (2016)
58. Zhou, P., Qi, Z., Zheng, S., Xu, J., Bao, H., Xu, B.: Text classification improved by integrating bidirectional LSTM with two-dimensional max pooling. [arXiv:1611.06639](https://arxiv.org/abs/1611.06639) (2016). <http://arxiv.org/abs/1611.06639>
59. Zhang, Y., Yuan, H., Wang, J., Zhang, X.: YNU-HPCC at EmoInt-2017: Using a CNN-LSTM model for sentiment intensity prediction. In: Proceedings of the 8th Workshop Computer Approaches Subjectivity, Sentiment Social Media Analysis, pp. 200–204 (2017)

60. Zhang, H., Wang, J., Zhang, J., Zhang, X.: YNU-HPCC at SemEval 2017 task 4: Using a multi-channel CNN-LSTM model for sentiment classification. In: Proceedings of the 11th International Workshop Semantic Evaluation (SemEval), pp. 796–801 (2017)
61. Sun, B., Tian, F., Liang, L.: Tibetan micro-blog sentiment analysis based on mixed deep learning. In: Proceedings of the International Conference on Audio, Language Image Process. (ICALIP), pp. 109–112 (2018)
62. Zhang, Z., Robinson, D., Tepper, J.: Detecting hate speech on Twitter using a convolution-GRU based deep neural network. In: Gangemi, A., et al. (eds.) The Semantic Web: 15th International Conference, ESWC 2018, Heraklion, Crete, Greece, June 3–7, 2018, Proceedings, pp. 745–760. Springer International Publishing, Cham (2018)
63. Abd El-Jawad, M.H., Hodhod, R., Omar, Y.M.K.: Sentiment analysis of social media networks using machine learning. In: Proceedings of the 14th International Computing Engineering Conference (ICENCO), pp. 174–176 (2018)
64. Yang, Z., Yang, D., Dyer, C., He X., Smola, A., Hovy, E.L Hierarchical attention networks for document classification. In: Proceedings of the Conference on North American Chapter Association Computing Linguistics, Human Language Technology, pp. 1480–1489 (2016)
65. Wang, L., Cao, Z., de Melo, G., Liu, Z.: Relation classification via multi-level attention CNNs. In: Proceedings of the 54th Annual Meeting Association Computing Linguistics, pp. 1298–1307 (2016)
66. Wang, Y., Huang, M., Zhu, X., Zhao, L.: Attention-based LSTM for aspect-level sentiment classification. In: Proceedings of the Conference on Empirical Methods Natural Language Processing, pp. 606–615 (2016)
67. Cheng, J., Zhao, S., Zhang, J., King, I., Zhang, X., Wang, H.: Aspect-level sentiment classification with HEAT (HiErarchicalATtention) network. In: Proceedings of the ACM Conf. Information Knowledge Management, pp. 97–106 (2017)
68. Ma, D., Li, S., Zhang, X., Wang, H.: Interactive attention networks for aspect-level sentiment classification, [arXiv:1709.00893](https://arxiv.org/abs/1709.00893) (2017). <http://arxiv.org/abs/1709.00893>
69. Han, H., Li, X., Zhi, S., Wang, H.: Multi-attention network for aspect sentiment analysis. In: Proceedings of the 8th International Conference on Software Computing Application (ICSCA), pp. 22–26 (2019)
70. Gao, Y., Liu, J., Li, P., Zhou, D.: CE-HEAT: An aspect-level sentiment classification approach with collaborative extraction hierarchical attention network. IEEE Assess **7**, 168548–168556 (2019)
71. Yuan, H., Zhang, X., Niu, W., Cui, K.: Sentiment analysis based on multi-channel convolution and bi-directional GRU with attention mechanism. J. Chin. Inf. Process. **33**(10), 109–118 (2019)
72. Socher, R., Perelygin, A., Wu, J., Chuang, J., Manning, C.D., Ng, A.Y., Potts, C.: Recursive deep models for semantic compositionality over a sentiment treebank. In: Proceedings of the Empirical Methods Natural Lang. Process, pp. 1631–1642. MIT Press, Cambridge, MA, USA (2013)
73. Tai, K.S., Socher, R., Manning, C.D.: Improved semantic representations from tree-structured long short-term memory networks. In: Proceedings of the 53rd Annual Meeting Association Computing Linguistics 7th International Joint Conference on Natural Language Processing, vol. 1, pp. 1556–1566 (2015)
74. Cho, K., et al. Learning phrase representations using RNN encoder-decoder for statistical machine translation. [arXiv:1406.1078](https://arxiv.org/abs/1406.1078) (2014). <http://arxiv.org/abs/1406.1078>
75. Rehman, A.U., Malik, A.K., Raza, B., Ali, W.: A hybrid CNN-LSTM model for improving accuracy of movie reviews sentiment analysis. Multimed. Tools Appl. **78**(18), 26597–26613 (2019)

76. Katarya, R., Arora, Y.: Study on text classification using capsule networks. In: Proceedings of the 5th International Conference on Advance Computing Communication System, pp. 501–504 (2019)
77. Bahdanau, D., Cho, K., Bengio, Y.: Neural machine translation by jointly learning to align and translate [arXiv:1409.0473](https://arxiv.org/abs/1409.0473) (2014). <http://arxiv.org/abs/1409.0473>
78. Jianqiang, Z., Xiaolin, G., Xuejun, Z.: Deep convolution neural networks for twitter sentiment analysis. IEEE Assess **6**, 23253–23260 (2018)
79. Al-Twairesh, N., Al-Negheimish, H.: Surface and deep features ensemble for sentiment analysis of Arabic tweets. IEEE Assess **7**, 84122–84131 (2019)
80. Xiao, Y., Zhou, G.: Syntactic edge-enhanced graph convolutional networks for aspect-level sentiment classification with interactive attention. IEEE Assess **8**, 157068–157080 (2020)
81. Dong, Y., Fu, Y., Wang, L., Chen, Y., Dong, Y., Li, J.: A sentiment analysis method of capsule network based on BiLSTM. IEEE Assess **8**, 37014–37020 (2020)
82. Ishaq, A., Asghar, S., Gillani, S.A.: Aspect-based sentiment analysis using a hybridized approach based on CNN and GA. IEEE Assess **8**, 135499–135512 (2020)
83. Kumar, A., Narapareddy, V.T., Aditya Srikanth, V., Neti, L.B.M., Malapati, A.: Aspect-based sentiment classification using interactive gated convolutional network. IEEE Access **8**, 22445–22453 (2020)
84. Tang, D., Qin, B., Liu, T.: Aspect level sentiment classification with deep memory network. In: Proceedings of teh EMNLP, pp. 214–224 (2016)
85. Gu, S., Zhang, L., Hou, Y., Song, Y.: A position-aware bidirectional attention network for aspect-level sentiment analysis. In: Proceedings of the Conference on Computing Linguistics (COLING), pp. 774–784 (2018)
86. Huang, B., Ou, Y., Carley, K.M.: Aspect level sentiment classification with attention-over-attention neural networks. In: Proceedings of the International Conference on Social Computing, Behavioral-Cultural Modeling Predict. Behavior Representation Modeling Simulation, pp. 197–206. Springer (2018)
87. Zhao, N., Gao, H., Wen, X., Li, H.: Combination of convolutional neural network and gated recurrent unit for aspect-based sentiment analysis. IEEE Access **9**, 15561–15569 (2021)
88. Cai, R., et al.: Sentiment analysis about investors and consumers in energy market based on BERT-BiLSTM. IEEE Access **8**, 171408–171415 (2020)
89. Aslam, N., Rustam, F., Lee, E., Washington, P.B., Ashraf, I.: Sentiment analysis and emotion detection on cryptocurrency related tweets using ensemble LSTM-GRU model. IEEE Access **10**, 39313–39324 (2022)
90. Cheng, Y., et al.: Sentiment analysis using multi-head attention capsules with multi-channel CNN and bidirectional GRU. IEEE Access **9**, 60383–60395 (2021)
91. Hameed, Z., Garcia-Zapirain, B.: Sentiment classification using a single-layered Bi-LSTM model. IEEE Access **8**, 73992–74001 (2020)



Fake Product Detection Using Blockchain with Encryption and AI

Shyam Pandit, Abhay Jakhere, Siddharth Tated^(✉), Omkar Bhakare,
and Chandan Prasad

Department of Information Technology, JSPM's Rajarshi Shahu College of Engineering, Pune,
Maharashtra 411033, India
Siddharthtated05@gmail.com

Abstract. The ubiquity of counterfeit goods in today's global market is a serious problem that can cause firms to suffer large financial losses as well as consumer health risks. This abstract presents a potential fake goods detection system that leverages blockchain technology's inherent security properties to address the shortcomings of existing approaches. Blockchain's openness gives customers access to detailed information about the goods they've bought, which increases their trust in the legitimacy of their purchases. A safe system for storing and retrieving data is ensured by the distributed, decentralised, and tamper-resistant nature of blockchain technology, which offers a practical way to detect and reduce the presence of counterfeit goods. The importance of using blockchain technology to improve the security and transparency of global supply chains is highlighted in this abstract, which will eventually protect consumers and businesses from the negative impacts of counterfeit goods.

Keywords: Fake Product Detection · Fake Product · Counterfeit Product · Blockchain · First Section

1 Introduction

Since counterfeiting has a negative impact on the economy and its population, it is a serious problem that requires appropriate solutions. It goes beyond monetary losses since purchasing fake items can seriously harm customers' health. Substandard makeup products can result in skin conditions, rashes, and allergic responses, while fake electrical parts in devices can fail and endanger users' safety as well as cause property damage. When counterfeited, even seemingly innocuous items like clothes and shoes can be uncomfortable and fall short of quality standards. The reputation of a business is also permanently harmed by counterfeiting. Consumer complaints, demands for reimbursement, and damage to a company's reputation result from consumers holding legitimate businesses responsible for flaws or malfunctions with counterfeit goods. Businesses now have to handle both shoddy imitations and client concerns, which puts them in a challenging situation. The relationships between distributors and retailers, as well as between customers and legal firms, are negatively impacted by individuals or entities engaged

in the production of counterfeit goods. The market's trust is weakened as a result of counterfeiters' actions. Customers, wholesalers, and retailers are therefore reluctant to conduct business with or lend support to legitimate companies since they are dubious about the legitimacy and calibre of the goods.

Effective mitigation strategies have shown to be successful in the fight against counterfeiting in international supply chains. Transparency in the network is essential for allowing traceability and visibility across the supply chain. It is imperative to proactively identify and resolve vulnerabilities that counterfeiters exploit by implementing measures such as cost control and pre-supply review. Good supplier connections promote cooperation and trust, which helps to mitigate the possibility of counterfeit goods entering the market.

A suggested approach combines artificial intelligence (AI) and blockchain technologies to address this widespread problem. This hybrid system combines AI's sophisticated pattern detection and learning capabilities with Blockchain's security and transparency advantages. Because of its openness, immutability, and decentralised consensus, the Blockchain component enhances trust and accountability by offering a secure environment for tracking the products' supply chain. By examining trends and abnormalities in product attributes, artificial intelligence (AI) plays a part in proactively identifying counterfeit items. This all-encompassing strategy gives stakeholders the ability to confirm product integrity, guaranteeing authenticity and shielding customers from the dangers of buying fake goods. In conclusion, the negative impacts of counterfeiting call for practical remedies. The suggested AI-Blockchain hybrid system provides a comprehensive strategy to improve trust, accountability, and consumer protection along the entire supply chain.

2 Literature Survey

See (Table 1).

3 Impact of Blockchain

The integration of blockchain technology in fake product detection can have significant implications for both industries and consumer.

3.1 Consumer Confidence

Because blockchain technology offers a visible and verifiable record of a product's origin and history, it can improve consumer confidence in false product identification. Through this verification procedure, customers are safeguarded from counterfeit goods and are given the power to make informed purchasing decisions. They also ensure that they receive real products.

Table 1. Literature on different methodologies for fake product detection and blockchain

Sr. No.	Paper Title	Publication Details	Author Name	Limitations	Challenges
1	Fake Product Detection System using Blockchain	IEEE 2022	Aadeesh Bali, Amrit Singh and Sunandan Gupta	Use of fake QR code	Managing Blockchain ledger
2	Fake Product Detection Using Blockchain Technology	IJARIIE 2021	Tejaswini Tambe, Sonali Chitalkar, Manali Khurud, Madhavi Varpe, S. Y. Raut	Technical Complexity	Efficiency
3	Blockchain based Fake Product Identification System	IRJMETS 2022	Swaroop Jambhulkar, Harsh Bhoyar, Shantanu Dhore, Arpita Bidkar,	Adaptability	Complexity of algorithm
4	Detection of Counterfeit Products using Blockchain	ITM Web of Conferences 2022	Kunal Wasnik, Isha Sondawle, Rushikesh Wani, Namita Pulgam	Fake RFID tags	Data Accuracy
5	Anti-Counterfeit Product System Using Blockchain Technology	IJRASET 2021	Ishaan Singhal, Himanshu Singh Bisht, Yogesh Sharma	Less Secure	Privacy and confidentiality
6	Anti-Counterfeiting Blockchain Using a Truly Decentralized, Dynamic Consensus Protocol	PDX Scholar	Naif Alzahrani, Nirupama Bulusu	Technical Complexity	Complexity of algorithm

(continued)

Table 1. (*continued*)

Sr. No.	Paper Title	Publication Details	Author Name	Limitations	Challenges
7	Research in Blockchain-Based Manufacturing Supply Chain	IEEE 2023	Alex Kibet Langat; Yu Guo	Cost	Scalability
8	Food Supply Chain Tracebility System using Blockchain Technology	IEEE2022	G Sai Radha Krishna;P Rekha	Lack of Standardization	Data Accuracy

3.2 Supply Chain Efficiency

The use of blockchain technology for product tracking and verification can simplify supply chain procedures by diminishing the intricacy of conventional paper-based paperwork and manual checks. The supply chain can operate more efficiently and cost-effectively by avoiding delays, mistakes, and disagreements thanks to the transparency and precision offered by blockchain technology.

3.3 Brand Protection

Blockchain gives brands the ability to actively fight counterfeiting, thereby safeguarding their reputation. Brands may show their dedication to authenticity and gain the trust of their customers by putting blockchain-based solutions into practice. Furthermore, blockchain technology can offer useful information and insights to spot trends in counterfeit activity and pinpoint the origin of phoney goods.

3.4 Smart Contracts

Smart contracts, which are self-executing contracts with predetermined terms and conditions, can be included into blockchain technology. Because they may automate some processes and trigger actions based on predefined criteria, smart contracts can be extremely useful in the detection of bogus products. For instance, a smart contract can be designed to immediately alert the appropriate parties in the event that it detects any differences and verify the legitimacy of a product using the data stored on the blockchain.

4 Use of AI

AI plays a crucial role in detecting fake products across various industries, helping to protect consumers from counterfeit goods. Here are some ways AI is used in fake product detection:

4.1 Image Recognition

Visual Inspection: Product photos can be analysed by AI-powered image recognition systems and contrasted with real ones. Any differences in details, branding, or packaging can be reported for additional research.

Watermark Analysis: Artificial intelligence is capable of detecting security features or hidden watermarks that are hard for counterfeiters to copy in product photographs.

4.2 Natural Language Processing (NLP)

Review Analysis: Artificial intelligence systems are capable of examining customer reviews and product reviews to look for trends that might point to fake goods. Unusual wording, a steady stream of unfavourable reviews, or an unexpected surge in reviews may indicate that more research is necessary.

4.3 Machine Learning Models

Pattern Recognition: To identify patterns and traits linked to counterfeit goods, machine learning models can be trained on big datasets of real and phoney products. Subsequently, these models can be employed to identify comparable trends in novel items.

4.4 Data Analytics

Market Analysis: Artificial intelligence algorithms are capable of examining big datasets from online marketplaces in order to spot patterns and trends linked to fake goods. Alerts may be set off by abrupt price reductions, peculiar seller behaviour, or large quantities of a specific product.

4.5 Deep Learning for Document Authentication

Verification papers: Deep learning models can be used to evaluate and authenticate papers like certificates, licences, and permits in sectors like pharmaceuticals where official documentation is essential for product verification.

5 Present System

The current method uses a traditional online database, accessible only with manufacturer logins, to store product data. Product information is entered into the database by manufacturers, and consumers then visit the website and enter particular product codes to confirm the legitimacy of the products. However, this method's susceptibility to manipulation and tampering is a major downside. This database structure is easily altered by unauthorised parties due to its simplicity. One significant worry is that fake goods would be able to get hold of and abuse legitimate product codes, which would compromise the verification process' efficacy. This flaw emphasises the requirement for a more reliable and impenetrable mechanism to support the integrity of the product authentication procedure (Fig. 1).

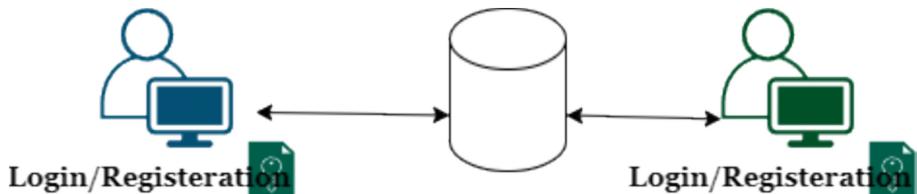


Fig. 1. Present System

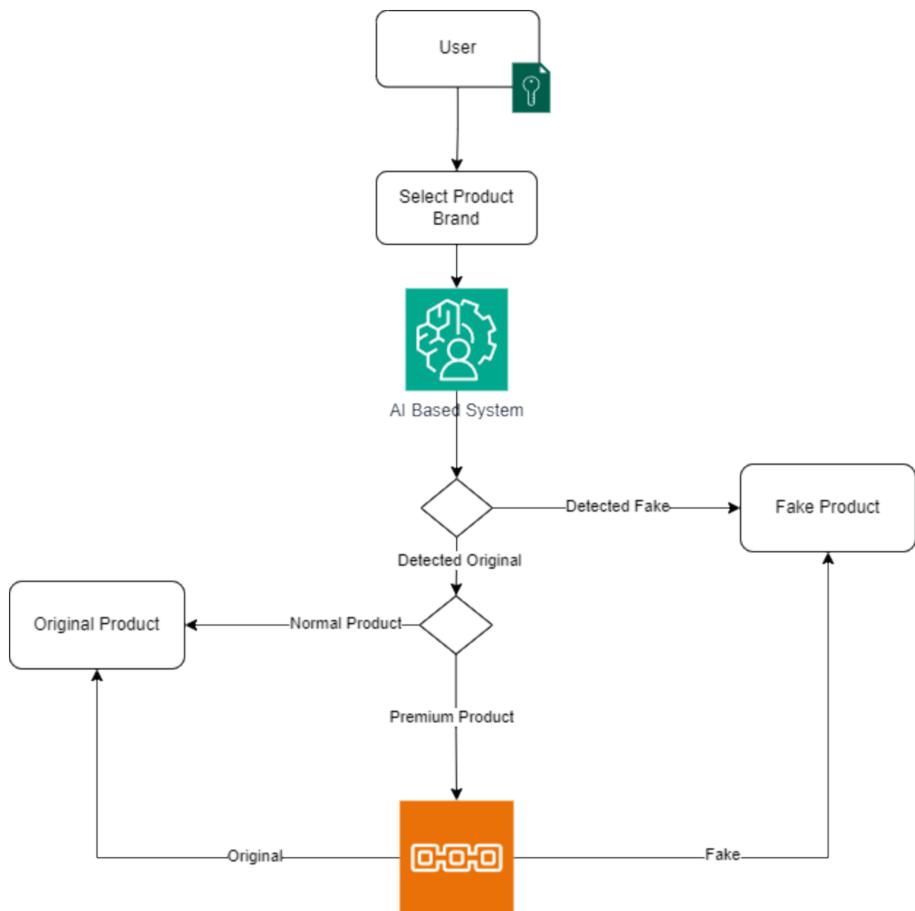


Fig. 2. System Architecture

6 Working Model

6.1 Proposed System

The suggested solution integrates Artificial Intelligence (AI) and Blockchain, two popular technologies for detecting bogus products. Blockchain by itself can solve this problem, but it has several limitations, such as bulkiness, high resource requirements, and the requirement for specialised devices in order to read QR codes (Fig. 2).

Artificial Intelligence (AI) has been added to Blockchain to improve system efficiency and lessen its limits. Algorithms and Deep.

Learning (DL) capabilities are utilised by AI, a flexible technology that is applied in a variety of fields like robotics, engineering, and medicine. These algorithms are capable of conducting tests, gathering data from various datasets, training models, and identifying patterns.

Even for manufacturers, it might be difficult to tell authentic products from counterfeit ones in situations involving high-end goods. By enabling the hash of the product to be added to a blockchain block, blockchain provides a solution and guarantees safe retrieval when needed.

The suggested system's architecture takes a hybrid approach, combining Blockchain's safe and transparent record-keeping capabilities with AI's skills in pattern identification and data processing. The goal of this integration is to build a strong system that combats the problems caused by counterfeit goods, especially in situations where traditional techniques like visual examination may not be sufficient.

6.2 AI Based Detection

See (Fig. 3).

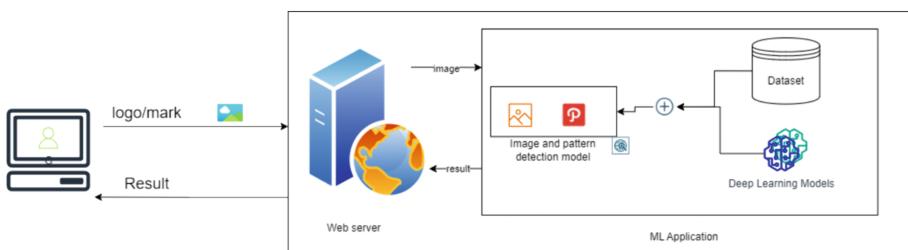


Fig. 3. AI Based Detection

Several essential elements make up the architecture of the AI-based Fake Product Detection system. First, a scan is performed on the product image or logo to uncover details that are only visible to the producers. These unique traits are present in the datasets used to train the model. The logo is sent to a web server after the training stage.

The ML programme then receives the image from the web server and is in charge of confirming the legitimacy of the pictures. An image and pattern detection model that has

been trained on various datasets of the corresponding products handles this verification process. With their ability to extract features and continuously learn from new photos, DL models actively assist this process.

The technology decides whether the image corresponds to a genuine or phoney product after successful recognition. After that, the data is sent back to the server, which notifies the client of the outcome. Through the analysis of hidden features and patterns inside product logos or photos, this integrated approach—which combines machine learning, deep learning, and image and pattern detection—aims to give a comprehensive solution for detecting counterfeit products.

A variety of specialised models are used by the AI-based Fake Product Detection system to provide thorough analysis. Convolutional Neural Networks (CNNs) are highly effective in identifying complex patterns and features in product photos, identifying the distinctive visual components linked to authentic products. Siamese Networks support one-shot learning scenarios by confirming whether an image matches the unique properties that were learned during training. Generative Adversarial Networks (GANs) are essential for creating fake product photos, which complement the dataset and improve the model's ability to identify minute.

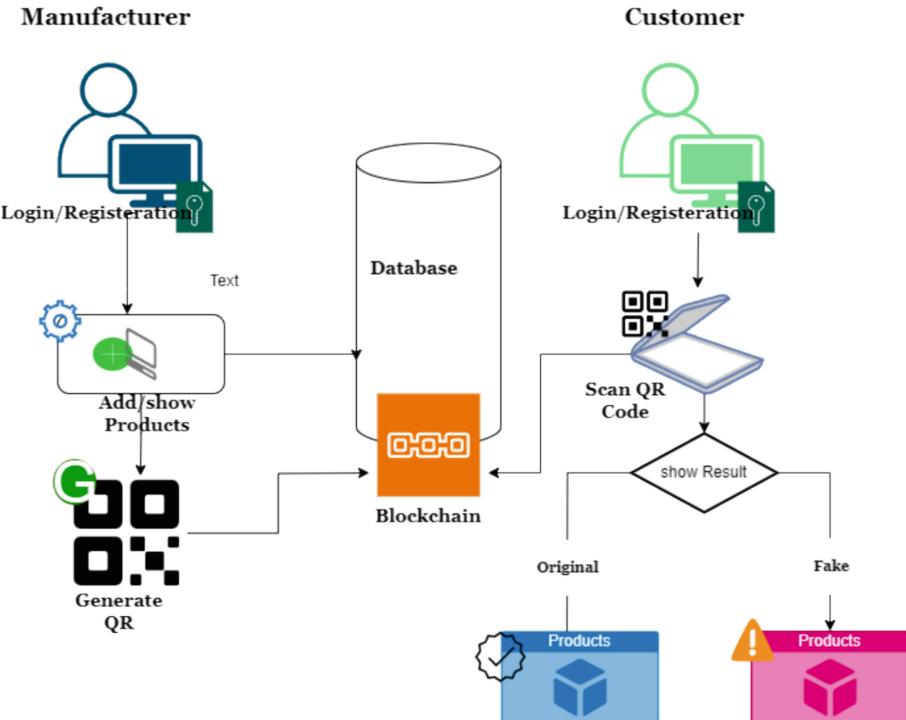
6.3 Blockchain Based Detection

See (Fig. 4).

Supply chain complexities mean that wholesalers and retailers must work together to contribute digital data to the blockchain as products move through the chain. Important information contained in these records includes the location of the goods, sales transactions, pricing, and relevant information about the seller and brand.

With the use of a specialised mobile app, customers may easily combine the process of verifying the legitimacy of a product. After they log in, they can quickly scan a QR code with their cellphones to see a complete history that is safely kept on the blockchain, confirming the product's legitimacy. If a counterfeit is discovered, the system quickly alerts the maker and provides vital location information. This user-friendly strategy protects brand reputation and revenue by empowering businesses to quickly combat counterfeiting and fostering customer confidence in the authenticity of products. The effectiveness of blockchain in creating a safe and open market is highlighted by the smooth integration of technology and real-time notifications.

Manufacturers log onto the system to start the process of adding products to the blockchain. The encrypted product code is sent to the system upon login, and it is carefully added to the blockchain. This encrypted identity adds to the stability of the blockchain-based system by acting as a distinct digital fingerprint for every product. With a tamper-resistant ledger guaranteed by the secure connection, manufacturers can be confident that their supply chain trip and product origins are accurately and permanently documented. Thanks to this simplified method, manufacturers can now proactively manage and validate their product information inside the blockchain ecosystem, greatly improving traceability and accountability.

**Fig. 4.** Blockchain based Detection

6.4 Key Generation and Encryption

See (Fig. 5).

**Fig. 5.** Key generation and Encryption

We encrypt the product code in the suggested system using the very effective symmetric key method, SHA-256. SHA-256 encryption generates hash codes that the system compares to verify authenticity instead than simply comparing product codes. These hash codes are safely kept inside the system and serve as distinct digital fingerprints. By using hash code comparisons to verify the accuracy of product information, this method greatly improves security and creates a strong, impenetrable mechanism. The product verification system is more secure and efficient overall since SHA-256 is incorporated, which guarantees a reliable and efficient encryption procedure.

6.5 User Registration

One strategic answer to the possible issue of counterfeit QR codes attached to counterfeit products is to register consumers to specific products on the blockchain. Through the safe storage of ownership information in an unchangeable ledger, this method creates a distinct relationship between every authorised user and their particular goods. The registration process reduces the possibility of tampering by ensuring that only authorised users can carry out ownership transfers.

By cross-verifying the ownership of a product against the registered user, the blockchain adds an extra degree of protection against fraudulent activity occurring inside the supply chain when customers authenticate it using a QR code. This strong method makes sure that product verification not only validates the authenticity of the product but also the legitimate owner, greatly strengthening the security framework overall and protecting against unauthorised changes or manipulations.

7 Conclusion

In summary, current studies highlight the potential of machine learning algorithms—more especially, Random Forests—to improve the efficacy of systems for detecting counterfeit items. By adding more layers of complexity to the identification process, this AI-driven method enhances the system's capacity to recognise minute patterns and anomalies in product attributes. A hybrid system is created by fusing blockchain technology, hash functions, and machine learning algorithms. This system provides a transparent and safe way to identify counterfeit items.

The system's incorporation of blockchain technology enables smooth verification, hence mitigating the likelihood of counterfeit goods entering the market. Hashing values can be compared to verify a product's legality, offering a reliable authentication method. Moreover, the blockchain ledger functions as an all-inclusive documentation, facilitating the tracing of the product's origin and the identification of any attempted manipulation.

This hybrid solution provides a thorough method of battling fake goods since it combines blockchain technology with AI-driven machine learning algorithms. By using AI's learning capabilities, it not only improves the accuracy of product verification but also makes proactive actions possible. This tactic is well-positioned to safeguard customer happiness and safety while successfully stopping the sale of fake goods. Adoption of this hybrid technology might, in essence, completely transform the way that we detect and prevent the sale of counterfeit goods in a variety of businesses.

References

1. Ma, J., Lin, S.-Y., Chen, X., Sun, H.-M., Chen, Y.-C., Wang, H.: A blockchain-based application system for product anti-counterfeiting. *IEEE Access* (2020)
2. Tambe, T., Chitalkar, S., Khurud, M., Madhavi Varpe, S., Raut, Y.: Fake Product Detection Using Blockchain Technology. *IJARIIE-ISSN(O)-2395-4396*
3. Jambhulkar, S., Bhoyar, H., Dhore, S., Bidkar, A., Desai, P.: Blockchain Based Fake Product Identification System. *Int. Res. J. Modernization Eng. Techno. Sci.* (2022)

4. Wasnik, K., Sondawle, I., Wani, R., Pulgam, N.: Detection of counterfeit products using Blockchain. *ITM Web Conf.* **44**, 03015 (2022)
5. Singhal, I., Bisht, H.S., Sharma, Y.: Anti-counterfeit product system using blockchain technology. *Int. J. Res. Appl. Sci. Eng. Technol.*
6. Vidhya Lakshmi, G., Gogulamudi, S., Nagaeswari, B., Reehana, S.: Blockchain based inventory management by QR code using open CV. In: International Conference on Computer Communication and Informatics (ICCCI -2021) Coimbatore, India. 27–29 Jan 2021
7. Sanghi, A., Aayush, Kata war, A., Arora, A., Kaushik, A.: Detecting fake drugs using blockchain. *Int. J. Recent Technol. Eng.* **10**(1) (2021), 100-109
8. Sanjay, K.S., Danti, A.: Detection of fake opinions on online products using Decision Tree and Information Gain. In: Third International Conference on Computing Methodologies and Communication (ICCMC 2019) (2019)
9. Montes, J.M., Ramirez, C.E., Gutierrez, M.C., Larios, V.M.: Smart Contracts for supply chain applicable to Smart Cities daily operations. In: 5th IEEE International Smart Cities Conference (ISC2 2019) (2019)
10. Hongekar, A., Jaju, A., Bhargade, P., Acharya, N., Atul Pawar, P.: A survey on fake product identification system. In: 18th International Conference e-society (2020)
11. Srikrishna Shastri, C., Vishal, K., Sushmitha, S., Lahari, Ashwal, R.S.: Fake product detection using blockchain technology. 2019 Dissertations and Theses. Paper 5038
12. Jambhulkar, S., Bhoyar, H., Dhore, S., Bidkar, A., Desai, P.: Blockchain based Fake Product Identification System. *Insight – Information*, 2. 10. 18282/ii. v2i2.365
13. Ma, J., Lin, S.-Y., Chen, X., Sun, H.-M., Chen, Y.-C.: A blockchain-based application system for product anti-counterfeiting. *J. Sensor Actuator Netw.* (2019)
14. Funde, A., Nahar, P., Khilari, A., Marne, N., Nerkar, N.: Blockchain based fake product identification in supply chain. *Int. J. Res. Appl. Sci. Eng. Technol.* (2021)
15. Mitbavkar, T., Pedamkar, S., Kuvalakar, S., Wasnik, K.: Fake product detection using Blockchain. *Int. J. Res. Appl. Sci. Eng. Technol.* (2021)
16. Kiruthika, C., Manaswini, H., Sherin, R., Subalakshmi, S., Padmapriya, M.: Block chain in fake product identification system using QR code. *Int. Res. J. Modernization Eng. Techno. Sci.* 2582–5208 (2021)
17. Nila, U., Luther, A., Vignesh, A.: Block chain in fake product identification system using QR code. *Appl. Sci.* **11**(12), 5585 (2021)
18. Panigrahi, S., Rai, A.K., Rajput, A.K., Bhardwaj, A.: Fake news detection using Blockchain. *J. Online Informatika* **5**(2), 239–244 (2020)
19. Kalpana Devi, S., Samy Durai, K., Shri Balaji, A.M., Ravi Kumar, J.: Fake product identification with the help of block chain technology. *J. Online Informatika* **5**(2), 239–244 (2020)



S-Defender: A Smishing Detection Approach in Mobile Environment

Ankit Kumar Jain¹ , Ankur Panday², and Diksha Goel³

¹ National Institute of Technology, Kurukshetra, India

ankitjain@nitkkr.ac.in

² Manipal University Jaipur, Jaipur, India

³ University of Adelaide, Adelaide, Australia

Abstract. This paper introduces an accurate smishing filtering model named S-Defender, designed to detect smishing messages. The model analyzes the text message content and employs a Naive Bayesian classifier to categorize the message as either smishing or ham. Additionally, our approach normalizes and converts Lingo language text messages into their standard form to improve classification accuracy. The model's validation, conducted using the English SMS dataset, resulted in an overall accuracy of 94.72%.

Keywords: Smishing · Mobile Phishing · Short Messaging Service · Smartphone

1 Introduction

Smartphones have become popular within a short period of time due to their small screen size, longer battery life and portability. The popularity of these gadgets has attracted the cybercriminals for performing various attacks on these devices, such as SMS spam, Phishing, Denial of service attack, social engineering attack, etc. [1]. Earlier email was widely used by the attackers to reach users but now attackers have found alternatives for email to reach the targeted users that includes messaging based services, such as SMS and WhatsApp. SMS is one of the most widely used text-based communication service. It is used for both personal as well as professional communication. Now a days, SMS is widely used by the organisation to communicate with their customers for promotions, delivering the products or for conducting a survey. According to Portio research report, worldwide mobile messaging revenue increased from \$128 billion to \$227 billion between 2012 to 2017 [2].

Word ‘Smishing’ is composed of two words – SMS and phishing [3]. When SMS is used to perform phishing attack then it is called as smishing attack. Smishing is a variant of phishing attack in which an attacker sends an SMS to the users that appears to be a genuine message and trick them into revealing their credentials such as credit card details, login ids and passwords.

In order to reach mobile device users, attackers are using SMS-based services. Moreover, trust level of users on SMS has attracted the attacker to choose SMS as a medium

of attack. Figure 1 presents the SMS traffic growth from 1999 to 2015 [2]. Now-a-days instead of email, attackers are preferring SMS for tricking users. There are 7 billion mobile subscribers worldwide, with only 2.5 billion email users [4]. Furthermore, SMS boasts a 98% open rate and a 45% response rate, while email struggles with a mere 20% open rate and a 6% response rate, consequently diminishing the success rate of attacks [4]. A report reveals that 33% of mobile users have received fake messages offering various attractive deals and discounts [5]. Additionally, one out of every four unsolicited SMS messages aims to pilfer sensitive user information [7]. However, the threat is somewhat reduced in the case of messaging applications. Twenty-six percent of messaging application users receive unwanted text messages daily, while 49% receive at least one message per week. A significant number of people trust SMS as a medium of text-based communication compared to other messaging applications [8]. Approximately 35% of users consider SMS the most reliable mode of communication, 28% feel messaging applications like WhatsApp are more secure, and 18% of users trust Facebook, Skype, and Yahoo [8].

In an SMS phishing scam, a consumer of the UK bank Santander lost £22,700 in 2016 [9]. Due to a lack of security mechanisms and awareness among users in detecting smishing messages, attackers actively use SMS as a medium for carrying out attacks [6]. Users are often unaware of the appropriate actions to take upon receiving smishing messages. The MEF report [8] states that approximately 70% of users do not take any action against unwanted messages. Upon receiving unsolicited messages, 54% of users delete them, 17% ignore them, 13% respond with “stop,” and only 16% of users report them (13% to the operator and 7% to the company), as shown in Fig. 1 [8].

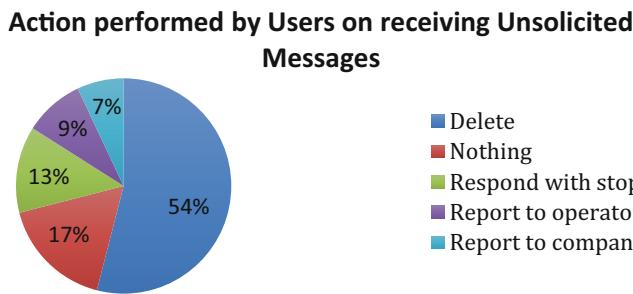


Fig. 1. Action performed by users on receiving unsolicited messages

This paper presents a novel smishing detection approach ‘S-Defender’ which is capable of defending the users against smishing attacks. It is a content-based analysis approach that analyses the URL and mobile number of the sender. Moreover, our approach filters the self-answering messages as per the requirements of the user by maintaining a ‘UserInterestList’. In addition, we have performed analysis using various machine learning algorithm but Naïve Bayes Algorithm gives maximum accuracy. Since now-a-days, abbreviations are used in text-based communication, so we have normalized and replaced the text in the message by its standard form. We have used two dictionaries

namely English and Lingo dictionary. Following are the key contributions of proposed method:

- Detection Accuracy: Our technique provides reasonable amount of accuracy in detection of smishing messages. False positive and false negative rates are also low.
- Platform independent: A large number of platforms are available for mobile phones. The smishing detection technique should be such that it can be used for all mobile device platforms. We have not considered any platform dependent features, due to which our approach is able to work for all the devices.
- Self-learning: Our approach is able to learn from its past decisions and update itself so that it is able to detect new patterns during further classification.
- Maintain privacy of user: SMS has become an important medium of communication and hence, it may also contain crucial information. Therefore, messages should not be revealed to any third party. The proposed approach maintains the privacy of the user.

The rest of the paper is structured as follows: Sect. 2 discusses the Smishing attack procedure. Section 3 presents our proposed smishing detection model ‘S-Defender.’ Sect. 4 covers the experimental work, including dataset collection, classification techniques, features used, and experimental details, including results. Finally, Sect. 5 concludes the paper.

2 Smishing Attack Procedure

Figure 2 shows the lifecycle of smishing attack.

Phase 1: Design phishing webpage or application – In the first step, the cybercriminal constructs a fake application or a webpage of a well-known organization that looks like the legitimate one. This malicious application or webpage is hosted by the attacker over the fake web server [10].

Phase 2: Send phishing SMS – Attacker composes a text message that may contain malicious link and send this text message to the users via SMS or other messaging applications. Attacker may use social engineering technique to reach a large number of mobile users.

Phase 3: Redirecting User – After receiving text message, the user opens and reads the message. When the user clicks on the embedded link contained in the message, he is directed to a malicious webpage containing the login form or a webpage containing the malicious application.

Phase 4: Installing malicious application or visiting phishing webpage – When the user clicks on the URL, a fake webpage appears. Now either user is asked to fill the form or is tricked into installing a malicious application.

Phase 5: Identity theft – The fake login page in which users have filled the credentials or the malicious application installed on the device, sends the credentials to the attacker. Attacker uses this sensitive information for performing malicious activities or for financial gain.

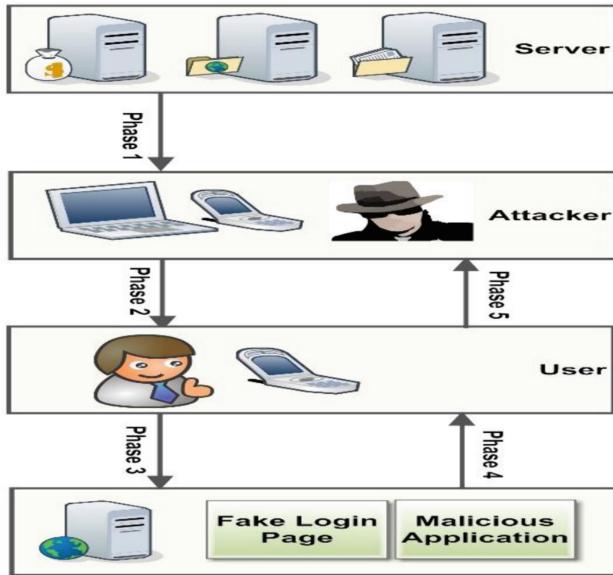


Fig. 2. Smishing attack procedure

3 The Proposed Framework: S-Defender

3.1 Architecture of the Proposed Framework

S-Defender consists of three phases: SMS analysis phase, SMS pre-processing phase, and SMS classification phase. Figure 3 represents the architecture of S-Defender.

Phase 1: SMS Analyzer

In this phase, basic filtering is performed based on the characteristics of smishing messages. This phase comprises three modules: Mobile Number Analyzer, URL Analyzer, and Self-Answering Message Analyzer.

Mobile Number Analyzer: It analyzes the mobile number of the sender from which the SMS is received. It checks whether the mobile number of the sender is blocked by the user or not. This feature addresses the vulnerability of smartphones where blocking a number often only blocks phone calls and not SMS messages.

URL analyser: It checks if the message contains any URL or not. Since now-a-days URL is used by the attackers for tricking users into downloading malicious applications. So if the message contains URL, then S-Defender checks if this URL is present in the blacklist or not. If blacklist does not contain this URL, then it checks if this URL leads to download of an APK file. The APK file may steal personal information of the users. It might be possible that URL redirects the user to a webpage, so we check for the presence of sensitive terms on the webpage. In both cases i.e. downloading of the APK file and presence of sensitive terms on the webpage, the message is categorised as smishing message.

Self-Answering Message Analyser: This module is used to avoid the spam messages. It blocks the messages which ask to subscribe any service. In most of the cases, self-answering messages are regarded as spam by the users, but there may be some services which are of interest to the users and these interests vary from user to user. S-Defender maintains a list ‘UserInterestList’ where users can define their own ham keywords and with the help of this list, self-answering message which are of user’s interest will not be categorized as spam and are delivered to the users.

Phase 2: SMS Pre-processing

In this phase, the SMS is pre-processed and normalized. This phase is basically used to enhance the classification accuracy of the proposed framework. Pre-processing includes identification of boundaries of sentences, division of text messages into tokens, conversion of SMS text message into lowercase characters and normalization of the text.

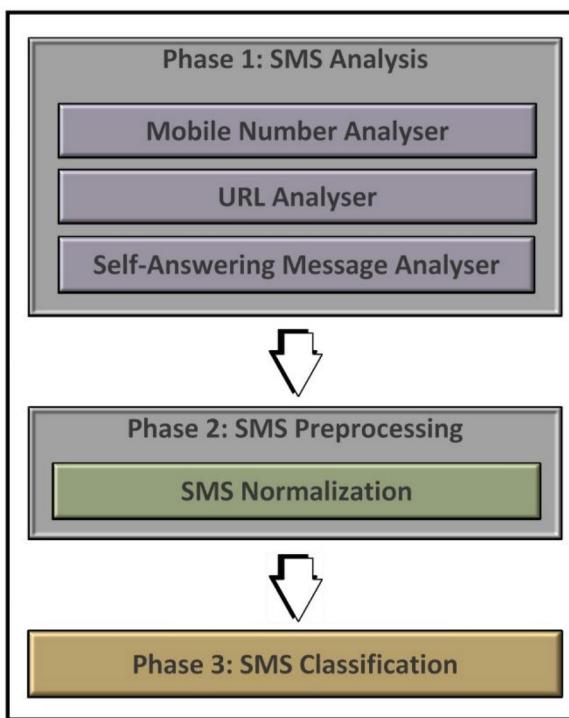


Fig. 3. Architecture of S-Defender

Normalization: It is the process of normalizing noisy text and replacing this text by its standard form. Nowadays, Lingo language like short forms and abbreviations are used in text-based communication. It is also popularly used in text based communication. In this phase, two dictionaries are used for normalizing the text – English dictionary [11] and Lingo dictionary [12]. Lingo dictionary is used to convert Lingo words present in the

message into their standard form. For example, ‘gm’ or ‘gudmrng’ will be converted to their standard form ‘good morning’. English dictionary is used to normalize each word into its root form. For example, ‘playing’ or ‘plays’ will be converted to ‘play’.

Phase 3: SMS Classification

In this phase, the pre-processed and normalized text message is provided as input to the NB classifier. It requires a dataset for training, where, during the training stage, the occurrence of each word in smishing and genuine messages is computed. For instance, words like ‘offers,’ ‘lottery,’ ‘password,’ and ‘won’ have a higher smishing probability due to their repeated occurrence in smishing messages. Thus, during the dataset training, we obtain the smishing probability of each word. Subsequently, a Machine Learning Classifier is employed to calculate the smishing probability of the complete message. The smishing probability of the message is then compared with the threshold value. If the smishing probability exceeds the threshold value, the message is considered a smishing message; otherwise, it is categorized as a ham message. S-Defender blocks smishing messages and delivers ham messages to the users.

3.2 Flowchart of the Proposed Framework

The proposed framework analyzes the URL, sender’s mobile number, and checks for self-answering messages. The flowchart of S-Defender is illustrated in Fig. 4. The following steps are involved in the detection of smishing messages.

Step 1: When a device receives an SMS, it first checks if the sender of the SMS is blacklisted. If the sender is blacklisted, then it is identified as a smishing message.

Step 2: If the sender is not blacklisted, the framework checks whether the SMS contains a URL or not.

Step 3: If the URL is not present in the blacklist, the URL is invoked to check if it downloads an APK file. If yes, then it is considered a smishing message.

Step 4: If it does not download an APK file, there is a possibility that the user is directed to a malicious webpage. The framework then checks for the presence of sensitive terms like login, password, and credit card on the webpage. If sensitive terms are detected, it is regarded as a smishing message.

Step 5: The SMS is further examined for the presence of a self-answering message. A self-answering message is one that asks users to reply ‘yes’ if they want to subscribe to services, such as health tips or jokes. While the majority of users consider self-answering messages as spam, some users might be interested in specific services. Users can list keywords of their interest in ‘UserInterestList,’ for example, health, jokes. Self-answering messages containing these keywords are not considered smishing messages, allowing users to filter messages based on their preferences.

Step 6: After basic filtering, the text is normalized and converted into its standard form using two dictionaries: the Lingo dictionary and the English dictionary. This normalization process enhances classification accuracy, resulting in unambiguous and normalized text.

Step 7: The normalized text is then input to the NB Classifier, which calculates the smishing probability of the message based on trained data.

Step 8: If the smishing probability of the message exceeds the threshold value (pre-defined risk value), the message is identified as a smishing message; otherwise, it is categorized as a ham message.

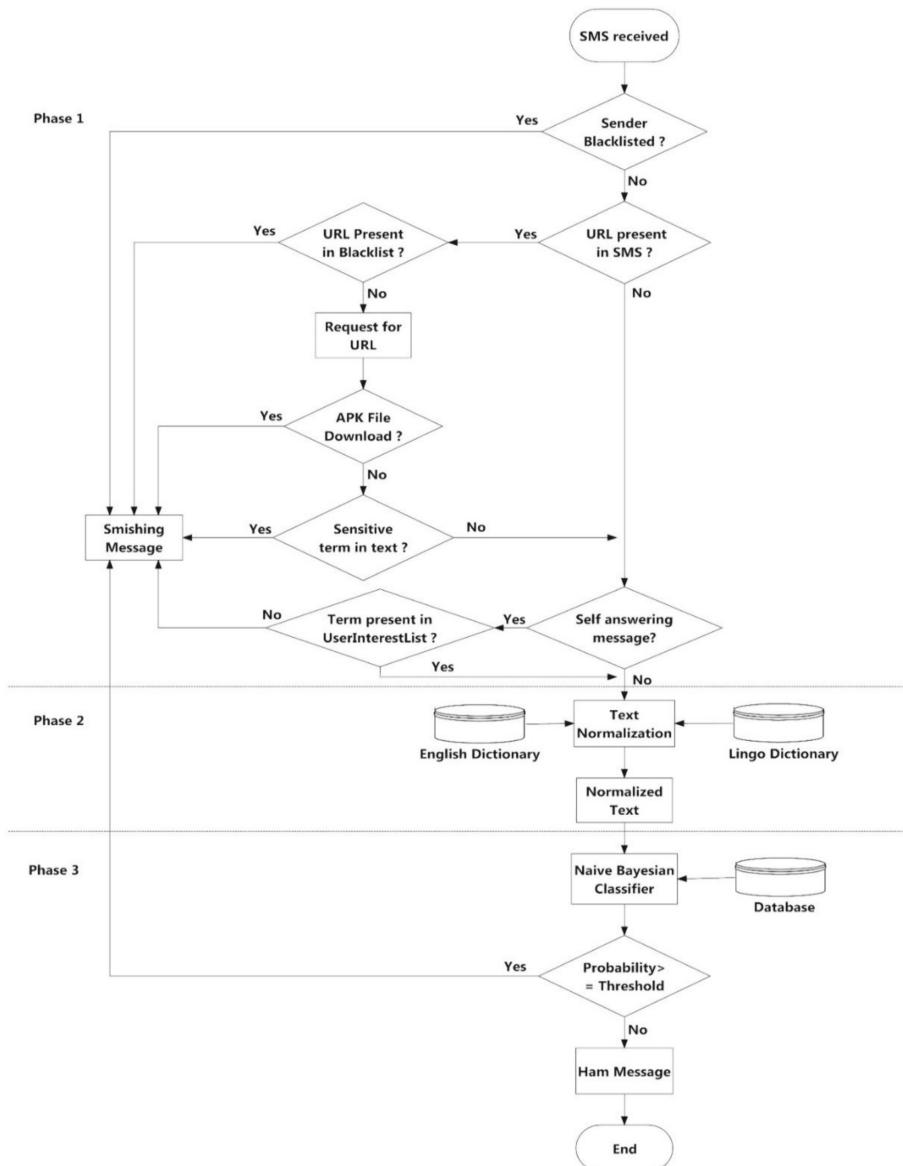


Fig. 4. Flowchart of S-Defender

4 Experimental Work

This section discusses the dataset used for experimental purpose, various features used and results obtained that validate our scheme.

4.1 Dataset Collection

The SMS spam and smishing dataset were sourced from various references, as detailed in [10, 13]. The consolidated dataset comprises a total of 5,169 messages, with 4,807 classified as ham messages and 362 as smishing messages.

4.2 Features Used

We have selected the feature set with great care so as to ensure correct classification of smishing and ham messages. We have identified a set of eight features – $F = \{F_1, F_2, F_3, F_4, F_5, F_6, F_7, F_8\}$, out of which two are ham oriented feature $\{F_1, F_8\}$ and six are smishing oriented features $\{F_2, F_3, F_4, F_5, F_6, F_7\}$. We have labelled the dataset with these features. Each feature represents a value either 0 or 1, where 0 represent ham and 1 represent smishing. The feature set is discussed below:

Greetings tokens (F₁): Greeting tokens are the keywords like ‘hi’, ‘hello’, ‘hey’, ‘good morning’, ‘good night’. Since only genuine users use greeting tokens during their conversations, so the presence of greeting tokens in the message indicates that it is a ham message.

Presence of URL (F₂): The attacker incorporates the URL into the message to redirect users to a malicious webpage. This feature assesses the presence of a URL in the message. We also analyze the URL to determine if it downloads an APK file or redirects users to a malicious login page. In both of these scenarios, the message is classified as a smishing message.

Message length (F₃): Message length refers to the overall length of the message, encompassing characters, whitespace, smileys, special symbols, etc. The maximum length limit for a text message is 160 characters. The length of a smishing message is typically larger compared to that of a ham message. Therefore, this feature yields a value of 1 if the message length exceeds 150 characters.

Currency sign(F₄): This feature checks if the Currency sign is present in the message or not. For example £ (Pound) symbols are extensively used in smishing messages in order to represent the cash prizes.

Smishing keywords (F₅): This feature checks if smishing keywords are presents in the text message or not. After analysing various smishing messages, we have come up with effective smishing keyword set. This set consists of nineteen smishing keywords. The presence of any of these keywords in the message indicates that it is a smishing message.

Presence of mobile number (F₆): It checks if the mobile number is present in the text message or not. The attacker uses this tactic to get sensitive information from the user by asking them to call on given number. Therefore, if the mobile number is present in the message, then it is considered as smishing message.

Presence of emails (F₇): It checks for the presence of email address in the message. Cybercriminals send email in the message in order to communicate with the user and to fetch some personal details that can future be used for financial gain. Presence of email address in the message indicates that it is a smishing message.

Presence of emotions (F₈): Emotions symbols are generally present in ham messages. Users generally use emotions while chatting. Example of emotions are love, sad, happy, hurt, sweet etc. Now a days, people are using symbols to represent emotion like:-) is used for happiness while:-(: is used for sadness. Presence of emotion symbols in message indicates that message is a legitimate message.

4.3 Results and Discussions

In this subsection, we describe the prototype of the proposed model ‘S-Defender’. The implementation is done in Python. We have evaluated our scheme on various classifiers named Naïve Bayes, J48, Decision Table and SVM. With the help of selected features and Naïve Bayesian Classifier, overall detection accuracy achieved is 94.72%. J48 algorithm achieves 93.45% classification accuracy, Decision table algorithm was able to achieve 93.39% classification accuracy and 94.01% classification accuracy was achieved by using SVM. Figure 5 shows the accuracy of our scheme on various classifiers. Maximum accuracy is achieved by Naïve Bayes Classifier and this is the reason we have used Naïve Bayes classifier in our scheme. Table 1 shows the confusion matrix. For our scheme, True Positive Rate (TPR) is 94.48% and True Negative Rate (TNR) is 94.74%. The scheme is capable of detecting zero-day attacks.

Table 1. Confusion Matrix

	Smishing Messages 362	Ham Messages 4807
Classified as Smishing	True Positive = 342	False Positive = 253
Classified as Ham	False Negative = 20	True Negative = 4554

4.4 Comparative Analysis

The proposed approach checks for the mobile number of the sender, URL, download of APK file, the presence of sensitive terms on the redirected webpage, self-answering messages. The content of the message is analyzed and text is normalized to its standard form. Comparative analysis of S-Defender with other solutions is shown in Table 2. It shows that our approach is able to achieve better results as compared to other approaches and the reason for this is that we have selected suitable classifier for our scheme. Naïve Bayes Classifier gives good results even when we have small dataset for training and testing. Although SVM achieves comparable performance as that of Naïve Bayes Classifier but Naïve Bayes classifier is preferred as it is computationally faster when it comes to taking decisions. In addition to Naïve Bayes Classifier, effective features set and text normalization technique have enhanced the classification accuracy of our scheme.

Results with different classifiers

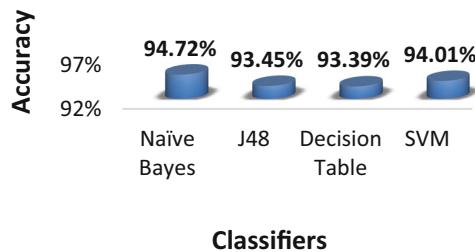


Fig. 5. Accuracy of proposed approach on different classifiers

Table 2. Comparison with existing work

Authors	Algorithm used	Results
Yadav et al. [14]	Bayesian algorithm, SVM	84.75%
Delany et al. [15]	SVM	93.31%
Eshmawi et al. [16]	SVM, CART, RF Naïve Bayes	88%
Idris et al. [17]	Particle swarm optimization	83.20%
Proposed framework	Naïve Bayes Algorithm	94.72%

5 Conclusion

With the surge in smartphone usage, the reliance on SMS for text-based communication has witnessed a corresponding increase. Exploiting this trend, attackers are directing smishing attacks at mobile phone users. Smishing, a security threat, involves an attacker sending an SMS to the user with the intent of pilfering personal information. The use of abbreviations and idioms in these messages adds to the challenge of detection. To address these limitations, we present a novel smishing security approach called ‘S-Defender.’ This approach involves a thorough analysis of the sender’s URL and mobile number. The model scrutinizes self-answering messages and maintains a ‘UserInterestList,’ filtering messages based on user preferences. Message content is examined to extract features, and prior to classification, we normalize the text to convert it into a standardized form. Furthermore, a comparative analysis with existing detection approaches demonstrates that our approach not only provides higher accuracy but also covers more security aspects.

References

- Choudhary, N., Jain, A.K.: Towards filtering of sms spam messages using machine learning based technique. In: Singh, D., Raman, B., Luhach, A.K., Lingras, P. (eds.) Advanced Informatics for Computing Research. CCIS, vol. 712, pp. 18–30. Springer, Singapore (2017). https://doi.org/10.1007/978-981-10-5780-9_2

2. Worldwide A2P SMS Markets 2014–2017. <https://www.xconnect.net/wp-content/uploads/worldwide-sms-markets-portio-strikeiron.pdf>. Accessed July 2017
3. Goel, D., Jain, A.K.: Mobile phishing attacks and defence mechanisms: state of art and open research challenges. *Comput. Secur.* **73**, 519–544 (2018)
4. 2016 Daily SMS Mobile Usage Statistics. <https://www.smseagle.eu/2017/03/06/sms-mobile-statistics-2/>. Accessed July 2023
5. Phishingpro. <http://www.phishingpro.com/>. Accessed April 2023
6. The Human Factor 2017. <https://proofpoint.com/us>. Accessed April 2023
7. The Social Engineering Framework. <https://www.social-engineer.org/framework/attack-vectors/smishing/>. Accessed July 2023
8. MEF Mobile Messaging Fraud Report 2016. https://mobileecosystemforum.com/wp-content/uploads/2016/09/Fraud_Report_2016.pdf. Accessed July 2023
9. Smishing. <http://resources.infosecinstitute.com/category/enterprise/phishing/phishing-variations/phishing-variations-smishing/>. Accessed July 2023
10. Jain, A.K., Gupta, B.B., Kaur, K., Bhutani, P., Alhalabi, W., Almomani, A.: A content and URL analysis-based efficient approach to detect smishing SMS in intelligent systems. *Int. J. Intell. Syst.* **37**(12), 11117–11141 (2022)
11. Freelng English dictionary. <http://devel.cpl.upc.edu/freeling/>. Accessed Sep 2023
12. NoSlang: Internet Slang Dictionary & Translator, <http://www.noslang.com/dictionary/full/>. Accessed Sep 2023
13. Jain, A.K., Yadav, S.K., Choudhary, N.: A novel approach to detect spam and smishing SMS using machine learning techniques. *Int. J. E-Serv. Mobile Appl.* **12**(1), 21–38 (2020)
14. Yadav, K., Kumaraguru, P., Goyal, A., Gupta, A., Naik, V.: Smsassassin: Crowdsourcing driven mobile-based system for SMS spam filtering. In: Proceedings of the 12th Workshop on Mobile Computing Systems and Applications, pp. 1–6. ACM (2011)
15. Delany, S.J., Buckley, M., Greene, D.: SMS spam filtering: methods and data. *Expert Syst. Appl.* **39**, 9899–9908 (2012)
16. Eshmawi, A., Nair, S.: Feature reduction for optimum sms spam filtering using domain knowledge. In: International Conference on Security and Management (SAM), The Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing (WorldComp) (2013)
17. Idris, I., et al.: A combined negative selection algorithm–particle swarm optimization for an email spam detection system. *Eng. Appl. Artif. Intell.* **39**, 33–44 (2015)



Mathematical Socio Analysis of Cybercrimes Preparedness a Simulation Odessey with R

S. Dheva Rajan^(✉)

University of Technology and Applied Sciences AlMusannah, Muladdah, Sultanate of Oman
technoinst@rediffmail.com

Abstract. This paper presents a unique mathematical model, SAEIQRS, to understand the dynamics of cybercrime propagation in response to the growing danger of cybercrime. The model combines epidemiological principles with cyber-attack dynamics to capture the complexities of cyber threats. Assessing the impact on vulnerable individuals, asymptomatic illnesses, exposures, and organizational vulnerabilities are among the goals. Questions about infection phases, workplace dynamics, and external sources are investigated. Simulations using deSolve, ggplot2, dplyr, and R libraries reveal detailed patterns that relate cyber-attacks to diseases and organizational weaknesses. The model's adaptability in scenario research is highlighted by key findings, which identify vital periods and sensitive sectors. The research yields useful insights for cybersecurity policy and intervention plans.

Keywords: Cybercrime · Epidemiology · DeSolve · Simulation ·
Cybersecurity · Vulnerability · dynamics · ODE

1 Introduction

1.1 Cybercrime

In this day of technological ease, cybercrime (CRM) is a serious danger. This lesson takes a thorough approach to covering different cyberthreats and practical defenses. Cybercriminals use a variety of sophisticated approaches, including identity theft, phishing, financial frauds, and diversionary tactics. Protection requires vigilance, skepticism, and proactive steps like password security and internet literacy. The public's education, particularly that of the younger age, evidence preservation, and timely event reporting are essential elements in the fight against CRM. Building a watchful and knowledgeable digital community is crucial in a world where convenience meets vulnerability. The digital technology brings unprecedented convenience, but it also brings a slew of hazards. Individuals may traverse this world safely with knowledge, skepticism, and proactive steps. The burden lies not just with law enforcement or cybersecurity (CS) professionals, but with every user who, equipped with information, may turn the tide against cyber dangers and contribute to a safer digital world.

The worrisome data highlight the fragility of India's digital infrastructure, requiring immediate and comprehensive CS initiatives. The significant growth in website intrusions highlights the crucial need for stronger security measures at both the individual and

corporate levels. Financial fraud's dominance in CRM highlights the need for financial organizations and individuals to strengthen their CS defenses. The exponential surge in reported cyber-crime necessitates a multifaceted strategy that includes legislative measures, public awareness initiatives, and coordination between law enforcement and CS agencies. [1]

1.2 CRM Status

The percentage of internet users' experiences CRM study is conducted in selected countries viz, India, US, Australia, New Zealand, France, UK, Germany, Japan. [2] Out of these countries, India holds its first position with 68% followed by US and Australia with 49% and 40%. Globally around 8 trillion US dollars financial lost happened due to CRMs in 2022. It is expected to reach around 14 trillion US dollars in 2028. [3] In India, the CRM complaint registrations holding its exponential increase year by year especially in India. [4, 5] For an instance, state wise CRM registration in India is given in Fig. 1. India has seen a remarkable increase in CRM, particularly in website hacking occurrences. In 2018, 17,560 sites were compromised, and this figure is expected to rise to 26,121 by 2020, highlighting the evolving threat scenario. CRM in the country increased from 208,456 in 2018 to an astonishing 212,485 in just the first two months of 2022, outnumbering the whole 2018 total. Financial fraud emerges as the leading CRM, accounting for 75% of occurrences between 2020 and 2023. A disturbing discovery is that 78% of Indian firms experienced ransomware attacks in 2021, above the global average of 66%. Furthermore, 80% of these attempts resulted in data encryption, highlighting the vital necessity for strong CS measures. [1]

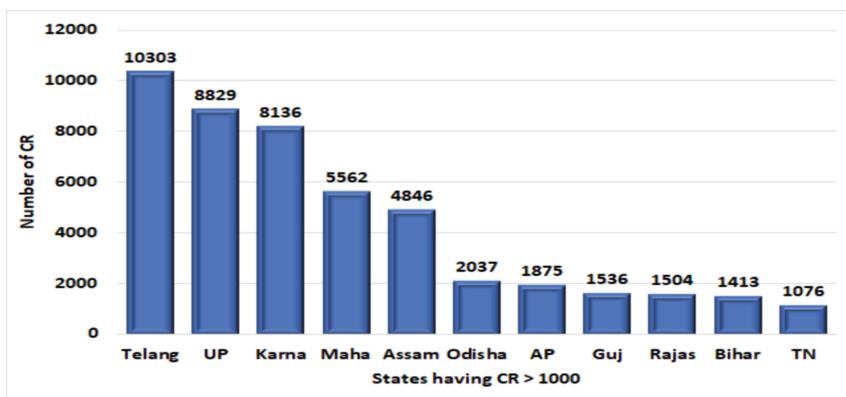


Fig. 1. CRM registration greater than 1000 in 2021 in India

The 78% frequency of ransomware attacks on Indian firms in 2021 serves as a stern warning, underlining the sophistication of cyber threats that businesses face. The high encryption rate caused by these assaults highlights the potential severity of data breaches and the urgent need for better encryption mechanisms. As the global average encryption rate is 65%, India's 80% rate indicates a heightened danger picture that

requires quick action. The rising CRM scenario, which is characterized by an increase in website intrusions, financial fraud, and ransomware assaults, needs immediate and thorough CS solutions. The findings are a wake-up call for individuals, corporations, and governments to address CS in the digital era. The changing nature of cyber threats needs ongoing adaptation and coordination to preserve the world's digital infrastructure and combat the growing tide of CRM. [1] The states with single digit CR are as follows: Nagaland, A&N Islands, D&N Haveli, and Daman & Diu+, Ladakh, Lakshadweep, Sikkim, Puducherry. Within the states, Manipur is having more variation over the years and Chandigarh having the least. The number of CRM police stations is given in Fig. 2. It is observed that there is a huge gap between the growth in the CRM increase and the number of CRM investigative special police stations.

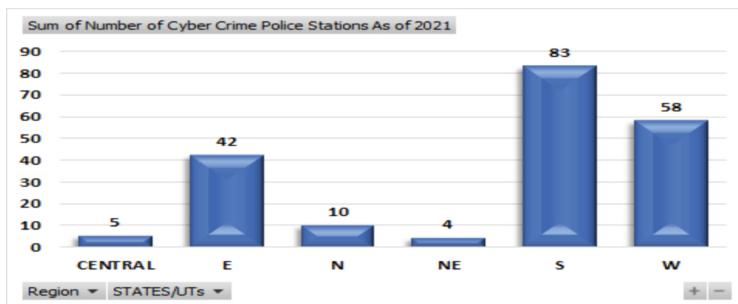


Fig. 2. Number of CRM police stations as of 2021 region wise

2 Background

Mishra and Jha, (2010) proposed SEIQRS technique for the transmission of malicious objects in computer network not for human oriented. Mishra and Jha, (2010) paves the way to propose to view the model in multiple dimensions. Kochedykov et al., (2019) proposed mathematical models for cyber-attacks on critical information system become the crown for the modellers. But the model by Kochedykov et al., (2019) lacks in CRM. Zhu et al., (2019) proposed SEIR model for smart grid Information Security Risk Propagation but not for CRM. Bjørnstad et al., (2021) proposed SEIRS mathematical modelling for infectious disease dynamics through which the authors inspired to develop the model for CRM using the proposed technique. Trenchev et al., (2023) proposed algebraic methodologies to convert CS from protoscience to science but only for computers. Chernikova et al., (2023) proposed self-dispersal malware using epidemiological techniques but not for CRM. Hence, it is identified, there is a gap in the literature in addressing CRM using Mathematical modelling techniques, especially using ODE by adopting the model of epidemiology. The current methodologies like GPS tracking is get in to the picture after event has happen, but the spread forecasting methods took place on past data. Here, it is proposed a universal SAEIQRS model for the spread of CRMs. Such mathematical modelling provides increased accuracy in such forecasting and is highly impactful for

the policy makers. A total of 16 states with 26 parameters have been proposed with inter-relations.

3 Modelling

Susceptible (S) the portion of the population that is susceptible to being infected. Asymptomatic Infected (A) represents individuals who are infected but asymptomatic, not showing apparent symptoms. Exposed (E) represents individuals who have been exposed to the virus but are not yet infectious. Infected (I) represents individuals who are infected and showing symptoms. Quarantined (Q) represents individuals who are in quarantine, separated from the general population to prevent further spread. Recovered (R) represents individuals who have recovered from the infection and gained immunity.

Dead (I_d) represents individuals who have succumbed to the infection. Asymptomatic Recovered (I_o) represents individuals who were asymptomatic but have recovered. Asymptomatic Dead (I_r) represents individuals who were asymptomatic but have succumbed to the infection. Reservoir of Infection in the Population (R_n) represents the reservoir of infection in the general population. Infection in the Workplace (W) represents the infection present in the workplace. Death (D) represents individuals who have died. External Infection (E_{ext}) represents infection introduced from an external source.

Infected by External Source ($I_{Infected_ext}$) represents individuals infected by an external source. Infected in Closed Environment (C) represents infection in a closed environment. Attack on Organization (AoO) represents an attack on an organization, possibly related to the CRM scenario and the respective derivatives constitute the set of equations. The proposed SAEIQRS model is given below.

$$dS = A - \sum (\beta_{Si} SI_i) - \beta_{SA} SA - \beta_{SQ} SQ - \beta_{SR} SR - \epsilon R \quad (1)$$

$$dA = \beta_{SA} SA - \beta_{AS} AS \quad (2)$$

$$dE = \sum (\beta_{Si} SI_i) + \beta_{SA} SA + \beta_{SQ} SQ + \beta_{SR} SR - \alpha E \quad (3)$$

$$dI = \alpha E - (d + \delta)I - \gamma I \quad (4)$$

$$dQ = \beta_{AS} AS + \beta_{SQ} SQ - \beta_{QS} QS \quad (5)$$

$$dR = \gamma I - \epsilon R \quad (6)$$

$$dI_d = \sum (\beta_{Si} SI_i) \quad (7)$$

$$dI_a = \alpha E - (d + \delta)I_a - \gamma I_a \quad (8)$$

$$dI_r = \gamma I_a - \epsilon I_r \quad (9)$$

$$dRn = \beta_{Rn} SR - \beta_{RnS} RnS \quad (10)$$

$$dW = \beta_W RnW - \beta_{WS} WS \quad (11)$$

$$dD = \beta_D WD - \beta_{DS} DS \quad (12)$$

$$dE_{\text{ext}} = \beta_E DE - \beta_{ES} ES \quad (13)$$

$$dI_{\text{infected_ext}} = \beta_I EI - \beta_{IS} IS \quad (14)$$

$$dC = \beta_C IC - \beta_{CS} CS \quad (15)$$

$$dAoO = \beta_{AoO} CAoO - \beta_{AoOS} AoOS \quad (16)$$

4 Parameters of SAEIQRS Model

A : Total Population

β_{Si}

Transmission rate from susceptible to asymptomatic infected or unknowingly fall in.

β_{SA} : Transmission rate from susceptible to asymptomatic infected

β_{SQ} : Transmission rate from susceptible to quarantined(suspected CRM)

β_{SR} : Transmission rate from susceptible to recovered(resolved CRM)

d : Natural death rate, people move without resolution.

ϵ : Rate of loss of immunity, lacking support from external sources

α : Rate of progression from exposed to infected

γ : Rate of recovery from infected

δ : Rate of disease - induced mortality, mortality, moving out of box due to CRM

β_{AS} : Transmission rate from asymptomatic infected to susceptible

β_{QS} : Transmission rate from quarantined or identified to susceptible

β_{Rn} : Transmission rate from reservoir to susceptible β_{RnS} :

Transmission rate from reservoir to susceptible

β_W : Transmission rate from workplace to susceptible

β_{WS} : Transmission rate from workplace to susceptible

β_D : Transmission rate from dead to workplace

β_{DS} : Transmission rate from dead to susceptible

β_E : Transmission rate from external source to exposed

β_{ES} : Transmission rate from external source to susceptible

β_I : Transmission rate from exposed to infected

β_{IS} : Transmission rate from infected to susceptible

β_C : Transmission rate from infected to closed environment

β_{CS} : Transmission rate from closed environment to susceptible

β_{AoO} : Transmission rate from closed environment to attack on organization

β_{AoOS} : Transmission rate from attack on organization to susceptible

5 Solution and Analysis

It follows Runge Kutta 4th (RK4) order method to solve the differential equation numerically. Also, it is used deSolve package in R program [12] to solve the system of equations with assumed initial conditions 1: $S = 10,000$, $A = 1000$, $E = 900$, $I = 800$, $Q = 700$, $R = 600$, $I_d = 700$, $I_a = 600$, $I_r = 500$, $Rn = 100$, $W = 50$, $D = 25$, $E_{ext} = 10$,

$I_{infected_ext} = 10$, $C = 10$, $AoO = 5$. Even though it is given as assumed initial conditions, the parameter values have been extracted from [1, 5–7]. Figure 3 shows the solution of the system for S , A , E and I . ggplot package in R programming used to generate the plots. [13]

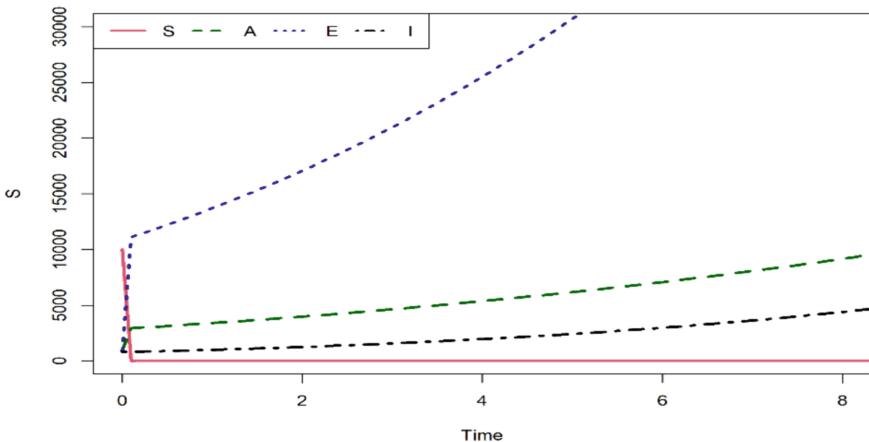


Fig. 3. Solution of CR model for S , A , E , I with assumed initial conditions 1.

Figure 4 shows the solution of the system for S , A , E and I assumed initial conditions 2: $S = 20,000$, $A = 1500$, $E = 1200$, $I = 1000$, $Q = 800$, $R = 700$, $I_d = 800$, $I_a = 700$, $I_r = 600$, $Rn = 200$, $W = 100$, $D = 50$, $E_{ext} = 20$, $I_{infected_ext} = 15$, $C = 15$, $AoO = 10$.

The susceptible population (S) shrinks with time, showing a decline in the number of people who are susceptible to infection. This decrease might be attributed to people being ill or relocating to different compartments. The number of asymptomatic infections (A) rises at first, then stabilizes or falls. This shows that asymptomatic infections spread initially, followed by either recovery or transfer to other compartments. Initially, the exposed (E) population grows, reflecting individuals who have been exposed to the virus but are not yet infectious. This number rises and then begins to fall, showing that individuals either become contagious or recover. The number of infected (I) people grows at first, representing the spread of the sickness. Depending on the model settings, the number of infected persons may decrease as a result of recovery or death. Figure 5 shows the solution of the proposed model for the states Q , R , W and D with respect to initial conditions 1.

Figure 6 shows the solution of the proposed model for the states Q , R , W and D with respect to initial conditions 2.

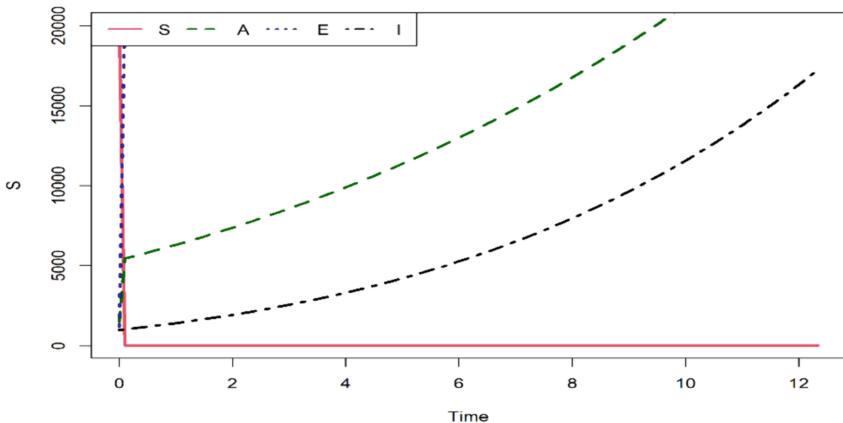


Fig. 4. Solution of CR model for S, A, E, I with assumed initial conditions 2

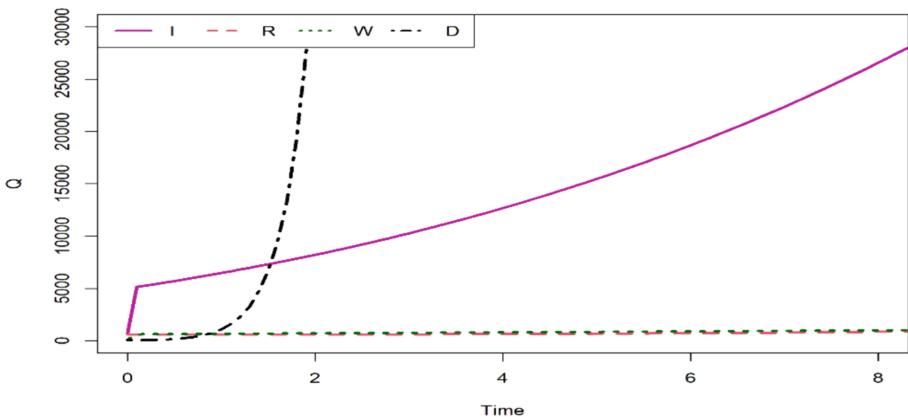


Fig. 5 Solution of CR model for Q, R, W, D with assumed initial condition 1

Quarantined (Q) cases increase with time, indicating that efforts are being made to contain the disease's spread. As people heal or move into other compartments, this number could decrease in the future. Over time, the number of recovered (R) persons increases, signifying those who have effectively conquered the illness. This is a positive outcome that strengthens population immunity. Occupational Infection (W) indicates the fraction level of occupational infection. To comprehend the impact on certain metrics, this compartment has to be observed. Death (D) denotes more deaths brought on by problems at work. Information on occupational mortality is provided by monitoring this compartment. Figure 7 shows the solution of the proposed model for C vs time and AoO vs time with respect to initial condition 1.

Figure 8 shows the solution of the proposed model for C vs time and AoO vs time with respect to initial condition 1.

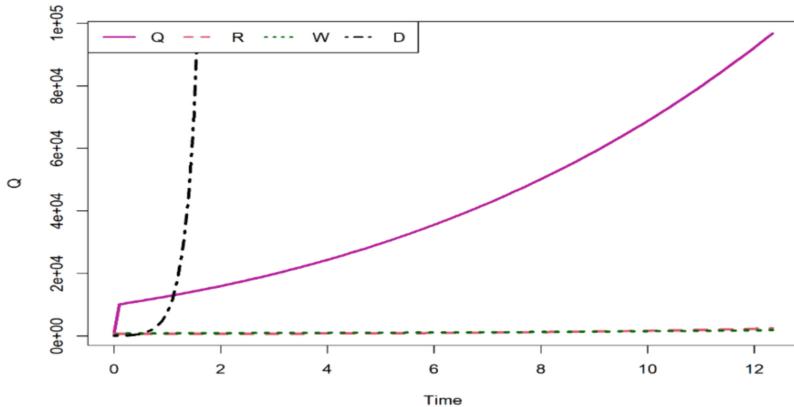


Fig. 6 Solution of CR model for Q , R , W , D with assumed initial condition 2

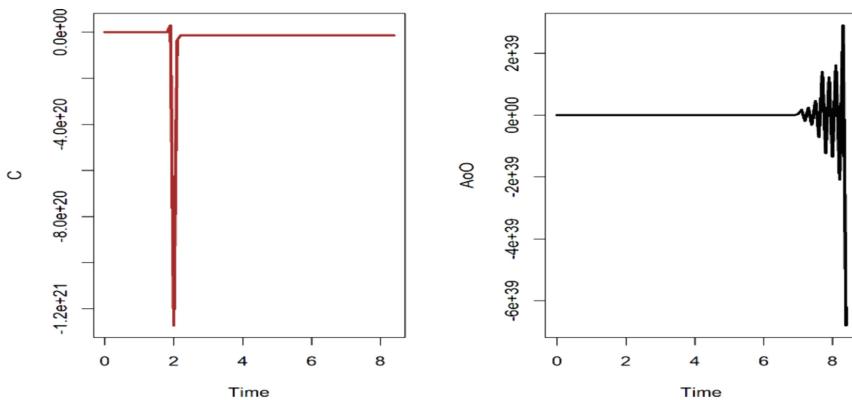


Fig. 7 Solution of CRM model of C and AoO vs time ‘ t ’ with assumed initial condition 1

The marker C signifies infection in a minor region. In certain circumstances, observing an eye on this compartment is necessary to realizing the dynamic forces of infection. A battering alongside an organization is described as an AoO in CRM scenarios. This section delivers perception into how administrative procedures are jammed by CRM. The solution of the proposed model comparison within the states A vs R , W vs R , A vs C , AoO vs E is given in Figs. 9 and 10 gives the solution with initial condition2.

The increasing curve in A vs. R graph confirms that the quantity of regaining improves in cycle with the prevalence of symptomless nature. This discloses that chosen asymptomatic persons can recover after a cyberattack. A downward tendency establishes concerns with place of work recuperation, where an increasing trend in the W vs. R plot indicates that place of work infections serves in overall recovery. This possibly will supply information on how resilient companies are to cyberattacks. The affirmative correlation between A and C shows that infections in confined conditions are caused by an asymptomatic nature. Determining the possibility of cyber hazards spreading across

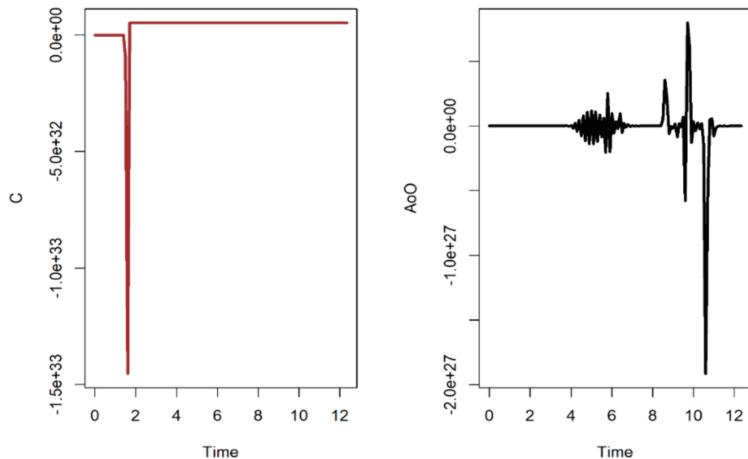


Fig. 8 Solution of CRM model of C and AoO vs time ‘ t ’ with assumed initial condition 2

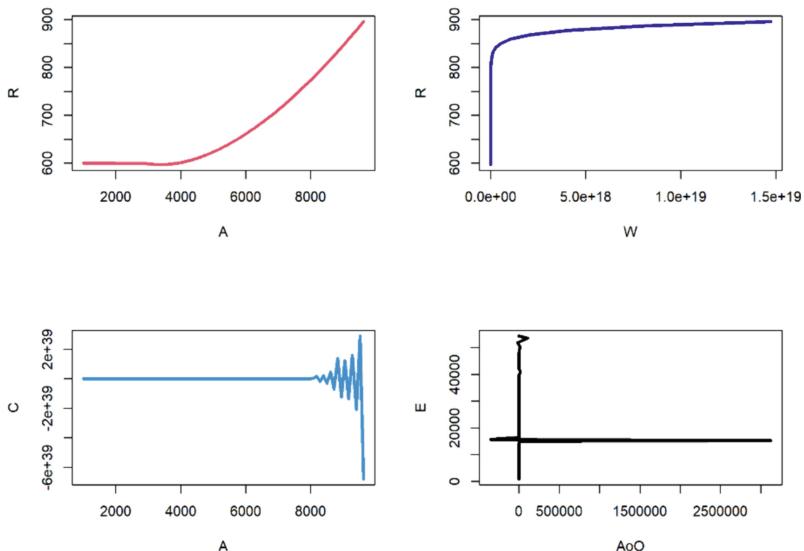


Fig. 9 Solution of CR model A vs R , W vs R , A vs C , AoO vs E with assumed initial condition 1

different environments requires an understanding of this relationship. The AoO vs. E in Fig. 10 shows a positive correlation, meaning that more people are exposed as a result of attacks on organizations. This highlights how cyberattacks are connected to human vulnerability to cyberthreats. The diagrams clarified the interdependent dynamics of the several compartments in the model. Plot patterns illustrate how some factors affect other aspects, which aids in understanding the wider ramifications of cyber hazards. Trends in these correlations can direct CS precautions, cyberattack mitigation strategies, and decision-making.

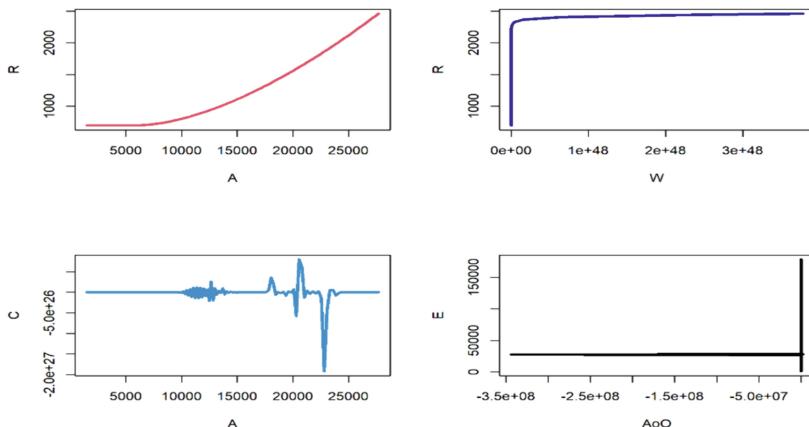


Fig. 10 Solution of CR model A vs R, W vs R, A vs C, AoO vs E with assumed initial condition2

6 Cybercrime Ecosystem for Proactive Combat

In a groundbreaking endeavor, a nationwide platform is suggested to evaluate the numerous dimensions of CRM, producing analytical threat intelligence reports. This program seeks to establish a multi-stakeholder environment by bringing together law enforcement and industry professionals for collaborative endeavors. Combating CRM against women and children is prioritized, with an automated reporting mechanism in place to ensure prompt action. The center's goal is to lead intelligence-led responses against critical cyber threats while supporting coordinated, multi-jurisdictional activities. It proposes a well-equipped National Cyber Forensic Laboratory and a Central Forensic Science Laboratory to conduct sophisticated forensic analyses. The construction of a National CRM Training Centre, which includes a Cyber Range, demonstrates a commitment to practical training and awareness. The program aims to standardize CS training, conduct proactive technology tracking, and form strategic alliances to strengthen the CRM combating ecosystem. This comprehensive strategy seeks to proactively address new risks and strengthen community resistance to CRM.

7 Summary, Conclusion and Future Work

The essay emphasizes the pervasiveness of CRM in today's technologically advanced society and its rising threat on a worldwide scale. With a noteworthy 68% of CRM incidents, India stands out and has a serious problem. The author points out a gap in the body of research on mathematical modeling of CRM and makes the case for the necessity of more thorough methods. To fill the vacuum in the literature, a brand-new SAEIQRS mathematical model with 26 parameters and 16 states is presented. The differential equations of the model are numerically solved, yielding a quantitative insight of the dynamics of CRM. The model's graphical representations provide visual insights for CS measures and decision-making by showcasing patterns in sensitive populations, infections, recoveries, and organizational repercussions. The unique SAEIQRS mathematical

model, which combines epidemiological ideas with cyber-attack dynamics, is introduced in the Socio ODE Analysis on Temporal Analysis of CRMs using R. This model depicts the spread of CRM in detail, considering infection phases, recoveries, mortality, external infections, and cyber-attacks. It excels in capturing the impact on workplaces, recognizing external sources, simulating closed environment dynamics, and introducing an infection reservoir. Because of the model's versatility, scenario exploration is possible, showing significant trends in the dynamic depiction of cyber-attack impact across time.

In today's technologically sophisticated world, CRM is becoming a greater concern. The investigation addresses this issue and stresses the importance of protection tactics and thorough understanding. It provides statistics information on CRM incidence worldwide, with India leading the way at 68%. In order to fill a vacuum in the literature on mathematical modeling of CRM, the author presents a brand-new SAEIQRS model with 26 parameters and 16 states. Numerical solutions to the differential equations shed light on the dynamics of the propagation of CRM. In order to support CS measures and decision-making, the model's graphical depiction shows trends in sensitive populations, infections, recoveries, and organizational implications. A compartment for infections from external sources, in particular, emphasizes the worldwide interconnection of cyber dangers. This study stands out for its comprehensive methodology, which provides nuanced insights for policymakers, researchers, and CS professionals, developing a deep awareness of the larger societal effects of CRM.

The study's assumptions and limitations may have led to an oversimplification of the intricate nature of cyber threats. The way that hackers are altering their techniques may not be fully captured by the model. Since the model's dependability analysis isn't examined in this study, CRM data from other geographical areas can be used to validate the model's demographic assumptions. Furthermore, as the RK4 method's numerical solution for the proposed SAEIQRS model depends on its assumptions, any breaches of those assumptions result in the proposed model's limitations. A limited awareness of the entire cyber threat environment may result from the incomplete consideration of external elements impacting the dynamics of CRM.

Subsequent investigations have to concentrate on enhancing the SAEIQRS model through the integration of more intricate factors and confirming its accuracy against actual CRM data. The effects of geopolitical events, technical breakthroughs, and socioeconomic issues on cyber dangers might be the subject of future research. Prediction accuracy may be improved by comparative research using different modeling strategies and the incorporation of machine learning methods. A more thorough knowledge of cyber dangers might be obtained by longitudinal assessments and cooperation with CS professionals, which could direct the creation of successful preventative measures. The limitations provided may be considered when the suggested model is developed in the future.

The SAEIQRS model bridges the gap between epidemiological concepts and cyber risks, providing a flexible tool for managing the developing CRM scenario. Its versatility and temporal representation help to inform policy makers and implement effective mitigation methods. The combination of network dynamics and behavioral elements, as well as real-time data, shows promise for future improvements, paving the way for a more secure and robust digital future. As people grapple with the intricacies of the

digital realm, the SAEIQRS paradigm encourages us to break down barriers and pave a path toward a secure and enlightened cyber future.

References

1. AAG IT: The Latest Cyber Crime Statistics (updated January 2024) | AAG IT Support. <https://aag-it.com/the-latest-cyber-crime-statistics/>. Accessed 30 Jan 2024
2. Statistica: Cybercrime rate by country 2022. Statista. <https://www.statista.com/statistics/194133/cybercrime-rate-in-selected-countries/>. Accessed 7 Dec 2023
3. Statistica: Global cybercrime estimated cost 2028. Statista. <https://www.statista.com/forecasts/1280009/cost-cybercrime-worldwide>. Accessed 7 Dec 2023
4. MHA: CYBER FRAUDS. <https://pib.gov.in/pib.gov.in/Pressreleaseshare.aspx?PRID=1988272>. Accessed 30 Jan 2024
5. PIB Delhi: Cyber Crimes and Frauds. <https://www.pib.gov.in/www.pib.gov.in/Pressrelease/share.aspx?PRID=1883066>. Accessed 8 Dec 2023
6. Mishra, B.K., Jha, N.: SEIQRS model for the transmission of malicious objects in computer network. *Appl. Math. Model.* **34**(3), 710–715 (2010). <https://doi.org/10.1016/j.apm.2009.06.011>
7. Kochedykov, S.S., Grechishnikov, E.V., Dushkin, A.V., Orlova, D.E.: The mathematical model of cyber attacks on the critical information system. *J. Phys. Conf. Ser.* **1202**, 012013 (2019). <https://doi.org/10.1088/1742-6596/1202/1/012013>
8. Zhu, B., Deng, S., Xu, Y., Yuan, X., Zhang, Z.: Information security risk propagation model based on the SEIR infectious disease model for smart grid. *Information* **10**(10), 323 (2019). <https://doi.org/10.3390/info10100323>
9. Bjørnstad, O.N., Shea, K., Krzywinski, M., Altman, N.: Author Correction: the SEIRS model for infectious disease dynamics. *Nat. Methods* **18**(3), 321–321 (2021). <https://doi.org/10.1038/s41592-021-01079-6>
10. Trenchev, I., Dimitrov, W., Dimitrov, G., Ostrovska, T., Trencheva, M.: Mathematical approaches transform cybersecurity from protoscience to science. *Appl. Sci.* **13**(11), 6508 (2023). <https://doi.org/10.3390/app13116508>
11. Chernikova, A., Gozzi, N., Perra, N., Boboila, S., Eliassi-Rad, T., Oprea, A.: Modeling self-propagating malware with epidemiological models. *Appl. Netw. Sci.* **8**(1), 52 (2023). <https://doi.org/10.1007/s41109-023-00578-z>
12. R Core Team: R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, 2023. <https://cran.r-project.org/manuals.html>. Accessed 8 Dec 2023
13. Hadley Wickham: ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York (2016). <https://ggplot2.tidyverse.org>. Accessed 8 Dec 2023



Revolutionizing Weather Forecasting: Harnessing Machine Learning and Big Data in Upcoming Technologies

Basetty Mallikarjuna¹ and Varun Tiwari²

¹ Department of Information Technology, Institute of Aeronautical Engineering, Dundigal,
Hyderabad 500043, India

b.mallikarjuna@iare.ac.in

² Manipal University Jaipur, Jaipur, India
drvvaruntiwari2020@gmail.com

Abstract. Weather forecasting gives a critical part in our daily lives, influencing our decisions ranging from what to wear to planning outdoor activities or even preparing for natural disasters. However, accurate weather prediction remains a complex and challenging task. Traditional forecasting methods have their limitations in dealing with the vast amount of data generated by weather systems. In upcoming technologies, there has been a growing interest in harnessing the potential of machine and deep learning to improve weather forecasting accuracy and efficiency. This article explores the role of weather forecasting in upcoming technologies, the benefits of leveraging big data, the integration of artificial intelligence in weather models, the evaluation of the advantages and limitations of machine learning, the future potential, and the transformative impact of machine and deep learning in revolutionizing weather prediction.

Keywords: Weather forecasting · Machine Learning · Deep Learning · Big Data · Internet of Things

1 Introduction

Weather forecasting plays a critical role in our daily lives [1]. From planning outdoor activities to ensuring our safety during severe weather events, accurate predictions are essential [1]. Without weather forecasts, you might end up caught in a sudden thunderstorm without an umbrella or wearing a heavy coat on an unexpectedly scorching day [2]. We rely on these forecasts to make informed decisions and adapt accordingly. While weather forecasting has come a long way, it still faces several challenges [3]. The atmosphere is a complex system with countless variables that interact with each other in intricate ways [4]. This complexity makes it difficult to accurately predict the weather beyond a certain time frame. The accuracy of predictions decreases as the forecast lengthens, often resulting in uncertainties [5]. Additionally, there are various climatic regions with unique weather patterns, making it challenging to develop a universal forecasting model. Furthermore, the availability and quality of data, computational limitations, and the need for real-time updates pose additional hurdles for meteorologists [6].

Machine and deep learning, a subset of artificial intelligence, empowers computers to learn and make predictions without being explicitly programmed. It involves the development of algorithms that learn patterns and relationships from data, enabling machines to make informed decisions or predictions based on new inputs [7].

1.1 Objectives

Machine and deep learning offers significant advantages in weather forecasting, the following objectives are as follows [8].

- By analyzing historical weather data allows for more accurate predictions and improved understanding of weather phenomena [1, 2].
- Machine learning also helps in handling the massive amounts of data collected, making sense of it, and extracting meaningful insights [3].
- Furthermore, upcoming machine learning and deep learning models can improving their accuracy over time [4].
- This adaptability enables meteorologists to continually refine and enhance their forecasting models, keeping up with the ever-changing nature of the atmosphere [5].

1.2 Big Data in Weather Forecasting

Weather forecasting heavily relies on vast amounts of data from various sources, including weather stations, satellites, radar systems, and climate models [9]. With the advent of technology, the volume, variety, and velocity of this data have increased exponentially. This data deluge, known as big data, has opened new avenues for improving weather predictions [10].

1.2.1 Advantages of Big Data in Weather Predictions

Big data analytics provides a wealth of opportunities for enhancing weather forecasting accuracy. By combining and analyzing diverse datasets, meteorologists can gain comprehensive insights into weather patterns, trends, and anomalies [11]. This enables them to make more nuanced predictions and better understand the factors influencing weather phenomena [12]. Moreover, big data facilitates the development of more sophisticated models that can handle complex interactions and capture micro scale weather variations [13]. By leveraging big data, weather forecasting can move beyond traditional methods and unlock new frontiers of accuracy [14].

1.3 Applying Machine and Deep Learning Algorithms to Weather Data Analysis

Machine and deep learning algorithms form the backbone of analyzing weather data allowing meteorologists to make predictions based on historical trends [15]. From decision trees to neural networks, there exists a rich variety of algorithms that can be applied to weather forecasting [16].

1.3.1 Upcoming Machine Learning and Deep Learning Algorithms

Some commonly used machine learning and deep learning algorithms can handle the complexity and non-linear relationships inherent in weather data [14, 17]. By training these models on historical weather data and associated outcomes, meteorologists can create predictive models that assist in accurate weather forecasts [18].

It's an exciting time where the combination of machine and deep learning and big data has the potential to revolutionize weather forecasting, equipping us with ever-improving predictions for the ups and downs of Mother Nature's mood swings. So, the next time you reach for your umbrella or choose your outfit based on the forecast, remember there's a witty and intelligent machine working tirelessly behind the scenes to keep you dry and stylish [19].

2 Literature Survey

Artificial Intelligence (AI) is revolutionizing the field of weather forecasting by providing innovative approaches to analyze the data and improve the accuracy of predictions [1]. AI concepts to apply historical weather data, satellite imagery, and atmospheric conditions, enabling meteorologists to make more informed forecasts [2].

2.1 Challenging Issue in Weather Forecasting

AI is being utilized in various aspects of weather forecasting, such as predicting severe weather events, improving hurricane tracking, and enhancing short-term and long-term weather predictions [3]. Deep learning concepts can identify subtle patterns that human forecasters might miss, it can accurate and timely warnings for extreme weather conditions. Additionally, AI can help optimize the allocation of resources and assist in decision-making processes during emergencies [5].

2.2 Justification of Weather Forecasting

Machine Learning offers several advantages in weather forecasting. It can handle large datasets and complex models, allowing meteorologists to analyze a different aspects of variables simultaneously [7]. Deep learning and machine learning algorithms can detect intricate relationships between different weather parameters, resulting in more precise predictions [8]. Moreover, AI-driven automated systems can process data more quickly, enabling faster updates and improved forecast accuracy [9].

2.3 The Limitations and Challenges of Machine Learning in Weather Forecasting

Deep learning and Machine Learning in weather forecasting has some limitations. Weather patterns can be highly nuanced, and there is always a risk of overfitting the models to the training data, leading to inaccurate predictions [10]. Additionally, unpredictable factors, such as sudden atmospheric changes or limited historical data for rare events, can pose challenges to Machine Learning algorithms. Ensuring the transparency, interpretability, and accountability of AI systems in weather forecasting is another significant challenge that scientists and researchers need to address [11].

2.4 Solving Weather Prediction

The Machine and deep Learning in weather forecasting looks promising. Researchers are exploring innovative techniques, such as deep learning networks and ensemble modeling, to enhance prediction accuracy and handle complex weather phenomena [13]. Integration of technologies AI and Internet of Things (IoT) and Big Data analytics can further revolutionize weather forecasting by providing real-time data and improved model performance [15].

Machine Learning and Deep Learning in the field of weather forecasting and prediction by enabling more precise and timely forecasts [17]. With continued advancements in AI algorithms and data availability, meteorologists can demand on Deep learning and Machine Learning to mitigate the impacts of severe weather events, enhance disaster preparedness, and improve public safety. By integrating AI into weather forecasting systems [18].

3 Proposed Methodology

Machine Learning is revolutionizing weather forecasting by offering innovative approaches to analyze and predict complex weather patterns. With the ability to process large datasets and identify subtle relationships, AI algorithms enhance forecast accuracy and provide timely warnings for severe weather events. This work overcome the existing limitations, the proposed research and advancements in Machine Learning and Deep learning techniques hold immense potential to transform weather prediction, ensuring a safer and more prepared future. So, next time you check the weather forecast, remember the role of Machine Learning in making those predictions a little bit more accurate (and maybe grab an umbrella, just in case!). This work aim the effectiveness of weather forecasts, which can be categorized as either satisfactory or unsatisfactory [1]. As per the meteorological data, it can be inferred that “the climate channel” mentioned by the authors is likely a source that provides real-time weather data. This could include information from weather stations, satellites, radars, and other sources that monitor atmospheric conditions. The use of deep and machine learning methods mentioned in the study indicates that the collected meteorological data is processed and analyzed to improve weather forecasting accuracy [2]. Machine learning algorithms such as Naive Bayes (NB), Logistic Regression (LR), Random Forest (RF), Support Vector Machines (SVM), and Neural Networks are all popular machine learning algorithms known for their ability to handle complex datasets and make predictions based on patterns and relationships within the data. Each of these algorithms has its own strengths and weaknesses. Naive Bayes is often used for text classification tasks but can also be applied to weather forecasting by considering various meteorological parameters as features. Logistic Regression is commonly used for binary classification problems and could potentially be utilized for predicting weather conditions such as rain or no rain. Random Forest, Support vector machine and neural networks is an ensemble and the process as shown in Fig. 1. This diverse range of meteorological factors plays a crucial role in understanding and predicting weather patterns [3].

Temperature is one of the fundamental variables measured in weather forecasting. It gives the current state of the atmosphere and helps identify temperature gradients that

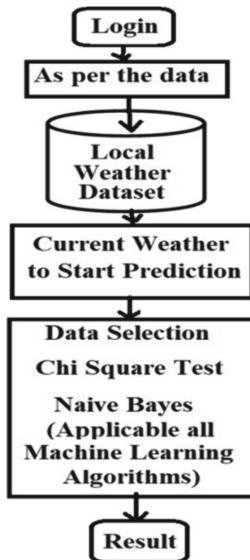


Fig. 1. The flow diagram of the methodology [1–3, 5]

influence air movement. By analyzing temperature data from various sources, meteorologists can detect patterns such as warm or cold fronts, which are essential for predicting changes in weather conditions. Precipitation data, including rainfall, snowfall, sleet, or hail, is vital for assessing water availability and distribution. Precipitation measurements help determine drought conditions, flood risks, and overall water resource management. By collecting information on precipitation from different locations, meteorologists can identify precipitation patterns and predict their movement over time. Evaporation is another critical factor that affects weather patterns. It refers to the process by which liquid water transforms into vapor and enters the atmosphere. Evaporation rates are influenced by temperature, humidity.

The Naive Bayes popular approach for data classification due to its efficiency. However, it still provides reasonably accurate results in many cases. In this work, the training dataset plays a crucial role in building the classification model. It consists of labeled instances that are used to train the Naive Bayes classifier. These instances contain both predictor (input) variables and target (output) variables. The model receives data from the training set and applies it to the test data for prediction scoring.

i. Naive Bayes Algorithm

Naive Bayes works with the probabilistic algorithm, it commonly for classification of tasks. It not only sufficient for weather forecasting, it can be useful to find many applications to find predictions. Weather forecasting have Temperature, and Humidity, and you want to predict whether it will rain or not. The dataset consist as follows.

```

Dataset_Weatherforecasting = {

    'Environment': ['Sunny_nature', 'Rainy_nature ', 'Rainy_nature ', 'Overcast_nature ',
    'Sunny_nature ', 'Sunny_nature ', 'Rainy_nature ', 'Sunny', 'Overcast_nature'],

    'Temperature': ['Hot_nature ', 'Hot_nature ', 'Hot_nature ', 'Mild_nature ', 'Cool_nature
    ', 'Cool_nature ', 'Cool_nature ', 'Mild_nature '],

    'Humidity': ['High_nature ', 'High_nature ', 'High_nature ', 'High_nature ', 'Normal_nature '], 

    'PlayTennis': ['No', 'No', 'Yes', 'Yes', 'Yes', 'No', 'Yes', 'No', 'Yes', 'No']

}

```

Evidence: It observing the given set of features, regardless of the class.

$P(A|B) = P(A|B) * P(A) / P(B)$, A means event and B means evidence.

In the Naive Bayes equation, $P(A|B)$ follows $P(\text{event}|\text{evidence})$ represents the probability, the of an event occurring given the observed evidence. This is what we are trying to predict or classify.

$P(\text{evidence}|\text{event})$ refers to the probability of observing the given evidence if the event has occurred. This term is often calculated by assuming that each feature or attribute in the dataset is conditionally independent of others, hence the term “naive.” Although this assumption may not hold true in all cases, it simplifies calculations and still provides reasonably accurate results.

$P(\text{event})$ represents the prior probability of the event occurring without considering any evidence. It can be estimated from historical data or domain knowledge.

$P(\text{evidence})$ denotes the probability of observing the given evidence, regardless of any specific event. It acts as a normalization factor and ensures that probabilities sum up to 1.

ii. Decision Tree

The making of predictions heavily depends on the decision tree generated using the training data. By selecting the most suitable attribute, the set of samples can be effectively segmented, the decision tree structure as shown in Fig. 2 [7].

The Fig. 2 provides the system to classify or predict outcomes based on the given attributes. In tree represents an attribute, the branches are conditions for that attribute. To make predictions using this decision tree, it start from the root node and follow the branches based on the attribute values of the samples being evaluated. Internal node, a decision based on whether a certain condition is met or not. This process continues until we reach a leaf node, which provides the final prediction or classification for that particular sample. The reliability of these predictions on how well the decision tree has been constructed from the training data. The training data serves as a foundation for building an effective decision tree by identifying which attributes are most relevant and informative in segmenting or differentiating samples.

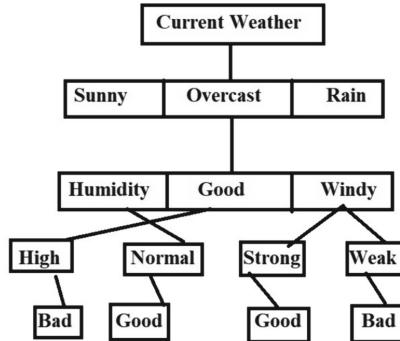


Fig. 2. The Structure of Decision Tree [7, 8]

4 Results and Discussions

Machine and Deep learning algorithms can identify patterns that may not be discernible forecasters. By learning from historical data, deep learning models make more predictions, leading to improved weather forecasting. Big data plays vital role in weather prediction allows meteorologists to collect and analyze an immense amount of weather-related information from various sources. By leveraging big data analytics, forecasters can gain deeper insights into weather patterns and improve the accuracy. While deep learning offers significant potential for weather forecasting, it does have limitations. Weather systems are complex and influenced by numerous variables, making it challenging to capture all factors accurately. Additionally, machine learning models may struggle with interpreting rare or unprecedented weather events that deviate from historical data.

The researchers collected a vast amount of meteorological data from multiple weather stations across different regions. This data included historical records of temperature, precipitation, evaporation, sunlight duration, wind speed and direction, cloud cover, and humidity. This study aimed to determine which machine learning algorithms could effectively analyze this complex and diverse dataset to provide accurate weather predictions. Naïve Bayes is a probabilistic nature gives the accurate results, which assumes independence between variables and its features. Artificial Neural Networks takes the images as input and can learn complex patterns in the data. The performance of these algorithms, and also worked for various metrics such as accuracy, precision, recall, and F1 score. The outcomes of our created models and data from the climate channel are summarized in the following Table 1.

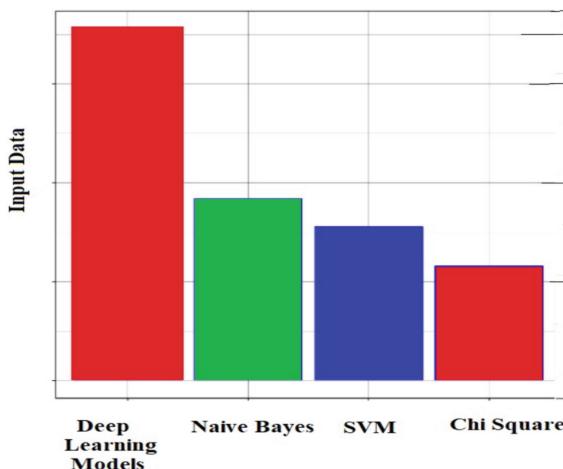
The absence of data from the climate channel in Model 4 is an important limitation to consider. This work explore the potential of machine learning algorithms and deep learning in weather forecasting using other meteorological data.

This work is a comprehensive approach to examine the efficacy of various machine and deep learning algorithms in weather prediction. By utilizing meteorological data such as temperature, precipitation, evaporation, sunlight, wind speed, wind direction, cloud cover, and humidity, this work got efficient accuracy and reliability of weather forecasts. Deep learning algorithms have shown promise in numerous fields by identifying patterns and making predictions based on large datasets. In weather forecasting,

Table 1. Machine Learning Algorithm Wise Climate Channel

Model	Climate Channel Data	Result
Model 1 (Multi Layer Neural Networks)	Yes	Hypothesis and obtained result
Model 2 (SVM)	YES	Supports hypothesis, improved training data
Model 3 (Naïve Bays)	YES	Enhanced aesthetics, neighbouring urban areas
Model 4 (Decision Tree)	NO	N/A
Model 5 (Chi Square Test)	YES	Promising outcomes, expanded dataset

this work can potentially analyze historical meteorological data and learn complex relationships between different variables. This work gives accurate predictions and a better understanding of how different factors contribute to weather patterns. The following Fig. 3 shows the comparison various models.

**Fig. 3.** Comparison of various machine learning

The upcoming technology gives the advances and more data becomes available, machine learning models will continue to improve. Additionally, the integration of artificial intelligence and advancements in computational power will enable more sophisticated and precise weather forecasting models in the future.

5 Conclusion

Upcoming models in deep learning are more powerful in revolutionizing weather forecasting. In big data and applying advanced algorithms, deep learning models are potential to enhance the accuracy and efficiency of weather predictions. However, this work improved the existing models and overcome the challenges and limitations in implementing machine learning in weather forecasting systems. Continued research, innovation, and collaboration between meteorologists and data scientists are crucial to harnessing the full potential of machine learning in weather prediction. As technology continues to advance, the future of weather forecasting looks promising, with machine learning paving the way towards improved forecasts and heightened resilience in the face of changing weather patterns.

References

1. Sen, A., Mazumder, A.R., Dutta, D., Sen, U., Syam, P., Dhar, S.: Comparative Evaluation of Metaheuristic Algorithms for Hyperparameter Selection in Short-Term Weather Forecasting. arXiv preprint [arXiv:2309.02600](https://arxiv.org/abs/2309.02600) (2023)
2. Verma, S., Srivastava, K., Tiwari, A., Verma, S.: Deep Learning Techniques in Extreme Weather Events: A Review. arXiv preprint [arXiv:2308.10995](https://arxiv.org/abs/2308.10995) (2023)
3. Du, P.: Ensemble machine learning-based wind forecasting to combine NWP output with data from weather station. *IEEE Trans. Sustain. Energy* **10**(4), 2133–2141 (2018)
4. Mallikarjuna, B., Addanke, S., Anusha, D.J.: An improved deep learning algorithm for diabetes prediction. In: Venkata Krishna, P. (ed.) *Handbook of Research on Advances in Data Analytics and Complex Communication Networks*, pp. 103–119. IGI Global (2022). <https://doi.org/10.4018/978-1-7998-7685-4.ch007>
5. Bochenek, B., Ustrnul, Z.: Machine learning in weather prediction and climate analyses—applications and perspectives. *Atmosphere* **13**(2), 180 (2022)
6. Mallikarjuna, B., Addanke, S., Anusha, D.J.: An improved deep learning algorithm for diabetes prediction. In: Venkata Krishna, P. (ed.) *Handbook of Research on Advances in Data Analytics and Complex Communication Networks*, pp. 103–119. IGI Global (2022). <https://doi.org/10.4018/978-1-7998-7685-4.ch007>
7. Slater, L., et al. (2022) Hybrid forecasting: using statistics and machine learning to integrate predictions from dynamical models. *Hydrol. Earth Syst. Sci. Discuss* [preprint]. <https://doi.org/10.5194/hess-2022-334> (in review).
8. Slater, L., et al. (2022) Hybrid forecasting: using statistics and machine learning to integrate predictions from dynamical models. *Hydrol. Earth Syst. Sci. Discuss* [preprint]. <https://doi.org/10.5194/hess-2022-334> (in review).
9. Li, M., Dai, L., Hu, Y.: Machine learning for harnessing thermal energy: from materials discovery to system optimization. *ACS Energy Lett.* **7**(10), 3204–3226 (2022)
10. Mallikarjuna, B., Srivastava, G., Sharma, M.: Blockchain technology: A DNN token-based approach in healthcare and COVID-19 to generate extracted data. *Expert. Syst.* **39**(3), e12778 (2022)
11. Camporeale, E.: The challenge of machine learning in space weather: nowcasting and forecasting. *Space Weather* **17**(8), 1166–1207 (2019)
12. Krzemińska, A., Miller, T., Kozłowska, P., Lewita, K.: Harnessing the power of random forest machine learning in global agriculture innovation. In: Collection of scientific papers «SCIENTIA», (July 28, 2023; Tel Aviv, Israel), pp. 59–65 (2023)

13. Mallikarjuna, B.: Osmosis machine learning load balancing of healthcare tasks in cutting edge technologies with smart grid. *Int. J. Smart Grid Green Commun.* **2**(2), 150–169 (2022)
14. Sahu, S., Kaur, A., Singh, G., Arya, S.K.: Harnessing the potential of microalgae-bacteria interaction for eco-friendly wastewater treatment: a review on new strategies involving machine learning and artificial intelligence. *J. Environ. Manage.* **346**, 119004 (2023)
15. Mallikarjuna, B., Sathish, K., Venkata Krishna, P., Viswanathan, R.: The effective SVM-based binary prediction of ground water table. *Evol. Intel.* **14**, 779–787 (2021)
16. Shamji, M.H., et al.: EAACI guidelines on environmental science in allergic diseases and asthma – Leveraging artificial intelligence and machine learning to develop a causality model in exposomics. *Allergy* **78**(7), 1742–1757 (2023)
17. Altameem, A., Mallikarjuna, B., Saudagar, A.K.J., Sharma, M., Poonia, R.C.: Improvement of automatic glioma brain tumor detection using deep convolutional neural networks. *J. Comput. Biol.* **29**(6), 530–544 (2022)
18. Mallikarjuna, B., Viswanathan, R., Naib, B.B.: Feedback-based gait identification using deep neural network classification. *J. Crit. Rev.* **7**(4), 2020 (2019)
19. Tasan, M., Ghorbaninasab, Z., Haji-Aghajany, S., Ghiasvand, Al.: Leveraging GNSS tropospheric products for machine learning-based land subsidence prediction. *Earth Sci. Inform.* **16**(4), 3039–3056 (2023)
20. Mallikarjuna, B., Kiranmayee, D., Saritha, V., Krishna, P.V.: Development of efficient ehealth records using IoT and Blockchain technology. In *ICC 2021-IEEE International Conference on Communications*, pp. 1–7. IEEE (2021)



Trends in Drowsiness Detection & Analysis of the Different Technologies Engaged

Sachin B. Honrao¹(✉) and U. D. Shiurkar²

¹ VDF GOI, Latur, India

honrao.sachin@gmail.com

² DIEMS, Aurangabad, India

Abstract. In past few years research is carried out worldwide to support driver while driving. The survey was carried out to understand causes for accidents. It is found that the major cause for occurring accidents on road is driver's drowsiness & fatigue. Therefore, detection of driver's drowsiness is an important aspect to avoid & prevent accidents occurring on road. With consideration of this issue this article undergoes a through survey of research for drowsiness detection carried worldwide.

Keywords: drowsiness · Transportation systems · PERCLOS · EAR · ADAS · HTC · EEG · ECG · EMG · HAR algorithm · ECR · MATHLAB · Support Vector Machine (SVM) · Artificial Neural Networks (ANN) · Human Machine Interface (HMI)

1 Introduction

The sleepiness or unconsciousness state of human being is considered as drowsiness. Transportation systems are now an integral part of human activities. Drowsiness can affect everyone, whether it's driving, lack of sleep, changing physical conditions, or traveling long distances. This leads to accident because of which, the number of deaths and fatal injuries increases by worldwide, each year. Amid 2019, a add up to 4,37,396 accidents were recorded within the nation. Street accident cases within the nation have diminished from 4,45,514 in 2018 to 4,37,396 in 2019. But death rate in street accidents have expanded by 1.3% (from 1,52,780 in 2018 to 1,54,732 in 2019). (STATE/UT – WISE TRAFFIC ACCIDENT CASES DURING 2019).

Insights have appeared that over 10% of accidents are due to fatigue and 2.6% accidents occurred because of driving under the influence of drug & alcohol, most of which happen on thruways or after driving on an expansive number of kilometres.

In this context, it is imperative to utilize unused advances to plan and construct frameworks that are able to screen drivers and to degree their levels of consideration amid the complete handle of driving.

In this manner to discover out inquire about research gap it gets to be vital to carry out review which is done through this review paper.

2 Distinctive Strategies Utilized in Drowsiness Detection in Recent years

There are a number of frameworks that have been developed to examine driver fatigue. They often need to be untangled in order to function partially or in unusual circumstances to find the research gap. These developed frame works are as discussed below.

- 1) *Marco Javier Flores et al.* (Flores et al., 2008) developed Advanced Driver Assistance System (ADAS) for automatic driver's drowsiness detection based on visual data & artificial inelegance. In this system a supervised classification method & support vector machine (SVM) Classification algorithm is used to find and monitor the face and the eyes in order to calculate a sleepiness index.
- 2) *Belal ALSHAQAQI et al.* (Alshaqaqi et al., 2013) developed Advanced Driver Assistance System (ADAS) for automatic driver's drowsiness detection based on visual data & artificial inelegance. Here they utilized the symmetry algorithm to detect eye, Hough transform for circles (HTC) on the captured image of the eye to find the iris and did ANDing on circles to obtain intersection level i.e. eye state's'.
- 3) *T. Vesselenyi et al.* (Maftukhaturrizqoh et al., 2017) create a system to detect drivers' drowsiness based on three different approaches: Driver image analysis and signal processing for EEG and EOG. Two types of artificial neural networks were used for this purpose: an auto encoder network and a single hidden layer network.
- 4) *Ratnarup Dey et al.* (Dey & Paulose, 2018) developed a system to detect drowsiness, Sensor parameters are used are detecting head movement, steering grasping and driving under influence of alcohol. Using a load cell, the driver's grip force on the

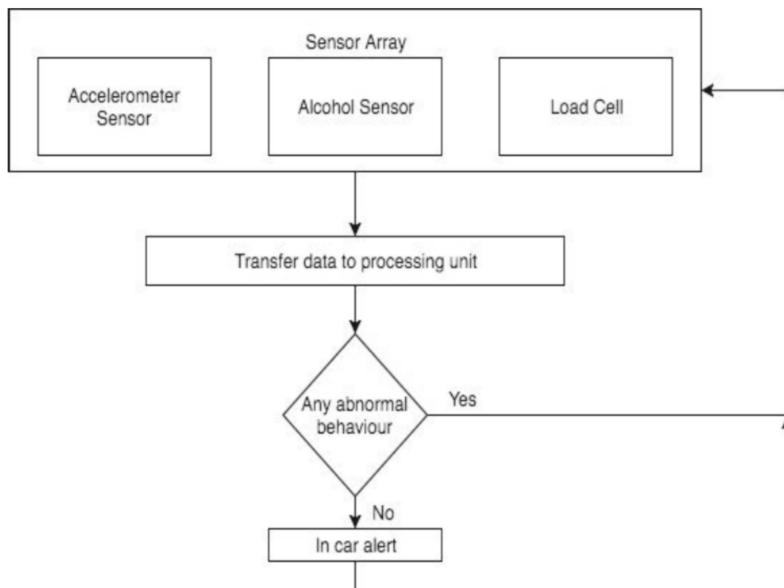


Fig. 1. Flowchart of proposed system for driver behavior detection system mechanism.

steering wheel is measured to determine whether the driver is drowsy. The head movement is examined using an accelerometer. *The algorithm for the development is as below (Fig. 1).*

- 5) *Eddie E. Galarza et al.* (Galarza et al., 2018) To identify and notify the driver of any sleepiness, they created a surveillance system. In order to construct the Human Computer Interaction System, a mobile phone is employed as a tiny computer with a mobile application running on the Android operating system. Face pictures are processed using image processing (Fig. 2).

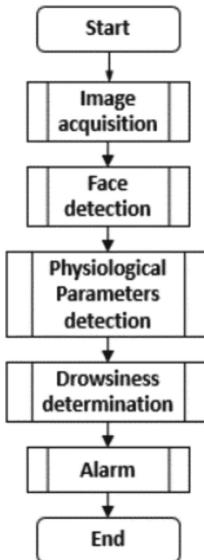


Fig. 2. Drowsiness detection methods

- 6) *Zhipeng Ma et al.* (Ma et al., 2018) created a technique that incorporates the four markers of eye that are movement, Electromyogram (EMG), Electrocardiograph (ECG), and grip force to identify driving weariness. It employed segmentation to find the eye, and PERCLOS eigen values were used to detect eye movement. Also, EMG is assessed using the median frequency, ECG using the standardized high-frequency power, and grip strength is assessed using the mean change rate.
- 7) *Maryam Hashemi et al.* (Hashemi et al., 2020) developed a system, Considering the twin goals of real-time applications of high accuracy and high speed. This used Convolutional Neural Networks (CNNs) to detect driver drowsiness as a sign of fatigue. Three networks were introduced as potential networks for classifying eye conditions. One is a fully engineered neural network (FD-NN) and the other uses transfer learning with additional engineered layers in VGG16 and VGG19 (TL-VGG).
- 8) *Sukrit Mehta et al.* (Mehta et al., 2019) developed a light weight real-time driver drowsiness detection system implemented through an Android application. The system recorded video and used image processing technology to recognize the driver's

face in each frame. The system recognized facial features, calculated eye aspect ratio (EAR) and eye closure ratio (ECR), and detected driver drowsiness based on adaptive thresholds. A machine learning algorithm was used to test the effectiveness of the proposed approach (Fig. 3).

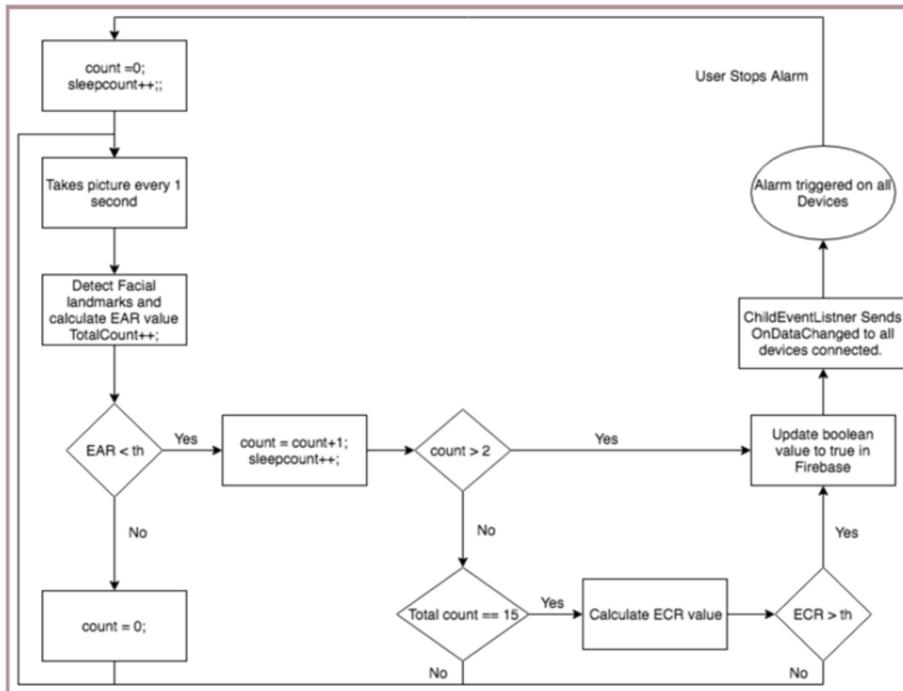


Fig. 3. Machine Learning algorithm was used to test the effectiveness of the proposed approach.

- 9) Reshma J et al.(Vennela, 2020) detect driver drowsiness while driving a vehicle and alert the driver using various applications of eye blinking, closing eyes and yawning. They used three approaches to detect fatigue: behavioral, physical, and vehicular as shown in Fig. 4.

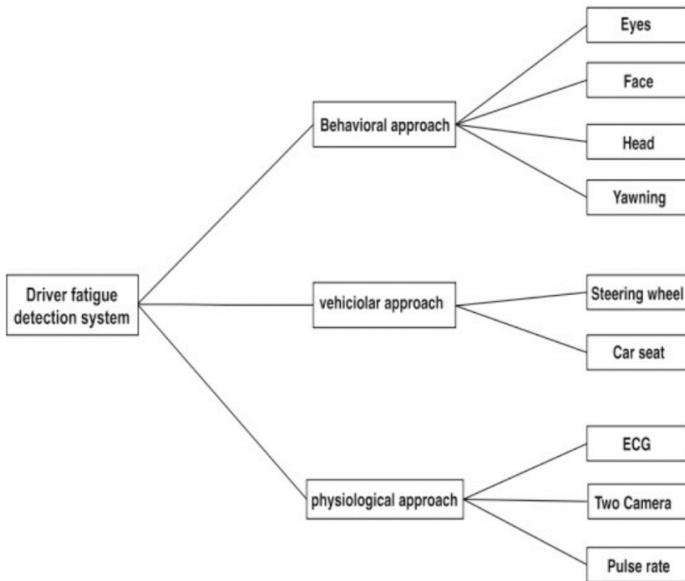


Fig. 4. Approaches to detect fatigue.

It utilized PERCLOS algorithm to examine eye closeness.

- 10) *Shivani Sheth et al.* (Sheth et al., 2020) developed a model based on the Haar Cascade algorithm with the OpenCV library for real-time driver vision tracking and driver eye detection. The system used the concept of eye aspect ratio (EAR) to determine whether the eyes were open or closed. A Raspberry Pi single-board computer with a camera module and an alarm system was also fabricated to simulate compact drowsiness detection.
- 11) *P. Sakthi et al.* (Sakthi et al., 2021) developed a model which works on an algorithm which has four stages: Haar Feature Selection, Creating Integral Images, Adaboost Training, and Cascading Classifiers. Architecture of system is as below (Fig. 5).

It completely focused on image processing to observe eye status. It involved MATLAB for programming.

- 12) *Prof.Ankita V. Karale et al.* (Karale et al., 2021) proposed an algorithm to detect, track and examine the face and eye positions of drivers to live PERCLOS, a scientifically valid measure of sleepiness associated with slow eye closure. In addition, they used a face recognition algorithm and a decision algorithm to detect face and eye.
- 13) *P.E. Kekong et al.* (Kekong et al., 2021) created and put into use software that records real-time driver behavior while they were driving and trained Convolutional Neural Networks (CNN); a deep learning algorithm, to anticipate the driver's conduct. A sleepy driver dataset, an intelligent video-based gadget, and CNN architecture and settings were developed to achieve this. The deep learning network designer program and MATHLAB were used to create the system. Using the necessary toolbox, such as

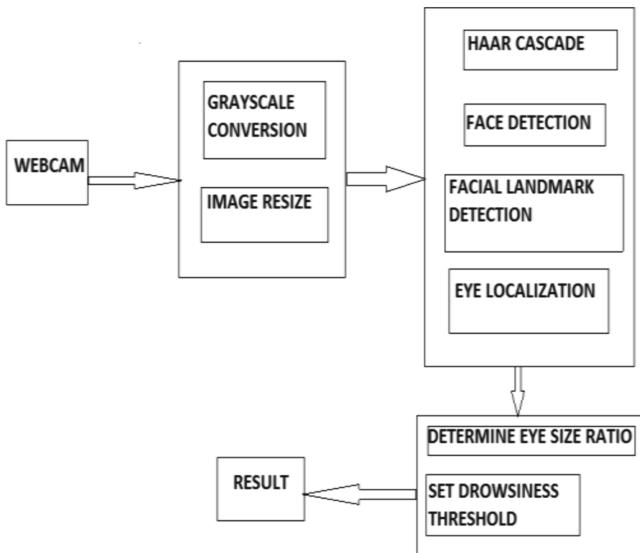


Fig. 5. Architecture of system

the image acquisition toolbox, computer vision toolbox, deep learning toolbox, and neural network toolbox to implement the designed models, the deep learning toolbox was used for the training process while the MATLAB was used to implement the algorithms as application software.

- 14) *Mahek Jain et al.* (Jain et al., 2021) developed a system where the ratio of the distances between the horizontal and vertical eye landmarks was used to calculate the EAR in order to identify drowsiness. A YAWN value was calculated using the distances between the upper and lower lips. This value was then compared to a threshold value for detecting yawns. The eSpeak module, a text-to-speech synthesizer, was used to provide appropriate voice alerts when the driver was drowsy or yawning. It utilized raspberry pi, & python to implement the codes.
- 15) *Singh Himani Parmar et al.* (Parmar et al., n.d.) developed a prototype module to alert driver's drowsiness that employed the four phases such as localizing the face, locating the eyes, tracking the eyes in succeeding frames, and detecting tracking failure. They employed a symmetry algorithm to locate the face & a raster scan algorithm to identify the eye and extracting the vertical location of the eyes. To identify if the eyes were closed or open, they utilized the horizontal histogram across the pupil.

3 Summary of Trends in Drowsiness Detection

The technologies evaluated in this article are summarized in Table xyz. To the best of the author's knowledge, patent and confidentiality requirements have prevented the manufacturers' systems' precise detection accuracy from being revealed. Regarding the detection and estimating techniques covered in this study, it is estimated that, depending

on the environment and the number of collaborators, the agreement rate with the degree of alertness based on subjective assessment is around 70% to 90%. If a certain technique is developed, it will probably be used to mass manufacturing. But as of now, no specific technique is employed in wakefulness estimation and detection technologies (Table 1).

Table 1. Summary of technologies engaged

Sr. no.	Authors	Methodology/Technology	Details
1	<i>Marco Javier Flores et al.</i>	Artificial inelegance, supervised classification method & support vector machine (SVM) Classification algorithm	Supervised classification method & support vector machine (SVM) Classification algorithm is used to find and monitor the face and the eyes in order to calculate a sleepiness indices
2	<i>BelaL AL SHAQAQI et al.</i>	artificial inelegance, symmetry algorithm, HTC algorithm	Symmetry algorithm to detect eye, Hough transform for circles (HTC) on the captured image of the eye to find the iris and did ANDing on circles to obtain intersection level
3	<i>T. Vesselenyi et al.</i>	EEG, ECG, artificial neural networks	An autoencoder network and a single hidden layer network along with image processing
4	<i>Ratnarup Dey et al.</i>	load cell, accelerometer, alcohol sensor	A load cell determined the driver's grip force on the steering wheelAlcohol sensor determined whether the driver is drowsy An accelerometer examined the head movement
5	<i>Eddie E. Galarza et al.</i>	Image processing, PERCLOS	Percllos calculated eye close time
6	<i>Zhipeng Ma et al.</i>	Accelerometer, Electromyogram (EMG), Electrocardiograph (ECG), and	An accelerometer examined the head movement Electromyogram (EMG), Electrocardiograph (ECG) measured voltage changes during sleepiness
7	<i>Maryam Hashemi et al.</i>	Convolutional Neural Networks (CNN)	Convolutional Neural Networks (CNNs) to detect driver drowsiness as a sign of fatigue

(continued)

Table 1. (*continued*)

Sr. no.	Authors	Methodology/Technology	Details
8	<i>Sukrit Mehta et al</i>	Android application, image processing, machine learning	calculated eye aspect ratio (EAR) and eye closure ratio (ECR), and detected driver drowsiness based on adaptive thresholds
9	<i>Reshma J et al.</i>	PERCLOS, image processing, camera and processing by MATLAB	PERCLOS algorithm to examine eye closeness
10	<i>Shivani Sheth et al.</i>	Haar Cascade algorithm with the OpenCV library,	Haar Cascade algorithm with the OpenCV library for real-time driver vision tracking and driver eye detection
11	<i>Prof. Ankita V. Karale et al.</i>	PERCLOS, image processing, camera and processing by MATLAB	Face and eye positions of drivers to live PERCLOS
12	<i>P.E. Kekong et al.</i>	convolutional neural networks (CNN); a deep learning algorithm, MATLAB	Convolutional neural networks (CNN); a deep learning algorithm
13	<i>Mahek Jain et al.</i>	raspberry pi, & python, EAR	Raspberry pi & python to implement drowsiness detection
14	<i>Singh Himani Parmar et al.</i>	Symmetry algorithm, ROSTER algorithm, histogram	Histogram used to detect eye closer

4 Outcomes of Analysis

Drowsiness detection trends are analyzed in a tubular form based on the methods and recording parameters given below (Table 2).

Table 2. Analysis of trends in drowsiness

Methods	Measuring parameters	Details	Accuracy
Contact methods	Driver's body parameters	Electromagnetic readings from wheel-mounted electrodes(Satti et al., 2021) Electrocardiographic observations using wearable sensors (Kundinger et al., 2020)	Up to 90%

(*continued*)

Table 2. (continued)

Methods	Measuring parameters	Details	Accuracy
Non-contact methods	Vehicular behaviour	HMI issues alerts when it notices changes in the behaviour of the vehicle. <i>(Jaguar. Jaguar Technology Is AWake-Up Call for Drivers. Available Online: https://Media.Jaguar.Com/News/2019/11/Jaguartechnology-Wake-Call-Drivers (), n.d.)</i> <i>(Volvo. Available Online: https://Www.Volvocars.Com/En-Th/Support/Manuals/V40/2017w17/Driver-Support/Driver-Alertsystem/Driver-Alert-Control-Dac (Accessed on 17 September 2021), n.d.)</i>	NA
		Use the steering angle as an input for ANFIS. [Arefnezhad] (Arefnezhad, S.; Samiee, S.; Eichberger, A.; Nahvi, 2019)	98%
		Utilizing inputs from the pedal pressure and steering, an ensemble network model estimates [Jeon] (Jeon, Y.; Kim, B.; Baek & Drowsy, 2021)	94%
	Driver's graphic information	Warnings for closed eyes (Toyota. <i>Toyota Enhances Pre-Crash Safety System with Driver-Monitoring Function. Available Online: https://Global.Toyota/En/Detail/248128 (Accessed on 17 September 2021), n.d.</i>)	NA
		Applied PERCLOS to visible images from a thermal imaging camera and used deep learning to classify “wakefulness,” “fatigue,” and “dozing.”	65 TO 70%

(continued)

Table 2. (*continued*)

Methods	Measuring parameters	Details	Accuracy
		Based on the temperature patterns of the forehead and cheeks, a support vector machine, a K-nearest neighbour method, and a decision tree were used to classify sleepiness. (Tashakori et al., 2021)	84%
		EAR (Chakravarthy., n.d.)	75%
		Detecting fatigue from driver's eye closure time, few blinks, and few yawns. (Li, 2020)	95%
		Learning gray scale face images with CNN. (Khanna et al., 2019)	98%

5 Conclusion

The primary drowsiness detection techniques and mass-produced technologies utilized today were described and analyzed in this study. Additionally, it was noted that all the researchers have focused only on using image processing as the primary technique to identify drowsy driving.

The use of radio frequency identification detection (RFID) to identify driver drowsiness received little attention. Therefore, there is a huge opportunity for studying on the efficient application of RFID for driver's drowsiness detection.

References

- Aishaqaqi, B., Baquaizel, A.S., El, M., Ouis, A.: DRIVER DROWSINESS DETECTION SYSTEM Laboratory signals and images (LSI) University of Sciences and Technology of Oran Mohamed Boudiaf (USTO-MB), pp. 151–155 (2013)
- Arefnezhad, S., Samiee, S., Eichberger, A., Nahvi, A.: Driver drowsiness detection based on steering wheel data applying adaptive neuro-fuzzy feature selection. Sensors **19**(4), 943 (2019). <https://doi.org/10.3390/s19040943>
- Praveen Chakravarthy, S.: Smart monitoring of the status of driver using the dashboard vehicle camera. Int. J. New Pract. Manag. Eng. **9**(01), 01–07 (2020). <https://doi.org/10.17762/ijnpme.v9i01.81>
- Dey, R., Paulose, J.: An improved algorithm for drowsiness detection for non-intrusive driving. Int. J. Appl. Eng. Res. **13**(2), 1219–1226 (2018)
- Flores, M.J., Armingol, J.M., De Escalera, A.: Real-Time Drowsiness Detection System for an Intelligent Vehicle, pp. 637–642 (2008)

- Galarza, E.E., Egas, F.D., Silva, F.M., Velasco, P.M.: Real Time Driver Drowsiness Detection Based on Driver's Face Image Behavior Using a System of Human Computer Interaction Implemented in a Real Time Driver Drowsiness Detection Based on Driver's Face Image Behavior Using a System of Human Computer Interaction Implemented in a Smartphone (2018). <https://doi.org/10.1007/978-3-319-73450-7>
- Hashemi, M., Mirrashid, A., Beheshti, A.: Driver safety development : real - time driver drowsiness detection system based on convolutional neural network. *SN Comput. Sci.* (2020)
- Jaguar. Jaguar Technology Is aWake-Up Call for Drivers. <https://media.jaguar.com/news/2019/11/jaguartechnology-wake-call-drivers>.
- Jain, M., Bhagerathi, B., Sowmyarani, C.N.: Real-time driver drowsiness detection using computer vision. *Int. J. Eng. Adv. Technol.* **11**(1), 109–113 (2021). <https://doi.org/10.35940/ijeat.A3159.1011121>
- Jeon, Y., Kim, B., Baek, Y.: Ensemble CNN to detect drowsy driving with in-vehicle sensor data. *Sensors* **21**(7), 2372 (2021)
- Karale, P.A.V., Patil, N.P., Sarvar, P.M., Sangle, P.M., Shinde, S.A.: Driver drowsiness detection system 1. **8**(5), 307–312 (2021)
- Kekong, P.E., Ajah, I.A., Chidiebere, U.: Real time drowsy driver monitoring and detection system using deep learning based behavioural approach. *Int. J. Comput. Sci. Eng.* **9**(1), 11–21 (2021)
- Khanna, A., Goyal, R., Verma, M., Joshi, D.: Intelligent traffic management system for smart cities. *Commun. Comput. Inform. Sci.* **958**(2), 152–164 (2019). https://doi.org/10.1007/978-981-13-3804-5_12
- Kundinger, T., Sofra, N., Riener, A.: Assessment of the potential of wrist-worn wearable sensors for driver drowsiness detection. *Sensors* **1**, 1–21 (2020). <https://doi.org/10.3390/s20041029>
- Li, K., Gong, Y., Ren, Z.: A fatigue driving detection algorithm based on facial multi-feature fusion. *IEEE Access* **8**, 101244–101259 (2020)
- Ma, Z., Yao, S., Zhao, J., Qian, J., Su, J., Dai, J.: Research on Drowsy-driving Monitoring and Warning System Based on Multi-feature Comprehensive. *IFAC-Papers OnLine* **31**, 784–789 (2018)
- Maftukhaturrizqoh, O., Nuryani, N., Darmanto, D.: Driver drowsiness detection using ANN image processing. *IOP Conf. Ser.: Mater. Sci. Eng.* (2017). <https://doi.org/10.1088/1757-899X/252/1/012097>
- Mehta, S., Dadhich, S., Gumber, S., Bhatt, A.J.: Real-Time Driver Drowsiness Detection System Using Eye Aspect Ratio and Eye Closure Ratio. 1333–1339 (2019)
- Parmar, S.H., Jajal, M., Brijbhan, Y.P.: Drowsy driver warning system using image processing. 78–83 (n.d.)
- Sakthi, P., Kiruthika, S., Surya, A., Venkatraj, C., Ks, V., Gokulakrishnan, S.: Detection of driver drowsiness using face recognition. *Turkish J. Comput. Math. Educ.* **12**(9), 2894–2900 (2021)
- Satti, A.T., Kim, J., Yi, E., Cho, H.: Microneedle array electrode-based wearable EMG system for detection of driver drowsiness through steering wheel grip. *Sensors* **21**(15), 5091 (2021)
- Ramalingam, V.V., Shivani, Aditya: Driver drowsiness detection system using machine learning algorithms. *Int. J. Recent Technol. Eng.* **8**(6), 990–993 (2020). <https://doi.org/10.35940/ijrte.F7514.038620>
- STATE/UT – Wise traffic accident cases during, pp. 117–128 (2019)
- Tashakori, M., Nahvi, A., Kiashari, S.E.H.: Driver drowsiness detection using facial thermal imaging in a driving simulator. *Proc. Inst. Mech. Eng.* **236**(1), 43–55 (2021). <https://doi.org/10.1177/09544119211044232>
- Toyota: Toyota Enhances Pre-Crash Safety System with Driver-Monitoring Function. Available online: <https://global.toyota/en/detail/248128>. Accessed on 17 Sep 2021. (n.d.)
- Vennela, G.S.: A Survey on Driver Drowsiness Detection Techniques. **IX**(Xi), 13–17 (2020)
- Volvo. <https://www.volvocars.com/en-th/support/manuals/v40/2017w17/driver-support/driver-alertsystem/driver-alert-control-dac>. Accessed on 17 Sep 2021. (n.d.)



Detecting Local Software Issues Using NSGA Multi-optimization

K. R. Jothi¹, Gireesh Kambala², Chetan Khemraj Lanjewar^{3(✉)}, Leena Jain⁴, Janjhyam Venkata Naga Ramesh⁵, and Pavitar Parkash Singh⁶

¹ Department of Computational Intelligence, School of Computer Science and Engineering,
Vellore Institute of Technology, Vellore, India
jothi.kr@vit.ac.in

² Web development. Teach For America, 2140 Hillside Derby Run, 30040 Cumming, Georgia
Gireesh.kambala@teachforamerica.org

³ KCC Institute of Management and Technology, Greater Noida, India
chetan.khemraj@gmail.com

⁴ Global Group of Institutes Amritsar, Amritsar, India

⁵ Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur, Andhra Pradesh, India
⁶ Lovely Professional University, Punjab, India

Pavitar.19476@lpu.co.in

Abstract. The primary goal of software testing is to identify and fix bugs. If a problem exists, it must be identified, tracked down, and fixed. As the adjacent tasks, fault detection and localization are difficult to merge. When it comes to fault localization, information helps to locate defects precisely, but when it comes to fault detection, test information is needed to find issues early. However, one thing that is necessary for both of these things is test information. Combining fault detection and localization becomes much easier with this prerequisite. This paper proposes a multi-objective approach that utilizes code and mutation coverage to provide a condensed test suite capable of identifying and pinpointing faults. The NSGA-II method was used for the optimization of the test cases. The findings from the analysis of the projects in the Defects4j repository demonstrate that the suggested approach is capable of producing smaller test suites that can effectively detect 94.16% of defects and accurately pinpoint the location of these discovered defects, with a fault localization score that is comparable to that of the whole test suite. They provided scores for problem identification and localization; the normal reduction percentage suite size is 77% which is an excellent reduction percentage.

Keywords: Test Information · NSGA · Fault Detection · Adjacent Tasks · Software testing

1 Introduction

There are many steps involved in creating software; nevertheless, debugging and testing are the most important. It is difficult for developers to create software without errors. Software difficulty is carried out after development to identify errors. The purpose of

debugging is to identify and eliminate the root cause of any defects that have been found. Software development is complete by first identifying and then fixing bugs. While fault localization entails identifying the suspect lines that caused a defect to arise, fault detection involves evaluating the system under test (SUT) to discover faults. About half of a software project's development effort should go into testing, and the other quarter should go into debugging because of how crucial error detection and localization are. When creating new software or updating existing software, there are two scenarios in which fault detection is necessary. The purpose of regression testing is to address changes made to previously released software. Reducing, choosing, or prioritizing test cases can make regression testing easier. Additionally, optimization methods may be used to decrease the number of test cases under the adequacy criteria that are considered, such as execution time, code coverage, mutation coverage, fault history coverage, and so on. An assortment of methods, including optimization techniques like NSGA-II, MOPSO, Harmony search, Genetic algorithm with diversity, and others, have been suggested for regression testing and fault identification. Conversely, fault localization relies on test cases that provide as much diagnostic data as possible. Two popular methods for finding errors are statement-based fault localization (SBFL) and mutation-based fault localization (MBFL). While MBFL methods employ test case pass/fail and killed-mutant coverage data, SBFL methods use test case statement coverage information to assess the suspiciousness of each line. When it comes to accurately localizing faults, MBFL procedures outperform SBFL techniques. A lot of work would be needed for fault localization if the suspiciousness value of code was calculated using all test cases. Optimization methodologies may be used in conjunction with fault localization methods to streamline the process. To optimize fault localization, refrain from eliminating test cases that provide the most diagnostic information. Various methodologies may be discovered in the literature for obtaining a condensed test suite that is appropriate for fault localization.

It is common practice to carry out fault detection and localization sequentially. To begin debugging, first use failed test cases to identify errors and then use the pass/fail data from those cases to pinpoint precisely where the problems are. Each task takes a lot of time, and optimization methods have been applied to each task separately to make them faster. This is because methods developed for detecting faults are only sometimes applicable when localizing them, and vice versa. The lack of sufficient diagnostic information in the test cases selected for fault detection makes it unable to provide valuable findings for problem localization, rendering it useless. Therefore, methods that save as much varied information from test cases as possible are necessary for defect localization. Each of the activities under consideration is also affected differently by test case reduction. Results showed that fewer tests were more effective in finding flaws than fault detection alone. Currently, there is a limitation in that defect detection and localization test case optimization are done independently. Minimizing effort is possible by merging the optimization processes of the two tasks. They need distinct test data, which makes this integration challenging. While problem localization necessitates the most diagnostic data available, such as pass/fail test case information, fault detection necessitates data that aids early defect detection. Therefore, it won't be accessible until regression testing is complete. Because the next version would have a different problem and the prior version would have fixed it, using pass/fail information from past versions will likewise not

be helpful. Regarding fault localization via test suite reduction, most methods are online and need pass/fail data on test cases. Integrating optimization of the best scenarios for defect identification and localization becomes challenging.

To disconnect, improper localization methods are used in conjunction with improper discovery techniques. The pass/fail status of test cases is not necessary for these methods. By combining these methods, a reduced test suite might be created before regression testing, which could aid in discovering and localising errors. It is possible to provide a combined method that reduces test cases for finding and detecting problems, as each job already minimizes test cases. Optimization of code attention and amount of code dividers led to the depreciation of test cases in their technique. Instead of developing separate sets of test cases for each action, combine them to create assessment cases suitable for finding and detecting errors.

Similarly, data on all test cases' coverage and partitions may be used to decrease the total quantity of test cases. Using the NSGA-II method, they maximized code coverage and code partitions simultaneously rather than combining test suites via union. The final NSGA-II reduced test suite successfully detected and localized flaws. It initially sorted test cases by fault detection, and test cases are re-prioritized for improved fault localization upon encountering a fault-detecting test case. The integration of defect detection and localization is the goal of all these studies. Another important step is choosing adequate criteria to identify and locate flaws. Similarly, choosing test cases with the most statement partitions aids fault localization. Partitioning aid's fault localization. See Sect. 2 for details.

To make things easier, optimization strategies and fault localization procedures may work together. When trying to improve fault localization, it's best to keep test cases that provide the most diagnostic information intact. Finding a streamlined set of tests suitable for fault localization is possible using a variety of approaches detailed in published works. Utilizing NSGA-II, these three parameters were fine-tuned. To find and locate mistakes for fault detection, the D-criterion, statement, and branch coverage are useful tools. According to the findings, the D-criterion for kill-mutants adequacy is better than the K-criterion. DAM-FL fault location technique based on D-criterion. Most Defects4j repository defects were solved better or as well by Metallaxis. Because it partitions mutants into distinct sets, the D-criterion effectively detects defects. One offline fault localization method is partitioning. Statements use it often, but mutations seldom do. Better MBFL approaches and mutant partitioning improve fault localization.

2 Methodology

This paper proposes a combined approach to improving problem identification and localization by decreasing the number of test instances. Statement coverage, branch coverage, and the diversity-aware mutation adequacy criterion (D-criterion) have been used to minimize test cases. The D-criteria for error detection is useful for both statement coverage and branch coverage. To address many facets of SUT(System under test) at once, this research considers multiple sufficiency criteria. Optimize these adequacy criteria simultaneously since they all need test case information. To clarify, the D-criterion asks for details about the mutations that the test suite eliminated. In contrast, statement and branch

coverage asks for details about the statements and branches that the test suite covered. MBFL approaches outperform SBFL ones because mutation testing allows for creating many bugs that do not exist in the code. Therefore, testers may understand the potential error types and the test cases that can identify them. Seeded mutant faults might cause the localization of genuine non-located faults if the behaviour of actual and seeded faults is identical.

One of the optimization capability criteria is the D-criterion, which is based on mutation testing and selects test cases that can partition and differentiate mutants. Covering the program's varied behaviour is made easier with its support. Statement partitioning has seen extensive application in both standalone and combined methods for finding and localizing errors. (Figs. 1 and 2) Unfortunately, fault localization has seen very little application of mutant partitioning. Therefore, using the suggested method, the authors have endeavoured to assess the efficacy of this declaration coverage and mutant partitioning combination for defect discovery and localization. The following three goals were maximized collectively using the NSGA-II multi-objective procedure to decrease the size of the test collection:

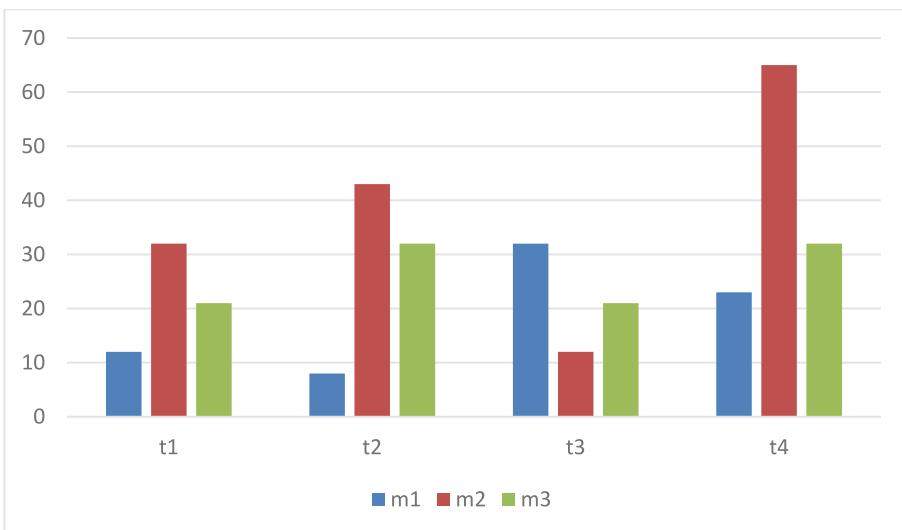


Fig. 1. Eliminating Mutants Via the Use of Test Cases.

To calculate the D-score, the diversity-aware mutation adequacy criteria are used. Pretend for a moment that it is possible to distinguish between all mutations. In such a case, it is calculated by adding up all the unique d-vectors and then dividing it by the number of unique d-vectors generated by the test suite. Maximizing the D-score involves running as few tests as possible while achieving the best possible D-score. The D-score

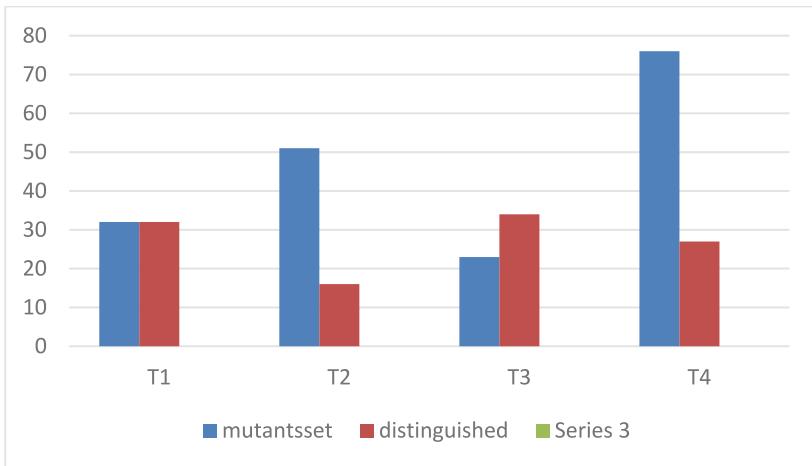


Fig. 2. The Test Suite generated Mutant Partitions.

may take on a range of values. Hence, the maximum D-score that is suggested is 1. Here is the equation to calculate the D-score using the original software p_o and a set of mutants M.

$$D - score(ts, m, p_o) = \{D(ts, p_{o,m}) | M \in m\}$$

3 Experimental Process

The section has covered the procedure used to execute the suggested technique and the artefacts needed to experiment.

3.1 Process

Figure 3 shows the experiment procedure. Defects4j repository's first projects were examined for data. Each test case's statement and branch coverage were calculated using Coverture. MAJOR was utilized to create mutants and analyze death data. Diversity-aware alteration adequacy criteria were utilized to divide and differentiate mutants. NSGA-II algorithm receives version-relevant test scenarios. NSGA-II generates subset test suites from these instances. When this occurs, the total number of unique d-vectors is added together and then divided by the total number of unique d-vectors produced by the test suite. This yields the value. To get the highest possible D-score with the fewest number of tests feasible is the goal of optimizing the D-score.

3.2 Assessment Method

Four projects were used to validate the suggested technique: chart, lang, time, and closure. The suggested methodology relies on mutation testing, known for its high cost. To

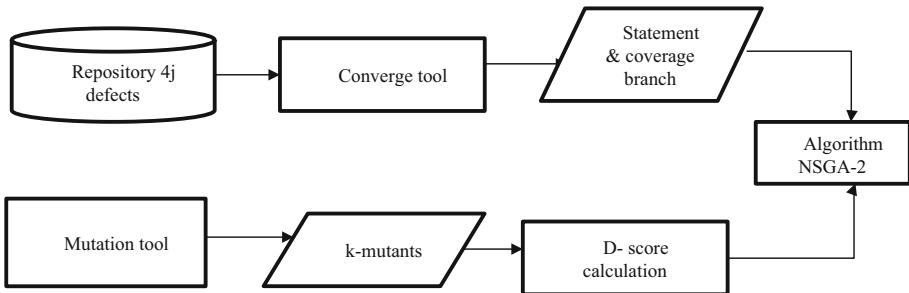


Fig. 3. Implementation Workflow for the Suggested Approach

apply the same strategy to regression testing, only the relevant classes of the projects that have been updated or altered and loaded are used. Modified classes refer to those that previously had errors and have now been rectified. During the execution of the tests that trigger them, certain classes are loaded. Triggering tests are useful for finding bugs. It should be noted that every project has its unique version. Reviewed all versions of the projects that are being considered. The versions without accessible data were eliminated. As a result, three variants, one from a chart and two from a timeline, were eliminated. The research assessed variations of all the projects under consideration.

3.3 Testing Suites

The tests that are linked to each version of a project are known as developer-authored tests. The term “relevant tests” refers to test cases in developer-written tests that load one or more pertinent classes, regardless of whether they have been changed or not. Writers are looking at the loaded and altered classes, hence the proposed technique considers all relevant test cases together. Both “relevant test group” and “full test suite” mean the same thing these days. The overall size of the test suite differs for each variety of the project provided the range of sizes for each project’s whole test suite. The goal is to minimize the number of tests to a least while achieving optimal results in terms of statement and branch coverage and, by doing so, separating as many potentially distinct mutants as possible.

4 Results

For each version of a project being examined, a reduced test suite is generated by running the NSGA-II procedure on the schemes. Then, count how many fault-detecting test cases are in each concentrated test suite of the examined project versions. Determining the mutants’ suspiciousness is also possible using the formulas. After that, find the most suspicious line of code by averaging the suspiciousness values of all its mutations. As mentioned in Sect. 2, rank is derived once each line’s suspiciousness has been calculated. After that, the reduced test suites are put through their paces regarding localization score and defect detection utilization. Because fewer apprehensive statements need to be inspected to identify and remove the problem, a high fault localization score from

a reduced test suite is encouraging. Additionally, the presence of even a single fault-detecting test case in the reduced test suite suggests that it can detect flaws. To determine the efficacy of NSGA-II in finding and localizing errors, compare the full test suites for each project version with the reduced test cases it generated. The results demonstrate that these condensed test suites were as effective as the relevant test suite in finding and localizing faults. The following subsections make up the results section.

4.1 Evaluation of the Power to Identify Errors

Following the procedure outlined, the NSGA-II algorithm's reduced test suites were assessed based on the proportion of mistakes detected. The minimum test suite includes a fault-detecting test case for every project version. There is a flaw in the version of the project being assessed if a fault-detecting test is included in the reduced test suite. Figure 4 shows the total number of defects and the missed reduced test suite. The minimized test suite failed to identify six versions in lang, three in time, and three in the closure project. The proportion of projects where the reduced test suite found flaws, was 95.16 per cent for all projects. It is possible to discover errors using only the statement, branch, and D-score metrics.

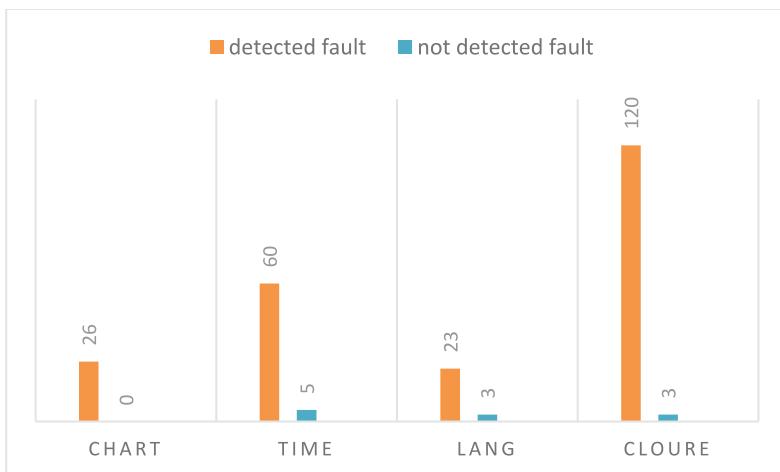


Fig. 4. Count of errors found in every project under consideration.

4.2 Analysing the Potential of Fault Localization

The mentioned fault localization score was used to analyse fault localization. Less statement inspection is required for fault localization when the score is high. Figures 5, 6, 7, 8, 9, and 10 show the fault localization scores for each project's complete and reduced test suites. These numbers illustrate the.

The dissimilarity between curves obtained from each version of the project under consideration when administered with either the whole or reduced test suite. Their smaller

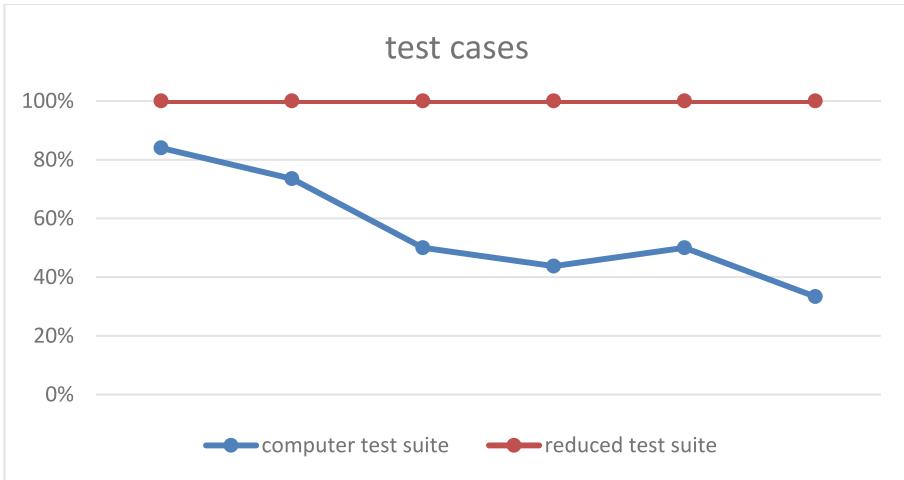


Fig. 5. Chart project fault localization score using full and reduced test suites.

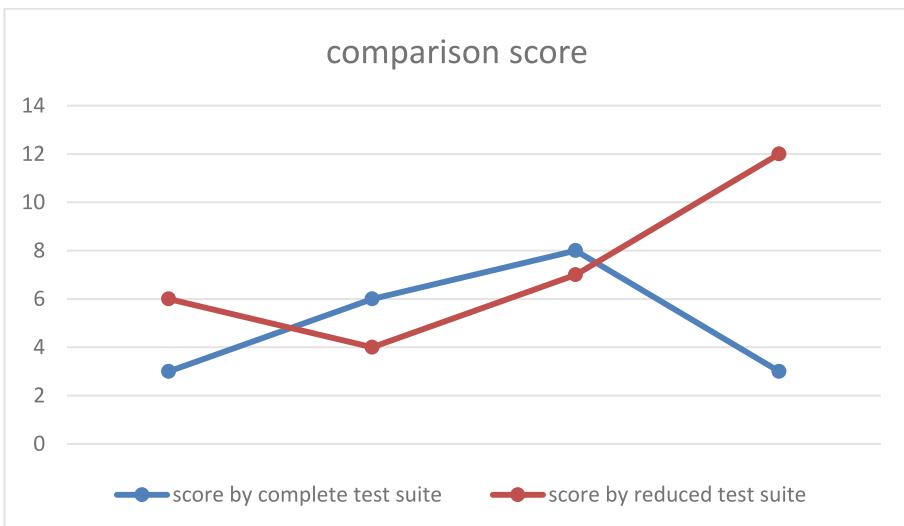


Fig. 6. Lang project fault localization score with full and reduced test suites.

test suites' fault localization capabilities are similar to those of the entire test suite. These numbers show that the reduced test suite successfully detected and removed non-mutable problems. A total of 38 errors occurred in lines that cannot be changed. Since the method employed for fault localization relies on mutants, they received a score of 0 in this area. This meant that it couldn't find bugs in these immutable versions. Out of all the mutable errors, the reduced test suite failed to discover four versions/faults in lang, two in time and two in closure. Consequently, the versions of the projects that were considered likewise did not have a fault localization score and the issue needed to be

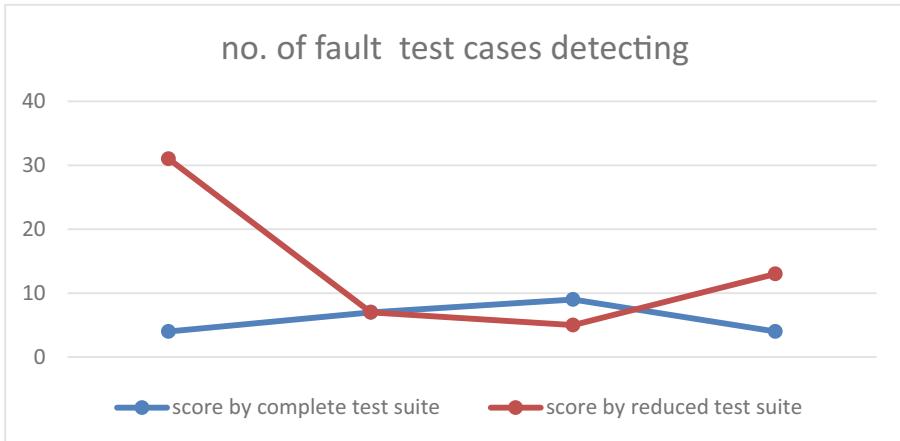


Fig. 7. Lang project fault-detecting test cases with entire and test-reduced suites.

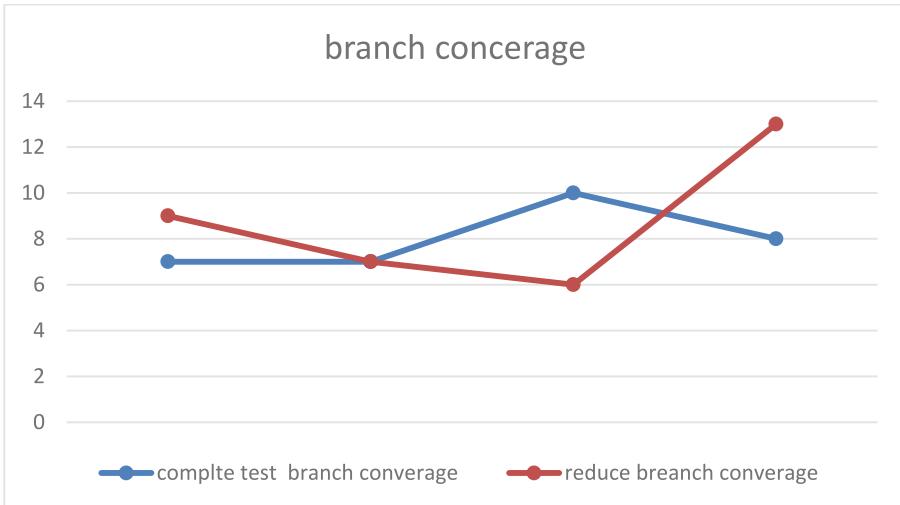


Fig. 8. Complete and test reduced suite coverage for time projects in branches.

located. On the other hand, for every other version of the project, the reduced test suite's fault localization score is almost identical to the complete tests. Despite the reduced size of the test suite, the suggested method still attempted to find the same number of faults. Agreed on the critical need for fault localization test scenarios. It is the most vulnerable as bypassing test cases do not cover this statement, but failing ones do. This is why it's crucial to include relevant defective test situations. The majority of the problematic test cases throughout the whole suite may be caught by the suggested technique. Display both the total and reduced amount of defective test cases beside the project chart, lang, time, and closing version. The results show that most fault-detecting tests were still part

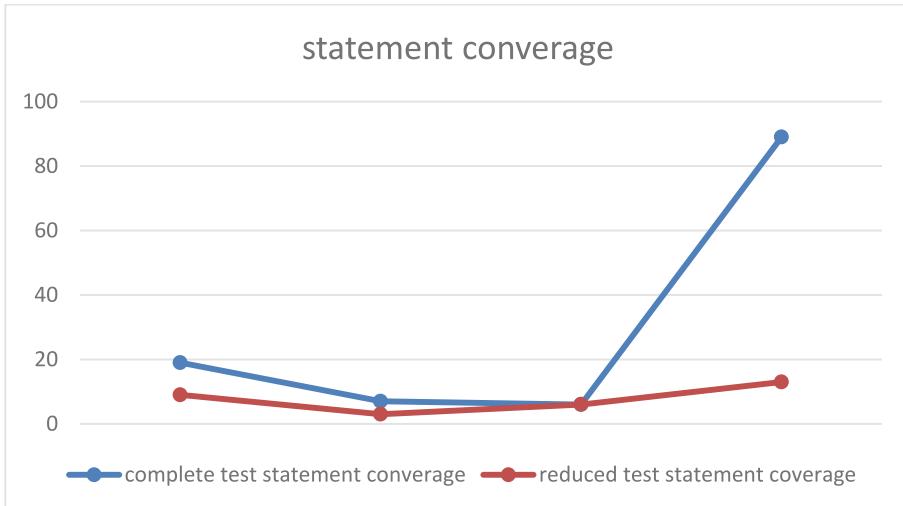


Fig. 9. Closure projects 1–60 statement coverage by full and reduced test suite.

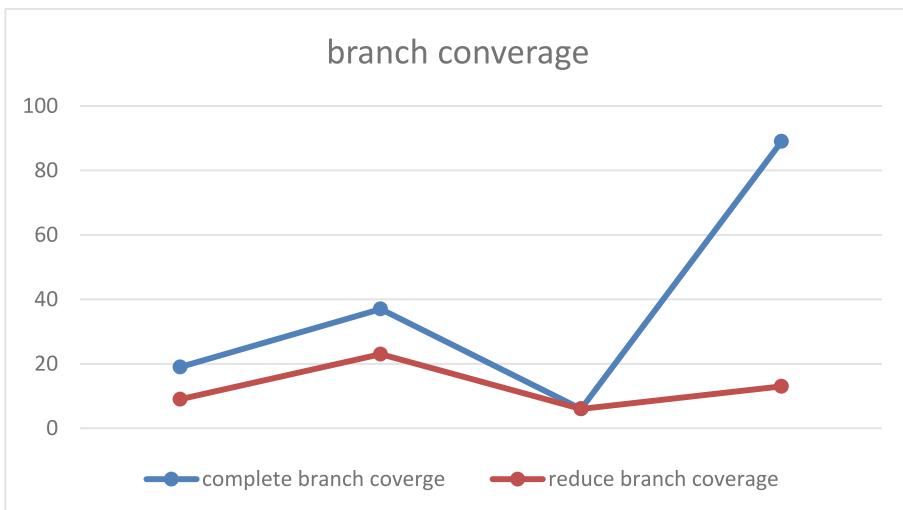


Fig. 10. Coverage of closure projects 1–60 by full and reduced test suite branches.

of the reduced test suite under the suggested method. The lines and mutants with the most suspiciousness values also indicated the real problems.

In the absence of partitioning, the suspiciousness grade of mutants eliminated by identical test cases would remain constant. The accuracy of MBFL is impacted by this. Mutant partitions should be relatively tiny on average, but there should be a large enough number to distinguish between mutants with defects. In the suggested technique, the D-criterion does this task. D-criterion can differentiate mutants and choose varied test cases

instead of similar ones. Thus, the fault localization score is excellent. The quantity of fault-detecting test cases has risen due to diversity. Thus, the D-score is an applicable fault localization criterion since it partitions test cases to keep maximal and diverse information.

5 Comparison with Existed Techniques

To the best of the information, there is a dearth of work on how to optimize test cases while simultaneously detecting and localizing errors. Enhanced test cases for better error detection and localization using statement partitioning and statement coverage metrics. When it came to defect localization and detection, they discovered that statement partitioning improved results. The concept of partitioning has also been used. Mutants, on the other hand, are what are used for partitioning. An improved and more efficient approach called DDU is suggested. The diagnosability measure recommends computing variety, uniqueness, and density metrics for the component under investigation before choosing test cases. According to their findings, the DDU-based method required less effort for analysis compared to the branch-coverage-based approach. The critical distinction is that their suggested method is limited to diagnostics, which rely on a failed test suite to pinpoint errors. Because the suggested approach employs an offline manner of fault localization, it can detect and localise faults. Therefore, test suite reduction does not need test cases' failure/passing information. All that is needed is the coverage data from the test cases. Also, although considered mutants in the analysis, their focus is only on code statements. To choose a test suite appropriate for reversion testing and fault detection, they have improved DDU and antique metrics using MOPSO. However, MOTSD's performance is far worse than ideal when choosing failed tests. Unlike those previous attempts, consider statements, branching, and mutations in addition to code statements. For both statement-based and mutation-based fault localization methodologies, it is best to choose test cases based on both adequacy requirements.

This study has a lot of room for improvement; for example, thus far, validation for defect localization has relied only on a mutant-based method. This resulted in poorly localized non-mutable defects. It is possible that in the future, a mix of declaration and mutation-based fault localization devices may be used to enhance the fault localization score. Since the suggested method reduces the number of test cases based on declaration coverage and mutation coverage, similar work in the future should be possible. The suggested method also requires an execution time study. Future studies should focus on developing a hybrid fault localization methodology that improves existing methods. Additionally, it is essential to conduct a temporal analysis of this strategy.

Using the mutant partitioning approach, the D-criterion has been influential in achieving excellent defect identification and localization scores in the present study. The controlled creation of mutants or the use of enhanced test suites may further increase the number of partitions. Finding the optimal method for mutant splitting is an area for potential future research. Additionally, only one component, namely the reduction of test suites for problem detection and location, has been implemented. Creating, selecting, and prioritizing test suites sufficient for defect identification and localization is another potential investigation avenue.

6 Validity Threats

The third avenue of investigation is that even with a reduced set of testing, several issues were discovered. The capacity of the D-criterion to distinguish between mutants makes the accurate results for fault localization available. Not only that but the test suite size has been reduced by an astounding 77% on average. There has yet to be an examination of the time and money needed to decrease the test group in this effort. Reducing the test suite size should make the suggested technique both time and cost-efficient, which is why it is crucial. Due to the time commitment involved, this investigation is necessary. Mutant-based testing is often reserved for mission-critical systems to ensure software integrity. As a result, it might find a home in mission-critical applications where speed is paramount above cost.

7 Conclusion

It is essential to have test case information to perform fault identification and localization. An oversized test suite is usually necessary for software testing due to the program's size. Both of these tasks involve reducing the number of test instances. Testers will have to do less work if they can incorporate test suite reduction for both tasks. To reduce test suites for fault localization and detection, the authors provide a method that incorporates diversity-aware mutation adequacy criteria, assertion, and branch coverage. With the test suite now smaller, regression testing and problem localization have been much easier thanks to the suggested method. On average, the size of the test suite is condensed by 77%, and the fault localization scores of most project versions are comparable to the whole suite. The rate of improper detection is 94.16 per cent. Therefore, by minimizing the test suite for both procedures, two tasks can be done simultaneously.

References

1. Cai, X., et al.: Unified integration of many-objective optimization algorithm based on temporary offspring for software defects prediction. *Swarm Evol. Comput.* **63**, 100871 (2021)
2. Khalili-Damghani, K., Abtahi, A.-R., Tavana, M.: A new multi-objective particle swarm optimization method for solving reliability redundancy allocation problems. *Reliab. Eng. Syst. Saf.* **111**, 58–75 (2013)
3. Ma, H., da Silva, A.S., Kuang, W.: NSGA-II with local search for multi-objective application deployment in multi-cloud. In: 2019 IEEE Congress on Evolutionary Computation (CEC). IEEE (2019)
4. Verma, S., Pant, M., Snasel, V.: A comprehensive review on NSGA-II for multi-objective combinatorial optimization problems. *IEEE Access* **9**, 57757–57791 (2021)
5. Marghny, M.H., et al.: A hybrid multi-objective optimization algorithm for software requirement problem. *Alex. Eng. J.* **61**(9), 6991–7005 (2022)
6. Almhana, R., Kessentini, M., Mkaouer, W.: Method-level bug localization using hybrid multi-objective search. *Inf. Softw. Technol.* **131**, 106474 (2021)
7. Singh, S.P., et al.: A soft computing based multi-objective optimization approach for automatic prediction of software cost models. *Appl. Soft Comput.* **113**, 107981 (2021)

8. Mkaouer, M.W., et al.: High dimensional search-based software engineering: finding tradeoffs among 15 objectives for automating software refactoring using NSGA-III. In: Proceedings of the 2014 Annual Conference on Genetic and Evolutionary Computation (2014)
9. Ouni, A., et al.: Maintainability defects detection and correction: a multi-objective approach. *Autom. Softw. Eng.* **20**(1), 47–79 (2013)
10. Jaeggi, D.M., et al.: The development of a multi-objective Tabu Search algorithm for continuous optimisation problems. *Europ. J. Operat. Res.* **185**(3), 1192–1212 (2008)
11. Vachhani, V.L., Vipul, K.D., Harshadkumar, B.P.: Improving NSGA-II for solving multi-objective function optimization problems. In: 2016 International Conference on Computer Communication and Informatics (ICI). IEEE (2016)
12. Guo, D., et al.: Chaotic-NSGA-II: an effective algorithm to solve multi-objective optimization problems. In: 2010 International Conference on Intelligent Computing and Integrated Systems. IEEE (2010)
13. Canfora, G., et al.: Defect prediction as a multiobjective optimization problem. *Softw. Test. Verif. Rel.* **25**(4), 426–459 (2015). <https://doi.org/10.1002/stvr.1570>
14. Seada, H., Deb, K.: U-NSGA-III: A unified evolutionary algorithm for single, multiple, and many-objective optimization. COIN report 2014022 (2014)
15. Alrezaamiri, H., Ebrahimnejad, A., Motameni, H.: Parallel multi-objective artificial bee colony algorithm for software requirement optimization. *Requirements Eng.* **25**, 363–380 (2020)
16. Mkaouer, M.W., et al.: Recommendation system for software refactoring using innovation and interactive dynamic optimization. In: Proceedings of the 29th ACM/IEEE International Conference on Automated Software Engineering (2014)
17. Pradhan, D., et al.: CBGA-ES+: a cluster-based genetic algorithm with non-dominated elitist selection for supporting multi-objective test optimization. *IEEE Trans. Softw. Eng.* **47**(1), 86–107 (2021). <https://doi.org/10.1109/TSE.2018.2882176>
18. Singh, C., Rao, M.S.S., Mahaboobjohn, Y.M., Kotaiah, B., Kumar, T.R.: Applied machine tool data condition to predictive smart maintenance by using artificial intelligence. In: Balas, V.E., Sinha, G.R., Agarwal, B., Sharma, T.K., Dadheech, P., Mahrishi, M. (eds.) Emerging Technologies in Computer Engineering: Cognitive Computing and Intelligent IoT: 5th International Conference, ICETCE 2022, Jaipur, India, February 4–5, 2022, Revised Selected Papers, pp. 584–596. Springer International Publishing, Cham (2022). https://doi.org/10.1007/978-3-031-07012-9_49
19. Hojjati, A., et al.: Application and comparison of NSGA-II and MOPSO in multi-objective optimization of water resources systems. *J. Hydrol. Hydromech.* **66**(3), 323–329 (2018)
20. Langdon, W.B., Harman, M., Jia, Y.: Efficient multi-objective higher order mutation testing with genetic programming. *J. Syst. Softw.* **83**(12), 2416–2430 (2010)
21. Chowdhury, S., Sesharao, Y., Abilmazhinov, Y.: IoT based solar energy monitoring system. *Int. J. Eng. Technol.* **9**, 1905–1908 (2021)
22. Jalili, A., et al.: Controller placement in software-defined WAN using multi-objective genetic algorithm. In: 2015 2nd International Conference on Knowledge-Based Engineering and Innovation (KBEI). IEEE (2015)
23. Jothi, K.R., Yawalkar, P., Mamatha, V.L.: Automatic Speech Assessment System for Aphasia Speech Disorder. Annals of the Romanian Society for Cell Biology 5382–5392 (2021). <https://www.annualofrscb.ro/index.php/journal/article/view/6425>
24. Zhang, G., et al.: Constraint handling in NSGA-II for solving optimal testing resource allocation problems. *IEEE Trans. Rel.* **66**(4), 1193–1212 (2017)
25. Hu, T., et al.: An efficient approach to robust controller placement for link failures in Software-Defined Networks. *Fut. Gener. Comput. Syst.* **124**, 187–205 (2021). <https://doi.org/10.1016/j.future.2021.05.022>



Deep Learning and IoT Based Robotics to Monitor the Traffic

V. Vishwa Priya¹(✉), Soumitra S. Pande², Md Ilyas³, R. Jayasudha⁴,
Janjhyam Venkata Naga Ramesh⁵, and D. Suganthi⁶

¹ Department of Computer Science, Vels Institute of Science, Technology and Advanced Studies, Chennai, India

vishwapriya13@gmail.com

² Ericsson Global India Private Limited, Bangalore, India

³ Prestige Institute of Engineering Management and Research, Indore, India
milyas@piemr.edu.in

⁴ Dr. N. G. P. Institute of Technology, Coimbatore, India

⁵ Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur, Andhra Pradesh, India

⁶ Saveetha College of Liberal Arts and Sciences, SIMATS, Thandalam, Chennai, India

Abstract. In densely populated areas, managing traffic laws has become more difficult due to the rising number of cars on the road. One of the most critical components of effective traffic management and improved mobility is real-time traffic monitoring systems. Consequently, precise and dependable real-time traffic information has always been essential for cars and drivers. Many other approaches have been suggested as potential answers to the issues plaguing traffic conditions recently. A different approach is VCC or vehicle cloud computing. In addition, an advanced model called an IoT-aided robotic (IoRT), which uses cameras and Internet of Things (IoT) detector nodes to gather actual traffic information, has been created. Two deep learning methods, one based on improved LeNet-5 for real-time signs for traffic identification and the other on the Inception-V3 model for detecting and identifying traffic lights, are the critical contributions provided by this research effort. Additionally, to decrease road accidents, the ideal distance between barriers and ultrasonic sensors was determined by analyzing the timing and speed of ultrasonic waves. Data acquired by cameras and sensors is analyzed using a variety of image processing algorithms before being uploaded to the cloud and made accessible to commuters and drivers via a mobile app. According to the test findings, the suggested models are far more accurate. The expanded GTSRB (EGTSRB) dataset achieved 98.89% accuracy while the German Traffic Sign Recognition Benchmark (GTSRB) dataset achieved 98.23% accuracy using the modified LeNet-5. The second model achieved a precision of 99.7 percent using data from the Laboratory for the Intelligent and Safe Automobiles (LISA). The results of this study beat previous research on comparable traffic monitoring systems by 2.13% for traffic light recognition and detection and by 4.89% for traffic sign identification.

Keywords: Traffic Management · Vehicle Cloud Computing · Internet of Things · Deep Learning Method · GTSRB

1 Introduction

A wide variety of electronic gadgets, such as cars, appliances, phones, smart wearable's, and other machinery, are part of what is known as the "Internet of Things" (IoT). Data may be sent between these devices over the Internet thanks to an application programming interface (API). Everyday problems are the target of the IoT [1–3]. Also, the transportation, healthcare, and industrial sectors are just a few that stand to benefit from IoT technology and its solutions [4, 5]. The term "Internet of Vehicles" (IoV) describes the phenomenon that occurs when cars connect to the web ad hoc manner [6]. This soon-to-be-popular technology aims to link automobiles to the web and one another, facilitating real-time data exchange and interaction. Internet of Vehicles (IoV) technology may revolutionize transportation by making roads safer, reducing traffic, and increasing mobility. Significant contributions to this research include these. Using modern technology like artificial intelligence (AI), 5G networks, and the IoT, IoV systems enable vehicle coordination and interaction with traffic setup like signals and signs for traffic [7–9]. Developing an intelligent and functional transportation system will improve each person's quality of life. VCC is a cutting-edge technological advancement that offers drivers pay-as-you-go services over the cloud. Consequently, VCC's primary goals are to lessen accidents, traffic jams, time spent travelling, and pollution by offering drivers affordable computing services. It is believed that the VCC enables the integration of pervasive technologies like Intelligent Transport systems (ITS), Wireless Sensor Networks (WSNs), and Mobile Cloud Computing (MCC) to enhance the security and intelligence of traffic systems for urban and make roadways safer [10, 11].

Improving mobility and providing commuters and drivers with real-time information is the goal of road monitoring, which entails collecting, analyzing, and exploiting data about roads [12]. Thanks to developments in related fields of study, it is now an essential part of ITS. Integrating many technologies, including sensors, cameras, and the Internet of Things, has also made it likely to collect information about traffic conditions and roadways in real time. According to sources, the Internet of Things (IoT), networks for communication, storage and processing of information systems, user interfaces, and road monitoring systems are the usual components. Vehicles and roadways with Internet of Things (IoT) sensors can monitor traffic and road conditions [13–15]. Communication networks, such as cellular networks or Wi-Fi, allow for real-time information dissemination among these gadgets and the central scheme [16]. Vehicles can enhance traffic flow, avoid crowded places, and choose the optimal route according to the processed data [17]. This is an excellent tool for drivers to update themselves on road conditions. The significance of road monitoring systems is not without its obstacles; they comprise but are not limited to, the following: the expense of installation and maintenance, the dependability and accuracy of data, the absence of system interoperability, and the sharing of personal data, such as the location of vehicles [17–19].

Many research have attempted to address roadway linked difficulties, but they have met various obstacles that still need to be resolved. The utilization of deep learning models, which are renowned for their higher reliability and precision, the combination of vision-based systems and IoT sensors for collecting information in real-time, and, maximum significantly, the absence of a dedicated information Centre or cloud-based systems for analyzing and storing the collected data are all factors to consider [20, 21].

Moreover, the gathered data is not accessible to drivers and commuters, which limits its ability to enhance mobility. Using a hybrid approach that draws on visual and sensor data, this research improves the precision of sensing and identifying traffic lights and signs using two modified deep-learning approaches. The data is collected in real time from the roads. After that, the data is wirelessly sent to the Firebase platform to be stored and analyzed further. Subsequently, a smartphone notification system tailored to this research disseminates this processed data in real time to commuters and drivers, including details on the times and places of accidents, traffic jams, road closures, and more. Improving road safety, decreasing traffic accidents, and alleviating congestion are just a few ways this method will enhance transportation [22, 23].

This work makes several significant contributions. (1) The clue behind the Internet of Things enhanced road monitoring is to use cameras and sensors to gather information about roads to track specific characteristics in real-time to improve mobility. (2) The Autonomous Robotic Car (ARC) is an integral part of the suggested system; it drives itself using Internet of Things (IoT) sensors, cameras, and two enhanced deep learning algorithms. Using robots, the system improves the capacity of Internet of Things devices to gather data on roads in real time. (3) The system is designed to enhance navigation and traffic management using real-time data analysis, Internet of Things devices, and robotics. Its purpose is to provide drivers, commuters, and vehicles with the most current information possible. (4) To minimize road accidents, the ideal distance between barriers and ultrasonic sensors is calculated by calculating the light weight duration and the speediness of ultrasonic waves.

2 Roadway Monitoring IoT Technologies

Rapid use of IoT tools in roadway monitoring is driving improvements in mobility. These technologies use Internet of Things sensors, cameras, and other linked devices to provide immediate information on road conditions. The information is sent to a cloud platform for storage and analysis. Then, drivers and vehicles may utilize the data to improve mobility in many ways, such as minimizing accidents, improving traffic flow, and more. In addition, Internet of Things (IoT) technology used for road monitoring will revolutionize transportation by making roads safer, reducing travel times, and increasing passengers' experiences [24, 25].

2.1 Road Monitoring IoT Sensor Types

Many people are increasingly interested in and using IoT sensor technology. Many sectors, such as healthcare, agriculture, road and vehicle monitoring, and many more, have begun using IoT sensors. Multiple traffic management, monitoring, safety, etc., applications are designed and developed with the help of Internet of Things sensors in the transportation sector. As previously said, a car has several sensors. However, when these vehicles gain intelligence, that number may increase to maximum number of sensors. Therefore, sensors connected to the Internet of Things might be crucial in fixing many transportation problems. For purposes such as commercial and emergency information

services, vehicle detection, real-time traffic management, and traffic monitoring, computerised traffic schemes have utilized a range of Internet of Things (IoT) detectors and sensors. As shown in Fig. 1, there are six different kinds of Internet of Things (IoT) sensors, each designed for a specific purpose to aid in traffic control [26–28].

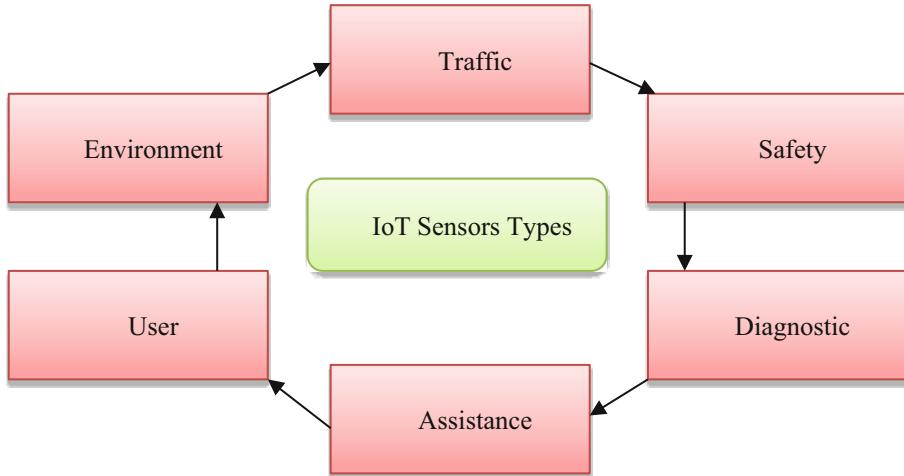


Fig. 1. Six Different Kinds of Internet of Things (Iot) Sensors

IoT measuring device in transport, especially for traffic observing, may be implemented on-board or remotely. Drivers, passengers, and the vehicle's surroundings are monitored via on board and remote sensors. IoT measuring device in transport, especially for traffic monitoring, may be implemented on-board or remotely. Vehicles have on-board sensors to monitor passengers, car driver, and surroundings, while distant detectors monitor roads, accidents, weather, and congestion.

2.2 Road Monitoring Iot Systems

A country's road system, transportation options, and travel policies influence every individual's day-to-day life. The World Health Organization found that a country's health and development rate are significantly impacted by the condition of its roads and traffic flow. Additionally, the research highlighted that 2.24 million individuals lost their lives, and 55 million were wounded in road accidents. Therefore, the large numbers show how quickly we need answers. Several methods exist to accomplish this goal, such as better cars, roads, road monitoring, and more driver and vehicle monitoring. The Internet of Things (IoT) is an essential part of this movement, which may improve efficiency, connection, and coherence. The IoT has shown its worth in a number of transportation-related areas, such as traffic and vehicle tracking, vehicle repair, navigation, and more. The latest data on road conditions in different locations may be provided via road monitoring systems that are Internet of Things (IoT) based. Consequently, these technologies are designed to improve road safety, efficiency, and sustainability.

First and foremost in an Internet of Things (IoT) traffic monitoring scheme are (1) IoT device, (2) communication device, and (3) data analysis software. Internet of Things (IoT) devices, such as cameras and sensors, may be installed in cars or along highways to collect real-time data on a selection of parameters, such as traffic flow, roadway conditions, weather, and others. The collected information is sent to a centralized scheme for processing and analysis to provide valuable insights via wireless communication technologies like WI-Fi, Bluetooth, or cellular networks. Through the use of actuators, sensors, and other Internet of Things (IoT) devices that gather and send data to a control Centre, the transportation sector is able to connect physical things into a unified network. Vehicles, highways, stops for buses, street lights, and more may all have sensors and beacons installed thanks to the Internet of Things. Consequently, there will be less traffic-related problems. An IoT road monitoring system will be productive for the sector.

2.3 Methods of Data Gathering, Storage, and Analysis

To guarantee the security and effectiveness of their roads, transportation organizations worldwide depend on road monitoring systems. To provide up-to-the-minute information on road-related problems, these systems use several data-gathering, storing, and analyzing methods. This research will examine how traffic monitoring systems gather, store, and analyze data.

The researchers utilized GPS tracking, video surveillance, and traffic sensors to gather roadway information. To top it all off, it uses sensors that are already installed on-board instead of using expensive and prone to malfunction distant sensors. Other methods are used to gather road statistics than those listed above, including computerized toll collection and citizen reporting.

3 Proposed IoT Road Monitoring System

An Internet of Things (IoT) system for current traffic and roadway condition monitoring is suggested in this research. The suggested system collects data about road conditions, traffic, and road closures and works network-wide using sensors and Internet of Things devices. With the system's real-time data, commuters, drivers, and autonomous cars can all plot their routes. The technology can also track information over time to find trends and patterns, which may improve traffic flow and mobility. The ARC has been created using various software and hardware components. The ARC uses a smartphone app to communicate with commuters, drivers, and stores and analyses real-time data using Google's Firebase.

3.1 Actual Road Information from IoT Vision and Sensors

There are two main ways that roadways may be monitored: using vision-based systems or Internet of Things (IoT) sensors. Deploying sensors to gather real-time data is essential for a sensor-based strategy. The following data analysis derives insights into road use and possible traffic flow problems. On the other hand, a vision-based solution uses cameras to record the road in real time. Computer vision algorithms are used to evaluate the film

and extract data about the road conditions and the behaviour of the vehicles. The two methods complement one another effectively and, when used together, provide thorough road monitoring solutions while mitigating some of the problems that could occur when using either method alone. The ARC collected real-time data utilising several sensor-based sensors, such as infrared, ultrasonic, and GPS. It is possible to use infrared sensors to detect weather conditions, obstacles, potholes, and patches. Contrarily, ultrasonic can detect and identify everything near the ARC, including people, cars, and more. The data might provide two potential advantages: improved vehicle safety and traffic mobility. The ARC can also share its precise position thanks to its built-in GPS sensor. Additionally, the ARC uses the surveillance camera to collect data in actual using a visual based technique. The camera can record road signs and lights because of the deep learning algorithms suggested for the ARC.

Consequently, 44 distinct traffic signs and light conditions may be identified. As a result, the ARC may collect and communicate any vital data it finds, such as a road condition indicator or other relevant notices. These two deep learning models use different information pre-treatment methods to prepare their input data for the model and make it more accurate. In addition, the data pre-treatment methods greatly enhance the accuracy of the models. Combining the two methods may enhance traffic monitoring and reduce a number of road-related problems, according to the results after rigorous testing.

3.2 Autonomous Robotic Vehicle

A self-driving vehicle, or ARC, does not need a human driver. It can sense its environment, navigate, and make real-time decisions using information from IoT sensors, software, and ML algorithms. According to, these cars might drastically improve transit accessibility, efficiency, and the number of accidents caused by human error. On top of that, they may coordinate their actions on the road with other cars and pedestrians. Several software and hardware components were used in the design and development of the ARC in this research to assess the study's goals.

3.2.1 Architecture for Hardware and Mechanisms

This research's ARC was built using an Arduino Mega 2560 Rev3 microcontroller and a Raspberry Pi 4B embedded system. As a bonus, it comes with a camera, a NEO-6M GPS, three UAV sensors, four IR sensors, two batteries, and a driver for direct current (H-bridge DC) motors, four tires, four DC motors, two 4WD acrylic plates, jumper wires, and a small number of other parts that need to be put together. The parts are displayed in the image below.

A power bank powers the Pi Raspberry and other linked gadgets, while the H-bridge driver for motors, Arduino Mega, the and all of the detectors are powered by two 18650 Li-ion batteries.

The camera, which captures data about the actual surroundings (such as neighboring cars, signal lights, and more), is attached to the Raspberry Pi via its USB connection, acting as the system's input. The Raspberry Pi performs several deep learning techniques on the collected real-time data before sending it to the Arduino Mega. In addition, the Raspberry Pi is equipped with the NEO-6M GPS, which allows for transmitting the

ARC's current position. An Arduino Mega regulates the motor driver of the ARC and is linked to all the detectors, including the infrared and ultrasonic ones utilized as inputs. Six jumper wires link the motor driver to the Arduino Mega. Out of these, four wires control the direction of rotation of the DC motors, and two wires control the speed of rotation using PWM (Pulse Width Modulation). Also, data processed on the Raspberry Pi may be sent to the Mega by the Arduino and processed on the Raspberry Pi again via serial communication between the two devices. The two 4WD acrylic panels are entirely covered with hardware. Figure 2 is a schematic diagram demonstrating the connections between the ARC's components.

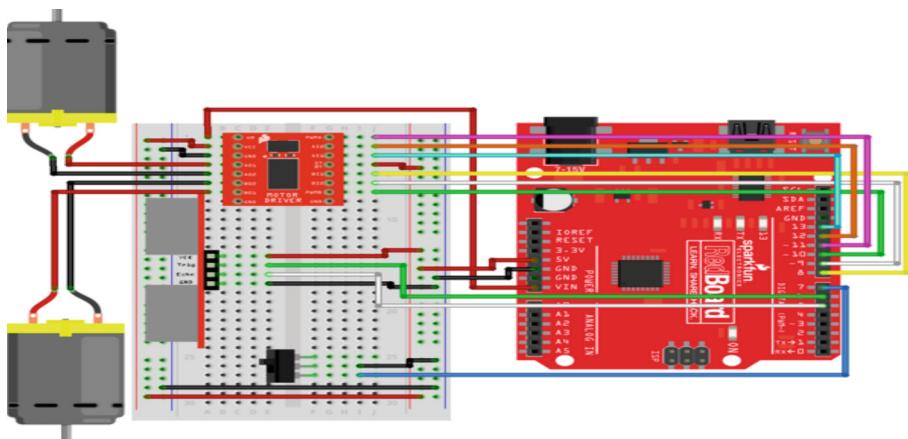


Fig. 2. ARC Schematic Diagram

3.2.2 Programming Elements

Python, Open CV, Dart, the Arduino programming language, and Google Collaborators are some of the software components necessary to complete the ARC.

3.3 Fire Base Actual Data Base Stores Information

A cloud-hosted information is usually meant when the phrase “Firebase actual database” is utilized in prior research using Firebase. Offering backend-as-a-service (BaaS), this real-time database service facilitates the development of various online and mobile apps. It doesn't employ queries to insert, remove, or modify data; instead, it uses the JavaScript Object Notation (JSON) format for storing information. Data management for iOS, Android, and web apps is easy with Google's mobile app development platform, which offers several features like storage, authentication, cloud messaging, reporting, notifications, and a real-time database. In addition, it offers data on how these services are used. There are several reasons to utilize Google Firebase, including its easy control interface, scalable and safe architecture for saving and sharing massive amounts of data, simplicity of usage and settings, and actual information that updates linked clients instantaneously.

with each modification. Additionally, this research uses a real-time database to store road information and share it with customers via an application for mobile; this is one of the most exciting services. Additionally, the client handled all code without database access or server-side layers. Clients may react to changes in database data by altering their user interfaces via triggered events in their code. The suggested solution uses “Firebase actual database safety rules” to organize information and specify user access.

3.4 Flutter Mobile App Real-Time Data Transfer Technology

Flutter is a well-liked framework for building mobile apps; it allows developers to make apps that operate well on iOS and Android. Developers may build dependable apps with the help of Flutter’s numerous communication protocols, which allow for real-time data transmission. Flutter may be used with several communication technologies for actual information transfer, such as Web Socket, Firebase actual database, SignalR, MQTT (Message Queuing Telemetry Transport), Socket.IO, Push Notification, etc. These communication methods create a solid foundation for building real-time transmission-capable Flutter apps. The ideal technology depends on the application’s specific needs and the type of information sent. Hence, this research used the Fire Base actual database techniques to send road information via a mobile application. One of the requirements of the proposed system was that all clients automatically get upgrades with the most current data while constructing cross-platform apps using Flutter and Firebase.

This research proposes ROSNOS, a customer’s mobile app built using Flutter that works on iOS and Android. So, motorists and commuters may use it to link up with the Fire Base, which has gathered all data related to road conditions, and get up-to-the-minute information. Reduced travel time, less energy use, and less congestion may be achieved using this data. The present whereabouts of an ARC may also be ascertained by its owner using ROSNOS.

4 Discussion and Results

An extensive analysis of several prior research confirms that mobility will unquestionably be improved using real-time road monitoring technologies. Similarly, several road problems will be lessened, such as accidents, congestion, and traffic flow as depicted in Fig. 3. The ARC initially uses a camera, infrared and ultrasonic sensors, and other technologies to gather data while monitoring highways. After sorting, the data is sent to the platform Google’s Fire base over Wi-Fi for actual storage. Firebase will take care of any further processing that may be required. In addition, the data is organized in the Firebase database in real-time according to three standards: the date and time of the information, the name of the street, and the status of the road. In conclusion, ROSNOS users may get up-to-the-minute road data using Firebase’s real-time database technology. Users who have provided this information will get notifications. So, by using this technology, they can help speed up their productivity, especially during peak hours, while minimizing energy usage and accidents. Since ROSNOS is a mobile app in a web browser, users will need an active Internet connection to access data saved in Firebase.

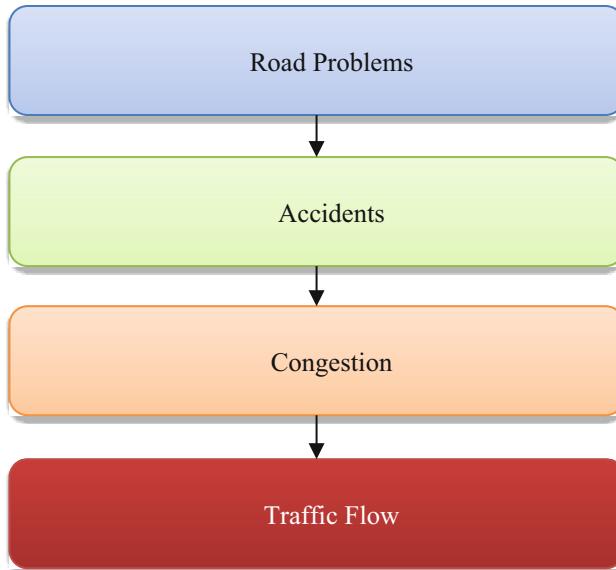


Fig. 3. Road Problem Types

This work uses two models based on deep learning and has been approved by two academic publications. The first model can identify traffic signals using a tweaked version of the Lenet-5 Convolutional Neural Network (CNN) architecture. Consequently, several traffic signals that impact traffic congestion may be identified and communicated with others, such as general caution, road closures, careful of snow/ice, and prohibited entrance. This model uses both the GTSRB and the EGTSRB for training and evaluation. To create an EGTSRB with balanced data across classes, the GTSRB and the Belgium TS databases were combined. Furthermore, several picture pre-processing methods were used to improve the quality of the images and get them ready to be fed into the model. The model's accuracy has been greatly enhanced by using these pre-processing processes compared to scenarios where they were not used. It is essential to follow these measures to enhance the quality of your photographs. Photographs of road signs, paired with different weather conditions, might result in blurrier or lower-quality images. Eighty percent of the information was used to train the model, twenty percent went into testing, and twenty percent came from the 85 percent used to validate the model. Hyper parameters and precision results of the model are shown in Fig. 4.

The second model can identify and distinguish between different types of traffic signals; it was developed utilizing Inception-V3 network-based transfer learning. Therefore, the ARC will alert other drivers and vehicles on this route of the congestion when it senses a red traffic light and a protracted standstill. A number of image preparation methods were used to ensure that the input photos were of high enough quality for this model. With the help of these methods, the model's accuracy has been enhanced. Furthermore, this strategy aims to lessen many road-related problems that arise from disobeying traffic signals. To train and test this model, the LISA information was used.

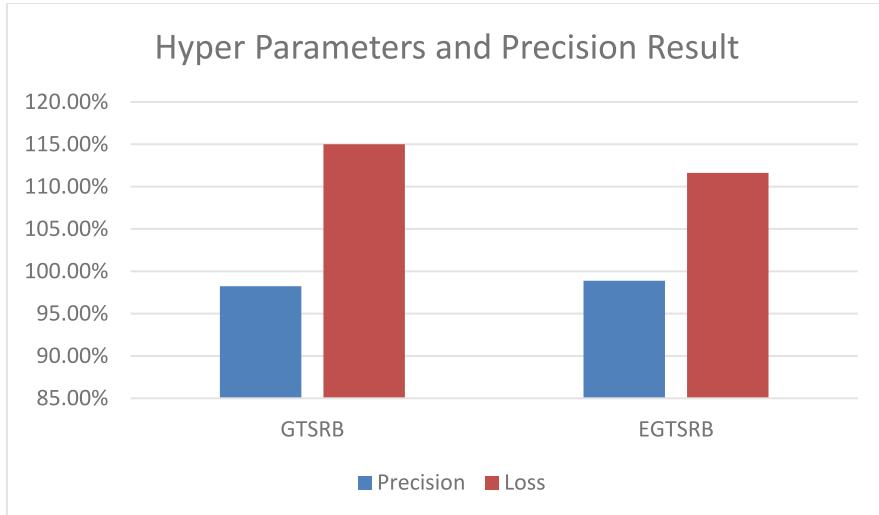


Fig. 4. Hyper Parameters and Precision Results of the Model

The suggested system uses vision-based and sensor-based methods to gather actual road information. Sensor-based technology makes use of infrared and ultrasonic sensors. Infrared sensors can detect weather conditions, road hazards like potholes and patches, and other road conditions. In addition, the ARC may be programmed to follow visual lines, such as white and black ones, using an infrared sensor. As shown in Fig. 5, the ARC is equipped with four infrared sensors—two on each side—to provide the most precise and dependable detection possible. Consequently, using two sensors to identify lines or other things provides much higher precision than other systems, enabling identification even if the initial sensor makes a mistake. These sensors can detect the line's reflection and send that information to the Arduino platform, which may modify the ARC. It tells the ARC to go in a specific direction, make a left or right turn, and then stop.

Furthermore, the ARC's ultrasonic sensor gathers information about nearby things, including cars, people, and more. The ARC can identify nearby things using three ultrasonic sensors, one on each side. If there are any obstructions from the ultrasonic sensor, the following calculations will determine how far away the item is from the ARC.

$$Time = \frac{Distance}{Speed} \quad (1)$$

In this hypothetical situation, an item is placed 26 cm from the ARC, while the airspeed of sound is 341 m per second or 1.145 cm per second. So, to cover that distance, the sound wave will need 736 μ s. The sensor then divides the observed distance by two because a sound wave travels in both directions, covering twice the distance. Thus, the formula must be divided by two.

$$Distance = \frac{Time * Speed}{2} \quad (2)$$

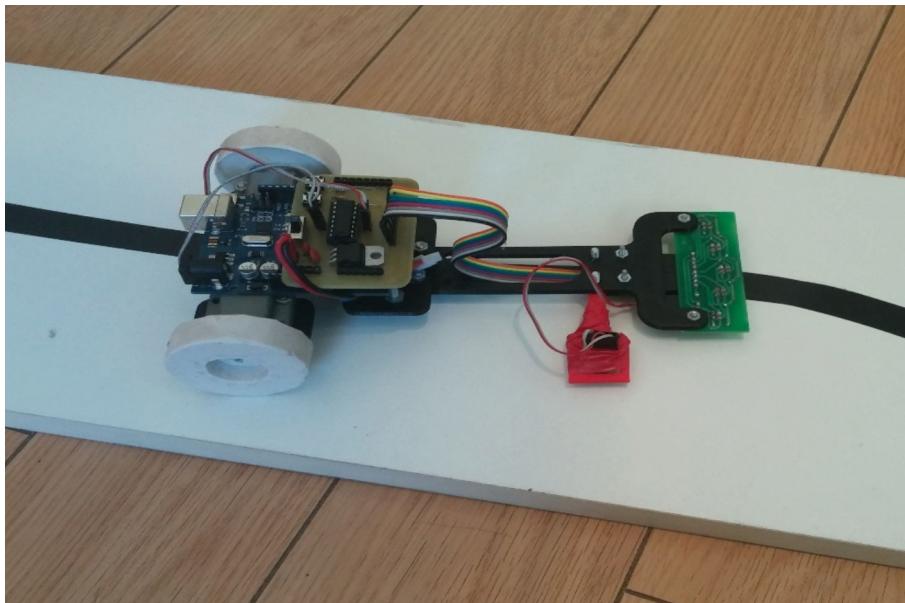


Fig. 5. ARC IR Sensors Detect Line When Traveling

Figure 6 shows the connection between real and calculated spaces, which shows how much the calculated distance deviates from the real distance. As a result, the measured distance is within permissible limits, and the deviation is not severe. Figure 7 shows the precision and error rate for ultrasonic sensors and the distance.



Fig. 6. Relationship between Actual and Calculated Distances



Fig. 7. Precision and Error Rate for Ultrasonic Sensors

At the accurate distance position of 26 cm, the system achieved a high degree of precision of 98.93%. Because of this, the findings show that the optimal detection distance for the ARC is between 26 and 51 cm. In addition, efficiency declines with increasing distance. Approximately 5.89% of the total errors occurred at the precise 200-cm distance. In contrast, the vision-based method collects real-time road data using a camera equipped with several vision algorithms, depending on both deep learning models.

The information obtained from these methods will assist the ARC in determining the best way to control movement. Immediate transmission of the first technique's data to the Arduino for ARC regulation and, in part, transmission to the Raspberry Pi system for further analysis will constitute the second approach. However, the data collected by the second method will all end up on the Raspberry Pi for analysis. To further improve mobility, the data is transferred to Google's Firebase for sharing. Through the system's Firebase, the ARC may communicate with other cars and share this data with them. Further, drivers and commuters with ROSNOS installed may get real-time road monitoring data that the ARC can provide. Consequently, the shared data used to address various road concerns might benefit drivers, commuters, and cars. Repeated testing has allowed us to analyze the proposed system, and confidently say that it will help reduce road-related concerns. The suggested system is one of several real-time systems that heavily depend on time. As discussed before, the system gathers data about the roads using cameras and sensors. Therefore, they will discuss time inference in three contexts: sensors, traffic signs, and traffic lights. A camera takes pictures of roadway signs and traffic lights, which are then sent to a Raspberry Pi to be processed in situations involving these elements. The modified LeNet-5 design has been suggested for detecting and recognizing traffic signals.

5 Conclusion

Due to the rise in vehicle traffic, managing traffic restrictions has grown increasingly complex, especially in densely populated metropolitan areas. This paper suggests an IoRT-based real-time traffic monitoring system to improve mobility and solve the difficulty. IoT nodes such as sensors and cameras capture real-time traffic data, which is analysed using image processing and deep learning algorithms. The data is wirelessly uploaded to the cloud and then made accessible to drivers and commuters via ROSNOS. Two suggested models use a modified LeNet-5 design for real-time recognition of traffic signs and transfer learning-based Inception-V3 for traffic light detection. The initial method is trained and evaluated using GTSRB and EGTSRB information sets, while the subsequent method uses the LISA dataset. Both models' accuracy tests show significant gains. Compared to the second model's 99.7 percent accuracy on the LISA dataset, the updated LeNet-5's performance on the GTSRB and EGTSRB datasets was 98.23% and 98.89%, respectively. Firebase, a real-time database, stores the processed data, and alerts are sent to mobile users. Further benefits include improved traffic management and reduced accident rates due to more accurate and trustworthy traffic monitoring. Internet of Things (IoT) technology in roadway monitoring is becoming crucial for improving mobility, traffic flow, accident prevention, and the travel experience. Therefore, the Internet of cars is critical for facilitating data exchange and real-time communication between cars. In addition, VCC uses the cloud to offer drivers a pay-as-you-go service, which helps decrease road concerns. The current traffic systems have limited options, but the proposed system could one day be used to intelligently control traffic by reducing accidents, saving time, and giving drivers real-time traffic information to choose the best lane or route.

References

1. Kheder, M.Q., Mohammed, A.A.: Real-time traffic monitoring system using IoT-aided robotics and deep learning techniques. *Kuwait J. Sci.* **51**(1), 100153 (2024)
2. Ponnusamy, M., Alagarsamy, A.: Traffic monitoring in smart cities using internet of things assisted robotics. *Mater. Today: Proc.* (2021)
3. Li, Y.: Constructing the intelligent expressway traffic monitoring system using the internet of things and inspection robot. *J. Supercomput.* **80**, 8742–8766 (2023)
4. Panda, S., Panda, G.: Intelligent classification of IoT traffic in healthcare using machine learning techniques. In: 2020 6th International Conference on Control, Automation and Robotics (ICCAR). IEEE (2020)
5. Banerjee, S., Chakraborty, C., Chatterjee, S.: A survey on IoT based traffic control and prediction mechanism. In: Balas, V.E., Solanki, V.K., Kumar, R., Khari, M. (eds.) *Internet of Things and Big Data Analytics for Smart Generation*, pp. 53–75. Springer International Publishing, Cham (2019). https://doi.org/10.1007/978-3-030-04203-5_4
6. Tran, M.-Q., et al.: Reliable deep learning and IoT-based monitoring system for secure computer numerical control machines against cyber-attacks with experimental verification. *IEEE Access* **10**, 23186–23197 (2022)
7. Liu, Y., et al.: Anomaly detection based on machine learning in IoT-based vertical plant wall for indoor climate control. *Build. Env.* **183**, 107212 (2020)

8. Kaur, J., et al.: Implementation of IoT in various domains. In: Sindhwan, N., Anand, R., Niranjanamurthy, M., Verma, D.C., Valentina, E.B. (eds.) *IoT Based Smart Applications*, pp. 165–178. Springer International Publishing, Cham (2023). https://doi.org/10.1007/978-3-031-04524-0_10
9. Iyer, S., et al.: Structural health monitoring of railway tracks using IoT-based multi-robot system. *Neural Comput. Applic.* **33**(11), 5897–5915 (2021)
10. Karmore, S., et al.: IoT-based humanoid software for identification and diagnosis of COVID-19 suspects. *IEEE Sensors J.* **22**(18), 17490–17496 (2022)
11. Jeena Jacob, I., Ebby Darney, P.: Design of deep learning algorithm for IoT application by image based recognition. *J. ISMAC* **3**(3), 276–290 (2021). <https://doi.org/10.36548/jismac.2021.3.008>
12. Sacco, A., et al.: An architecture for adaptive task planning in support of IoT-based machine learning applications for disaster scenarios. *Comput. Commun.* **160**, 769–778 (2020)
13. Patro, P., Azhagumurugan, R., Sathya, R., Kumar, K., Kumar, T.R., Babu, M.V.S.: A hybrid approach estimates the real-time health state of a bearing by accelerated degradation tests, Machine learning. InL 2021 Second International Conference on Smart Technologies in Computing, Electrical and Electronics (ICSTCEE), pp. 1–9. Bengaluru, India (2021). <https://doi.org/10.1109/ICSTCEE54422.2021.9708591>
14. Uma Maheswari, B., et al.: Internet of things and machine learning-integrated smart robotics. In: Habib, M.K. (ed.) *Global Perspectives on Robotics and Autonomous Systems: Development and Applications*, pp. 240–258. IGI Global (2023). <https://doi.org/10.4018/978-1-6684-7791-5.ch010>
15. Bahşı, H., Nömm, S., La Torre, F.B.: Dimensionality reduction for machine learning based iot botnet detection. In: 2018 15th International Conference on Control, Automation, Robotics and Vision (ICARCV). IEEE (2018)
16. Bharadwaj, H.K., et al.: A review on the role of machine learning in enabling IoT based healthcare applications. *IEEE Access* **9**, 38859–38890 (2021)
17. Uganya, G., et al.: A novel strategy for waste prediction using machine learning algorithm with IoT based intelligent waste management system. *Wireless Commun. Mobile Comput.* **2022**, 1–15 (2022). <https://doi.org/10.1155/2022/2063372>
18. Mewada, S., et al.: Smart diagnostic expert system for defect in forging process by using machine learning process. *J. Nanomater.* (2022)
19. Uddin, M.I., et al.: Ai traffic control system based on deepstream and iot using nvidia Jetson nano. In: 2021 2nd International Conference on Robotics, Electrical and Signal Processing Techniques (ICREST). IEEE (2021)
20. Sriram, S., et al.: Network flow based IoT botnet attack detection using deep learning. In: IEEE INFOCOM 2020-IEEE conference on computer communications workshops (INFOCOM WKSHPS). IEEE (2020)
21. Bhanu, K.N., Jasmine, H.J., Mahadevaswamy, H.S.: Machine learning implementation in IoT based intelligent system for agriculture. In: 2020 International Conference for Emerging Technology (INCET). IEEE 2020
22. SenthamilSelvan, R., Wahidabanan, R.S.D., Karthik, B.: Intersection collision avoidance in dedicated short-range communication using vehicle ad hoc network. *Concurrency Comput.: Pract. Experience* **34**(13), e5856 (2022)
23. Moshayedi, A.J., et al.: A secure traffic police remote sensing approach via a deep learning-based low-altitude vehicle speed detector through UAVS in smart cities: algorithm, implementation and evaluation. *Future Transport.* **3**(1), 189–209 (2023)
24. Ghazal, T.M., et al.: IoT for smart cities: Machine learning approaches in smart healthcare—A review. *Future Internet* **13**(8), 218 (2021)
25. Sharma, V., Tripathi, A.K.: A systematic review of meta-heuristic algorithms in IoT based application. *Array* **14**, 100164 (2022)

26. Tuli, S., et al.: HealthFog: an ensemble deep learning based smart healthcare system for automatic diagnosis of heart diseases in integrated IoT and fog computing environments. *Future Gener. Comput. Syst.* **104**, 187–200 (2020)
27. Andronie, M., et al.: Big data management algorithms, deep learning-based object detection technologies, and geospatial simulation and sensor fusion tools in the internet of robotic things. *ISPRS Int. J. Geo-Inform.* **12**(2), 35 (2023)
28. Zhang, M., et al.: Machine learning techniques based on security management in smart cities using robots. *Work* **68**(3), 891–902 (2021). <https://doi.org/10.3233/WOR-203423>



Development of Lightweight and Cheaper 5G Mobile Communication System to Analyze the Performance of Espresso Ciphers and Grain Family

C. Kamalanathan¹(✉), J. Balamurugan², Neelam Sharma³, A. Basi Reddy⁴, and R. Senthamil Selvan⁵

¹ Department of Electrical, Electronics and Communication Engineering, GITAM Deemed to Be University, Bengaluru Campus, Bengaluru, Karnataka, India
kchandrap@gitam.edu

² Department of Master of Business Administration, St.Joseph's College of Engineering, OMR, Chennai, Tamilnadu, India

³ Department of Physical Education, Lovely Professional University, Punjab, India
neelam.sharma@lpu.co.in

⁴ Department of Computer Science and Engineering, School of Computing, Mohan Babu University, Tirupati, Andhra Pradesh, India

⁵ Department of Electronics and Communication Engineering, Annamacharya Institute of Technology and Sciences, Tirupati, Andhra Pradesh, India

Abstract. A significant difficulty in a resource-constrained setting is keeping an optimal sense of balance among safety and other presentation parameters such as memorial need, energy requirement and throughput. In its most recent study on lightweight cryptography, the National Institute of Standards and Technology (NIST) recommended symmetric ciphers for devices with little resources. This work presents statistical security studies of six modern stream ciphers: Espresso, Grain v1, Sprout, Fruit, Plantlet, and Lizard. The analyses were conducted using the following tests: avalanche, randomness, structure, and autocorrelation. After porting the improved code to a microcontroller in low-cost, that is the ATmega 328P, they have also conducted a thorough performance study of these ciphers. A device's suitability for use in an IoT-based network is a deciding factor in its selection. According to statistical security, performance measurements, and comparative analysis, the chosen ciphers are good choices for situations with limited resources.

Keywords: Lightweight Cryptography · NIST · Ciphers · Espresso · Grain v1

1 Introduction

There has been a recent uptick in studies on the formidable problem of offering security services in settings with limited resources [1]. Promising applications requiring secured communication among heterogeneous distributed devices with limited resources, such

as innovative city applications, health surveillance mechanisms, disaster management systems, etc., have contributed to increasing research interests in this area [2]. The main obstacle to successfully implementing these apps on a larger scale is resolving security challenges in environments with limited resources [3]. Regarding memory, energy, and other metrics, the cryptographic primitives used to safeguard these networks should be as light as possible [4]. As a bonus, these primitives should be hardware-implementation-friendly. Many academics have focused on one particular security issue: keeping computational costs down while providing adequate protection in environments with limited resources [5]. Due to their high computational, memory, and energy requirements, asymmetric key cryptography algorithms are inappropriate for the limited environment [6, 7]. Due to its ability to provide high-level security in constrained situations such as the Internet of Things (IoT), lightweight symmetric cryptography has recently garnered much interest [8–10].

The United States National Institute of Standards and Technology (NIST) launched a “lightweight cryptography” initiative after hearing that symmetric ciphers may provide high-grade security on budget equipment [11–13]. Observe that two kinds of lightweight symmetric ciphers—stream cipher and block cipher—aim to fix the issue of data secrecy when sent across an unfriendly environment. Among the several contenders for the ECRYPT eSTREAM project, Grain stream cipher emerged victorious after a series of tests [14]. Some low-tech block ciphers, such as Hight, Present, Katan, Midori, Gift, and Klein, appeared after the eSTREAM competition and were able to compete with Grain cipher [15–17]. Researchers have been developing several lightweight block ciphers to retain the trade-off between information security and performance metrics, intending to improve performance [18]. Also, the block ciphers supposedly work for passively low-end devices as they need a chip area of around 1000 gate equivalents (GE). Lightweight block ciphers have so recently surpassed stream ciphers in terms of security for lightweight applications [19, 20].

The author introduced the lightweight stream encryption ‘Sprout’ by minimizing internal state size. Chip area may be reduced by about 800 GE due to smaller internal states [21]. The introduction of Sprout significantly enhances the acceptance of stream ciphers in lightweight applications. Inspired by the Sprout cipher, other lightweight LFSR-based stream ciphers, including Fruit, Lizard, and Plantlet, were developed [22]. Recently, Espresso cipher was introduced for 5G mobile communication systems. While several cryptanalytic high end assaults on the Sprout procedure are documented, the security of lightweight ciphers should be assessed using statistical tests such as randomness, autocorrelation, and the avalanche effect [23]. The discussion inspired us to concentrate on lightweight stream ciphers. This paper’s primary contributions are:

- Statistical tests were conducted on chosen ciphers of stream (Fruit, Lizard, Plantlet, Grain, Sprout and Espresso) to determine their appropriateness for cryptography.
- Stream cipher codes are tailored for limited environments. Additionally, assessed the act of chosen ciphers in a resource constrained ATmega328P microcontroller and low-cost.

Several current light weight ciphers of block, including Kasumi, Simeck, Present, Klein, Spec, and Hight, have also had their performances evaluated alongside those of the chosen stream ciphers. They have adhered to the project’s most current research

on block ciphers. Results from experiments and comparisons show that some stream ciphers work well on low-powered computers [24–27] (Fig. 1).

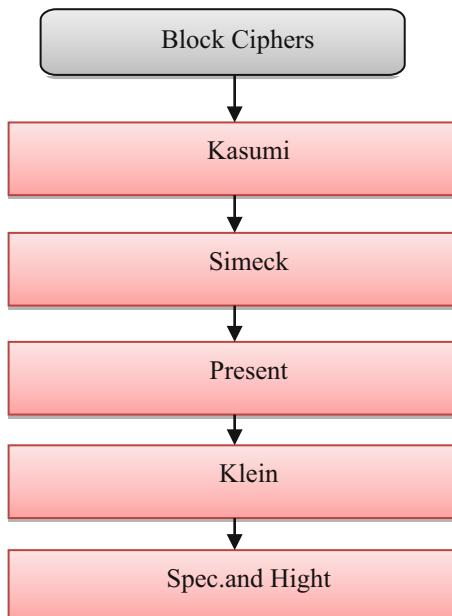


Fig. 1. Several Current Lightweight Block Ciphers

2 Probability Test Outcomes

Lightweight stream ciphers' pseudorandom generators are built using non-linear feedback shift registers (NFSR) and an LFSR, which includes a non-linear filtered function. With minimal computational cost, most modern stream ciphers promise to provide high-quality pseudorandom key streams. Because it finds the vulnerabilities of an outcome sequence quickly, the statistical test is a vital aspect of the cryptosystem design.

The following approach was used to conduct the experiments.

The randomness test was measured using the NIST statistical examination suite and AIS 20/31. SageMath 7.0, IBM SPSS, Arduino Studio, Mathematica 10, and GCC-4.8 were used in all experiments. The experiment system setup considered similar to this: Intel Core i7-6700 CPU, 3408 MHz, 3.40 GHz, 4 Cores, 8 Logical Processors.

Sprout, Espresso, Fruit, Plantlet, Lizard, and Grain v1 were the stream ciphers that were statistically tested. The findings of the test are summarized in the following subsections. They do the statistical tests using the following methods: avalanche, autocorrelation, and structural, key stream, and NIST randomness.

2.1 Analysis of Randomness

It is a challenging task to produce high-quality random numbers deterministically. When it comes to creating pseudorandom numbers, stream ciphers are perfect. This section discusses the NIST test suite's randomness test for a subset of lightweight stream cipher keystreams.³ They used GCC to develop these stream ciphers and produced 107 key stream bits for the randomness test. Remember that the Fruit cipher's creator suggested multiple variants. Fruit v2 is the name of the latest edition. The article will focus on the original Fruit edition.

In this experiment, the sequence is evaluated for suitability by calculating the test for randomness result using the probability value (P-value). This test assumes that the key stream information is random once the P-value is below 0.01. Ten distinct statistical tests make up the NIST test suite. This article provides a concise overview of our investigation into the outcomes of randomness tests conducted on 107-bit pseudo-random sequences produced by six different stream ciphers.

A new evaluation approach for random number generators, AIS 20/31, was established by the German Federal Office for Information Security (BSI). Evaluators of generators may benefit from it, and designers can use it to check for security flaws in their work. The eight statistical tests T0–T7 comprise this approach of randomness; they include, among others, the poker assessment, the long-term assessment, and the uniformity of distribution assessment. Additionally, only five statistical methods are used to verify randomization out of eight tests. Notably, no test has yielded a fail or reject assessment result. The previously indicated technique has become the gold standard for random number generator creation and assessment in Europe. The aforementioned lightweight stream ciphers have been found to pass the randomness tests according to the results of the AIS 20/31 and NIST test suite.

2.2 Analysis of Structural Tests

They described four distinct kinds of structural assessments: the key or stream key correlation test, the initialization vector (IV)/the frame correlation assessment, key stream correlation assessment, and the diffusion assessment. A structural test was also presented. Here, they check if the key stream is random by applying structural randomness tests methodically. For the structural test, they utilize the produced pseudorandom binary key streams of the chosen ciphers. They have adhered to the construction procedures to conduct the structural test. Please take notice of the structural test results. The chi-square goodness-of-fit assessment is used to evaluate the specified test. Two sets of autocorrelation indices for the 64-bit key stream arrangement of the chosen ciphers such as plantlet, lizard, fruit, sprout, and espresso and grain v1 are shown visually in Fig. 2.

For randomly produced key streams, stream ciphers are found to pass autocorrelation tests. Nevertheless, it demonstrates that partial knowledge of the inner workings may be used for key recovery in the Sprout cipher.

2.3 Analysis of Autocorrelation

Stream ciphers that rely on LFSRs are the ones that benefit the most from autocorrelation analysis. The stream cipher's key stream sequence is expected to be visually identical

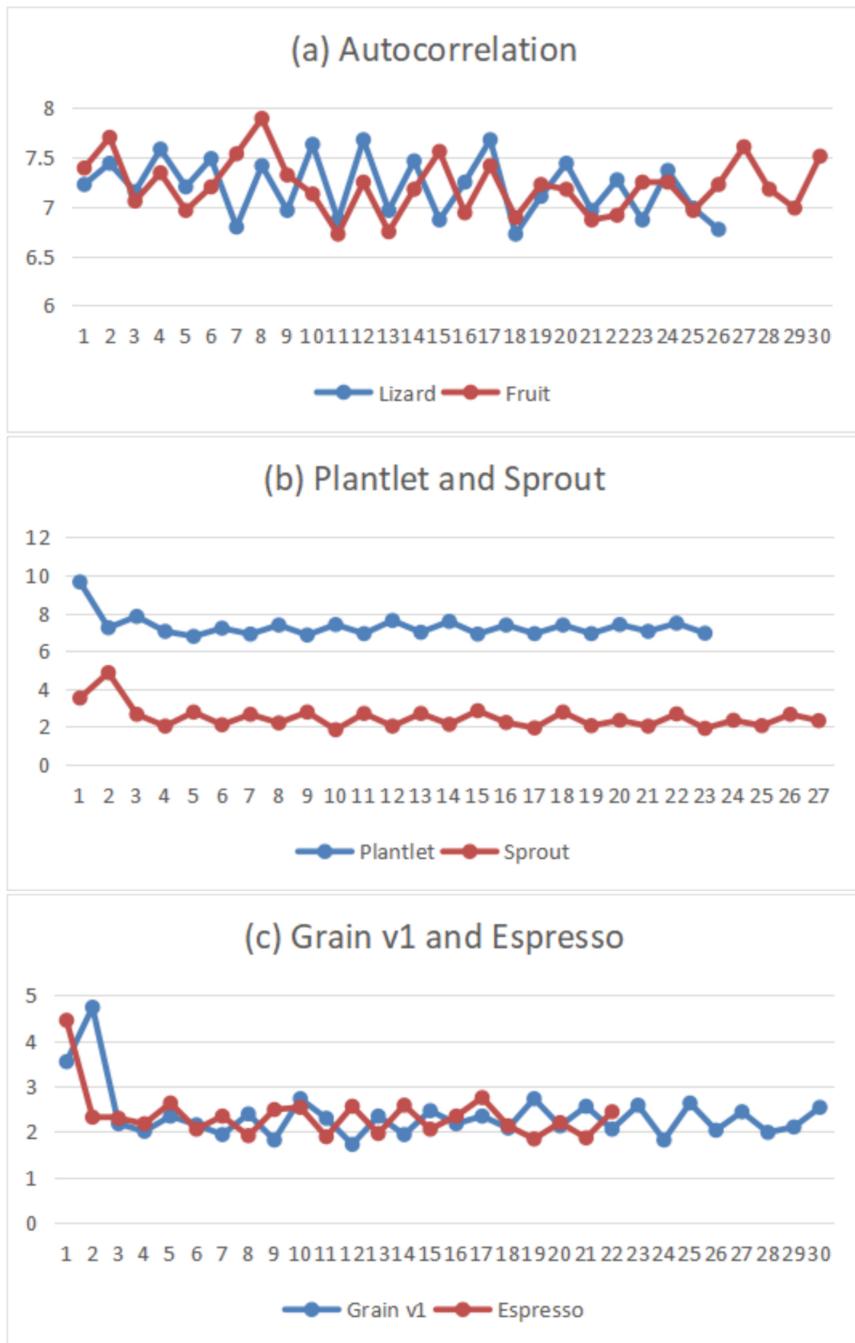


Fig. 2. (a), (b) and (c) Autocorrelation Indices

to random sequences. A key part of secure applications is binary strings with the right autocorrelation properties. The adversary attempts to ascertain a probable connection among the key bits and the amount produced within this framework. Consider the infinite sequence $S^N = (S_0, S_{N-1})$. Next, define the autocorrelation relationship of a series S^N as follows:

$$AC(\tau) = \frac{1}{N} \sum_{i=1}^{N-1} (-1)^{(S_0 + S_{i+\tau})}, 0 \leq \tau \leq S - 1 \quad (1)$$

The autocorrelation ACs quantifies the degree to which S^N and the change of S^N by “S” are comparable. Autocorrelation guarantees that an attacker cannot detect the correlations between shifted copies of the same bit stream, which is essential from a cryptographic standpoint.

2.4 Analysis of Key Stream Data

The statistical features of the key stream generated by the ciphers above are the main emphasis of this section. The encryption produces a series of 40 hex bits, such as d0 ; 1 ; 2 ; e ; $f\text{D}$. Let us examine these sequences. The first step is to compute the frequency distribution using the output data. Afterwards, the statistical characteristics of the key stream data were evaluated using metrics like the coincidences index of values, normalized Shannon entropy (NSE), variance, standard deviation, and standard error of the mean (SEM \pm). According to the data, the key stream sequence is robust against statistical assaults and follows a normal distribution.

2.5 Test for Avalanche

The avalanche effect’s key value is its ability to assess resistance to brute force and other real-time assaults. For cryptographic structures to work, it’s ideal if even little modifications to the input result in noticeable shifts in the output. Following is an example of the avalanche effect as it pertains to a map function from the n-bit to the m-bit binary sequence:

$$\sum_{0 \leq j \leq 2^{n-1}} D_j = m \times (2^{n-1}) \quad (2)$$

3 Analysis of Performance

Many new areas of networking, such as the Internet of Things (IoT), devote considerable attention to the limited space available. As previously mentioned, the IoT gadgets are designed to function for an extended period and have limited resources. One inexpensive open-source microcontroller is Arduino. A lot of low-level programming frameworks utilize it. Even more so, ZigBee, WiFi, Bluetooth, Z-Wave, ANT, and other limited networks use Arduino. Using the Arduino ATmega 328P microcontroller’s limited resources, we

evaluated the chosen ciphers' performance. They used the programmability of the low-cost microcontroller, CPU of Arduino's single core CPU to manage the parallel unit. It manages the encryption of both the input and the output in an asynchronous manner, allowing for the execution of many tasks simultaneously. Additionally, this gadget has two possible connections: a USB connector for connecting to a computer and an AC to DC converter and battery. Results and discussion about the stream cipher's performance are presented in the following subsection. In addition, Table 1 describes the performance data in detail.

Table 1. Performance Data

Metrics	Descriptions
Size of code	Storage space for cipher code and constants
Size of RAM	It usually stays in flash memory The memory size for storing intermediate stages during encryption code execution
Throughput	No. of encrypted second (kbps)/ bits
Velocity	No. of encrypt/decrypt sets per block per byte

3.1 Results

This research includes stream ciphers (Fruit, Plantlet, Lizard, Sprout, Grain v1, RC4, Espresso, Mickey and Trivium,) and block ciphers (Klein, Present, Hight, Kasumi, Speck and Simeck).

Using and deploying cryptographic algorithms is one of the primary challenges to attaining performance with Arduino. Compared to computer-based systems, Arduino's 32 KB of flash memory, 2 KB of static random-access memory (SRAM), and 1 KB of electrically erased programmable read-only memory (EEPROM) could be much better. Under these circumstances, a 128-bit key AES encryption requires ten processing cycles when implemented on an Arduino board. Each processing round includes the following steps: a row-wise permutations stage, a column wise mixing stage, a round key adding, and a single-byte replacement step. The encryption technique requires a certain amount of memory to store both temporary and final encrypted results on flash memory. A secret key for the cypher is kept in the EEPROM. In the event of trying to move a big program code onto devices with limited resources, the situation becomes worse.

Before porting to ATmega328P, they optimized the C codes of the chosen stream ciphers. To facilitate specific input-output tasks, the C code for Ciphers has been modified to reduce size, memory consumption, and execution time. At first, they focused on finding the most time-consuming portion of the code. To make stream cipher codes work better in this limited setting, they have briefly discussed the optimization techniques in the subsequent passages.

As previously mentioned, the time a microcontroller uses for () loop is often high. They have previously used the term volatile immediately before the integer type declaration int to preserve variables across loops in Arduino. Using this method, shrink the program's memory footprint by 145 bytes. The second step was to reduce memory use by achieving quicker input-output control using port manipulation. The three registers used for measuring the states of various input to output ports are B, C, and D.

It takes less processing time to generate output from any low-end device. Our analysis focuses on the computing requirements for generating 256-byte output data using stream ciphers and certain blocks. Fruit, Lizard, Sprout, Plantlet, Espresso, Grain v1, Trivium, RC4, Present, Klein, Mickey, Hight, Kasumi, Speck and Simeck, are some of the stream and block ciphers that are shown in Fig. 3a, which shows the results of the calculation time. Conventional blocks and stream ciphers take more time to compute than lightweight ciphers, as seen in Fig. 3a. The calculation time required by Lizard, Sprout, and Plantlet is much lower than that of other ciphers, as seen in Fig. 3a.

Energy needs are among the most essential criteria for the Internet of Things sector. Here, they have determined the overall energy needed to generate the key stream for the chosen cipher (as specified in the specification); the cipher operated at 5 V with a resistance level of 150 X. The outcome of the simulation may be seen in Fig. 3b. It shows that when the size of the key stream bit rises, the energy demand also increases. As shown visually in Fig. 3c, calculated the size of the program (in kilobytes) of the chosen ciphers of stream that are portable in the microcontroller. Regarding processing time, program size, and energy required, Grain v1 demands approximately the most resources, while Lizard requires the least.

Low computing capacity in lesser end devices has implemented decryption and encryption techniques that are more complicated in resource-constrained environments. Therefore, the throughput statistic is critical for the scheme's success. The act of the chosen cipher stream has been confirmed by creating a sizable key stream. Concerning encryption and decryption, throughput is defined as the average quantity of data (in bytes) processed per unit of time and is represented thus:

$$\text{Throughput} = \frac{\text{No. of bytes Encrypted/Decrypted}}{\text{Time}} \quad (3)$$

Streaming and block cipher throughput have been calculated using the previously stated equation. Figure 4 shows the throughput values of the chosen ciphers relative to the file size. Throughput has been shown in kbps in this work. Figure 5 shows the end outcome of Kasumi, Klein, Simeck, Hight, Present, and throughput Speck's. They do several tests using the Arduino studio platform to test how well the chosen and block and ciphers stream handle the throughput. It is clear from the throughput charts that the chosen stream cipher outperforms the block cipher with increasing file sizes. The Lizard cipher stood out, consistently outperforming them at every interval.

The numerical study of the speed of chosen stream ciphers has been completed in this section. In most cases, the efficacy of the ciphers is determined by presenting speed evaluations. Here is an equation that may be used to represent speed S:

$$\text{Speed } S = \sum s_i \pi_i \quad (4)$$

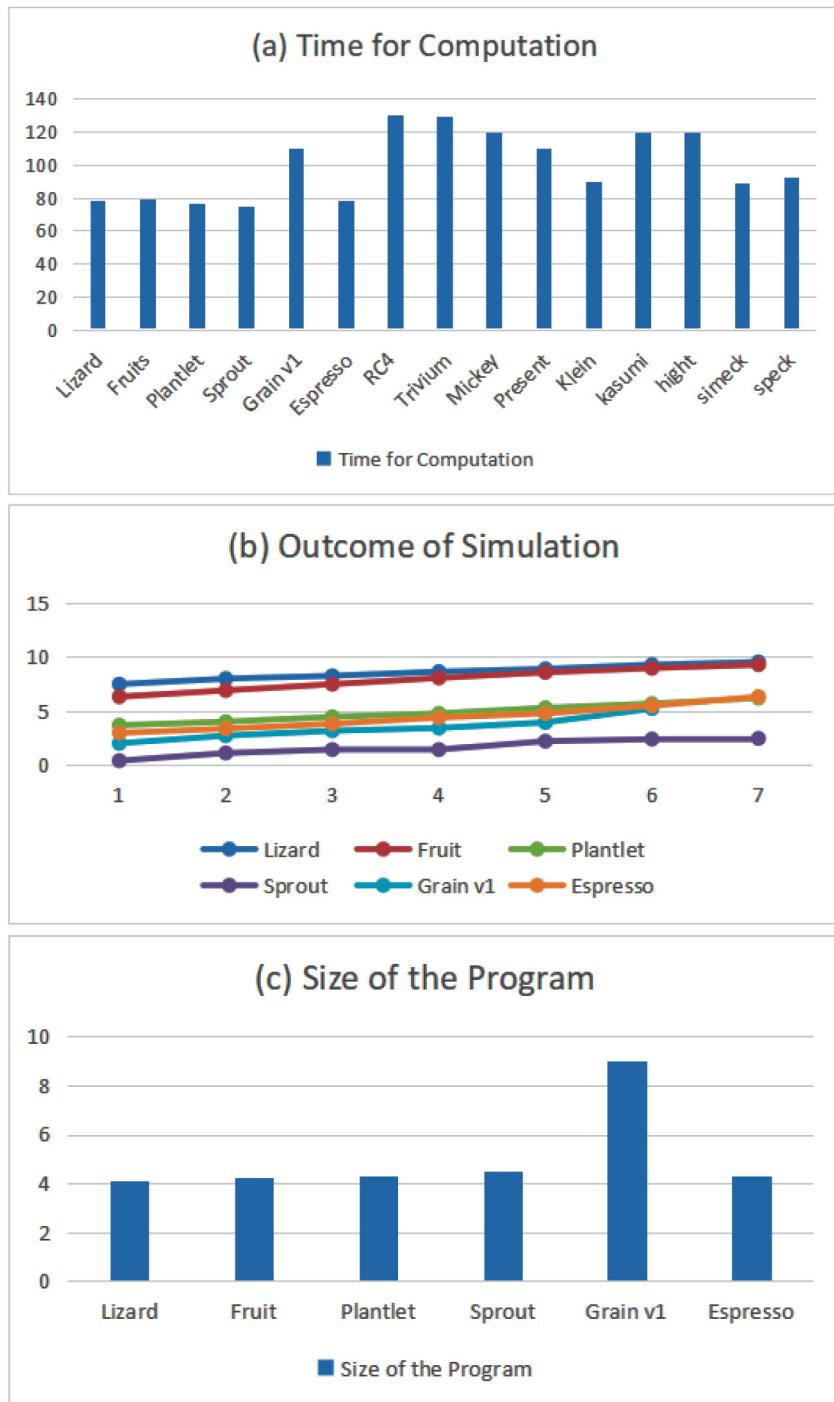
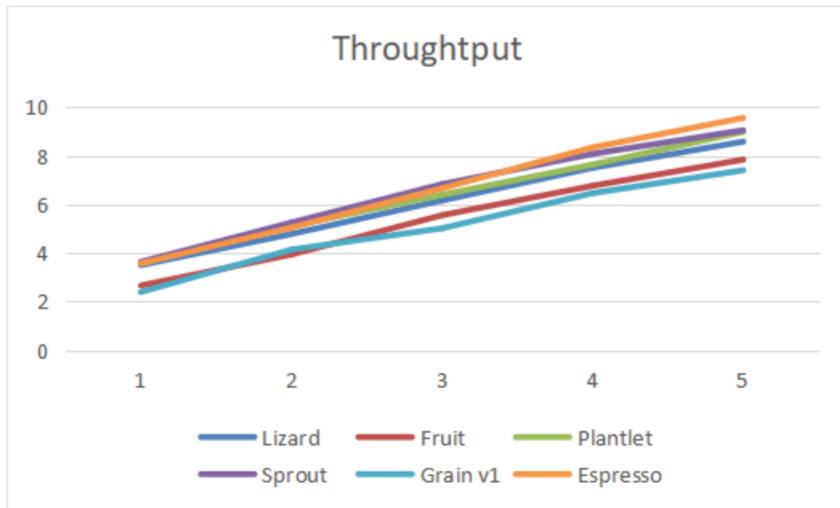
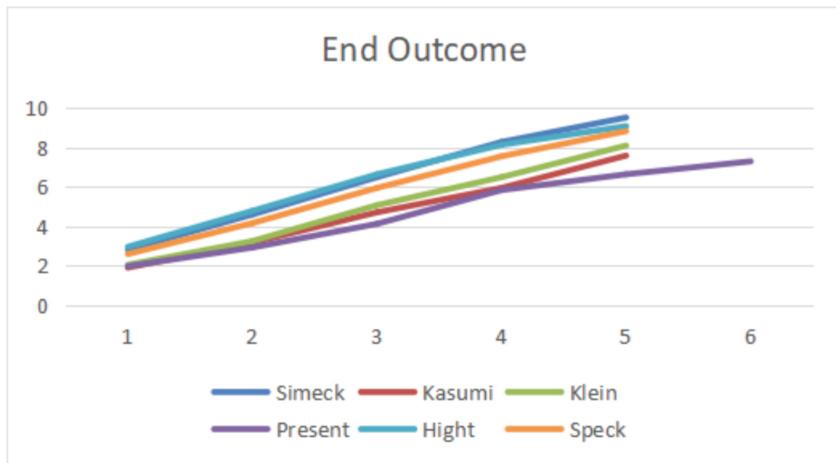


Fig. 3. (a) Block and Stream Ciphers Computational Time. (b) Outcome of the Simulation. (c) Program Size

**Fig. 4.** Throughput Values**Fig. 5.** End Outcome values for Throughput

S_i Is the cipher velocity for i bytes of plaintext, and π_i is the fraction of i byte to encrypt.

A CPU with a lot of processing power will make the cipher run faster. In this case, the encrypted packet rate per unit of time has been used to determine the speed. In this scenario, three hundred fifty packets, or 40 bytes, will be used for encryption. The chosen stream ciphers' speed performance is shown in Fig. 6a in bytes/ls and packets/ls. See Fig. 6b for an illustration of the results of a speed performance comparison with state-of-the-art block ciphers. The results obtained by our chosen stream ciphers are comparable to those of the lightweight block ciphers that were previously stated. The

speed assessment graphs for the stream and block ciphers show that Present, Hight, and Grain v1 have somewhat slower speeds, whereas Lizard consistently performs best throughout all trials.

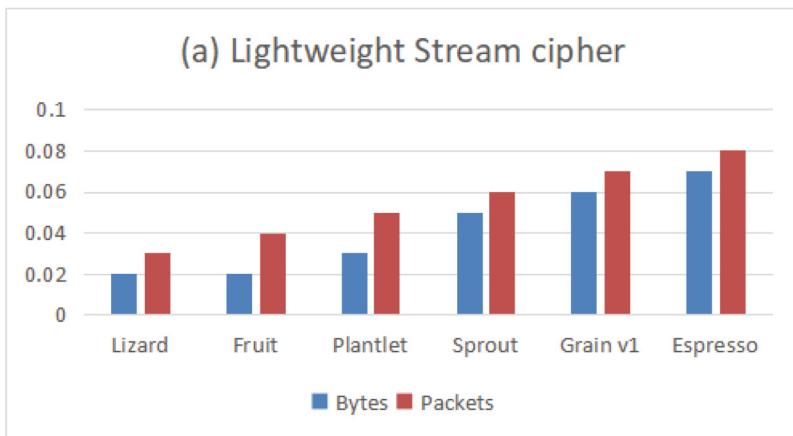


Fig.6. (a) Ciphers' of Stream Speed Performance

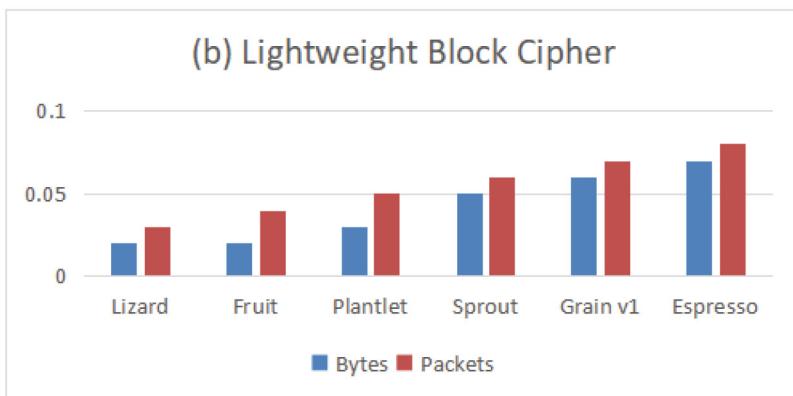


Fig.6. (b) Block Ciphers' Speed Performance

Fig. 6. (a) Ciphers' of Stream Speed Performance. (b) Block Ciphers' Speed Performance

3.2 Discussion

At first, symmetric cryptography was the main emphasis of the NIST lightweight cryptography study. Lightweight stream ciphers are also the subject of this research. Researchers have struggled to maintain the trade-off between the varied resource restrictions (e.g., power, prices, etc.) and the desired security level when securing devices with varying degrees of resource constraint.

As mentioned earlier, the sprout cipher offers a fresh perspective on constructing a stream cipher with a reduced internal state. Numerous high-level attacks against the Sprout stream cipher have been discovered in the literature. According to NIST, lightweight encryption may not account for high-end attacks because of the resources they consume. For lightweight cryptographic systems, 80-bit safety is often measured the bare least possible. They found that Sprout passed a lot of the arithmetical examinations used in the investigation. Lightweight cryptography cannot use Sprout because, despite its statistically solid features, there are attacks that can retrieve the secret key in less than 250 tries.

A side-channel attack, which uses partially exposed information via encryption, may crack stream ciphers. The majority of stream cipher inventors, however, have asserted that their ciphers can withstand side-channel assaults because they use shorter internal states.

The following finding is derived via the application of several block and stream ciphers through various approaches.

3.2.1 First Finding

It takes a lot of work to accurately and comprehensively analyse the nature of the chosen stream cipher. Although they use complicated random number generation procedures, tiny state stream ciphers have a relatively more minor internal state.

The chosen ciphers produce more reliable results for randomness. Also, compared to well-known DES algorithms, the chosen ciphers provide results comparable to those of the latter (proving the assertions via implementation). On top of that, the tiny computers may run an underlying challenge-response authentication protocol using carefully chosen lightweight stream ciphers. Complete testing has shown that the ciphers of Lizard design choices are superior to other lesser stream state cyphers. Lightweight applications do not need Lizard's $(2n/3)$ -bound security versus TMDTO critical recovery assaults; 80 bits of protection is more than enough.

Selecting cryptographic primitives in a resource-constrained setting may benefit from the general discussion.

4 Conclusion

In response to the proliferation of lightweight applications, academia and industry have developed lightweight cryptographic methods. A lightweight cryptographic method must find a happy medium between security and efficiency. An efficient security solution for devices with limited resources is shown in this study by the use of the stream cipher. They found that the chosen lightweight stream ciphers—Espresso, Grain v1, Sprout, Fruit, Plantlet, and Randomness—offer very high levels of statistical security. Performance parameters such as calculation time, energy consumption, memory demand, and throughput are used to evaluate the chosen ciphers, which are then programmed into a low-cost Arduino ATmega328P microcontroller. According to the comparison, the chosen ciphers are more suitable for devices with limited properties.

References

1. Deb, S., Bhuyan, B.: Performance analysis of current lightweight stream ciphers for constrained environments. *Sādhanā* **45**, 1–12 (2020)
2. Bîrleanu, F.G., Bizon, N.: Lightweight cryptography for Internet of Things using FPGA-based Design with Partial Reconfiguration. In: 2020 12th International Conference on Electronics, Computers and Artificial Intelligence (ECAI). IEEE (2020)
3. Dhanda, S.S., Singh, B., Jindal, P.: Lightweight cryptography: a solution to secure IoT. *Wireless Pers. Commun.* **112**(3), 1947–1980 (2020)
4. Philip, M.A.: A survey on lightweight ciphers for IoT devices. In: 2017 International Conference on Technological Advancements in Power and Energy (TAP Energy). IEEE (2017)
5. Fatima, R., Manal, R., Tomader, M.: Cryptography in e-Health using 5G based IOT: a comparison study. In: Proceedings of the 4th International Conference on Big Data and Internet of Things. (2019)
6. Philip, M.A., Vaithianathan, V., Jain, K.: Implementation analysis of rectangle cipher and its variant. In: 2018 3rd IEEE International Conference on Recent Trends in Electronics, Information and Communication Technology (RTEICT). IEEE (2018)
7. Majhi, S., Mitra, P.: Lightweight cryptographic techniques in 5g software-defined internet of things networking. In: *Lightweight Cryptographic Techniques and Cybersecurity Approaches*. IntechOpen (2022)
8. Lata, N., Kumar, R.: Analysis of lightweight cryptography algorithms for IoT communication. In: Sharma, H., Saraswat, M., Yadav, A., Kim, J.H., Bansal, J.C. (eds.) *Congress on Intelligent Systems: Proceedings of CIS 2020*, vol. 2, pp. 397–406. Springer Singapore, Singapore (2021). https://doi.org/10.1007/978-981-33-6984-9_32
9. Majhi, S., Mitra, P.: *Lightweight Cryptographic Techniques in 5G Software-Defined Internet of Things Networking*.
10. Shi, Z., et al.: Design space exploration of galois and fibonacci configuration based on espresso stream cipher. *ACM Trans. Reconfigurable Technol. Syst.* **16**(3), 1–24 (2023)
11. Jebrane, J., Lazaar, S.: A performance comparison of lightweight cryptographic algorithms suitable for IoT transmissions. *Gen. Lett. Math.* **10**(2), 46–53 (2021)
12. Kumar, P.K., Mondal, B.: Lightweight stream cipher for health care IoT. In: 2023 IEEE 2nd International Conference on Industrial Electronics: Developments & Applications (ICIDEA). IEEE (2023)
13. Biryukov, A., Perrin, L.: State of the art in lightweight symmetric cryptography. “*Cryptology ePrint Archive*” (2017)
14. Sfar, A.R., et al.: A roadmap for security challenges in the Internet of Things. *Digit. Commun. Netw.* **4**(2), 118–137 (2018)
15. Navarro, M., Luis, J., Sabater, A.F.: Review of binomial decomposition-based algorithms for efficient linear complexity computation. *Mathematics* **9**(5), 478 (2021)
16. Martin-Navarro, J.L., Fúster-Sabater, A.: Review of the lineal complexity calculation through binomial decomposition-based algorithms. *Mathematics* **9**(5), 478 (2021)
17. Dubrova, E.: Energy-Efficient cryptographic primitives. *Facta universitatis-series: Electron. Energ.* **31**(2), 157–167 (2018)
18. Ebrahimabadi, M., Younis, M., Karimi, N.: A PUF-based modeling-attack resilient authentication protocol for IoT devices. *IEEE Internet Things J.* **9**(5), 3684–3703 (2021)
19. Rehmani, M.H.: *Blockchain Systems and Communication Networks: From Concepts to Implementation*. Springer (2021)
20. Hamann, M., et al.: The DRACO stream cipher: A power-efficient small-state stream cipher with full provable security against TMDTO attacks. *IACR Trans. Symm. Cryptol.* (2022), 1–42

21. Dictionary, Macquarie. Macquarie Concise Dictionary Seventh Edition. Macmillan (2017)
22. Jebrane, W., Bouasria, I., Akchioui, N.E.: Autonomous vehicles motion planning techniques: Recent advances and future trends. In: AIP Conference Proceedings, vol. 2814. No. 1. AIP Publishing (2023)
23. Sami, T.M.G., et al.: A state-of-the-art survey for IoT security and energy management based on hashing algorithms. *Int. J. Recent Innov. Trends Comput. Commun.* **11**(10), 257–267 (2023)
24. Thakor, V.A., Razzaque, M.A., M, R. A. Khandaker,: Lightweight cryptography algorithms for resource-constrained IoT devices: A review, comparison and research opportunities. *IEEE Access* **9**, 28177–28193 (2021). <https://doi.org/10.1109/ACCESS.2021.3052867>
25. Buchanan, W.J., Li, S., Asif, R.: Lightweight cryptography methods. *J. Cyber Secur. Technol.* **1**(3–4), 187–201 (2017)
26. Mewada, S., et al.: Smart diagnostic expert system for defect in forging process by using machine learning process. *J. Nanomater.* (2022)
27. Patro, P., Azhagumurugan, R., Sathya, R., Kumar, K., Kumar, T.R., Babu, M.V.S.: A hybrid approach estimates the real-time health state of a bearing by accelerated degradation tests, Machine learning. In: 2021 Second International Conference on Smart Technologies in Computing, Electrical and Electronics (ICSTCEE), pp. 1–9. Bengaluru, India (2021). <https://doi.org/10.1109/ICSTCEE54422.2021.9708591>



Development of Light Weight Authentication Protocol Based on Cryptography to Access the IoT Device

Sameer Yadav¹(✉), Surepalli Venkataratnam², P. Balaji Srikaanth³, Jetti Madhavi⁴, A. Basi Reddy⁵, and R. Senthamil Selvan⁶

¹ Department of Commerce and Business Administration, University of Allahabad, Prayagraj, Uttar Pradesh, India

samuraisamu12112@gmail.com

² Department of History, Nizam College, Osmania University, Hyderabad, India

³ Department of Networking and Communications, SRM Institute of Science and Technology, Kattankulathur Campus, Chennai, India

Balajis7@srmist.edu.in

⁴ Department of Mathematics, Malla Reddy Engineering College, Medcha, India

⁵ Department of Computer Science and Engineering, School of Computing, Mohan Babu University, Tirupati, Andhra Pradesh, India

⁶ Department of Electronics and Communication Engineering, Annamacharaya Institute of Technology and Sciences, Tirupati, Andhra Pradesh, India

Abstract. The Internet of Things (IoT) has swept the linked globe. Internet of Things devices are growing exponentially globally due to their great practicality. Securing such gadgets, however, has received surprisingly little attention. Network creators and administrators have no option but to install these devices with inadequate or non-existent security due to their many limitations. 2016, during the notorious Mirai botnet assault, saw similar devices being used to launch distributed denial-of-service attacks. For this reason, provide a simple authentication procedure for these devices. When developing our authentication system, we considered several factors, including portability, user and device registration, scalability, and more. The design was based on a three-tiered paradigm with devices at the cloud, fog, and edge levels. Additionally, suggested other post-quantum cryptography-based cypher suites for review and implementation. Additionally, provide a fail-safe approach to ensure deployed IoT devices may continue delivering services autonomously if an authenticating server fails. When ring learning with mistakes is used, our protocol performs quickly. Using a tool for automated assessment of Internet security protocols and submissions, we demonstrate that our authentication system is secure. Finally, for Internet of Things device authentication in a fog computing setting, provide a secure, hybrid, and speedy protocol.

Keywords: IOT · Internet security protocols · Scalability · Speedy Protocol · Network Designers

1 Introduction

Many mobile gadgets have been put into operation since the Internet of Things (IoT) was integrated into households, industries, logistics, etc. These gadgets could be constantly moving about and might need the means to authenticate themselves or let users check their legitimacy properly [1, 2]. Scalability should be fine for these devices, and making choices quickly is crucial for minimizing transmission and propagation delays [3–5]. Not only will popularly cypher suites of the past be incompatible with the soon-to-be-released quantum computers, but new algorithms must be tested to ensure they are secure against both quantum and classical computer techniques [6, 7]. The security methods should be lightweight to be readily handled by restricted IoT devices [8, 9]. To solve these problems, we provide a hybrid authentication scheme in our study [10]. Moving cloud computing closer to the network's periphery is known as fog computing [11]. As mentioned, fog computation offers several benefits over a centralized cloud server [12]. To alleviate the computational burden on limited Internet of Things devices, fog computing shifts time-dependent processing to the network's periphery, allowing quicker response, improved mobility, and increased scalability. Data collected by Internet of Things (IoT) devices may be further processed [13], stored [14], and analyzed [15] by connecting fog servers to a central cloud server. Data that is not time-sensitive is stored on the cloud server and is only made available when needed by the expedient or fog server [16, 17]. Many different cloud-fog-edge architectures have been suggested over the years, with three- and four-layer designs being the most popular. As you can see in the following part, our study uses a three-layered architecture. An emerging concern in cryptography is Post-Quantum Cryptography (PQC). The user's text is empty [18]. Recent progress in the manufacturing of quantum computers has shown that Shor's algorithm may effectively break several public key encryption methods that depend on integer factorization and the discrete logarithmic problem, when run on a quantum computer of significant size [19, 20]. Many groups and societies have proposed public-key cryptography-based algorithms to counter algorithms executed by quantum computers and other non-traditional electrical computers [21]. The lattice-based encryption family is one example. Following its 1996 proposal by Miklos Ajtai, lattice-based cryptanalysis has been subject to analysis by several scholars [22]. Ring-Learning with Errors (Ring-LWE) and NTRU are integral parts of this cryptography family [23]. Another kind of cryptography that Robert McEliece first presented in 1978 is code-based cryptography [24, 25]. Code-based encryption makes use of error-correcting codes to provide confidentiality. Devote our protocol to the study of lattice-based cryptography [26, 27].

2 Methodology

2.1 Initial Steps

Here, describe in depth the cryptography suites for our study. Before NASA and D-Wave showed off a fully operational quantum computer in 2015, everyone assumed quantum computers only existed in theory. In contrast to qubits, which may exist in a superposition of states, classical electrical computers employ bits, which can only take on two distinct values 0 or 1. This is where quantum computers diverge from

conventional computers [27, 28]. With k qubits, a quantum processor may exist in 2^k states at once. A big enough quantum processor executing Shor's algorithm can crack the code for integer factoring and elliptic curve separate logarithm-based encryption. As a result, several groups and individuals have put forth other cryptographic suites dubbed post-quantum cryptography (PQC) that may one day supersede RSA and ECC. In the age of quantum cryptography, several researchers have proposed various suites of algorithms, as seen in Table 1. Attacks based on quantum computers cannot crack lattice-based encryption. In terms of efficiency and speed, lattice-based cryptography is superior. In contrast to code-based encryption, there is no issue with big key sizes with lattice-based cryptography. A brief overview of the underlying concepts of these two branches of public-key cryptography follows.

Table 1. Quantum-Post Emergence Encryption

Sl no	Family	Procedure
1	Lattice-based	NTRU
		BLISS
		Ring LWE
2	hash- based	Rainbow
3	multivariate	Merkle Name
		Lampert Name
4	Code-based	McEliece
		Niedermeyer

2.2 Secure Communication Using Lattices

Suggested lattice-based encryption structures. However, lattice-related mathematics has a long study history, with notable figures like Carl Friedrich Gauss and Joseph Louis Lagrange. Solving a worst-case lattice issue is at minimum as complex as an average-case lattice problem, a Short Integer Solution (SIS). As an alternative, they suggested NTRU, a system for public key encoding. Here, finds the exact definition of a lattice structure. Arrays of Any y -dimensional Euclidean space \mathbb{R}^y has a discrete subgroup L . A collection of points in \mathbb{R}^y is the best way to represent L . A linear combination of basis vectors b_1, b_2 , with real integer coefficients uniquely represents every element of L .

$$l = \left\{ \sum_{i=1}^Y D_i b_i : d_i \in \mathbb{Z} \right\}$$

2.3 NTRU

Several researchers have subsequently investigated NTRU. Its accuracy is grounded on the gathering characteristics of the sums of random variable quantity and its security in the

effort of evaluating the output of polynomial calculation modulo two unrelated moduli. The authors launched a lattice reduction-based assault. Over the years that followed, several researchers successfully launched many assaults against NTRU. Bernstein put forth the idea of NTRU prime in 2016 with the hope that it could guard against known assaults on NTRU by its robust algebraic structure. What follows is a more in-depth explanation

$$r = [X]/(x^p - X - 1)$$

2.4 Learning Rings with Mistakes

The writers presented Ring-LWE, which is based on Regev's premise that learning across polynomial rings is possible via errors. As part of LWE, distinguish between random linear comparisons perturbed by a little amount of noise and uniform linear equations. Bring Ring-LWE's difficulty down to the Short Vector Problem's worst-case hardness. Scientific investigations have made use of Ring-LWE in a variety of settings, such as homomorphic encryption, public key encryption, digital signatures, etc. During the intra-fog authentication phase, users or devices do not need to re-register when moving from FS 1X to a neighbouring server FS 1Y inside the same dispersed fog. To prevent re-authentication whenever device N is transferred from one fog attendant FS 1X to another, an inter-fog authentication phase is necessary. The inter-fog verification cannot be completed without the cloud server. This is the most basic explanation of Ring-LWE. Here you may find a realistic implementation of Ring-LWE that we utilize in our research. Following these methods will resolve the Ring LWE problem.

3 Procedure Proposed

3.1 Prototypical System

As seen in Fig. 1, our network model comprises a fog architecture with three layers. The most crucial server in this arrangement is a cloud attendant located at the very top. The fog server mediates communication among the edge plans and the central cloud attendant, as shown in Fig. 1. Consequently, the cloud server is linked to the edge devices associated to the server fog. The dispersed server's fog is linked to this central server cloud to analyse, process, store, etc., data. Our model predicts that the cloud server will get updates on the fog network's status at regular intervals.

One must first examine and comprehend time-dependent facts to make a time-dependent choice. To alleviate some of the strain on the cloud server, the fog server may handle the authentication of these devices.

In addition to lowering latency, it would improve scalability, mobility, and heterogeneity. Therefore, the whole network would benefit from enabling authentication for the edge devices using fog servers. Consideration is given to the possibility that a fog user or device may transition between different fogs. Additionally, our model will authenticate the transferring device if this scenario occurs. Figure 2 depicts the whole situation. All the symbols and notations for our procedure may be found in Table 2. According

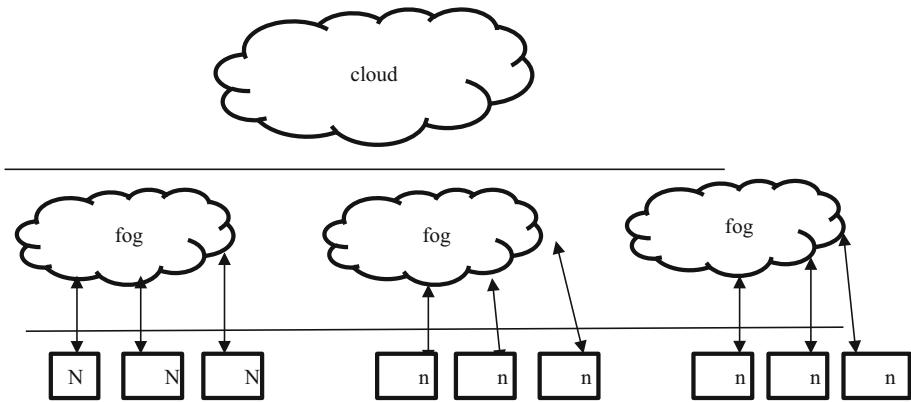


Fig. 1. Prototypical network

to our research, no existing protocol with a fail-safe mechanism specifies what to do if an authenticating fog server fails. Additionally, we need help locating studies that offer authentication for devices moving within or between fogs. Some studies use inter-fog communications, but that's purely for computational reasons. Figure 2 provides a bird's-eye perspective of our suggested authentication technique. The graphic shows that a new expedient N registers with fog attendant FS 1X, which subsequently portions the identifications with the server cloud, allowing for universal authentication.

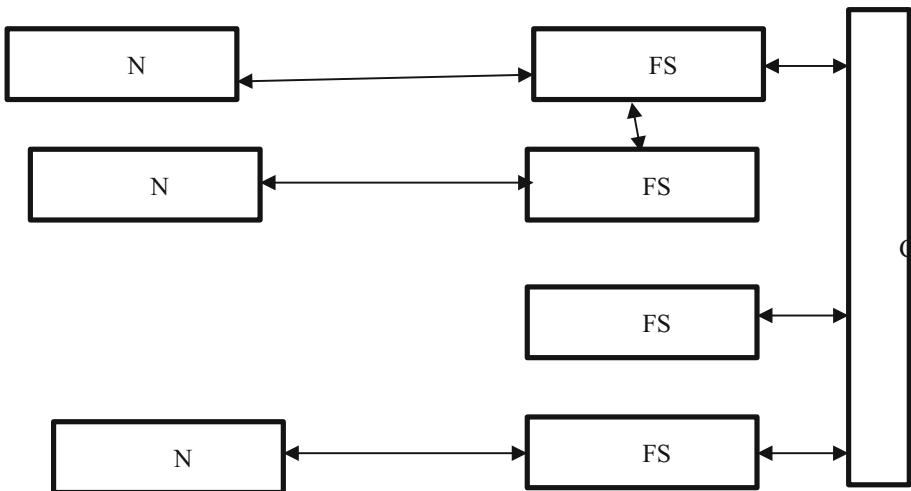


Fig. 2. Viewing our Protocol from Every Angle

Moving from FS 1X to a neighbouring server FS 1Y inside the same dispersed fog does not need re-registration during the intra-fog authentication phase for users or devices. There should be an inter-fog authentication phase to avoid re-authentication

when expedient N changes from the control of one server fog FS 1X to additional. For the inter-fog authentication to be completed, the cloud server is needed.

3.2 Registration of Nodes

An Internet of Things node (N) registers with a fog server (FS) at this stage. The process starts with Node N communicating with the fog-based authentication server via an encrypted message including T S, FS id, Nid, and noncook. Upon receiving a registration request, the authentication server in the fog verifies the node's credentials and records the nonce and time stamp. The fog-based verification server sends an encrypted acknowledgement message in the format after it confirms the node's credentials. The node decrypts the message and sends an ACK to confirm receipt after receiving it. The format the node uses to transmit its password to the authentication server based in the fog. This password is then associated with the node's identity, Nid.

3.3 The Authentication of Nodes

Now, when a node and an authentication server in the fog authenticate one another, the authentication server based on fog is used by the node to issue an access request. The fog-based verification server records the nonce and checks the node's identity (Nid) and password (PWD) when it receives the access request. The protuberance receives an acknowledgement message if the password and identity match. At this point, the node can communicate with the fog server. Additionally, the procedure for authenticating nodes is shown.

4 Results

Here, demonstrate that our authentication process can withstand several threats. Figures from earlier sections show that employed nonces and time stamps to prevent replay assaults in most cases. Also, that everyone participating in the authentication has each other's credentials. The cypher suites cannot be cracked using classical or quantum computers, as used public key coding from the PQC domain. Our protocol will take advantage of lattice-based encryption, namely NTRU and Ring-LWE.

Depending on their needs, the user can use either public-key coding or private-key coding for encoding and decryption. Our recommendation goes to Ring-LWE because, as shown in Fig. 3, it is the fastest encryption and decryption method. The parameters, such as the ECC curve used or the machine's processing setup, determine the values of the execution time. The precise figure is less important than the ratio of the execution times' differences. Code-based encryption and lattice-based cryptography have somewhat different execution times. It should be noted that code-based encryption has the problem of substantial vital sizes, and more research needs to be done to considerably lower these key sizes. To prevent distributed denial of service attacks, a fog server should only allow a certain number of devices to authenticate at once and should not accept any more requests. The fog server will prevent resource exhaustion attacks for security by blocking IP addresses that send too many authentication requests. Except for attacks

founded on quantum computers, it is clear that many authentication mechanisms are safe. For that reason, our methodology is unique validate our authentication protocol using the AVISPA program, which stands for Automatic Authentication of Internet Security Procedures and Applications. Our protocol is defined in AVISPA using the High-Level Protocol Requirement Language (HLPSL), and their safety is checked in OFMC and CLAtSe modes. The OFMC mode combines two approaches to the demand-driven and lazy examination of Internet security protocols. It uses symbolic approaches and optimizations to generate actions based on user demand, including a sluggish Dolev-Yao invader model. An effective on-the-fly model checker for procedures with vast state spaces may be built using lazy data types. One efficient and effective automated analyser developed specifically for cryptographic protocols is the CL-AtSe mode. Following the guidelines of the Dolev-Yao intruder model, this attack searcher uses limited logic to simulate all possible states of the protocol's participants. Fairness, authentication, and privacy are just a few examples of state-based security properties it can simulate.

Additionally, it can simulate exponentiation and the exclusive OR (XOR) operation, two operators with algebraic characteristics. Limitations like inequality and typing may also be efficiently analyzed by it. Thus, demonstrate that our authentication mechanism is legitimate in these two cases.

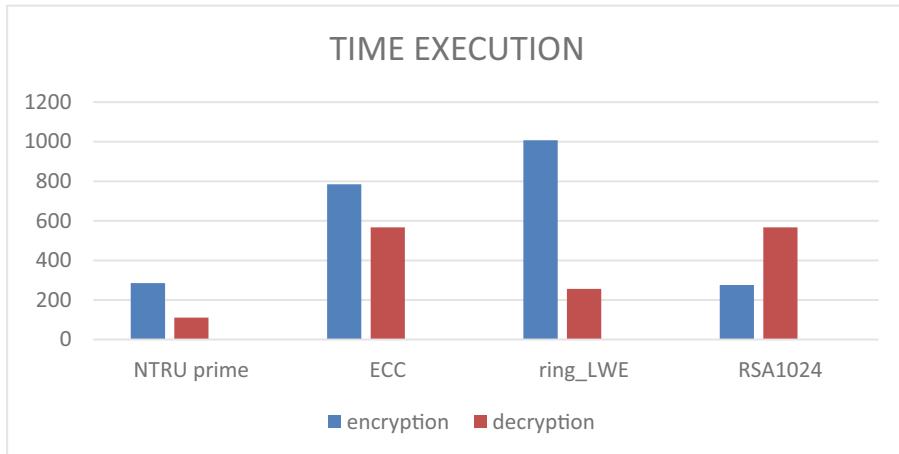


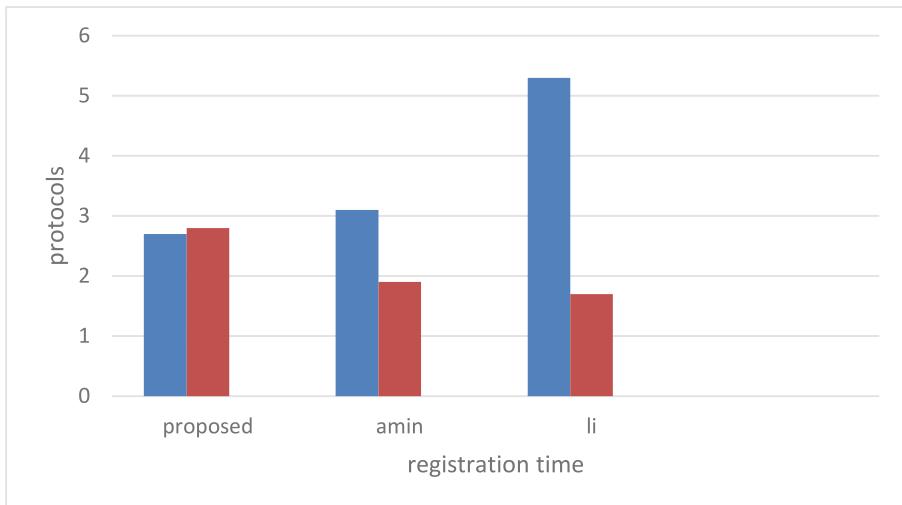
Fig. 3. Time Execution.

Regarding calculation time, we also evaluate our procedure compared to others, as shown in Figs. 4a, 5 and 6b. Compared to protocols that use public key cryptography, such as those shown in Figs. 4a, 5a, and 6a, ours is the speedier option. Lazing authentication based on hashes requires hardly any processing time at all; the only way to show that our protocol is safe against algorithms executed on both quantum computers and conventional electronic computers something that other protocols don't offer was to make this trade-off between computational expense and demonstrable refuge against occurrences based on quantum processors. Compares the execution times of several public vital cryptographies. In a Linux system with 4 GB of RAM and an Information

Core i5 (4310 U, 1.7 GHz), the execution time for 128-bit encryption and decryption is achieved. Even if a new study predicted that a quantum computer big enough to crack RSA-20481 won't be built in the next decade. On the other hand, RSA's very lengthy execution time makes it unsuitable for use in an IoT landscape. This highlights the need for PQC's applicability, as RSA-2048 is appropriate for other networks but not IoT networks. Here, provides an unofficial review of our authentication method.

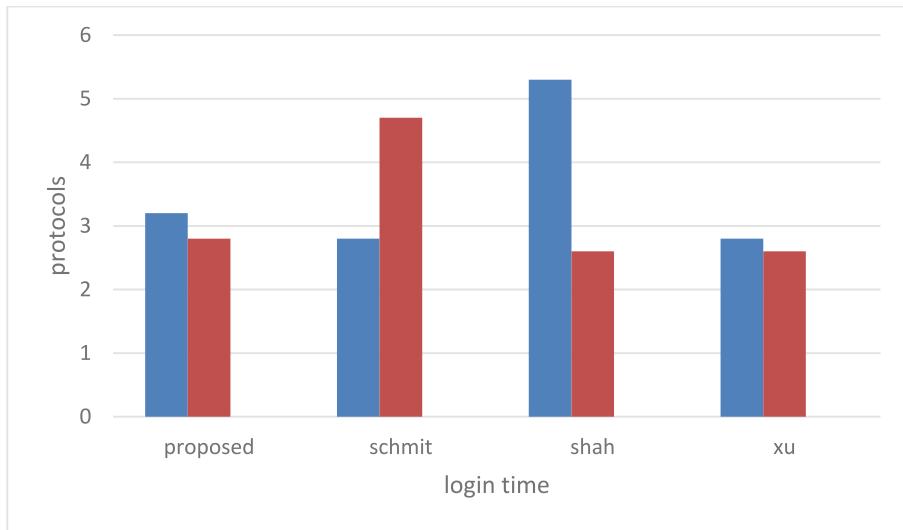


(a)Using PKC with protocols

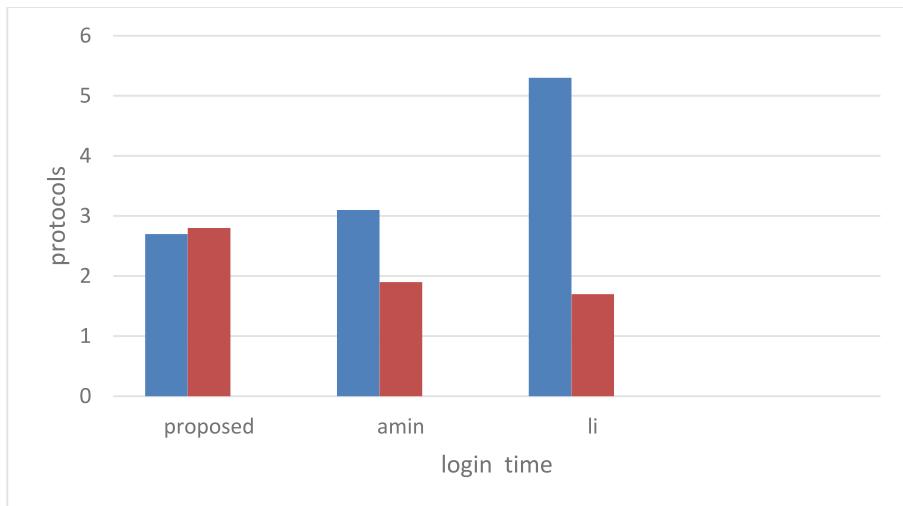


(b) Using protocols that include Hash

Fig. 4. Time comparison of registration



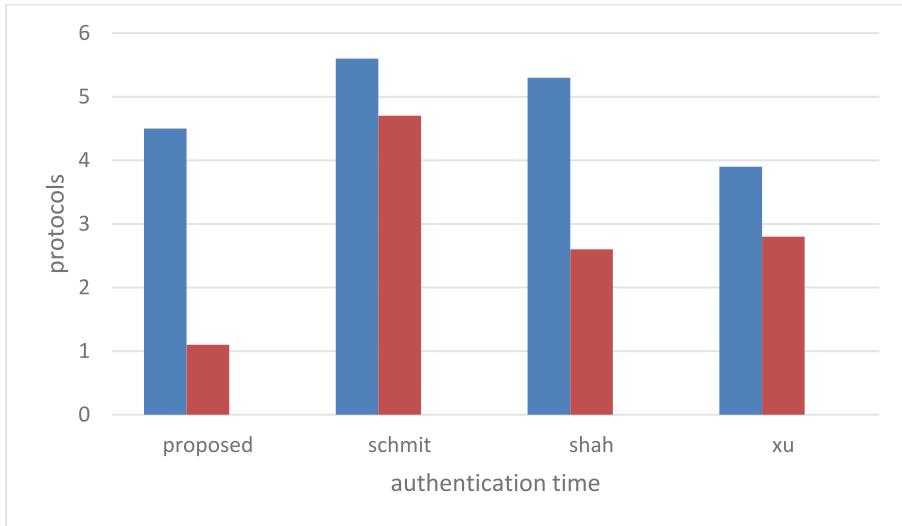
(a)Using PKC with protocols



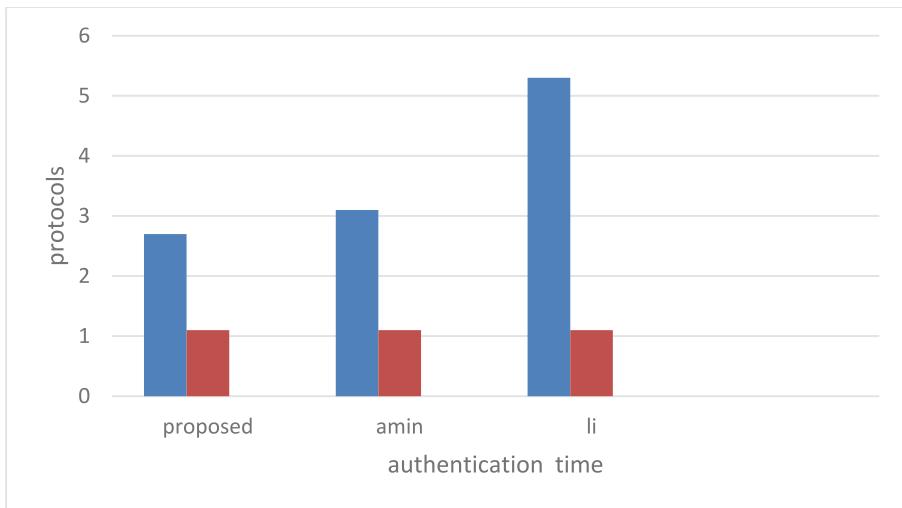
(b)Using protocols that include Hash

Fig. 5. Time Comparison of Login

Offline Password Guessing: To protect themselves against brute force password guessing attacks, users should choose robust passwords that include letters, numbers, and special characters. It would help if you took precautions to prevent a replay attack when a hacker attempts to pose as a genuine user by resending a message that the actual user has already transmitted and recorded. The verification protocol should prevent such an attack. For this reason, we often protect ourselves from this kind of assault by using



(a)PKC using protocols



(b) Using protocols that include Hash

Fig. 6. Time comparison of authentication

nonces and time stamps. In mutual authentication, both parties to the communication take part in the verification process. Each side verifies the other's authenticity to ensure that the other is a genuine entity. An attacker may get harmful access to a network by impersonating a genuine user; it's a crucial feature. As a result, verify that each participant has successfully authenticated the others. An insider with sufficient access

credentials may launch an impersonation attack on another network by gathering information about another user and attempting to utilize it in their system. This kind of attack should be able to bypass the authentication mechanism. So, it's recommended to keep the password and user ID encrypted.

One Man's Assault on Another: In this classic technique, the eavesdropper pretends to be a regular user and switches roles between sending and receiving messages. Such an attack should not be able to penetrate a secure authentication process. Using the AVISPA tool, we double-check that our protocol resists this attack. An adversary launches a resource depletion assault when they flood the server with requests. In a rapid sequence, the enemy would attempt to transmit the same message. Shortly after receiving the ciphertext, the server would compare it to the one it had previously received. The user's Internet Protocol address will be prohibited if it is determined to be identical.

5 Conclusion

This article provides an authentication mechanism for the Internet of Things (IoT) in a fog computing setting. The proximity of fog servers to edge devices reduces scalability, latency, and communication costs, making fog computing an attractive alternative to cloud computing. As a result, Internet of Things (IoT) networks have fog servers. Cryptosystems that are impervious to quantum computer assaults are a part of PQC. While quantum computers now do not compromise network security, that will soon change. In both OFMC and CL-AtSe modes, we show that our authentication protocol's six phases are secure using HLP SL in the AVISPA tool. When we compare several public-key cryptosystems, we find that Ring-LWE is the best, followed by NTRU and Niedermeyer, which outperform RSA and ECC regarding execution time. Through comparison with existing protocols, we confirmed that our IoT verification technique is secure against most known attacks. Moving forward, it is essential to include a fail-safe device that can construct a fog network on the fly if a fog server goes down.

References

1. Amin, R., et al.: A light weight authentication protocol for IoT-enabled devices in distributed Cloud Computing environment. Future Gener. Comput. Syst. **78**, 1005–1019 (2018). <https://doi.org/10.1016/j.future.2016.12.028>
2. Tewari, A., Gupta, B.B.: A lightweight mutual authentication protocol based on elliptic curve cryptography for IoT devices. Int. J. Adv. Int. Paradigms **9**(2–3), 111–121 (2017)
3. Yilmaz, Y., Gunn, S.R., Halak, B.: Lightweight PUF-based authentication protocol for IoT devices. In: 2018 IEEE 3rd International Verification and Security Workshop (IVSW). IEEE (2018)
4. Patro, P., Azhagumurugan, R., Sathya, R., Kumar, K. , Kumar, T.R., Babu, M.V.S.: A hybrid approach estimates the real-time health state of a bearing by accelerated degradation tests, Machine learning. In: 2021 Second International Conference on Smart Technologies in Computing, Electrical and Electronics (ICSTCEE), pp. 1–9. Bengaluru, India (2021). <https://doi.org/10.1109/ICSTCEE54422.2021.9708591>

5. Mewada, S., et al.: Smart diagnostic expert system for defect in forging process by using machine learning process. *J. Nanomater.* (2022)
6. Kumar, V., et al.: Light weight authentication scheme for smart home iot devices. *Cryptography* **6**(3), 37 (2022)
7. Tewari, A., Gupta, B.B.: A novel ECC-based lightweight authentication protocol for internet of things devices. *Int. J. High Perform. Comput. Networking* **15**(1–2), 106–120 (2019)
8. Atiewi, S., et al.: Scalable and secure big data IoT system based on multifactor authentication and lightweight cryptography. *IEEE Access* **8**, 113498–113511 (2020)
9. Chuang, Y.-H., et al.: A lightweight continuous authentication protocol for the Internet of Things. *Sensors* **18**(4), 1104 (2018)
10. SenthamilSelvan, R., Wahidabu, R.S.D., Karthik, B.: Intersection collision avoidance in dedicated short-range communication using vehicle ad hoc network. *Concurrency Comput.: Pract. Experience* **34**(13), e5856 (2022)
11. Aman, M.N., Basheer, M.H., Sikdar, B.: Data provenance for IoT with light weight authentication and privacy preservation. *IEEE Internet Things J.* **6**(6), 10441–10457 (2019). <https://doi.org/10.1109/JIOT.2019.2939286>
12. Chen, Y., Xu, W., Peng, L., Zhang, H.: Light-weight and privacy-preserving authentication protocol for mobile payments in the context of IoT. *IEEE Access* **7**, 15210–15221 (2019)
13. Aydin, Ö., Dalkılıç, G., Kösemen, C.: A novel grouping proof authentication protocol for lightweight devices: GPAPXR+. *Turk. J. Electr. Eng. Comput. Sci.* **28**(5), 3036–3051 (2020)
14. Rao, V., Prema, K.V.: Light-weight hashing method for user authentication in Internet-of-Things. *Ad Hoc Networks* **89**, 97–106 (2019)
15. Banerjee, S., et al.: Physically secure lightweight anonymous user authentication protocol for internet of things using physically unclonable functions. *IEEE Access* **7**, 85627–85644 (2019)
16. Shahidinejad, A., et al.: Light-edge: A lightweight authentication protocol for IoT devices in an edge-cloud environment. *IEEE Consumer Electron. Mag.* **11**(2), 57–63 (2022). <https://doi.org/10.1109/MCE.2021.3053543>
17. Hammi, B., et al.: A lightweight ECC-based authentication scheme for Internet of Things (IoT). *IEEE Syst. J.* **14**(3), 3440–3450 (2020)
18. Alzahrani, B.A., Mahmood, K., Kumari, S.: Lightweight authentication protocol for NFC based anti-counterfeiting system in IoT infrastructure. *IEEE Access* **8**, 76357–76367 (2020)
19. Suganthi, S.D., et al.: End to end light weight mutual authentication scheme in IoT-based healthcare environment. *J. Reliable Intell. Environ.* **6**, 3–13 (2020)
20. Jadhav, S.P.: Towards light weight cryptography schemes for resource constraint devices in IoT. *J. Mobile Multimed.* **15**(1), 91–110 (2020)
21. Shen, J., et al.: A lightweight multi-layer authentication protocol for wireless body area networks. *Fut. Gener. Comput. Syst.* **78**, 956–963 (2018)
22. Li, N., Liu, D., Nepal, S.: Lightweight mutual authentication for IoT and its applications. *IEEE Trans. Sustain. Comput.* **2**(4), 359–370 (2017)
23. JiHyeon, O., et al.: A secure and lightweight authentication protocol for IoT-based smart homes. *Sensors* **21**(4), 1488 (2021). <https://doi.org/10.3390/s21041488>
24. Rao, V., Prema, K.V.: A review on lightweight cryptography for Internet-of-Things based applications. *J. Ambient Intell. Humanized Comput.* **12**(9), 8835–8857 (2021)
25. Idriss, T.A., Idriss, H.A., Bayoumi, M.A.: A lightweight puf-based authentication protocol using secret pattern recognition for constrained iot devices. *IEEE Access* **9**, 80546–80558 (2021)
26. Amanlou, S., Hasan, M.K., Abu Bakar, K.A.: Lightweight and secure authentication scheme for IoT network based on publish–subscribe fog computing model. *Comput. Netw.* **199**, 108465 (2021)

27. Khan, M.N., Rao, A., Camtepe, S.: Lightweight cryptographic protocols for IoT-constrained devices: a survey. *IEEE Internet of Things J.* **8**(6), 4132–4156 (2021)
28. Bendavid, Y., et al.: Iot device security: Challenging “a lightweight rfid mutual authentication protocol based on physical unclonable function.” *Sensors* **18**(12), 4444 (2018)



Development of Elliptical Cryptography Technique to Watermark Embedded and Extrusion for Healthcare Records

N. Prajwal Hegde¹(✉), R. Sivaraman², G. Dharmamoorthy³, Ankit Kumar Dubey⁴, Pramoda Patro⁵, and S. Suma Christal Mary⁶

¹ Department of Artificial Intelligence and Data Science, NMAM Institute of Technology, Nitte
Deemed to be University, Karkala, India
prajwal.hegde@nitte.edu.in

² Department of Mathematics, Dwaraka Doss Goverdhan Vaishnav College,
Arumbakkam, Chennai, Tamil Nadu, India

³ Department of Pharmaceutical Analysis, MB School of Pharmaceutical Sciences, Mohan
Babu University, Rangampeta, Tirupati, India

⁴ Department of Computer Science and Engineering, Baderia Global Institute of Engineering
and Management, Jabalpur, MP, India

⁵ Department of Mathematics, Koneru Lakshmaiah Education Foundation, Hyderabad, India

⁶ Department of Information Technology, Panimalar Engineering College Poonamalle, Chennai,
India

Abstract. An application of E-healthcare watermarking method based on a hybridization of encryption as well as compression algorithms is proposed as the primary goal of this work. One step involves inserting a watermark into the image, and the other involves removing it. These are the two main components of the suggested system. Before beginning the embedding procedure, use a region-growing method to segment the tumour independently. Hash Secure Algorithm-256 encrypt the electronic health record (EHR) and Elliptical curve cryptography and its associated area of interest. Mathematical coding algorithms are then used to concatenate and compress the data. Lastly, the packed down bit is embedded into the original picture. When extracting a substance, the same steps are conducted again. Using the maximum signal ratio to noisy and normalized correlation, the experimental outcomes for various medicinal photos using EHR are shown, and the efficacy of the suggested method is evaluated.

Keywords: Water Marking · Embedding Procedure · Tumour · Electronic Health Record · Maximum Signal Ratio

1 Introduction

Protecting patients' privacy when dealing with their medical information has always been an issue. In order to ensure that patients' medical records remain private, the federal government issued regulations under the Accountability Act and Health Insurance

Portability. Healthcare treatment and diagnosis, research, education, and various commercial and non-commercial uses of medical data make them highly valued by both public and private enterprises. Recent developments in multimedia information management have led to digital watermarking (DWM). This technique subtly incorporates data into a signal host, such as an audio file, image, or video. There has been an ongoing effort to provide security in three areas: (i) the inaccessibility of transmitted medical images to unauthorized individuals (confidentiality), (ii) the preservation of the original quality of received images (integrity), and (iii) the verification that the intended recipients have received authentic images (authentication). Many visually appealing aspects of watermarking techniques are relevant to the healthcare industry. To include the watermark, they covertly alter the picture data. Water mark includes medical records of the patient's along with the doctor's and patient's data. Watermarking is a method that may be used in medical photos regardless of the assistance.

Digital watermarking is a way to add data to any media file electronically. To assert ownership of the material or picture, digital watermarks are used to identify it uniquely. Users may use text or images as watermarks. To insert a code word into a text, tiny structural changes are made, such as adjusting the width of lines and inter word spacing or changing the typefaces used for the characters. No amount of lossy image processing including using a low-pass filter, resampling, or lossy JPEG compression will remove the watermark from the original picture. The word "watermarking" refers to the process of subtly changing data to include metadata. In most cases, a digital watermark will take the form of embedded code inside a picture. The picture gains ownership or credibility due to its use as a digital signature. Diagnostic reports, pictures, and vital sign signals are just a few examples of the many shapes that electronic health record (EHR) technological advances take, which has supplanted chiefly the antiquated paper record paradigm. Also included may be the patient's susceptible medical records, including information about their demographics, physical exam, lab results, treatment plans, and medications. Due to the critical nature of these patient records, their storage on a shared network requires extra precautions to prevent their loss, which would compromise accurate diagnosis. Therefore, while transferring sensitive patient information electronically, the utmost care must be taken to ensure its integrity, security, and confidentiality. In this kind of situation, DWM is crucial. DWM approaches are used to ensure the privacy and security of medical photographs. Digital Rights Management aims to improve copy controls and tamper detection of digital material; DWM may achieve these goals. The picture quality must remain unaltered when adding digital watermarks or supplementary electronic patient records (EPRs) to medical images. The use of DWM approaches might lead to many benefits for medical data dissemination and administration, making them a one-stop shop for various issues.

This research provides a method for watermarking medical images in E-healthcare applications that combines compression and encryption. Here, begin by applying a region growing (RG) procedure to the input tumour picture to separate the ROI portion. After that, use the Secure Hash Algorithm (SHA)-256 to encode the ROI portion. They also use the elliptical curve cryptography (ECC) technique to encrypt the electronic health report. Merge the picture with the EHR data, then use an arithmetic coding (AC) approach to compress the data for better security. The last step is to include the

compressed bit streams in the initial picture. During the extraction process, the same steps are repeated. Watermarking is completed by integrating lossless information compression with encrypting methods and then randomly inserting watermark bits into the embedding zone. Because of all these features, the suggested strategy is a fresh approach to medical picture watermarking that works.

2 Research Contribution

This paper's main contribution is an efficient and effective watermarking scheme for E-healthcare records and medical images that can handle all four of the subsequent challenges:

- The suggested system integrates encryption and image hashes into the medical image, lossless information density to embed EHR and protecting the data from harmful assaults;
- An algorithm of RG segments accurately the input image to separate the ROI from it;
- Patient data should be more securely protected using the suggested approach without alterations;
- An attacker might try many different things to destroy the embedded watermark. Hence, the suggested solution has to be strong enough to withstand them all.

3 Proposed Watermarking Methods

The importance of medical data in healthcare diagnosis and treatment, research, teaching, and other commercial and non-commercial uses makes it vital for private and public organizations. Watermarking is a difficult task in these regions. They tried many ways to solve the problems of watermarking, as explained below.

3.1 Segmenting Medical Images

Medical images are most valuable when the ROI is significant. This part of the medical picture should be kept the same since it includes the most essential info. There are several ways to define discontinuous ROIs in medical imaging, and there might be many. Use the RG method to extract the region of interest from a picture described in this article. The input picture has dimensions 256×256 . First, the input picture is segmented to get the ROI. The RG projected method segments the image input relative to a seed opinion. The primary goal of RG segmentation is to control the first seed spots. Choosing a point seed, the starting point for RG is crucial for the segment key. The method of morphing in mathematics is utilized to get a starting point. How to segment images using projected RG is explained in full below.

Step 1: Compute a gradient of the source picture I on both the x and y axes I_{Rx}^2 and I_{Ry}^2 .
 Step 2: Calculate gradient vectors I^G by getting the hybrid gradient values using the following equation:

$$I^G = \frac{1}{1 + (I_{Rx}^2 + I_{Ry}^2)} \quad (1)$$

Step 3: To determine orientation values, convert gradient vector values from radians to degrees.

Step 4: Split the picture into networks of G^i .

Step 5: Set power and thresholds orientation.

Step 6: Repeat step 7 for each grid G^i until the image has the desired amount of grids.

Step 7: (a) Determine the H histogram for every G^i pixel.

Step 7: (b) Represent the G^i grid's maximum histogram commonly as F^H .

Step 7: (c) Choose a pixel based on F^H and assign it as a point seed with orientation OR_p and intensity IN_p .

Step 7: (d) Choose the neighboring pixel with orientation OR_n and intensity IN_n .

Step 7: (e) Compare p as well as n's intensity and orientation.

In other words,

$$D_{IN} = \|IN_p - IN_n\| \quad (2)$$

$$\text{and } D_{OR} = \|IN_p = IN_n\| \quad (3)$$

Step 7: (f) To extend the area, add an identical pixel if $D_{IN} < T_{IN}$ as well as $D_{OR} \leq T_{OR}$. Otherwise, go to step 7(h).

Step 7: (g): Verify whole area pixel addition. Step 6 then 7(h) if true.

Step 7: (h): Measure the area again, find fresh seed points, then repeat step 7 (a).

Step 8: End the process.

Segmenting the input pictures is accomplished with the use of this RG technique. The outcome of the segmentation process is displayed in Fig. 1 (a) and (b).

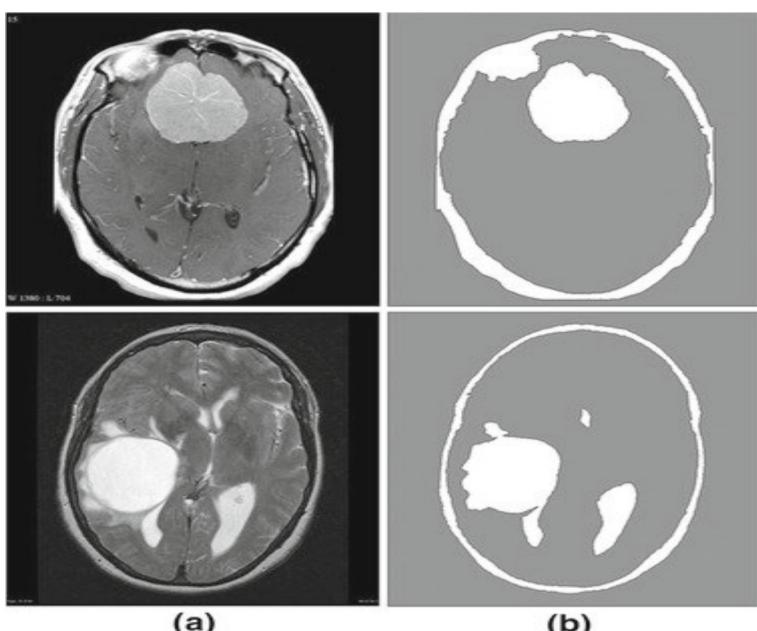


Fig. 1. Outcome of Segmented Images (a) Input image and (b) Tumour Segment Part

3.2 ROI Hash

Using the RG algorithm, they extract the ROI area from the input picture during the segmentation step. Calculate the ROI hash value after segmentation. This study uses SHA-256 to calculate the ROI hash, producing a 256-bit message digest with 64 characters. Since this technique is unidirectional, it generates a distinct code for each input. The ROI hash value is used for authentication purposes. While this study focused on a single return on investment (ROI), the predicted approach also applies to other ROIs. The SHA-256 hash value computation is shown in Fig. 2.

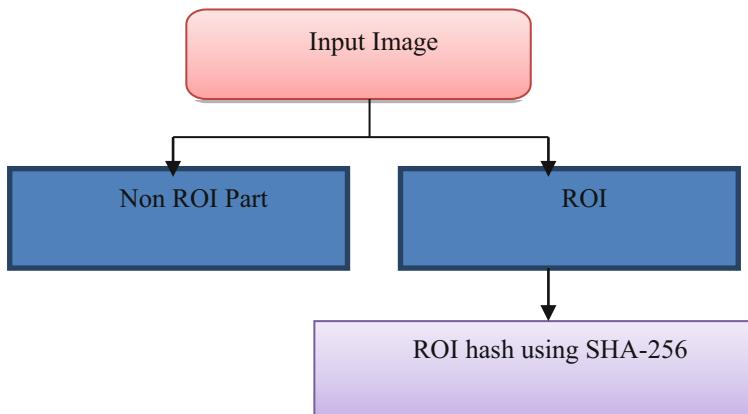


Fig. 2. SHA-256 Hash Value Computation

3.3 ECC

ECC uses elliptic curves across finite fields as an algebraic foundation for cryptography. Each consumer has a public key and private key and a set of actions associated with the keys to execute cryptographic operations according to public key cryptography. The design uses a four-step process to verify user validity. First, establish the connection, create an account, authenticate, and change data. ECC needs a better computing speed compared to linear methods. A further advantage is its sub-exponential temporal complexity, making it challenging to break. Sensitive data is encrypted and sent to the cloud. Keys allow users to explain sensitive information with the owner's authorization.

These are the stages that make up the ECC process.

Stage 1: A digital signature uses a hashing method to compress large amounts of data or documents into smaller ones, termed message digests.

Stage 2: To generate a digital signature, the private key is used to calculate the message digest.

Stage 3: The ECC algorithm encrypts the digitally signed document using the public key of the user.

Stage 4: The data owner may use a public key to decode the digital signature to a message digest and then use a private key to transform the encrypted text into plain text.

An elliptic curve's ethics form the basis of the elliptical curve cryptosystem. In modern cryptography, the points that fulfil the equation are enclosed by a planar curve known as an elliptic curve. Our study predicts the following equation for a curve with an elliptic shape over a field K:

$$x^3 = y^3 + ay + b \quad (4)$$

The coordinates are x and y, and the components of K are a, b.

The procedure consists of three stages: creating keys, encrypting data, and decrypting it.

3.3.1 Creating Keys

It is a substantial operation since they need to provide both the public and private keys during key creation. A developer's public key is used to encrypt messages, while a data owner's private key is used to decode them. At this point, they must choose a value f from the set of all possible integers m. using this formula, they may offer the public key:

$$H = f * q \quad (5)$$

H is an open key, f is the key that is private, q is a point on the curve, and f is a randomly chosen value between 1 and m-1.

3.3.2 Encrypting Data

The sensitive message is denoted by p. Put p on the curved line E and imagine it as point M. Determine k randomly from the interval [1 – (m – 1)]. After the encryption process, two cipher texts, R_1 and R_2 , will be generated.

$$R_1 = k * q \quad (6)$$

$$R_2 = M + k * H \quad (7)$$

3.3.3 Decrypting

The equation below represents the transliteration of the transmitted message M:

$$M = R_2 - f * R_1 \quad (8)$$

As a result, only authorized users may access the encrypted data via the linked query. A user may also use the data owner's credentials to access sensitive data if necessary.

3.4 EC

Bit stream compression is achieved using AC. When it comes to lossless data compression, statistical coder AC is your best bet. Finding the optimal length of code words is the primary objective of AC. They need to know the likelihood of the individual symbols appearing for this entropy coder, just like any other. According to data theory, the median code length is almost the smallest feasible value. The AC assigns a size to each symbol's interval, representing the likelihood of that symbol's presentation. A complying interval rational number may serve as a symbol's code word. A rational integer, always included inside the interval for every symbol, characterizes the whole data set. As more data is supplied, the number of crucial numbers keeps growing.

4 Method for Watermarking

This study combines compression and encryption techniques to provide a lossless watermarking solution for medical images. Watermarking is often used in privacy applications to prevent unauthorized individuals from accessing confidential communications. The two main components of the suggested system are the embedding and extraction processes. The patient's information and the medical imaging data are both watermarked here. This article describes embedding and extraction.

4.1 Process of Embed

The embedding procedure described in this research combines compression and encryption. Figure 5 is the schematic depicting the embedding process as a whole.

Step 1: Consider a 256 x 256 input picture I. Begin by segmenting the ROI area (I_{ROI}) from the input picture using the RG method.

Step 2: Calculate ROI hash function ($Hash_{ROI}$) using SHA-256 after segmentation.

Step 3: Calculate $Hash_{ROI}$ (B_{ROI}) binary and convert it to H_{ROI} .

Step 4: Improving the watermarking quality is achieved by immediately converting the I_{ROI} to binary format ($R_{B_{ROI}}$) and then converting the binary image to a value in hexadecimal (RH_{ROI}).

Step 5: As a further piece of feedback, think about the EHR. This is when the ECC algorithm comes into play, encrypting the EHR. The electronic health record includes the following elements: patient reference number, patient's name, doctor's name, patient's age, and date of addition. Protected health records are obtained using the ECC algorithm (Encrypt).

Step 6: The next step is to decrypt the data and then convert it to binary format. Translate the binary data into its hexadecimal counterpart, H_{EHR} , for analytical purposes.

Step 7: Concatenate picture and EHR after encryption. CON_I Is obtained by concatenating H_{ROI} , H_{EHR} , and RH_{ROI} .

Step 8: For better embedding efficiency, compress CON_I using the AC technique. Get the compressed C_{BS} bit stream here.

Step 9: Follow these procedures to build the first random matrix R.

$$I_{Seed} = \sum_{i=1}^n \sum_{j=1}^n I \quad (9)$$

A randomized matrix R is created using the input picture size and pseudo-random matrix creation using I_{Seed} :

$$R = PRMG[R_{Seed}]_{(2 \times 2)} \quad (10)$$

Step 10: Then, construct the last randomized matrix RM using R using these steps:

After subtracting 0.5 first created randomized matrix R, multiply the resulting matrix by 2. The resulting matrix is R_t .

$$R_t = (R - 0.5) \times 2 \quad (11)$$

- (ii) The desired randomized matrix RM is constructed utilizing the pseudo-randomized matrices creator, utilizing R_t matrix as the value of the seed.

$$RM = PRMG[R_t]_{(2 \times 2)}^* \quad (12)$$

Step 11: Insert C_{BS} into picture I. Finally, receive the watermarked picture I_W . See below for the embedding procedure.

Since the water mark bit is 0, multiply the randomized matrix R_M by the embedded strong point β and add it to the initial image I.

$$[I_M] = [I_M] + (\beta * [RM]), \text{ where } \beta = 2 \quad (13)$$

- (ii) Watermark bit 1 signifies no operation.

Step 12: Repeat procedures 10 and 11 to embed all watermark pixels. The first random matrix R is produced from the seeded PRMG for each cycle I_{Seed} .

4.2 Extraction Method

To compare the original picture I with the compressed bit C_{BS} and the information EHR, the extraction method involves operating in reverse of the embedding operation on a suspected watermarked image I_W . Starting with the same process as the embedding method, the extraction algorithm finds the original information. To begin, the extraction procedure is fed the embedded watermark image I_W .

5 Result and Discussion

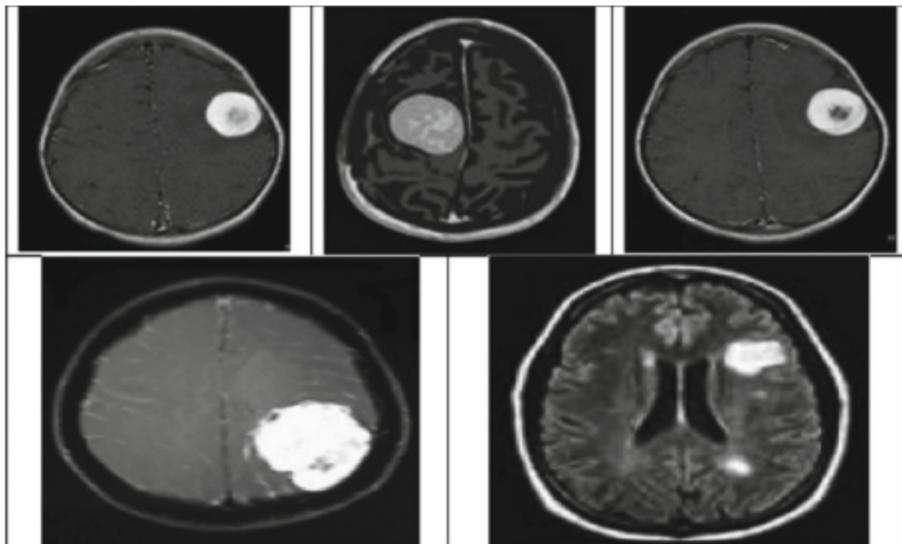
This part discussed the outcomes of our suggested technique and analysed our presentation. The following settings are used on a Windows PC to run the proposed watermarking system: System requirements include a 3.20 GHz Intel (R) Core i5 CPU, 4 GB of random-access memory (RAM), and Windows 7 Professional. Mat lab is the program that is used for the implementation. They used five distinct kinds of brain tumour pictures in our study. Table 1 displays the input photographs (embedded algorithm) and Example of EHR utilized in experimental result as depicted in Fig. 3.

Table 1. Embedded Algorithm

```

Input: Original picture I, EHR hidden key K_5
Output: Watermarked picture
Start
1. User input 2. RG algorithm segments ROI 3. SHA-256 calculates ROI hash
4. Encrypt EHR using ECC algorithm 5. Determine ROI binary value
6. Concatenate picture and EHR data 7. Compress data using arithmetic coder
8. Embed bit into original picture.
End Output
Watermarked picture I_W

```

**Fig. 3.** EHR Results

5.1 Assessment Methods

Equations below give the following measures for evaluating the proposed approach.

5.1.1 PSNR Peak

The peak signal-to-noise ratio (PSNR) is a metric for evaluating the watermarked image's quality. The PSNR measures how well the watermarked picture compares to the original. Using the mean squared error (MSE), the PSNR is determined. By squaring the difference between the highest power of the signal and the corrupting noise, the MSE calculates the total squared error. Lower MSE values and higher PSNR values show improved watermark quality.

5.1.2 NC-Normalized Correlation

The degree to which the attacked image's watermark resembles the original watermark is determined by NC.

5.2 Proposed Methodology Experimental Results

They provide a watermarking method that combines lossless data reduction with encryption to include electronic health records (EHRs) and image hashes into medical photographs. Information security is enhanced by the hybrid approach. In this case, use the RG technique to extract the ROI region from the original picture. Next, use the ECC method to encode the EHR and SHA-256 to encrypt the ROI. The data is then compressed using the AC technique after being concatenated. The medical picture incorporates the compressed bit stream. At last, the watermarked picture is in the possession. Medical photos at a resolution of 256×256 were used.

In this part, they thoroughly discuss the practical implications of the suggested strategy. Within the input picture, conceal the ROI and EHR in this study. Once the encryption is complete, the watermarked picture is processed to get the ROI and EHR. Implement two tiers of security to make the system more efficient. The first one deals with encryption, while the second one deals with compression. There is reduced computing complexity with this hybrid method. Figure 4 displays the results of adjusting the threshold for the suggested PSNR-based segmentation approach. Apply the RG algorithm for segmentation in the suggested technique. The RG algorithm adjusts its performance according to the threshold. The results show that when the value of the threshold rises, the PSNR value also rises. With a threshold value of 1.27, the PSNR is low, while with a threshold value of 1.28, it is high. The highest possible PSNR of 43.3472 dB is achieved by the recommended technique in this case.

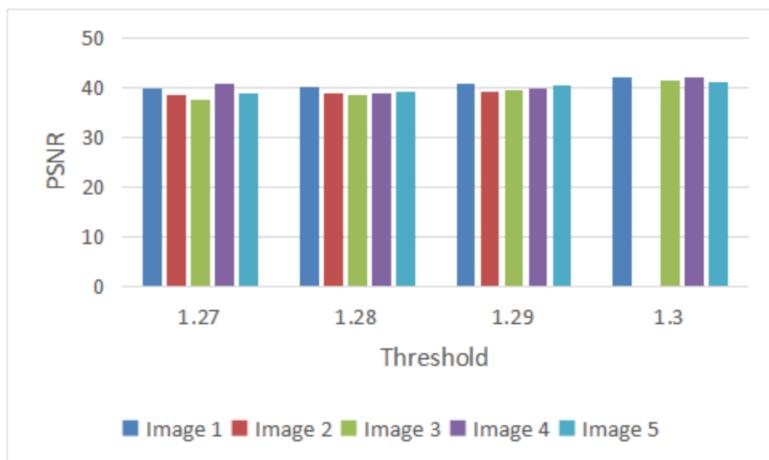


Fig. 4. Results of Adjusting the Threshold

Figure 5 compares the extracted ROI picture and the segmented ROI image. From this, they may deduce that the maximum accuracy for images 1, 2, 3, 4, and 5 is 99.20%, 97.5%, 98.89%, 97.35%, and 98.36% respectively. The effectiveness of NC metrics is seen in Fig. 6. The suggested strategy achieves the greatest NC. The suggested method's performance under various attacks is also shown in Fig. 7. Attacks are implemented during the encryption process. The first attack is making five-pixel changes, applying the encryption method, and then measuring the outcome. Almost the same result is obtained here. Consequently, our output is unaffected by this assault. The same goes for testing the performance with different pixel values. Even after the assault is applied, the strategy still produces superior outcomes.

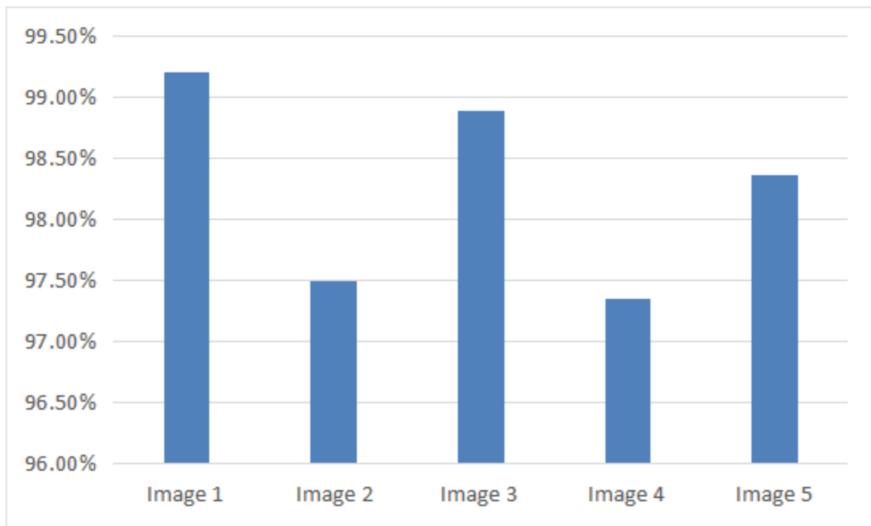


Fig. 5. Watermark Maximum Accuracy for Images

5.3 Comparative Analysis

Use of PSNR and NC allows for an analysis of the suggested watermarking technique's performance. The usefulness of the suggested methodology is shown by comparing the suggested method's matching results with those of current approaches. Their two watermarking algorithms were distinct from one another. Their first step was to include the EPRs and digital watermarks in the ROI and RONI. In addition, the digital watermark and EPRs were concealed using the RONI. In this suggested study, the RG algorithm was used at the segmentation step. Compared to other approaches, this algorithm ranks high as a top choice for segmentation. A k-means clustering technique is compared to the suggested approach here.

The results of the accuracy-based performance analysis are shown in Fig. 8. Contrast the k-means clustering process with the suggested RG approach. The segmentation

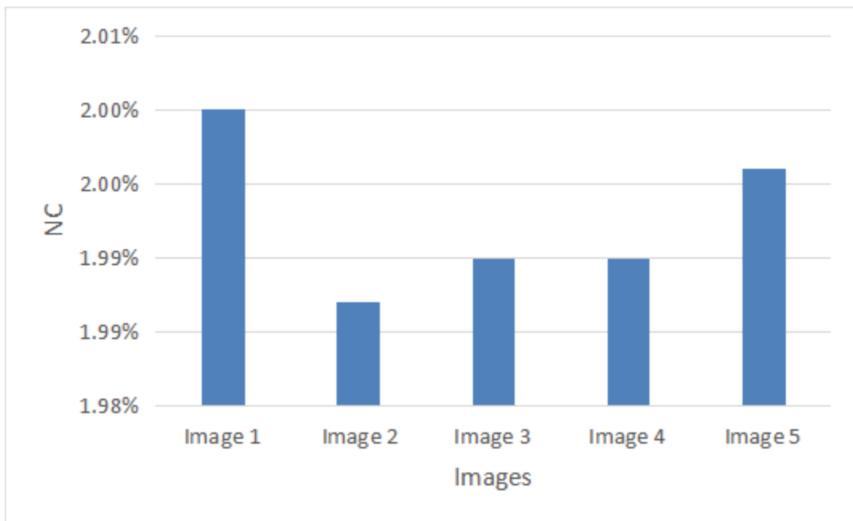


Fig. 6. Effectiveness of NC Metrics

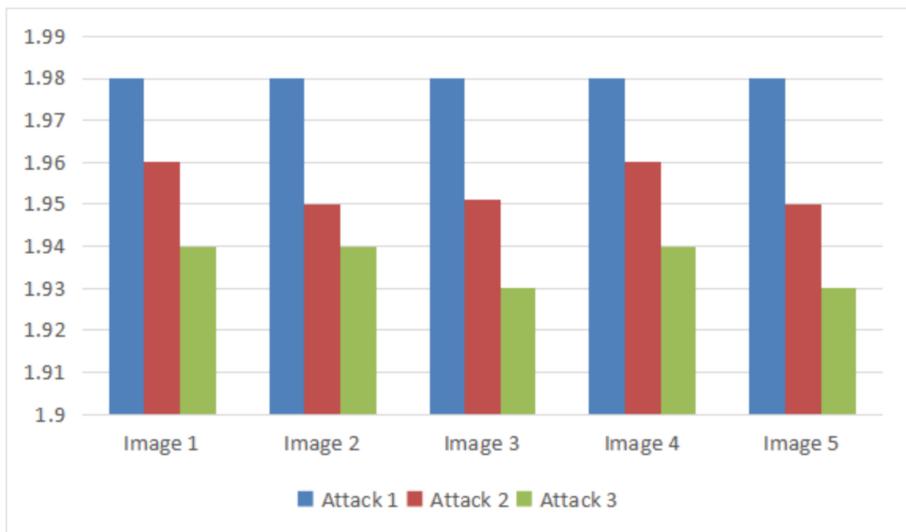


Fig. 7. Suggested Method's Performance under Various Attacks

algorithms that divide the ROI and RONI also include this one. Results for images 1, 2, 3, 4, and 5 reveal that the suggested method achieves maximum accuracy of 96.4%, 98.2%, 93.65%, 94.73%, and 95.78%, respectively, as shown in Fig. 8. The greatest PSNRs were 81.54%, 82.54%, 84.65%, 82.60%, and 83.10% when the k-means clustering approach was used. The suggested segmentation method outperforms the k-means algorithm based on the results.

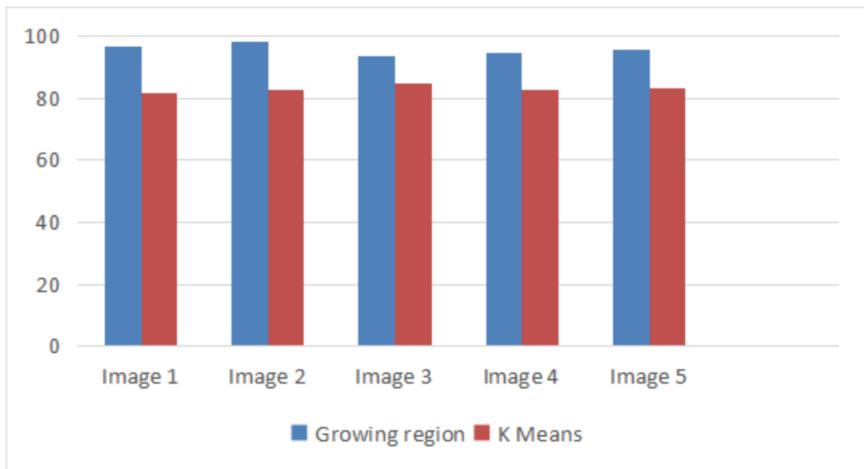


Fig. 8. Accuracy-Based Performance Analysis

Figure 9 also displays the results of comparing the suggested method to the current PSNR metric. Compared to the current algorithm's highest PSNR of 39.64 dB, the suggested method attains a maximum of 43.34 dB (Fig. 13). The results demonstrate that the suggested strategy outperforms competing approaches. Figure 10 compares the current NC metric with the suggested one. Compared to other methods, the suggested one produces the highest NC value.

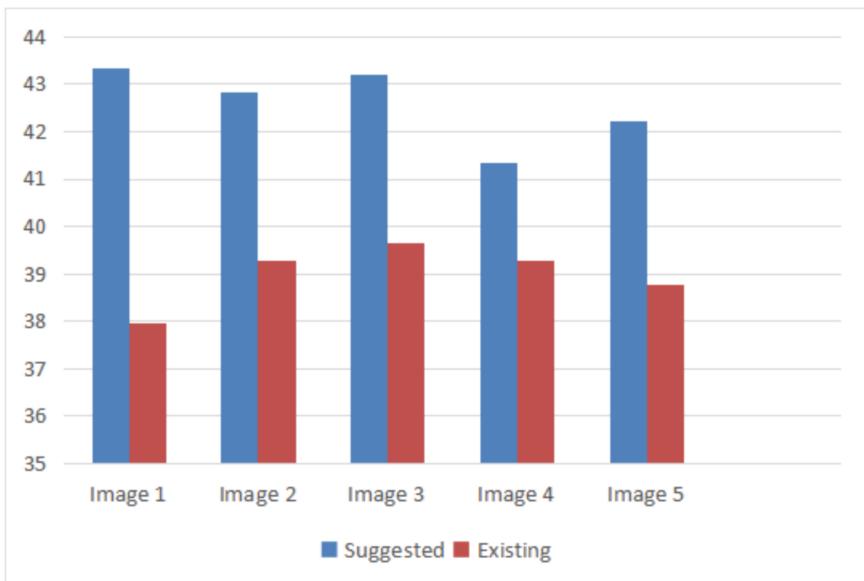


Fig. 9. Compare the Suggested Method to the Measure of PSNR

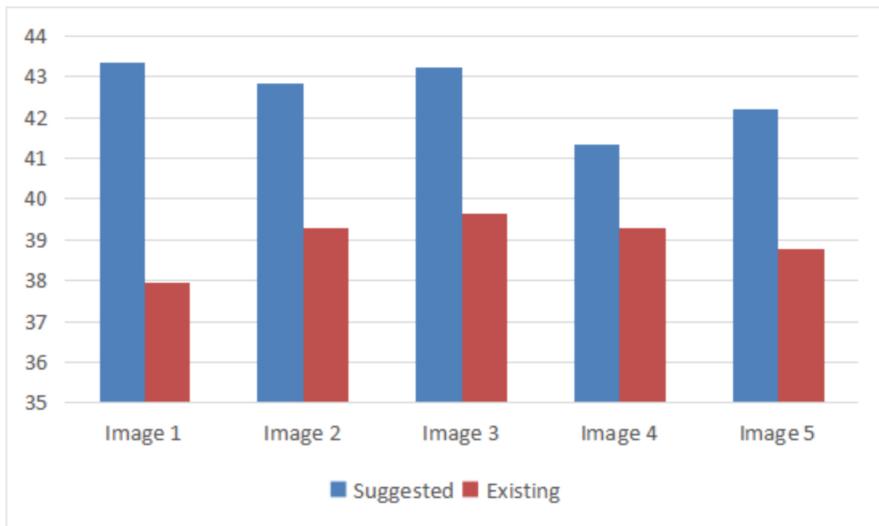


Fig. 10. Proposed Method vs. NC Measure Comparison

6 Conclusion

Develop new watermarking methods for digital image processing. Four algorithms—SHA-256, RG, AC and ECC—were examined in this research. The suggested approach employs simple mathematical computations to recover the original ROI and provide authentication and recovery data, reducing computing complexity. With an average PSNR of 43.34 dB, an embedded capability of 73,385 bits, an extracting accuracy of 99%, and an NC of 1, the suggested approach retains the watermarked picture quality. In addition to improving embedding efficiency, the testing findings show that the suggested strategy improves watermarked picture quality. To ensure medical picture integrity, system authentication, and secrecy, the suggested approach may be realistically integrated into medical information systems in future work.

References

1. Ullah, F., et al.: Data exfiltration: a review of external attack vectors and countermeasures. *J. Netw. Comput. Appl.* **101**, 18–54 (2018)
2. Shirey, R.: Internet security glossary, version 2. No. rfc4949 (2007)
3. Swarm, Foraging-Based Oriented By Particle, And Prediction Of Mining Subsidence. “1001 Mining Network Traffic Data.”
4. Gallez, Bernard, and Biomedical Magnetic Resonance Unit. “Information and Communication Technologies at Ucl.”
5. Varga, P., et al.: Converging telco-grade solutions 5G and beyond to support production in industry 4.0. *Appl. Sci.* **12**(15), 7600 (2022)

6. Belleville, Bill. Losing it all to sprawl: How progress ate my cracker landscape. University Press of Florida, 2006
7. Sheppard, C.: Coral Reefs: A Very Short Introduction. Oxford University Press, Oxford (2021)
8. Model, Driven Volatility, and Varying Electrical. "Part I Oral Sessions."
9. Chomsky, N., Lange, F.: The Mental and the Natural. Biolinguistics and Philosophy: Insights and Obstacles 111 (2012)
10. Mahaboob, S., et al.: Issn-2455–6300 October 2018, Spl Issue 10.1.
11. Bhamare, P.A., Naik, A., Bhushan, S.D.: Recent trends in 3D print organs. KJ Somaiya Inst. Eng. Inform. Technol. **11**(1), 74 (2018)
12. Ahmad, I., et al.: Communications Security in Industry X: A Survey (2023)
13. Takács, I.-C.: Conceptual metaphors: the fundamental grounding for language and thought. In: 13th Conference on British and American Studies: Language Diversity in a Globalized World. Cambridge Scholars Publishing (2017)
14. Holme, R.: Grammatical metaphor as a cognitive construct. Amst. Stud. Theory Hist. Linguist. Sci. Ser. **4**, 391–416 (2003)
15. Steen, G.: "gj steen@ vu. nl To be published in Companion to Cognitive Linguistics Editors John Taylor and Jeanette Littlemore London: Continuum Publishing."
16. Deng, Y.: Genius and Genus: How to Name Things with Metaphors. Chinese University of Hong Kong (2014)
17. Lakoff, G.: 10 The contemporary theory of metaphor. The Cognitive Linguistics Reader 267
18. Harbus, A.: Thinking in metaphors: figurative language and ideas on the mind. Sydney Stud. English **30**(2004), 3–20 (2004)
19. Cameron, L.J.: Andrew Goatly, The language of metaphors. London: Routledge, 1997. Pp. xvi+ 360. Hardback£ 60, ISBN 0 415 12876 5; paperback£ 17.99, ISBN 0 415 12877 3. *English Lang. Linguis.* 2(1), 162–165 (1998)
20. Janeke, C.: Language, cognition, and metaphor. Language Matters **26**(1), 129–146 (1995)
21. Stern, J.: Metaphor, semantics and context. The Cambridge Handbook Metaphor and Thought **564**, 564 (2008)
22. Vanparys, J.: Andrew Goatly. the language of metaphors. Funct. Lang. **6**(1), 159–161 (1999)
23. Wood, G.M.. The Feeling of Metaphor. Diss. University of Melbourne (2019)
24. Staton, S.: Ordinary Language Metaphors: A Plan for a Corpus-Based Research Project. Ordinary Language Metaphors 1000–1019
25. Deignan, A.: Corpus linguistics and metaphor. The Cambridge Handbook of Metaphor and Thought **280**, 290 (2008)



Step-by-Step Image Encryption Using UACI and PixAdapt

J. Balamurugan¹(✉), Mali Yadav², Jetti Madhavi³, A. Basi Reddy⁴, and R. Senthamil Selvan⁵

¹ Department of Master of Business Administration, St. Joseph's College of Engineering, OMR, Chennai, Tamilnadu, India

drjbalamuruganpdf@gmail.com

² Department of Physical Education, Lovely Professional University, Punjab, India

³ Department of Mathematics, Malla Reddy Engineering College, Medchal, Hyderabad, India

⁴ Department of Computer Science and Engineering, School of Computing, Mohan Babu University, Tirupati, Andhra Pradesh, India

⁵ Department of Electronics and Communication Engineering, Annamacharaya Institute of Technology and Sciences, Tirupati, Andhra Pradesh, India

Abstract. A new cutting-edge method recently garnered interest is image encryption utilizing a genetic approach. Currently, most genetically based picture encryption algorithms use fixed parameters to encrypt images, ignoring any image-representative properties. This research led to PixAdapt, a novel adaptive picture encryption technique. It is being re-engineered such that the fitness of the encrypted picture can be determined using UACI, and the relevant parameters may be adjusted using hereditary hill climb or pretend annealing. Pseudorandom numbers have been generated using a variety of chaos-based maps, including the Logistic, Rossler, Henon, and Tent maps, as well as the linear feedback shift register. The PixAdapt algorithm uses the theory of confusion and diffusion to further ensure that the plaintext and ciphertext images are completely dissimilar. This is the first instance of optimizing picture encryption settings using metaheuristic search algorithms. Finding the best possible UACI value, the hereditary hill climb algorithm encrypts all of the photos. Fitness enhancement, limit evolution, arithmetical analysis, and encryption quality testing have all been conducted on the method. When it comes to effectively encrypting images, PixAdapt is one of a kind and has shown that the UACI encryption parameter is a good match.

Keywords: UACI · Encryption Technique · Confusion Algorithm · Quality testing · Encryption Quality Testing

1 Introduction

Images come in various formats, each tailored to a particular need. In their way, each of these pictures is in a broader sense, an adaptive system designed to handle and adjust to environmental changes while maximizing performance goals [1, 2]. If this is the case, the suggested system adapts to its surroundings, consisting of different kinds of pictures

it sees on each pass [3–5]. The adaptive system determines whether the parameters need additional evolution based on their fitness for each picture when the environment changes. The overarching goal of the procedure is to provide a solution that falls within the permissible fitness range to sustain performance [6, 7]. Compared to standard picture formats like JPEG and PNG, medical field and satellite photographs are much bigger and have more dimensions. Ensuring the security of images during media transfers is crucial, regardless of the image's properties. Using picture encryption, the original image is transformed into an unintelligible format. The use of chaotic sequences to encrypt images has been the subject of much study. These classifications are produced using a fixed set of parameters that fail to consider the qualities of the accurate picture [8–10]. Currently, few encryption systems can adjust to diverse kinds of photos. This work aims to re-engineer the procedure of picture encryption by using a metaheuristic examination algorithm to discover the optimal pairings of parameters [11]. The absence of an adaptive mechanism prompted interest and led to this endeavor [12–14]. When faced with a problematic optimization issue, metaheuristic search algorithms may find a solution quickly. An adaptive system must consist of different entities, maintain continuity, and achieve success that differs from that of other entities to be subject to natural selection. The suggested mechanism meets these criteria [15, 16].

After each iteration, the values of the parameters used to construct chaotic sequences change. Depending on the approach, these values are continuously developed or derived from one another. A new entity is found after every pass to maximise the fitness value. The suggested processes, to be more precise, might be thought of as an adaptive system that modifies its behaviour to accommodate changes in its surroundings [17, 18]. The present state undergoes evolution into a new one when the prerequisites for adaptation are satisfied. This explanation pertains to the encryption system lying dormant until it encounters a collection of parameters that fail to generate the necessary fitness. Whenever the system detects that the fitness is below a certain threshold, it will not interfere [19, 20]. The system's adaptability is a result of metaheuristic search strategies such as replicated annealing and genetic hill climbing. Adaptability evolves parameters by algorithms when the system recognizes it is not producing the appropriate fitness [21–23]. This technology has been further used, such as developing picture resolutions and providing a feedback system to decrease or increase parameters based on favourable or negative feedback. One such use is encrypting a sequence of photographs [24, 25]. In conclusion, the study aims to achieve the following, which may be seen as the benefits and new features offered by the algorithm: First and foremost, want to create PixAdapt, an innovative adaptive picture encryption technique that employs hereditary hill climb or replicated annealing.

Second, provide a fitness function that will use genetic hill climbing or replicated annealing to maximize the UACI value of a picture. The same has been confirmed via the development of suitable experiments. Thirdly, PixAdapt employs chaos-based maps and a linear feedback shift register to produce suitable pseudorandom sequences. One alternative for producing the pseudorandom sequence is to use a switching mechanism. Creating a confusion-diffusion process for PixAdapt is the fourth objective. Providing evidence that the PixAdapt algorithm was optimized using metaheuristic search approaches [26, 27].

PixAdapt, short for “Pixel Encryption through genetic adaption,” is a principle. The method takes a genetic approach, which allows it to adapt to the optimal parameter values (such as UACI) and generates accurate fitness. The fitness function aims to adapt by evolving parameters until they provide a satisfactory fitness score. The suggested approach may soon be the first to use adaptiveness to encrypt images. PixAdapt uses the substitution-confusion-diffusion technique to encrypt genetic images symmetrically. Binary sequence regeneration controls the confusion-diffusion process, whereas utilizing the algorithm’s many maps control the replacement process.

2 Methodology

2.1 Image Description Example

Table 1 displays the picture sets utilized as samples in the experiments. The suggested method is tested with ten photos. Use 8-bit grayscale photos. Five symmetrical photos and one asymmetrical image were used for the photographs [2, 28]. The suggested approach’s shape and size variety may be tested on several picture kinds. Thus, the picture sizes span smaller and bigger photos, ranging from 257×257 to 1991×1343 .

Table 1. Specifics of encrypted picture samples.

Name	Size	Congruity
Images_1	(257, 257)	Regular
Images_2	(300,300)	Regular
Images_3	(300, 300)	Regular
Images_4	(512, 512)	Regular
Images_5	(800, 600)	Irregular
Images_6	(512, 512)	Regular
Images_7	(492, 600)	Irregular
Images_8	(594, 670)	Irregular
Images_9	(1991, 1343)	Irregular

2.2 Mechanism of Adaptation of the Algorithm Suggested

Adaptive systems are those that can change in response to their surroundings. Although adaption methods differ depending on the application, the end aim is always the same: to accomplish the set target and return to normal. An additional adaptive system tracks its performance and changes its settings to get better over time. This study looks at two ways that picture encryption might be improved. The first strategy uses genetic algorithms, namely the hill ascent algorithm, to change its parameters. An enhanced variant of the hill climb algorithm, Simulated Annealing, is the second technique used.

These adaptive processes aim to bring the UACI assessment parameter closer to 33.46. An illustration of how the adaptive picture encryption system operates may be seen in Fig. 1. After receiving the picture, the encryption system begins by encrypting it using a predetermined set of parameters. Using the UACI assessment limit, the fitness of the encrypted picture is tested. Iteration termination occurs if fitness is within the predetermined tolerance limit. The parameters are input into an evolutionary system, which adjusts those parameters when the fitness falls outside of the permitted range. Next, the picture is re-encrypted using the parameters that have been developed. This practice is repeated until the fitness level approaches perfection.

Because of the challenges encountered while trying to encrypt a picture and achieve an ideal UACI score of 33.46, UACI is being used as a fitness indicator. Nearly all UACI values around 33.46 other systems of measurement in the near-ideal range, and UACI is the most encryption-sensitive measure by far, much more so than entropy and NPCR to conduct experiments, a range of 34.47 ± 0.01 was chosen.

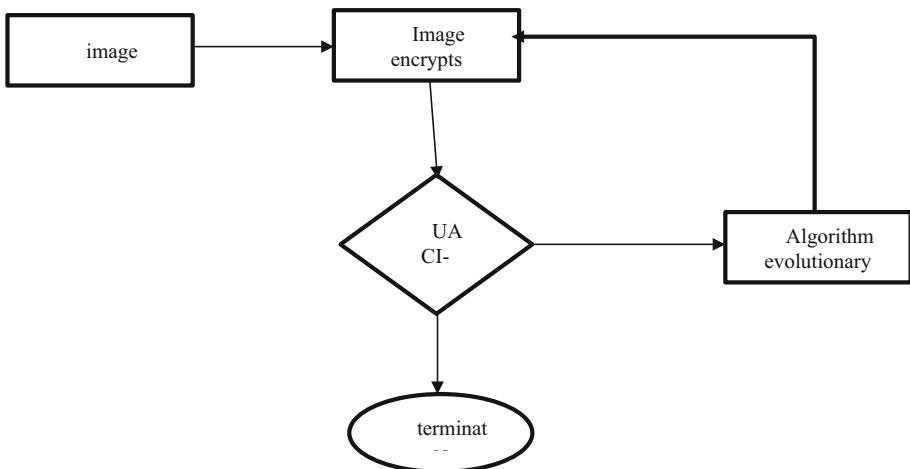


Fig. 1. An Algorithm for Adaptive Picture Encryption

2.3 Algorithm for PixAdapt Suggested

Two adaptive methods have been investigated and used for picture encryption. First, encrypt images using a Hill Climb genetic algorithm; then, employ collection and change based on a fitness metric called UACI to enhance the solution. Regardless of the encryption technique, most other picture encryption metrics, including entropy, NPCR, and correlation, do not exhibit considerable variance. The second method evolves the solutions derived from the cypher picture using Simulated Annealing.

2.4 Disorientation and Dispersion Algorithm

A new diffusion-confusion mechanism has developed in the PixAdapt algorithm. Before starting the binary decomposition procedure, the original picture array and the chaotic

array of pseudorandom orders are both taken into consideration. To further facilitate the division of the binary array into its eight channels, the two arrays are then converted to binary format. There are a total of eight channels, four in each pair. Given that “bi There are four sequences—A11, A21, a1, and a2—used in this method, and each binary sequence has a size of N iterations. The eight channels in both the original picture and the chaotic sequence array (C1 and C2) aid in the computation of their corresponding cypher image arrays. Finding the initial value of the C1 array may be achieved by XORing the first and last values of array A11 and array a1. The initial array value of C2 is determined using a comparable approach. The ith values of a1, the (i-1) the worth of A11, and the ith values of it are combined in an XOR operation to get the remaining array values.

Similarly, the remaining values of collection C2 are computed by XORing the ith and (i-1)th values of A21 with the ith worth of a2. According to sequences 1 and 2, the C1 and C2 arrays are transformed into four planes. To create the final encryption picture, arrays C1 and C2 are combined. This is done an infinite number of times, denoted as n. “nary decomposition” means this procedure. Then convert the four channels to the second arrangement, producing A1, A2, a1 and a2. After adding the other sequence to A1 and A2, the result is A11 and A21, the right-shifted versions of the original sequences.

The whole process of confusion and diffusion is shown in Fig. 2 and Algorithm 1.

```

Function BINARY Decomposition
height ← binary_imageheight
width ← binary_imagewidth
temp_array ← zeros
counter_1 ← height
counter_2 ← width
counter_3 ← 7
for I ← 0, counter_1 do
  for j ← 0, counter_2 do
    var ← binary_image(I,j)
  end for
  end for
  return temp_array
end function
if i==0 then
  c(i)-a11(i-1)a1(i)
  c-2(i)-a21(i)a2(i)
end if
end for
c1- reshape back to 4 planes according to seq_1
c2- reshape back to 4 planes according to seq-2
c-merging c1 and c2
end for
return
end function

```

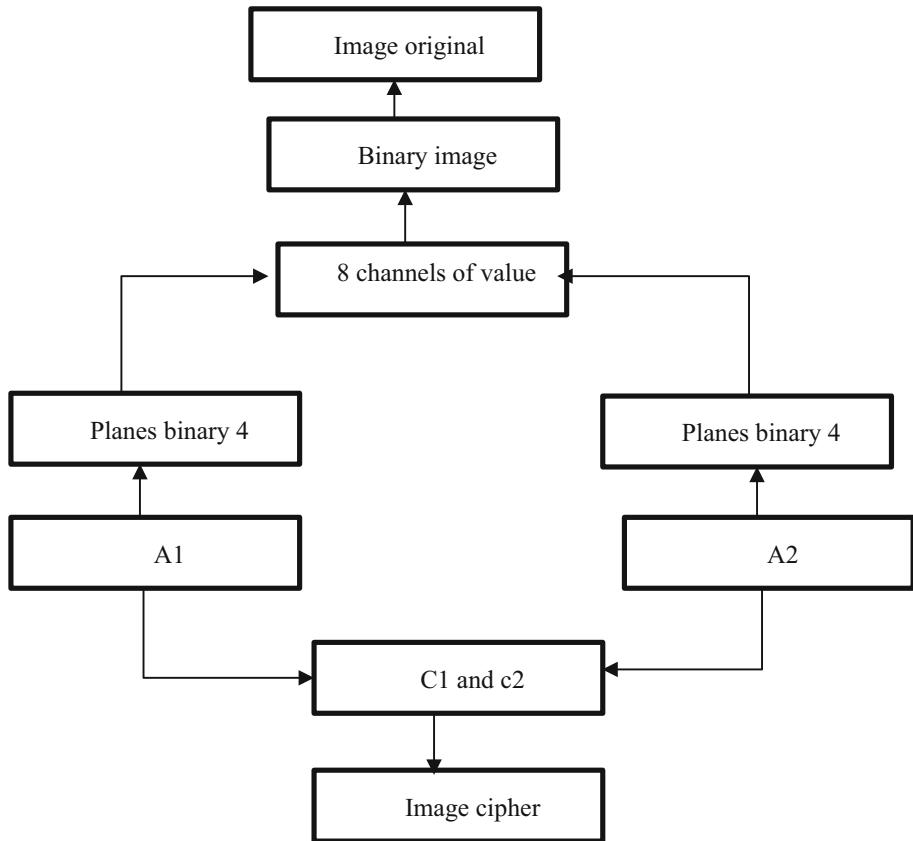


Fig. 2. Diffusion and confusion process

3 Results

Instead of considering each picture's unique characteristics, traditional image encryption methods encrypt them using a fixed set of values. By using an adaptive method to picture encryption, these limitations were circumvented. Using the pictures detailed before, tested the suggested algorithms, and tried the genetic hill climb algorithm with a 100-person population. in progress with a temperature of 20° and a reduction step size of 10 in the Simulated Annealing method.

The non-adaptive method relied on setting the parameters for pseudo-random sequence creation to their most arbitrary or chaotic states. Their values are as follows: $a = 1.4$, $x = 0.01$, $y = 0.01$, and $r = 3.99$ for the logistic map; 10001001 for the linear feedback register; 0.01, 0.01, and 1.99 for the tent map; 9 for the Rossler map c ; and 0.01, 0.01, and 0.01 for the Henon map x and y , respectively. Discussions around this approach centre on its fitness improvement, parameter development, statistical aspects, encryption quality, and critical space.

3.1 Enhanced Physical Fitness

Images' fitness as a function of epoch counts for the hereditary hill climb method (Fig. 3) and pretend annealing (Fig. 4), respectively, are shown. Most of the time, the fitness number needs to be higher or higher. The main criterion of a fitness function is that it stabilizes to its most ideal value of UACI, and it does so gradually. For instance, the fitness function exhibits linear performance (refer to Fig. 3). This indicates that the ideal UACI value for maximized as it approaches its maximum value. The ideal UACI value is immediately reached when using simulated annealing on the same picture. The Progress of Parameters The pseudo-random sequences have been generated using an overall of eight limits designated in the preceding section. Five more parameters have been used to toggle the sequences on and off.

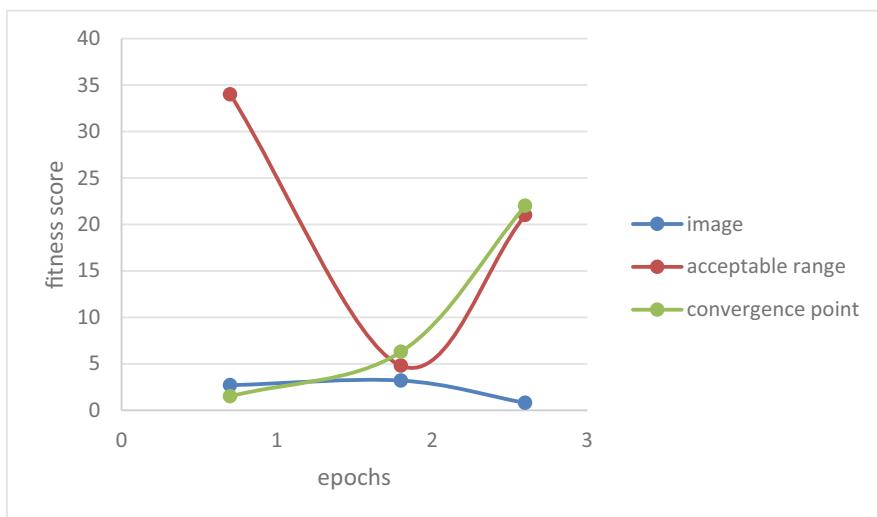


Fig. 3. Genomic hill climbs algorithm fitness enhancement.

3.2 The Rossler Map

Figure 5 shows that when the Rossler map was rotten, the UACI worth was inside the permitted variety, and that c values around 13 did not function well in the genetic hill climb strategy. Rossler's map did not help to improve fitness for the majority of pictures. When the map was turned on, the simulated annealing method achieved the ideal value of UACI.

3.3 Map Tent

Pay close attention to the ideal UACI value for the simulated annealing method, as shown in Fig. 6. The measured UACI value was low once the map was turned on but it was within the allowed variety when the map was turned off (using the simulated annealing technique).

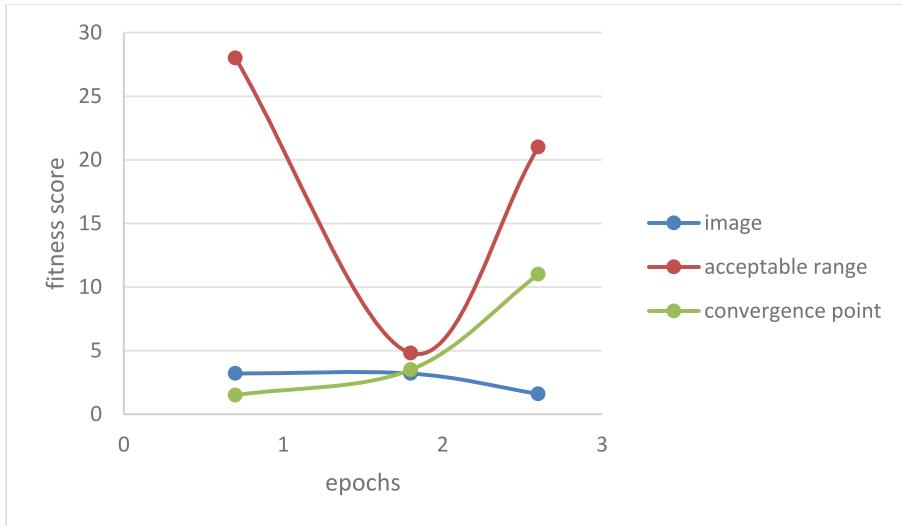


Fig. 4. Enhancing fitness using simulated annealing

4 Discussion

The findings indicate that the encryption quality was enhanced by using an adaptive approach. Investigated two metaheuristics for parameter evolution—the evolutionary hill climbs and the replicated annealing approach—and found that an adaptive system would have produced better encryption results than a non-adaptive one. Compared to the non-adaptive encoding system, both methods were much more successful. Hereditary hill climb, simulated annealing, and non-adaptive photo encoding were all surpassed by all three methods when it came to entropy, correlation, and contrast. Bypassing adaptive picture encryption resulted in a poor UACI score. When compared to the replicated annealing approach, the genetic hill climb method performed better in chi-square tests and NPCRs. According to the introduction, PixAdapt is an adaptable system in every sense. Optimal encryption quality and resilience to statistical, brute-force, plain-text, and cypher-text assaults are the primary goals of both of these adaptive techniques. The fitness of the encoded picture was measured using UACI as a metric. The acquired data provided further confirmation that UACI was a suitable fitness metric. As the fitness parameter that optimized the other parameters, UACI was shown by the histogram, correlation, contrast, entropy, and non-parametric probability density functions (NPDs). The results for adaptive image and sequence encryption algorithms were almost optimal when using these picture evaluation criteria. Further evidence that the adaptive technique worked for all image kinds was the cryptosystem's large critical space and key sensitivity. Results in terms of UACI, entropy, and connection values were similar for both the Genetic Hill Climb and the Imitation Annealing techniques.

Although Simulated Annealing produced superior results overall, the Genetic Hill Climb algorithm outperformed it for metrics including NPCR, chi-square, and contrast. Statistical, histogram, and encryption quality analyses revealed that the two adaptive

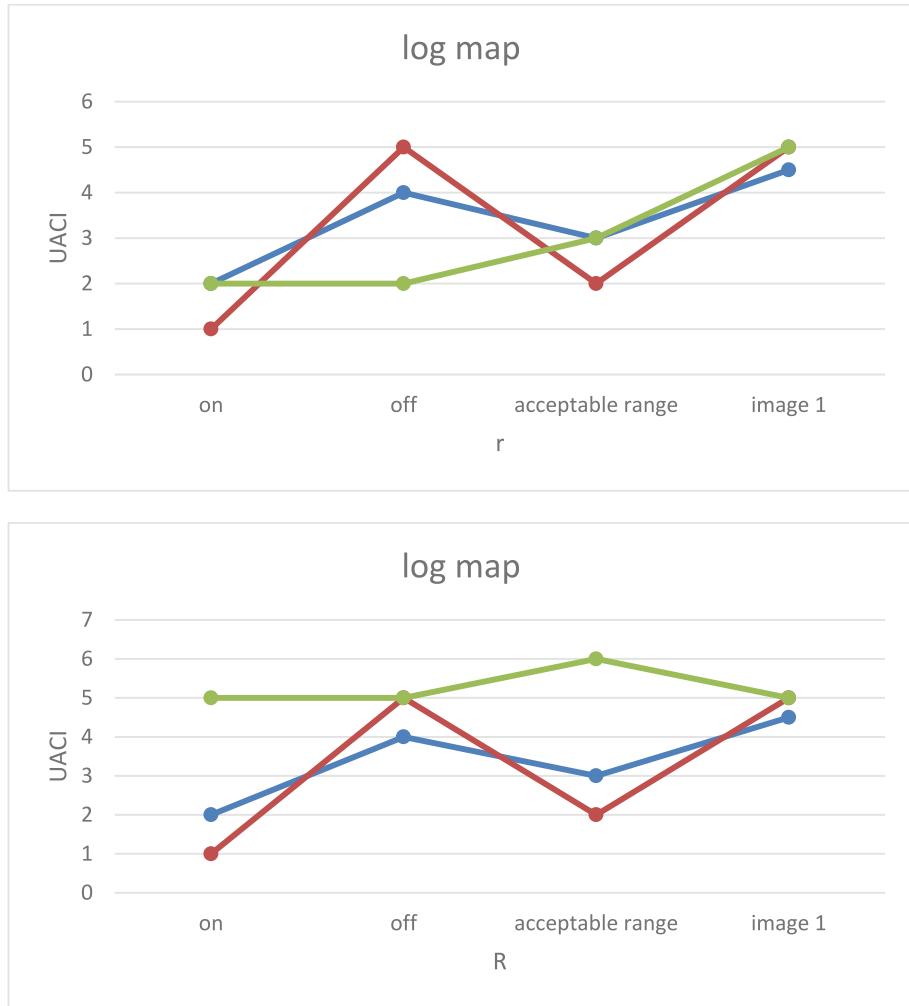


Fig. 5. Map criteria for logistics the settings for r and the seed value.

methods outperformed the static settings. The experiments aimed to optimize fitness by testing how well the adaptive mechanism worked. Parameter evolution, fitness optimization, statistical analysis, and encryption quality were assessed using non-adaptive, simulated annealing, and genetic hill climb techniques on PixAdapt. Among all the methods, genetic hill climbing yielded the best results. So, by using the initial image and enabling and disabling a few image-based pseudo-random sequences, the system became more flexible. The algorithm's overall adaptability was enhanced by including a switching mechanism. This method enhanced the search space for genetic and meta-heuristic algorithms by delving more into the behaviour of cypher images. The adaptive mechanism firmly established the non-adaptive picture encryption sequence because it failed to generate results within the acceptable range when used alone. After optimising

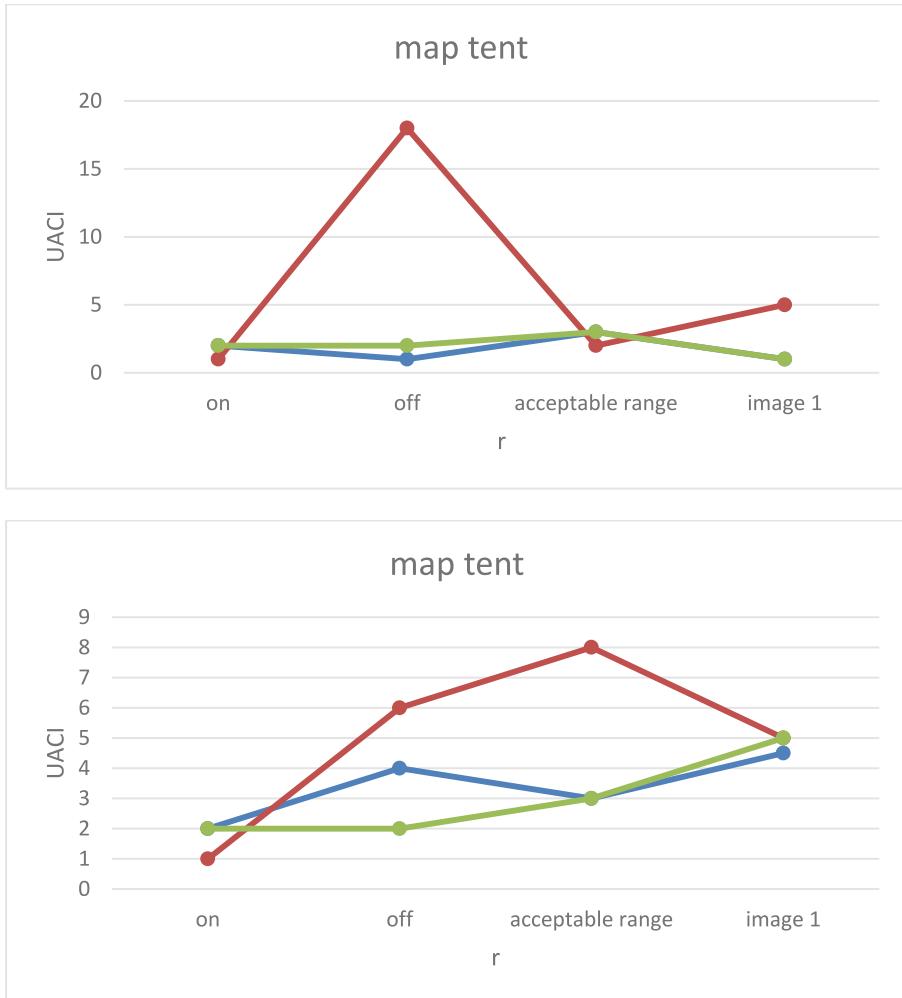


Fig. 6. R and seed value tent map parameters.

them throughout testing, it was determined that using a single set of settings for picture encryption is inefficient. Additionally, it was shown that generating a key for optimal picture encryption may need more than one pseudo-random classification.

Several parameters in the suggested picture encryption method produce chaotic maps, and a loop brings the system to convergence (acceptable fit). Parallel processing, map reduction, and graphics processing units (GPUs) are ways to address these problems, which contribute to the overall algorithm's high temporal complexity. Furthermore, the permutation procedure may be divided into sections, and the picture might be encrypted using the optimal range of values.

5 Conclusion

When encrypting a picture, most algorithms only employ a fixed set of parameters without considering the image's properties. Adaptive picture encryption, or PixAdapt, was developed and used in this article. By optimizing settings within a predetermined range of values, PixAdapt employs an evolutionary algorithm to encrypt images. If the ideal fitness is not achieved, the evolutionary or metaheuristic algorithm will activate to evolve the limits. Research in this area has shown that UACI is a suitable fitness function parameter. This article used a new method to activate and deactivate the pseudorandom sequences. Results were improved with this procedure compared to non-adaptive parameters that did not use the switching instrument. The switching process added a degree of adaptability, which was also noted. Testing the same technique on other parameters for evaluating picture encryption is possible. The approach may be used utilizing parallel processing to encrypt photos effectively. Incorporating new genetic algorithms and metaheuristic search approaches and running the algorithm on additional parameters may increase the adaptive mechanisms. Adding support for medical and colour picture encryption to PixAdapt is also possible.

References

1. Tuli, R., Soneji, H.N., Churi, P.: PixAdapt: A novel approach to adaptive image encryption. *Chaos Solit. Fractals* **164**, 112628 (2022)
2. Wang, Y., et al.: Adaptive fast image encryption algorithm based on three-dimensional chaotic system. *Entropy* **25**(10), 1399 (2023)
3. Zou, C., Wang, L.: A visual DNA compilation of Rössler system and its application in colour image encryption. *Chaos Solit. Fractals* **174**, 113886 (2023)
4. Korayem, Y.: Image Encryption Using Novel Variations of Chaotic Functions
5. Feng, W., et al.: Cieasp: A Novel and Efficient Chaotic Image Encryption Algorithm Based on a New Simplified Quadratic Polynomial Hyperchaotic Map and Pixel Fusion Strategy. Available at SSRN 4385127
6. Li, Z., et al.: Verifiable Multi-Image Encryption Scheme Based on Time Concomitant Sequence and Unbiased Random Exchange Permutation-Diffusion. Available at SSRN 4584180
7. Meng, X.F., et al.: Two-step phase-shifting interferometry and its application in image encryption. *Opt. Lett.* **31**(10), 1414–1416 (2006)
8. Li, S., Zheng, X.: On the security of an image encryption method. In: Proceedings International Conference on Image Processing, vol. 2. IEEE (2002)
9. Jun, W.J., Fun, T.S.: A new image encryption algorithm based on single S-box and dynamic encryption step. *IEEE Access* **9**, 120596–120612 (2021)
10. Zhang, G., Liu, Q.: A novel image encryption method based on total shuffling scheme. *Opt. Commun.* **284**(12), 2775–2780 (2011)
11. Özkaynak, F.: Brief review on application of nonlinear dynamics in image encryption. *Nonlinear Dyn.* **92**(2), 305–313 (2018)
12. Pareek, N.K.: Design and analysis of a novel digital image encryption scheme. arXiv preprint arXiv: 1204.1603 (2012)
13. Mirzaei, O., Yaghoobi, M., Irani, H.: A new image encryption method: parallel sub-image encryption with hyper chaos. *Nonlinear Dyn.* **67**(1), 557–566 (2012)

14. Wang, X.-Y., et al.: A chaotic image encryption algorithm based on perceptron model. *Nonlinear Dyn.* **62**, 615–621 (2010)
15. Han, Z., et al.: A new image encryption algorithm based on chaos system. In: IEEE International Conference on Robotics, Intelligent Systems and Signal Processing, Proceedings, vol. 2. IEEE (2003)
16. Patro, P., Azhagumurugan, R., Sathya, R., Kumar, K., Kumar, T.R., Babu, M.V.S.: A hybrid approach estimates the real-time health state of a bearing by accelerated degradation tests, Machine learning. In: 2021 Second International Conference on Smart Technologies in Computing, Electrical and Electronics (ICSTCEE), pp. 1–9, Bengaluru, India (2021). <https://doi.org/10.1109/ICSTCEE54422.2021.9708591>
17. Mewada, S., et al.: Smart diagnostic expert system for defect in forging process by using machine learning process. *J. Nanomater.* **2022** (2022)
18. Pisarchik, A.N., Zanin, M.: Image encryption with chaotically coupled chaotic maps. *Physica D* **237**(20), 2638–2648 (2008)
19. Babaei, M.: A novel text and image encryption method based on chaos theory and DNA computing. *Nat. Comput.* **12**(1), 101–107 (2013)
20. Kanso, A., Ghebleh, M.: A novel image encryption algorithm based on a 3D chaotic map. *Commun. Nonlinear Sci. Numer. Simul.* **17**(7), 2943–2959 (2012)
21. Li, H.: Image encryption based on gyrator transform and two-step phase-shifting interferometry. *Opt. Lasers Eng.* **47**(1), 45–50 (2009)
22. Pujari, S.K., Bhattacharjee, G., Bhoi, S.: A hybridized model for image encryption through genetic algorithm and DNA sequence. *Proc. Comput. Sci.* **125**, 165–171 (2018)
23. Laiphakpam, D.S., Khumanthem, M.S.: Medical image encryption based on improved ElGamal encryption technique. *Optik* **147**, 88–102 (2017)
24. Enayatifar, R., Abdullah, A.H., Isnin, I.F.: Chaos-based image encryption using a hybrid genetic algorithm and a DNA sequence. *Optics Lasers Eng.* **56**, 83–93 (2014)
25. Arab, A., Rostami, M.J., Ghavami, B.: An image encryption method based on chaos system and AES algorithm. *J. Supercomput.* **75**, 6663–6682 (2019)
26. Ye, G., Huang, X.: An efficient symmetric image encryption algorithm based on an intertwining logistic map. *Neurocomputing* **251**, 45–53 (2017)
27. Li, C., et al.: Cryptanalyzing an image encryption algorithm based on autblocking and electrocardiography. *IEEE Multimed.* **25**(4), 46–56 (2018)
28. SenthamilSelvan, R., Wahidabalu, R.S.D., Karthik, B.: Intersection collision avoidance in dedicated short-range communication using vehicle ad hoc network. *Concurr. Comput. Pract. Exp.* **34**(13), e5856 (2022)



Investigation of Post-Quantum Cryptography to Secure the Functionality of Vehicle Hardware Architecture

K. R. Jothi¹, Chetan Khemraj Lanjewar², R. Sivaraman³, Bramah Hazela⁴,
P. R. Sivaraman⁵, and A. Azhagu Jaisudhan Pazhani⁶✉

¹ Department of Computational Intelligence, School of Computer Science and Engineering,
Vellore Institute of Technology, Vellore, India

jothi.kr@vit.ac.in

² KCC Institute of Technology and Management, Greater Noida, India

³ Department of Mathematics, Dwaraka Doss Goverdhan Doss Vaishnav College,
Arumbakkam, Chennai, India

⁴ Amity School of Engineering and Technology, Amity University Uttar Pradesh, Lucknow
Campus, Lucknow, India

⁵ Department of Electrical and Electronics Engineering, Rajalakshmi Engineering College,
Chennai, India

⁶ Ramco Institute of Technology, Rajapalayam, Tamilnadu, India

alagujaisudhan@gmail.com

Abstract. A Hardware Secure Module (HSM) is often installed in the Electronic Central Unit (ECU) to include security features in the in-vehicle architecture, which has become more critical due to the fast expansion of IT in the automotive sector. Hence, the HSM's built-in hardware accelerators may perform secret cryptographic calculations, enforcing secure communications. No universally accepted method exists for developing the framework of an automotive HSM. The car industry hopes for a standard automotive HSM architecture to meet consumers' rising performance demands and ward against future assaults by malicious actors using large-scale quantum computers. Researchers are looking at post-quantum cryptography (PQC) to ensure that future HSMs will still be secure in anticipation of the advent of quantum computers. They provided new sets of hardware accelerators for the next generation of automotive HSMs after analysing the candidates in NIST's PQC process for standardization. Based on our testing, it is possible to construct a post-quantum safe automobile HSM that satisfies the stringent standards set by today's electronic central unit.

Keywords: Hardware Secure Module · Electronic Central Unit · Vehicle Architecture · Post-quantum Cryptography · Hardware Accelerators

1 Introduction

More and more automobiles have been linked in recent years, together in terms of communications among other automobiles and among various embedded sub-systems inside a vehicle, thanks to the advancements in information technology within the automotive

industry over a decade [1]. An extensive system of interconnected embedded systems with capabilities is installed in a vehicle using Electronic Central Units (ECUs). While in-vehicle systems that include connection services improve usability, they also provide more opportunities for cybercriminals to launch attacks. Integrating a Hardware Security Module (HSM) is fundamental to protecting automobiles from harmful attackers [2, 3]. An HSM is a safe enclave that encapsulates security based functionality inside an ECU [4].

HSM solutions from top chip manufacturers typically feature a microcontroller processor, blocks of memory (e.g., ROM, RAM, and flash), hardware acceleration for cryptographic operations, and safe user interfaces among the host processor and HSM in the ECU. ECUs often employ hardware accelerators in the HSM for secret-dependent cryptographic operations [5, 6]. A hardware security module (HSM) secures sensitive data (e.g., keys) and speeds up cryptographic processes using specialized hardware accelerations [7].

The automobile industry still needs to establish an official HSM architecture or functionality standard. As a result, more research into the future difficulties and insights into the present patterns in the contemporary automobile industry are required to construct the HSMs of the future [8–10]. On the one hand, contemporary automobiles’ ever-increasing performance demands need HSMs with more power. However, with a “quantum era” on the horizon, HSM architectural design may need to undergo even more dramatic shifts. Some popular asymmetric cryptography algorithms, including RSA and ECC, are susceptible to quantum computer attacks; for example, Shor’s algorithm can solve the RSA and ECC’s fundamental problematic issues in polynomial time. Modern HSMs’ symmetric cryptographic accelerators and hash functions—the core of AES—will still be functioning in the forthcoming with a critical size doubled, thanks to the algorithm of Grover’s that provides a brute force search quadratic speedup; however, all asymmetrical accelerators of cryptographic within HSMs will be broken entirely [11]. A new area of study called PQC has emerged in the past ten years in response to the fast development of quantum computers. PQC proposes algorithms for cryptography that are thought to resist assaults utilizing quantum computers [12, 13].

For automotive HSMs to remain safe in the “quantum era,” post-quantum secure algorithms will be required [14, 15]. Compared to modern cryptographic algorithms, these PQC techniques often use much larger key sizes and require significantly more time to complete key operations [16–18]. Furthermore, the mathematics necessary for PQC algorithms differs from the mathematics behind modern cryptographic methods [19–21]. Consequently, future HSMs will need rethought hardware architectures [22].

None of the existing general-purpose software has addressed automotive HSM use cases, and needs code signs for PQC schemes; moreover, all of these strategies concentrate on a single family of the schemes PQC [23, 24]. Various PQC candidates that have advanced to the third round of the NIST PQC certification process¹ have been thoroughly examined in this study about their potential use in ECU automotive designs. The study, provide post-quantum secure autonomous designs of HSM, focusing on hardware accelerators to speed up the suggested PQC algorithms [25–27]. This study makes the following contributions:

- Four choices for PQC are proposed for common automotive use cases.

- Candidate software profiling data aid hardware accelerator selection and design.
- Performance findings for post-quantum secure automotive HSM architecture generated on an Artix-7 FPGA are outlined.

2 Preliminaries

First, this section should serve as an introduction to the various PQC algorithm classes. Secondly, to illustrate the typical applications of contemporary HSMs, provide a comprehensive case study based on a real-world automobile situation.

2.1 Post-quantum Cryptography

Cryptography based on hashes, codes, lattices, multivariate, and isogeny is the most common PQC algorithm. Modern computers and quantum systems need help solving each of the courses' underlying mathematical problems. Key and message sizes, efficiency, confidence in security analysis, etc., vary between systems.

2.1.1 Cryptography by Code

The first proposal for code-based encryption was the McEliece cryptosystem, and its implementation utilizing binary Goppa codes is still considered safe today. The McEliece cryptosystem's huge public key is a downside. Even with the Niederreiter encryption system, which compresses the public key, the 128-bit post-quantum level of security public key surpasses 1 MB. Minimizing key sizes by code structure incorporation, such as quasi-cyclic codes, has been the subject of some studies.

2.1.2 Cryptography Based on Harsh

The hash-based signature security systems depends entirely on the well-understood features of the underpinning hash function, making them relatively mature. As part of the development of post-quantum cryptography, NIST is considering standardizing XMSS and LMS, two prominent techniques for stateful hash-based signatures. Given this, hash-based signatures seem like a good bet for safe signatures after quantum computing.

2.1.3 Cryptography Based on Lattice

One of the most well-liked PQC families is lattice-based cryptography. Numerous high-dimensional lattice-defined complex problem types, such as “shortest vector problem (SVP)” and “learning with errors (LWE)”, form the basis of its security. Much like code-based programs, lattice-based schemes built on ideal (organized) lattices often have higher performance and more minor keys, whereas generic lattice-based schemes tend to inspire greater trust. However, as the security of lattice-based methods against quantum-computer attacks is still poorly understood, selecting appropriate security settings has proven to be a formidable challenge.

2.1.4 Multi-variant Encryption

The multivariate system quadratic solutions over a limited field is a hard NP problem; multivariate cryptography is built on this. Despite much research on multivariate cryptosystem security, quickly and safely building such a scheme remains a formidable challenge. Several plans are vulnerable within the last ten years. Today, just a small fraction of those who relied on signature methods are still safe.

2.1.5 Cryptography Using Isogeny

Among the many potential PQC alternatives, isogeny-based cryptography is the most recent, having originated as an encryption system. A high-degree super singular isogeny involving two elliptic curves is notoriously difficult to detect, and this difficulty forms the basis of an isogeny-based system. One benefit is that schemes based on isogeny may borrow some of the mathematics used by standard ECC systems. When contrasted with structured lattice-based systems, the efficiency of candidates based on isogeny is noticeably lower. Also, more faith in these systems still needs to be built up since isogeny-based challenges are new.

2.2 SOTA Case Study

The software-over-the-air (SOTA) update allows car manufacturers to update their cars with the latest software upgrades by downloading them remotely. Additionally, it serves as a great illustration of how many fundamental security use cases interact at the ECU level. So, they utilise SOTA to find out what an HSM needs regarding security.

Executed inside an HSM, all SOTA updates and SecOC are services that rely on security. Therefore, digital signatures, hash functions, and symmetric cryptography are the building blocks of an HSM. Typically, key encapsulation (KEM) and public key encryption (PKC) are not implemented on the ECU level in modern HSMs. Nevertheless, crucial exchanges between electronic central units (ECUs) when the vehicle is in operation may necessitate the inclusion of KEM schemes in subsequent HSM generations. As a result, we also consider KEM schemes when constructing future HSMs.

3 Automotive HSMS PQC

3.1 PQC Schemes to Consider

The NIST PQC standardizing process has advanced ten proposals to the third round. Pick three PQC schemes from the list of ten that aim for security levels 3 medium of NIST and 5 high, respectively, to meet the needs of various automotive applications while balancing performance and security. Out of the NIST PQC competition, choose a digital signature technique that is well-known and understood in addition to these three options.

All four of the selected schemes satisfy the stringent running and memory storage criteria laid forth in Section II-B. The variety of the theoretically challenging issues underlying these four PQC schemes was another crucial consideration in our selection

process, alongside critical (or communication) sizes and performance measures. Our proposal for automobile HSMs included three schemes: one based on code, two on lattice, and one on hash. Additionally, one of the most essential criteria for our decision is the ease of software and hardware implementations.

3.1.1 Medium-Security Schemes

An ideal-lattice-based signature technique is suitable for applications that need a medium degree of security. Small key sizes and high efficiency led to the selection of CRYSTALS-Dilithium and code-based KEM system BIKE. The MLWE issue, on which Dilithium is established, in a version of a well-studied problematic. In addition to its high performance, Dilithium's operations are simple to implement in software and hardware. The QC-MDPC (Quasi-Cyclic Moderate Density Parity-Check) codes are the basis of BIKE, a code-based KEM. Its well-rounded performance makes it an attractive option for widespread usage. NIST deems BIKE a potential code-based option. Compared to HQC, BIKE has reduced the public key and cipher text sizes and better bandwidth metrics in the 3rd round. BIKE is a better contender for automobile use cases because of its efficiencies and key sizes. Dilithium and BIKE complement each other for applications with severe performance and key size constraints but do not have significant security needs.

3.1.2 High-Security Schemes

The high-security XMSS stateful hash-based signature method and the ideal-lattice-based CRYSTALS-Kyber KEM system and their well-studied mathematical problems were examined. NIST is now considering two stateful hash-based signature schemes for early standardization, whereas the Internet Engineering Task Force (IETF) established a standard for XMSS. The security of XMSS relies on efficiently implementing the underpinning hash function, which must withstand powerful quantum computers. Since Kyber's security is based on the same issue as Dilithium's, it has also been the subject of much research. Key operations for XMSS and Kyber, which feature short keys, can be efficiently accelerated using specialized hardware accelerators. As a result of their convincing security analyses, XMSS and Kyber may be integrated for usage in applications aiming for a high degree of protection.

3.2 Results of Software Profiling

Since the mathematical issues underpinning post-quantum cryptography (PQC) techniques are often quite different from present cryptography, a fresh class of hardware accelerators must be identified before a post-quantum safe HSM can be designed. As part of the hardware-software process for co-design, accelerators of hardware in HSMs are created to speed up the maximum computationally demanding activities. They assessed the software reference implementations provided in their submissions to the third level of the PQC NIST competition using the performance evaluation tool Gprof to identify which operations in the suggested schemes are the most compute-intensive. While all processes (e.g., key encapsulation, key generation, and critical DE encapsulation) are

comprised in the BIKE and Kyber profile KEM schemes (see Fig. 1), only the signature verification operation has been included for signing schemes XMSS and Dilithium.

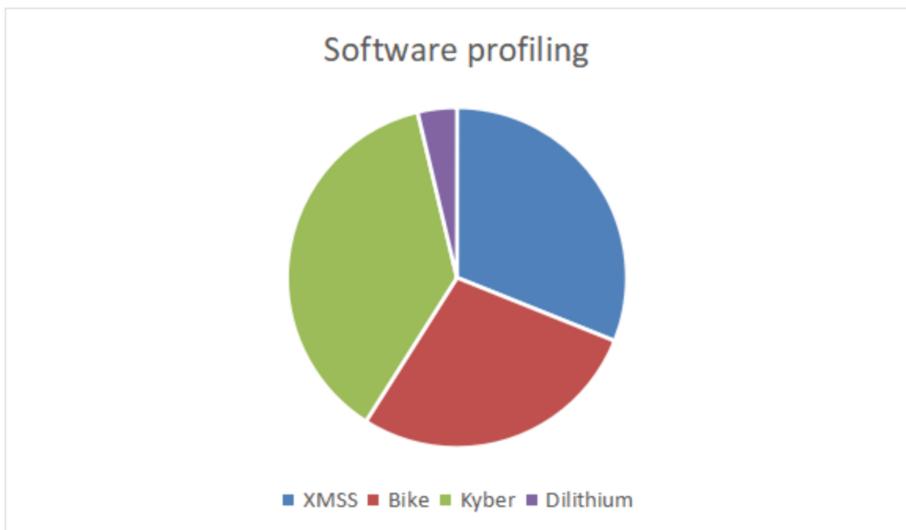


Fig. 1. Software Proofing

The following are the show results:

3.2.1 XMSS

The fundamental hash function in XMSS takes up most of the verification procedure time. They may use the SHA2-512 or SHAKE-256 hash algorithm for high security, as stated in the XMSS IETF protocol.

3.2.2 BIKE

In BIKE, all matrix operations are seen as polynomial operations due to using quasi-cyclic coding. According to the profiling data, BIKE's most costly calculation is polynomial multiplication. On top of that, the hash function SHA-384 and AES-256 are often utilized in software implementations; they are both instantiated during BIKE's construction.

3.2.3 Kyber and Dilithium

Since the foundations of Dilithium and Kyber are identical, the profile findings for the two procedures are also quite comparable. One utilizes SHAKE to extend the randomness and sample the random numbers, while the other uses AES-256. All of them have multiple implementation versions. In each of these implementations, the SHAKE/AES calculation and the subsequent NTT-based polynomial multiplications are the time-consuming parts.

4 PQC HSM Hardware Architectures

Designing an HSM with automotive applications in mind is the central topic of this section. The layout structure is based on the HSM description provided by the EVITA project 4. Figure 2 shows the standard design of an HSM similar to EVITA. This analysis may be used the same way to EVITA light and EVITA moderate as they are subsections of an EVITA complete HSM. Full EVITA HSMs will be referred to as modern HSMs going forward.

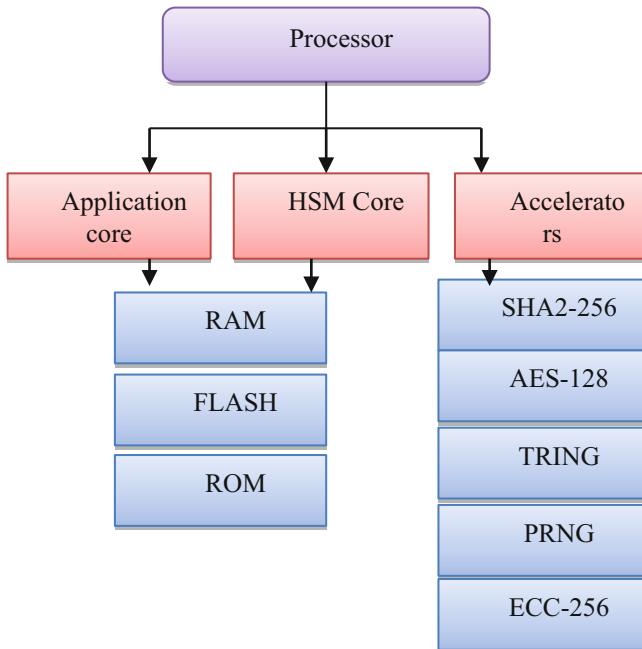


Fig. 2. Standard Design of an HSM Similar To EVITA

4.1 The Latest HSMs

The latest HSMs as graphically shown in Fig. 3 includes SHA2-256, AES-128, TRNG and PRNG etc.

- SHA2-256 is a hash algorithm used for general purposes instead of WHIRLPOOL.
- AES-128 for symmetric cryptography, including: key creation, encryption, and decryption are explained. SecOC must needed this block.
- TRNG is a real number generator for random. Hardware variations, such as jitter between digital ring oscillators in a TRNG, determine the electronic entropy.
- (AES-128, enc), a pseudo-random number generator. TRNGs seed PRNGs and employ an AES engine to increase randomness, intended for encryption exclusively.

- In an asymmetric cryptosystem, ECC-256 is utilized for 256-bit elliptical curve arithmetic.

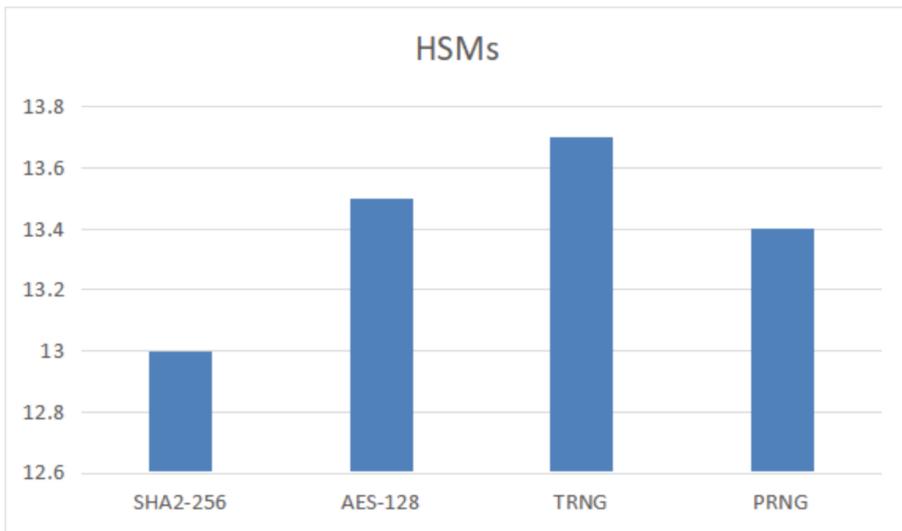


Fig. 3. Latest HSMs

The availability of large-scale quantum computers will render such an HSM insecure: While Grover's method halves the security of SHA2-256, Shor's technique ultimately compromises the security of ECC-256. It is recommended to replace ECC 256 with post-quantum secure asymmetric primitives and to utilize symmetrical primitives and hash functions with doubled security to make sure that an automobile HSM remains secure even when faced with quantum computers.

4.2 HSMs PQC

For high and medium security levels, post-quantum secure HSMs (PQC HSMs) use functions of hash and symmetrical primitives of cryptographic that can withstand quantum computers with a 128-bit security level, allowing for more conservative approaches. Significantly more extensive modifications are required for symmetric cryptographic primitives. In contrast to current HSMs, PQC HSMs need whole new cryptographic building pieces. Figure 4 both provide high-level and medium-level summaries of the HSM solutions that we suggested. Recommendations for cryptographic blocks for a PQC HSM are:

- SHA3-512, purpose a general function of hash.
- NIST's SHA-3 standard includes SHA3-512, which differs from SHA-2 in structure. SHA-3, a subclass of Keccak, employs sponge construction as its inner structure, whereas SHA-2 uses MD5-like structure.

- Symmetric encrypted communication (AES-256, all). The objective is the similar as (AES-128, altogether), but with double security. AES-256, enc, for construction of PRNG.

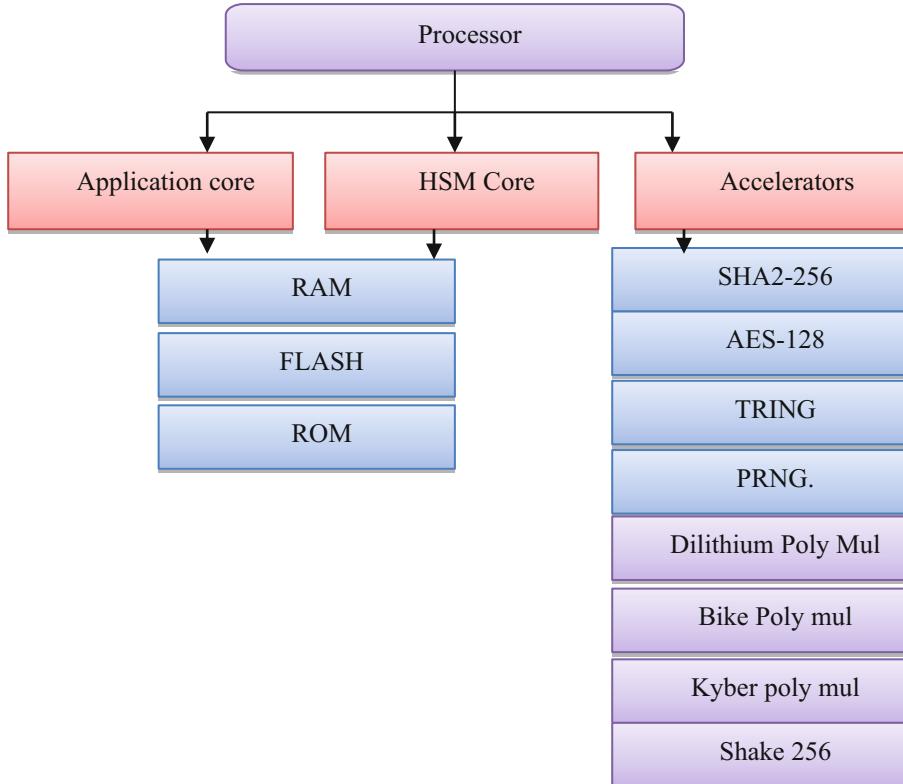


Fig. 4. High-Level and Medium-Level Summaries of the HSM Solutions

4.2.1 Medium-Security PQC HSMs

Dilithium should replace ECC-256 as the signature scheme for a medium-level PQC HSM. To further enable operations for KEM, which might be required for specific applications in upcoming HSMs automotive, they also suggested BIKE. Two of the most computationally costly operations in Dilithium are random expansion and polynomial multiplication. Hence, to speed up Dilithium, you should use the following cryptographic blocks:

- Using number theoretic transform (NTT), the Dilithium Poly Mul polynomials multiplier handles security settings and multiplication operations.
- Dilithium's AES-based implementation requires an AES-256 module to increase unpredictability. PRNG and Dilithium may share an AES-256 core. The following

module is developed for BIKE because polynomials multiplying is the most expensive operation:

- BIKE-Poly-Mul accelerates density and sparse polynomial operations in BIKE.

4.2.2 High-Security PQC HSMs

For high-security XMSS and PQC HSMs, is suggested as the autograph system, substituting ECC 256, and Kyber as the KEM system. A core for cryptographic is also required:

- A specialized NTT-based polynomial multiplier, much like Dilithium, is required to accelerate Kyber-Poly-Mul. Remember that an additional set of polynomial multiplies is required for Kyber and Dilithium, respectively, because the security parameters for the two schemes are distinct.
- It is selected because it is possible to build XMSS and Kyber with a high degree of security using SHAKE-256 because of this, signatures and KEM schemes inside an HSM may share resources.

The following section provides details on the hardware acceleration for these cryptographic blocks.

5 Comparison and Evaluation

This section delves into the synthesis and performance findings of the double suggested PQC HSMs, one with a degree of security for medium and the other with a high level of security. For the Artix-7 FPGA platform, the designs are synthesized. In addition, the synthesis outcomes of the PQC HSMs are contrasted with those of a conventional current HSM design.

5.1 Evaluation

Standard blocks such as ECC-256, AES-128/256, SHA3-512, SHA2-256, and TRNG9 are selected for their relatively high performance and reliance on open-source designs; these accelerators are aimed for resource-constrained embedded devices. Examples of algorithms that rely on repeated block calculations are AES-128/256 and SHA2-256. The chosen TRNG uses the jitter between many digital ring oscillators to generate its digital entropy. This randomness may be put into PRNG again to increase the unpredictability. A Murax system on a chip (SoC) represents an HSM's internal CPU. This open-source RISC-V processor can keep up with the performance of an ARM Coretex-M3 while using a fraction of the system's resources. Achieving high time-area efficiency was the objective when choosing all of these cores.

As a foundational element for lattice-based systems, SHAKE-256 is a parametric design. The SHAKE module, designed to be space-efficient, is the basis of this design. One way to accomplish a trade-off between area and performance is by dynamically choosing the total number of perpendicular slices inside the SHAKE core during synthesis. Sixteen slices are used in tandem in our endeavours.

The modern hardware application of BIKE includes the BIKE Poly Mul, a parametric as well as time constant polynomial multiplier. A performance parameter may be used during synthesis to adjust the size of the calculation data block, allowing the user to accomplish a trade-off between area and time, liable on the presentation.

The parametric and pipelined NTT-based polynomial multiplier design, used to accelerate lattice-based schemes, assesses both Dilithium-Poly-Mul and Kyber-Poly-Mul. Use Dilithium's medium security and Kyber's strong security as the synthesis criteria.

5.2 Modern vs. PQC HSM Synthesis Results

Space is used for three different HSM configurations: modern, PQC, and PQC, with high-security levels, and medium and respectively. At a medium security level, PQC HSM uses the same amount of space as current HSM; at a security level of high, however, a minor above is added in terms of area consumption; for example, when it comes to LUT usage, there is an overhead of around 13%. Since the DSP-intensive memory ECC-256 asymmetrical crypto core is removed from the design of HSM, both PQC HSM configurations are noticeably lighter, as seen from the memory and DSP utilization chart as depicted in Fig. 5.

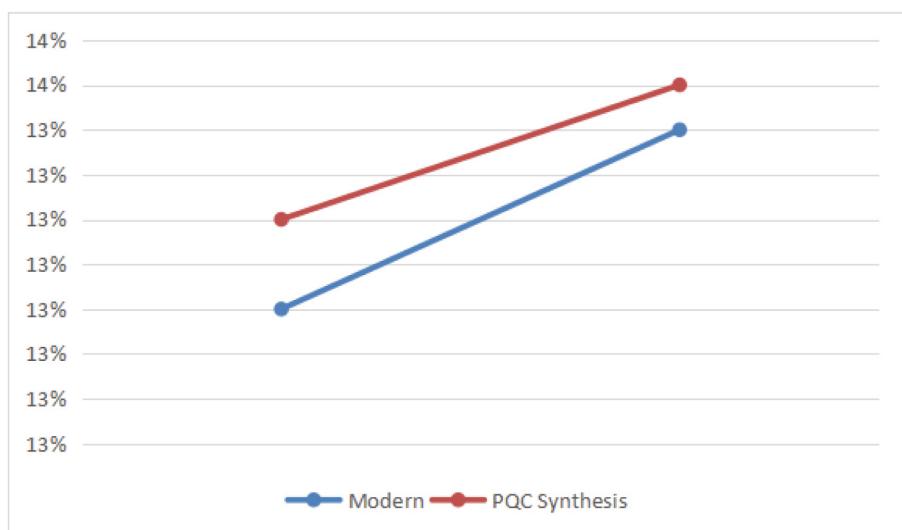


Fig. 5. Modern vs. PQC HSM Synthesis

The study should have included the additional space required by the usual components of an embedded system, such as the logic for communicating with the interface, controlling memory, and the logic at the highest level of control. Additionally, it should be noted that various optimizations may be used to fine-tune the act results for an actual implementation of HSM because these cryptographic cores' geographical metrics are implementation-specific.

To protect common automotive use cases against large-scale quantum computers, the study aims to be the first to investigate the possibility of building a time-area efficient HSM. The results show that medium- and high-security PQC HSM hardware designs are doable without requiring much extra space.

6 Conclusion and Further Work

Developing a post-quantum security HSM for applications in the automobile sector was the primary focus of our investigation in this work. After reviewing the everyday use cases of an automobile HSM, proposed signature and KEM methods that considered the performance needs and memory limitations and of these usage cases. They further profile the standard computerized applications of these systems and suggest novel hardware accelerators for medium and high-security level PQC HSMs, depending on the profiling results. Using post-quantum algorithms to guarantee their security, automobile HSMs may stay future-proof without considerable space overhead in developing cryptographic hardware accelerators, according to assessment findings of the PQC HSMs compared with contemporary HSMs. In this study, they investigated the potential for a hardware-software co design to create a post-quantum safe HSM with little space overhead. The performance measure is also significant for study when learning about genuine implementation based on hardware for a HSM post-quantum security. The effectiveness and area metrics of post-quantum security HSMs might be further studied using a prototype running on actual hardware.

References

1. Wang, W., Stöttinger, M.: Post-quantum secure architectures for automotive hardware secure modules. *Cryptology ePrint Archive* (2020)
2. Malina, L., et al.: Post-quantum era privacy protection for intelligent infrastructures. *IEEE Access* **9**, 36038–36077 (2021)
3. Sepúlveda, J., et al.: Post-quantum cryptography in mpsoc environments. In: *FIP/IEEE 29th International Conference on Very Large Scale Integration (VLSI-SoC)*. IEEE (2021)
4. Lohachab, A., Lohachab, A., Jangra, A.: A comprehensive survey of prominent cryptographic aspects for securing communication in post-quantum IoT networks. *Internet of Things* **9**, 100174 (2020)
5. Campos, F., et al.: Post-quantum cryptography for ECU security use cases. Ruhr-Universität Bochum (2019)
6. Bernstein, D.J., Lange, T.: Post-quantum cryptography. *Nature* **549**(7671), 188–194 (2017)
7. Liu, Z., Choo, K.-K.R., Grossschadl, J.: Securing edge devices in the post-quantum internet of things using lattice-based cryptography. *IEEE Commun. Mag.* **56**(2), 158–162 (2018)
8. Dam, D.-T., et al.: A survey of post-quantum cryptography: start of a new race. *Cryptography* **7**(3), 40 (2023)
9. Fernandez-Carames, T.M., Fraga-Lamas, P.: Towards post-quantum blockchain: a review on blockchain cryptography resistant to quantum computing attacks. *IEEE Access* **8**, 21091–21116 (2020)
10. Paul, S., Scheible, P., Wiemer, F.: Towards post-quantum security for cyber-physical systems: Integrating PQC into industrial M2M communication 1. *J. Comput. Secur.* **30**(4), 623–653 (2022)

11. Fritzmann, T., et al.: Towards reliable and secure post-quantum co-processors based on RISC-V. 2019 Design, Automation & Test in Europe Conference & Exhibition (DATE). IEEE (2019)
12. Taiber, J.: Unsettled topics concerning the impact of quantum technologies on automotive cybersecurity. No. EPR2020026. SAE Technical Paper (2020)
13. Saarinen, M.-J.O.: Mobile energy requirements of the upcoming NIST post-quantum cryptography standards. In: 8th IEEE International Conference on Mobile Cloud Computing, Services, and Engineering (MobileCloud). IEEE (2020)
14. Fritzmann, T., Vith, J., Sepúlveda, J.: Strengthening post-quantum security for automotive systems. In: 23rd Euromicro Conference on Digital System Design (DSD). IEEE (2020)
15. Asif, R.: Post-quantum cryptosystems for Internet-of-Things: a survey on lattice-based algorithms. *IoT* **2**(1), 71–91 (2021)
16. Barreto, P.S.L.M., et al.: qscms: Post-quantum certificate provisioning process for v2x. Cryptology ePrint Archive (2018)
17. Joseph, D., et al.: Transitioning organizations to post-quantum cryptography. *Nature* **605**, 7909, 237–243 (2022)
18. Althobaiti, O.S., Dohler, M.: Cybersecurity challenges associated with the Internet of Things in a post-quantum world. *IEEE Access* **8**, 157356–157381 (2020)
19. Bos, J.W., Dima, A., Kiening, A., Renes, J.: Post-Quantum Secure Over-the-Air Update of Automotive Systems
20. Chowdhury, S., Sesharao, Y., Abilmazhinov, Y.: IoT based solar energy monitoring system (2021)
21. Kumar, V.B.Y., et al.: Post-quantum secure boot. In: 2020 Design, Automation & Test in Europe Conference & Exhibition (DATE). IEEE (2020)
22. Alagic, G., et al.: Status report on the third round of the NIST post-quantum cryptography standardization process. US Department of Commerce, NIST (2022)
23. Labrado, C., Thapliyal, H.: Hardware security primitives for vehicles. *IEEE Consumer Electr. Mag.* **8**(6), 99–103 (2019)
24. Heyse, S.: Post quantum cryptography: implementing alternative public key schemes on embedded devices. For the degree of Doktor-Ingenieur of the Faculty of Electrical Engineering and Information Technology at the Ruhr-University Bochum, Germany (2013)
25. Singh, C., Rao, M.S.S., Mahaboobjohn, Y.M., Kotaiah, B., Kumar, T.R. (2022). Applied Machine Tool Data Condition to Predictive Smart Maintenance by Using Artificial Intelligence. In: Balas, V.E., Sinha, G.R., Agarwal, B., Sharma, T.K., Dadheech, P., Mahrishi, M. (eds) Emerging Technologies in Computer Engineering: Cognitive Computing and Intelligent IoT. ICETCE 2022. Communications in Computer and Information Science, vol. 1591. Springer, Cham. https://doi.org/10.1007/978-3-031-07012-9_49
26. Dang, V.B., et al.: Implementing and benchmarking three lattice-based post-quantum cryptography algorithms using software/hardware codesign. In: 2019 International Conference on Field-Programmable Technology (ICFPT). IEEE (2019)
27. Chowdhury, M., Islam, M., Khan, Z.: Security of connected and automated vehicles. arXiv preprint arXiv: 2012.13464 (2020)



Secure Data Management Using BlockChain

Akshay Tupetewar, Shivprasad Chaudhari, Shashikant Deshmukh^(✉),
and Sonali. V. Shinkar

Department of Electronics and Telecommunication Engineering, SCTR's Pune Institute of
Computer Technology, Pune, India
shashikantdeshmukh736@gmail.com

Abstract. Blockchain innovation, characterized by its consecutive, connected square structure, sets up a decentralized record framework, guaranteeing straightforwardness and security in recording exchanges. This paper investigates the application of blockchain within the instructive space, particularly inside Understudy Data Frameworks (SIS). By leveraging a decentralized, agreeable blockchain arrange, all intelligent among members are solidified inside a single record, advertising a dependable store for crucial instructive information. In conclusion, this paper traces determined challenges and promising bearings within the combination of blockchains and databases for information administration. It envisions an progressively imperative part for combination frameworks in future information administration assignments. This overview is balanced to advise both scholarly and mechanical communities, upgrading their understanding of the capabilities and imperatives of blockchain-related information administration frameworks. It moreover advocates for the advancement of combination frameworks custom fitted to a differing cluster of viable needs.

Keywords: Blockchain Integration · Student Data Security · Student Data Security · Decentralized Records · Blockchain-based Student Records · Education Data Privacy

1 Introduction

Blockchain technology offers a paradigm shift in securing data integrity by establishing decentralized and well managed data repositories. Its potential implementation within the automated management systems of individual universities or groups of educational institutions is a transformative approach in the education sector. Education, being a realm where various stakeholders constantly exchange and transform shared knowledge, stands to benefit significantly from blockchain technology. One of the key advantages is the ability to edit records at different security levels, catering to various access needs within the educational ecosystem. The impetus behind incorporating blockchain into Student Information Systems (SIS) is driven by the ever-growing necessity for high security and trust within critical systems. The fundamental principle of a decentralized, reliable, and highly secure ledger system for storing vital information resonates profoundly within

educational frameworks. Despite the allure of this technology, the overarching need for reliability and data integrity continues to escalate. A compelling aspect is that students gain complete independence regarding their personal data through blockchain, fostering autonomy independent of the institution while upholding the immutable nature of data. Blockchain's architecture involves permanently storing encrypted data in decentralized blocks. The cryptographic processes employed to generate and link these blocks significantly enhance the security of each transaction within the blockchain. This results in immutable records that cannot be altered once created, thereby ensuring security, flexibility, and irreversibility. The promise of blockchain in education extends far beyond mere data security. It grants students the ability to add a layer of anonymity to their personal data, ensuring independence from educational institutions, securing official documents and certificates' immutability, and instilling complete confidence in the accuracy and infallibility of information, courtesy of the network's design.

The proposed models emphasize the accessibility of information, empowering students to retrieve their data whenever needed. Blockchain technology presents an alternative avenue for education, sparking interest among various institutions, organizations, and companies to investigate its advantages in the educational realm. In the realm of education, student information is both critical and sensitive. Within traditional centralized educational systems, vulnerabilities in managing and extracting information from the general management framework, learning, and research persist. This article aims to explore alternative accounting methodologies through blockchain technology, providing a more secure and trusted archive of records. The proposed approach using blockchain for a fully functional SIS meticulously maintains students' records, course registrations, and marks. It offers a robust, secure, and transparent method to establish a global educational learning system. The potential benefits of blockchain technology in education span across various dimensions, ensuring not only data security but also revolutionizing the ways in which educational records are managed, accessed, and secured for the benefit of all stakeholders involved. This innovative approach not only mitigates vulnerabilities in the current system but also lays the foundation for a more transparent, secure, and accountable educational framework. Components of blockchain: A blockchain is a linked list of blocks that contains sets data. Each block has a cryptographic hash in its transaction list and another hash of the previous block. Transaction hashes are stored in a cryptographically protected data structure called a Merkle tree. Below is the blockchain terminology:

- Node:** A computer system linked to the blockchain, possessing either a full or partial copy of the blockchain and fulfilling various roles within the system. Nodes can be categorized as full nodes, which contain the complete blockchain, or light nodes, which hold a partial list of blocks. The presence of more nodes enhances the decentralization of the system.
- Blockchain:** A decentralized network of nodes maintaining a connected series of blocks, constituting a distributed ledger copied across all nodes based on their types. Each full node retains a valid copy of the entire ledger.
- Block:** A compilation of records containing a sequence of events, accompanied by hash values reflecting the current and previous block states. The genesis block serves as the initial block in a blockchain, lacking an upper limit.
- Event List:** Tailored to the specific domain, representing sets of data to be transferred. In Student Information Systems (SIS), events encompass data pertaining to students or institutional members, such as test scores for

courses. Status: Domain-specific, indicating aggregate data such as totals or amounts derived from transaction lists. The block's state must alter with each addition of an event. Merkle Tree: A binary tree structure wherein event identifiers are paired, hashed, and the resultant hashes concatenated until reaching the root of the tree. Consensus: A protocol governing the acceptance of newly generated blocks among blockchain nodes without centralized authority. Consensus mechanisms are crucial for appending new blocks to the blockchain.

2 Literature Review

In recent years, the incorporation of blockchain technology into numerous sectors has attracted considerable interest, and the education sector is no different. The decentralized, immutable, and secure nature of blockchain presents a promising solution for managing sensitive student data. Scholars have delved into the potential applications of blockchain in the educational context, with a particular focus on enhancing security measures. Svetinovic et al. (2018) argue that blockchain's cryptographic techniques and consensus mechanisms make it exceptionally resistant to data tampering or unauthorized access. This is crucial in safeguarding academic achievements and personal information, ensuring they remain confidential and unaltered. Furthermore, the issue of data security is of paramount concern in educational institutions. With the increasing prevalence of cyber threats and data breaches, safeguarding student information has become a top priority. The decentralized ledger of blockchain removes the necessity for a central authority, thereby mitigating the risk associated with a single point of failure. As noted by Hernandez et al. (2019), this decentralization ensures that student data is not stored in a vulnerable, centralized database that could be a target for malicious attacks. Instead, data is distributed across a network of nodes, making it significantly more challenging for unauthorized parties to compromise the system.

Additionally, blockchain's use of cryptographic hashes and digital signatures provides an extra layer of security. Each block in the chain contains a unique cryptographic hash of the previous block, creating a chain of blocks that are interlinked and secure. Any attempt to alter the data in a block would necessitate recalculating the hash of that block and all subsequent blocks, a computationally infeasible task. This cryptographic integrity check deters malicious actors from attempting to manipulate student records, offering a level of security that surpasses traditional centralized databases. Despite these advantages, it is imperative to acknowledge that blockchain technology is not entirely immune to security vulnerabilities. While the underlying blockchain protocol is robust, the implementation and design of specific systems can introduce potential points of failure. For instance, improper configuration of smart contracts, which are self-executing contracts with terms written directly into code, can lead to vulnerabilities. Furthermore, the human element in the deployment and maintenance of blockchain systems introduces a potential vector for security breaches. Therefore, it is crucial for educational institutions to adopt best practices in blockchain implementation and security protocols. In protocols. A foundation of the proposed models is the accentuation on information accessibility, enabling undergraduates with unlimited access to their records at any given time. This paper presents three fastidiously outlined models for executing completely

useful SIS through blockchain innovation. These models include the secure administration of understudy and staff records, course enrollment information, and scholarly execution measurements. Here, honest to goodness certificates can be consistently issued and spread to interested parties without dependence on centralized organization. In expansion to its application in instruction, the success of blockchain within the domain of cryptocurrencies underscores its potential in information administration. This study digs into the combination of blockchain and conventional databases, a developing slant within the database community. The classification of existing blockchain-related information administration advances, based on their positions along the blockchain-database range, lays the establishment for a comprehensive examination. Three particular sorts of combination frameworks are presented, each subjected to a nitty gritty investigation of plan contemplations and trade-offs. By scrutinizing normal frameworks and methods inside each combination framework, this study gives broad experiences into the qualities and confinements of each demonstrate.

Conclusion, the application of blockchain technology in student data management offers a robust solution to enhance security measures. The decentralized and immutable nature of blockchain mitigates the risks associated with centralized databases, providing a more secure environment for sensitive student information. The use of cryptographic techniques and distributed ledger technology adds an extra layer of protection, making it exceedingly difficult for malicious actors to compromise the integrity of student records. While not without its potential vulnerabilities, blockchain technology, when implemented and managed correctly, represents a significant step towards fortifying data security in educational institutions.

3 Existing Techniques

Proposed Model [6] Designed for sharing, verifying, and learning achievements, the innovative infrastructure of Blockchain technology proves highly suitable for this purpose. This paper introduces a university-level Blockchain model aimed at implementing a fully operational student management system (refer to Fig. 1). The focus lies specifically on student grades, given their pivotal role in the educational domain. The model serves as a functional prototype intended to explore the potential of Blockchain in this context. It operates as a ledgerbased system, capturing and managing all university activities using Blockchain principles. Each block within the Blockchain can contain student records and their corresponding hashes, utilizing roles for data retrieval. The decentralized nature of Blockchain ensures that any alteration to one block renders the entire Blockchain invalid. Each block comprises elements such as a nonce, header, hash code of the preceding block, hash value, and timestamp, with variations depending on the application. The model illustrates data and interactions among all participants involved. It encompasses three main functions: student registration, professor registration, and course enrollment. Both students and professors are required to register for courses. Upon student registration, a new Blockchain address is generated, with the administration assigning a student ID and forwarding this information to the student. Student data, including ID, registered courses, and subject information, are stored in a centralized database by the administration. Professor registration follows a similar process to that of students.

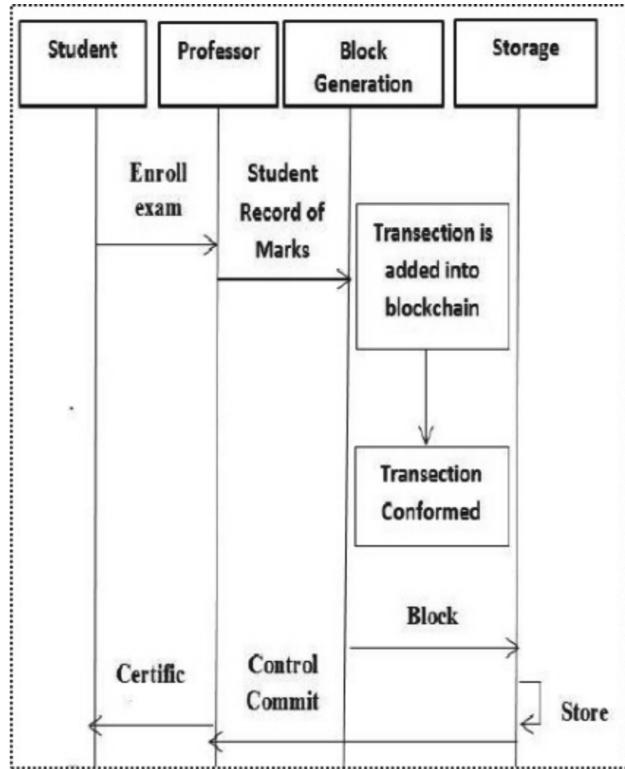


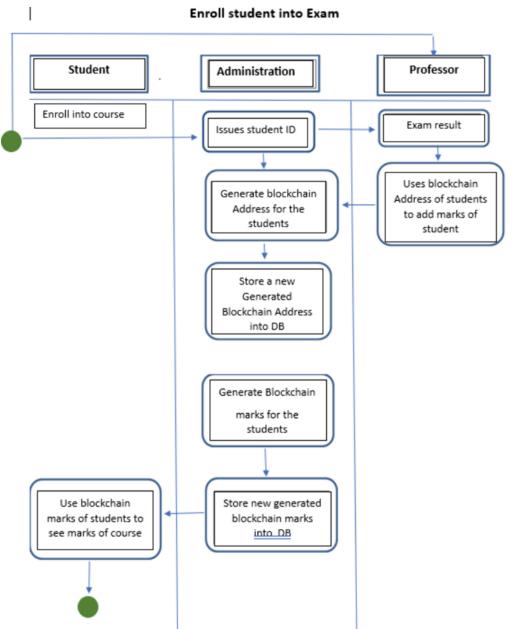
Fig. 1. Diagram for storing blocks of marks in Blockchain

In first model Professors and administrators are parties to the blockchain, when a student enrolls in a course, the first block or “Genesis Block” is generated, and all parties have access to the new student after enrollment. The model stores student course registration data, their grades (provisional grade, course grade, final exam) in one blockchain for each semester. Generally, the relevant institution organizes a midterm exam and a final exam due to the evaluation process. After registering for the exam, professors [6] enter the student’s grade into the created blockchain. The university control committee receives a report on grades and exam results every semester, which is repeated every semester. In the next phase, the respective authorities will release the marks of all the students. This result determines the status of the students room for copies of certificates, grades and each student receives a unique code printed on the certificate. Student system described involves recording academic information, such as grades and course notes, on a blockchain for transparency and security. Faculty can add course-related notes to the system, which are securely stored on the blockchain. They can also classify student subjects through transactions on the blockchain. Students can view their grades, which are recorded on the blockchain, and receive notifications for updates to their academic records. After exams, professors verify results and update student grade information on the blockchain. Transactions for updating grades are confirmed on the blockchain,

ensuring data integrity. The ledger-based system provides a transparent and immutable record of academic achievements. Unique course symbols are assigned to each course in each semester to reduce plagiarism and ensure accurate data recording.

The number of events depends on the number of courses and students for students to be assessed. All transaction data is stored in the blockchain model. No items can be changed or deleted. In this second model, when a student enrolls in the university, the administration assigns them a unique student ID and generates a new blockchain address for them. This blockchain address serves as a secure repository for storing crucial student information, including their name, phone number, address, identification card, and social security number. Additionally, a separate blockchain, referred to as the “Marks blockchain,” is established to record student course registration data. After a student completes an exam, the professor responsible for the course submits the results directly to the student’s blockchain address. Simultaneously, a token representing the completion of the exam is stored in the central database, providing an additional layer of verification. Students have the privilege of accessing their personal Marks blockchain, which displays their performance in various subjects. This transparent view allows them to track their academic progress effectively. Furthermore, as part of the control committee’s process, the student’s ID is integrated into the model as a prerequisite for taking the final semester exam. This ensures that only eligible students are able to participate in the examination process. Overall, this model leverages blockchain technology to create a secure and transparent system for managing student information, course registrations, and exam results. By utilizing blockchain addresses and tokens, the model establishes a robust framework that enhances data security and integrity while providing students with direct access to their academic records.

In third model records every transaction and stores student information and, of course, registration information in one blockchain. All events are recorded on the blockchain, such as exams, courses, lecturers and grades after the student registration process, all events are added to some blockchain. All posted updates and additions will be posted to everyone involved. In this scenario, the student registration mechanism controls the schedule of courses offered by the university. A list of blockchain transactions generates a token list document. When students register for courses, their data is sent to the control committee, creating a student profile that includes information about all enrolled courses. This information is accessible to everyone involved, and each student has a personal event unit containing ID, courses, intermediate grades, course grades, and final exam results. All transmitted data is stored on the blockchain, ensuring its immutability. The administration downloads and stores student records on the blockchain network, where transactions are confirmed and cannot be altered. University staff enter student details, and a unique hash is retrieved from the database and added to the blockchain. The university administration also stores information in its centralized database, including student ID, user ID, and registered courses. Each student has a unique username and password to access and manage their account. By employing blockchain technology in university management systems, the handling of student events is streamlined. Each approach has its strengths and weaknesses, contributing to the overall efficacy of the system while minimizing plagiarism concerns; let’s see it in the table (Figs. 2, 3).

**Fig. 2.** Storing blocks of marks in blockchain [6]

Features	Model1	Model2	Model3
Network Size	One blockchain supply demand network	Multi-blockchains supply demand network	One blockchain supply demand network
Performance	Improving the model cycle time decrease order updates	More confirmation time because more number of block	Storage data in one record as one data Changing the block continuously in real time
Security	Flexible and secure because of their storage data is limited so data is more reliable	Verified each transaction data is reliable	Storage capacity resilient ,flexible and secure resource sharing on a public

Fig. 3. Table 1: Comparing Modes

The three blockchain models presented offer innovative approaches to revolutionizing university management systems, each with its own set of strengths and potential drawbacks. Blockchain technology holds immense promise for revolutionizing education, particularly in the management of student data within SIS. While the road to implementation requires diligence and careful consideration, the benefits, including enhanced data control, security, and immutability, outweigh the challenges. Blockchain has the potential to reshape how educational institutions manage student information, ushering in an era of increased trust, transparency, and security in education. The choice among these models should be driven by the unique needs and priorities of the higher education institution. Model 1 is ideal for those institutions seeking transparency and real-time access, but careful scalability planning is essential. Model 2 suits institutions valuing data security and efficient record-keeping, though it may entail centralization concerns. Model 3 is the model of choice for universities seeking comprehensive, transparent, and automated record-keeping, despite potential privacy and scalability challenges. Ultimately, the integration of blockchain technology offers an exciting opportunity for universities to modernize their administrative processes, enhance data security, and better serve students and stakeholders in the digital age. The selection of the most suitable model should align with the institution's vision, scale, and commitment to privacy and transparency the integration of blockchain into mainstream applications, fostering a more robust and sustainable technological ecosystem (Table 1).

Table 1. Comparison of research reviewed

Sr. No.	Title of Work	Year	Type of Chain	Features
1	EduCTX: a blockchain-based higher education credit platform [1]	2018	Public	Secure, transparent credit recognition in education sector
2	Application of Blockchain-based Technology in Chemistry Education Students' Data Management [2]	2018	Private	Ensures security and integrity of student information
3	Security and privacy in decentralized energy trading through multi-signatures, blockchain and anonymous messaging streams [3]	2015	Hybrid	Enhances trustworthiness of the system
4	A Reliable Storage Partitioning for Permissioned Blockchain [4]	2020	Consortium	Improves performance, efficiency while maintaining security

(continued)

Table 1. (*continued*)

Sr. No.	Title of Work	Year	Type of Chain	Features
5	Blockchain Architecture to Higher Education Systems [5]	2019	Public	Enhances management and operation of educational institutions
6	DEFYING THE CERTIFICATION DIPLOMA FORGERY WITH BLOCKCHAIN PLATFORM [6]	2019	Hybrid	Ensures authenticity and integrity of academic credentials
7	Decentralized access control for IoT data using blockchain and trusted oracles [7]	2019	Consortium	Enhances security and reliability of data access in IoT
8	Usurping double-ending fraud in real estate transactions via blockchain technology [8]	2017	Consortium	Establishes a secure and transparent system to safeguard against fraudulent practices
9	A survey of blockchain security issues and challenges [9]	2017	Public	Provides valuable insights into securing blockchain-based systems
10	A Novel Blockchain-based Education Records Verification Solution [10]	2019	Public	Enhances the trustworthiness of educational certificates and credentials
11	Blockchain in Space Industry - Challenges and Solutions [11]	2014	Public	Enhances various aspects of the space sector
12	Blockchain: a secure, decentralized, trusted cyber infrastructure solution for future energy systems [12]	2018	Hybrid	Highlights potential benefits and trustworthiness in energy infrastructure
13	Blockchain Technology as an enabler of service systems: a structured literature review [13]	2017	Public	Insights into facilitating development and operation of service-oriented systems
14	BFT-Store: Storage Partition for Permissioned Blockchain via Erasure Coding [14]	2020	Consortium	Aims to enhance efficiency and reliability of blockchain systems

(continued)

Table 1. (*continued*)

Sr. No.	Title of Work	Year	Type of Chain	Features
15	ACE: asynchronous and concurrent execution of complex smart contracts [15]	2020	Private	Addresses challenges in executing intricate smart contracts
16	The Security of Student Information Management System based upon Blockchain [16]	2022	Private	Strengthens integrity and confidentiality of student data
17	Blockchain in industries: A survey [17]	2019	Public	Insights into adoption and leveraging in industry-specific challenges
18	A Blockchain-based framework for secure Educational Credentials [18]	2018	Public	Aims to enhance trustworthiness of educational certificates
19	Blockstack: A Global Naming and Storage System Secured by Blockchains [19]	2016	Public	Presents decentralized approach to naming and storage for enhanced security and trust
20	Blockchain Based Student Information Management System [20]	2021	Private	Designed to enhance security and integrity of student records in education
21	Hawk: The blockchain model of cryptography and privacy-preserving smart contracts [21]	2016	Consortium	Addresses privacy and security challenges in smart contract execution, novel approach to enhance these aspects
22	A Blockchain-Based Decentralized Data Storage and Access Framework for PingER [11]	2018	Hybrid	Aims to enhance data integrity and accessibility for monitoring and analyzing the global Internet in PingER project
23	Blockchain: a secure, decentralized, trusted cyber infrastructure solution for future energy systems [22]	2018	Public	Highlights potential benefits and its contribution to the trustworthiness of energy infrastructure
24	Enabling blockchain innovations with pegged sidechains [23]	2014	Public	Provides an online resource for further exploration of this concept
25	SoK: Research Perspectives and Challenges for Bitcoin and Cryptocurrencies [24]	2015	Public	Offers a comprehensive overview of key issues in the domain, valuable resource for researchers and practitioners

4 Conclusion

The surveyed papers collectively demonstrate the diverse applications of blockchain technology across various domains. These include secure higher education credit platforms, student data management, decentralized energy trading, permissioned blockchain storage solutions, and more. Additionally, the papers explore critical issues such as security, privacy, and fraud prevention. Notably, blockchain is leveraged to fortify data integrity and accessibility, offering transparent credit recognition and protecting against certification forgery. The proposed frameworks exhibit potential to revolutionize sectors like education, energy, and real estate, emphasizing trustworthiness and efficiency. Nevertheless, to further enhance blockchain's impact, future research could focus on scalability, interoperability, and regulatory frameworks to ensure seamless integration into existing systems. While the surveyed papers offer valuable insights, there is room for improvement in several areas. Firstly, more empirical studies and real-world implementations could provide concrete evidence of blockchain's effectiveness in practical scenarios. Additionally, addressing scalability challenges and exploring interoperability solutions would extend the technology's applicability. Moreover, future research could delve into the legal and regulatory aspects of blockchain adoption, ensuring compliance with existing frameworks.

Acknowledgements. We extend our heartfelt gratitude to Prof. Sonali Shinkar for their invaluable support throughout our project. Their guidance not only enriched our project but also contributed significantly to our personal growth as students. We are dedicated significantly to our personal growth as students. We are dedicated to applying the knowledge and teachings they have imparted to us in our future endeavors, with the ultimate aim of making them proud.

References

1. Turkovic, M., Holbl, M., Kosic, K., Hericko, M., Kamisalic, A.: EduCTX: a blockchain-based higher education credit platform. *IEEE Access*. **6**, 5112–5127 (2018). <https://doi.org/10.1109/access.2018.2789929>
2. Ezeudu, F.O., Eya, N.M., Nworgi, H.I.: Application of blockchainbased technology in chemistry education students' data management. *Int. J. Database Theory Appl.* **11**(2), 11–22 (2018)
3. Aitzhan, N.Z., Svetinovic, D.: Security and privacy in decentralized energy trading through multi-signatures, blockchain and anonymous messaging streams. *IEEE Trans. Depend. Sec. Comput.*, to be published. [Online]. Available: <http://ieeexplore.ieee.org/document/7589035/>
4. Qi, X., Zhang, Z., Jin, C., Zhou, A.: Areliable storage partition for permissioned blockchain. In: *IEEE Transactions on Knowledge and Data Engineering*, vol. 33, no. 1, pp. 14–27, 2021. [Online]. Available: <https://doi.org/10.1109/tkde.2020.3012668>. Palanivel, Blockchain architecture to higher education systems. *Int. J. Latest Technol. Eng. Manag. Appl. Sci.*, vol. VIII, no. II, pp. 124–138, (2019). itemrefurl1
5. Tolbatov, V.A.: Using Blockchain Technology for E-Learning. *Vimryuvalna ta Obchysliuvalna Tekhnika v Tekhnologichnykh Protsesakh* **61**(1), 110–113 (2018)
6. Chandra, Y.U., Surjandy, S., Fernando, E., Prabowo, H.: Defying the certification diploma forgery with blockchain platform: a proposed model (2019). <https://doi.org/10.33965/ict2019>

7. Al Breiki, H., Al Qassem, L., Salah, K., Habib Ur Rehman, M., Sevtinovic, D.: Decentralized access control for IoT data using blockchain and trusted oracles. In: Proc. IEEE Int. Conf. Ind. Internet Cloud, ICII 2019, pp. 248–257, (2019). <https://doi.org/10.1109/ICII.2019.00051>
8. Mashatan, A., Lemieux, V., Lee, S.H.M., Szufel, P., Roberts, Z.: Usurping doubleending fraud in real estate transactions via blockchain technology. *J. Database Manag.* **32**(1), 27–48 (2021). <https://doi.org/10.4018/JDM.2021010102>
9. Lin, I.-C., Liao, T.-C.: A survey of blockchain security issues and challenges. *Int. J. Network Security* **19**(5), 653–659 (2017). [https://doi.org/10.6633/IJNS.201709.19\(5\).01](https://doi.org/10.6633/IJNS.201709.19(5).01)
10. Torky, M., Gaber, T., Hassani, A.E.: Blockchain in space industry challenges and solutions. arXiv Prepr.arXiv2002.12878.2014
11. Dong, Z., Luo, F., Liang, G.: Blockchain: a secure, decentralized, trusted cyber infrastructure solution for future energy systems. *J. Modern Power Syst. Clean Energy* **6**(5), 958–967 (2018). <https://doi.org/10.1007/s40565-018-0418-0>
12. Seebacher, Schuritz, R.: Blockchain technology as an enabler of service systems: a structured literature review, pp. 12–23. Exploring Services Science, Cham (2017)
13. Qi, X., Zhang, Z., Jin, C., Zhou, A.: BFT-Store: Storage Partition for Permissioned Blockchain via Erasure Coding. In: IEEE 36th International Conference on Data Engineering (ICDE), pp. 1926–1929 (2020) [Online]. Available: <http://arxiv.org/abs/2002.12878>
14. Wust, K., Matetic, S., Egli, S., Kostiainen, K., Capkun, S.: ACE: asynchronous and concurrent execution of complex smart contracts. In: Proceedings of the 2020 ACM SIGSAC Conference on Computer and Communications Security, Virtual Event USA, pp. 587–600, Oct. (2020). <https://doi.org/10.1145/3372297.3417243>
15. Yang, M., Wang, J.: The security of student information management system based upon blockchain. *J. Electr. Comput. Eng.* (2022) [Online]. Available: <https://www.hindawi.com/journals/jece/2022/6746027/>
16. Al-Jaroodi, J., Mohamed, N.: Blockchain in industries: a survey. *IEEE Access* **7**, 36500–36515 (2019)
17. Shadab, G., Abdullah, H., Abdulhaq, R., Hussen, A.H.: A Blockchain-based framework for secure Educational Credentials
18. Ali, M., Nelson, J.C., Shea, R., Freedman, M.J.: Blockstack: a global naming and storage system secured by blockchains. In: 2016 USENIX Annual Technical Conference, USENIX ATC 2016, Denver, CO, USA, June 22–24, pp. 181–194 (2016)
19. Sudha, V., Kalaiselvi, R., Sathy, D.: Blockchain based student information management system. In: 2021 International Conference on Advancements in Electrical, Electronics, Communication, Computing and Automation (ICAEC), pp. 1–4 (2021). <https://doi.org/10.1109/ICAEC52838.2021.9675515>
20. Kosba, A., Miller, A., Shi, E., Wen, Z., Papamanthou, C.: Hawk: the blockchain model of cryptography and privacy-preserving smart contracts. In: Security and Privacy (SP), IEEE Symposium on, pp. 839–858 (2016)
21. S. Ali, G. Wang, B. White, and R. L. Cottrell, "A Blockchain-Based Decentralized Data Storage and Access Framework for PingER," in 17th IEEE International Conference On Trust, Security And Privacy In Computing And Communications / 12th IEEE International Conference On Big Data Science And Engineering, TrustCom/BigDataSE 2018, New York, NY, USA, August 1–3, 2018, pp. 1303–1308
22. Back, A., et al.: Enabling blockchain innovations with pegged sidechains. URL: <http://www.opensciencereview.com/papers/123/enablingblockchaininnovations-with-pegged-sidechains> (2014)
23. Bonneau, J., et al.: SoK: research perspectives and challenges for bitcoin and cryptocurrencies. In: 2015 IEEE Symposium on Security and Privacy, SP 2015, pp. 104–121, San Jose, CA, USA, 17–21 May (2015)

24. Dai, M., et al.: A low storage room requirement framework for distributed ledger in blockchain. *IEEE Access* **6**, 22970–22975 (2018)
25. Gorenfo, C., Golab, L., Keshav, S.: XOX fabric: a hybrid approach to blockchain transaction execution. In: IEEE International Conference 6 on Blockchain and Cryptocurrency (ICBC), pp. 1–9, IEEE, Toronto, ON, Canada (2020)
26. Hong, A., Sun, C., Chen, M.: A survey of distributed database systems based on blockchain. In: 2020 3rd International Conference on Smart BlockChain (SmartBlock), pp. 191–196, IEEE, Zhengzhou, China (2020)
27. Liu, J., Liu, Z.: A survey on security verification of blockchain smart contracts. *IEEE Access* **7**, 77894–77904 (2019)
28. Chang lin, I., Chen, Y.H.: Blockchain based smart contract for bidding system. In: ICASI 2018, IEEE, pp. 1–5 (2018)
29. Yu, G., et al.: Survey: sharding in blockchains. *IEEE Access* **1**, 1–1 (2020)
30. Gueta, G., et al.: SBFT: a scalable and decentralized trust infrastructure. In: 49th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN), pp. 568–580 (2019)
31. Zheng, Z., et al.: An overview of blockchain technology: architecture, consensus, and future trends. In: Proceedings – IEEE 6th International Congress on Big Data, BigData Congress 2017, pp. 557–564 (2017)
32. Meyliana, H., Hidayanto, A.N., Budiardjo, E.K.: Evaluation of social media channel preference for student engagement improvement in universities using entropy and TOPSIS method. *J. Ind. Eng. Manag.* **8**(5), 1676–1697
33. Prieto, J., Pinto, A., Das, A.K., Ferretti, S.: A brief review of database solutions used within blockchain platforms. *Blockchain Appl.* **1238**, 121–130 (2020)
34. Karafiloski, E., Mishev, A.: Blockchain solutions for big data challenges: a literature review. In: 17th IEEE International Conference on Smart Technologies, EUROCON 2017 – Conference Proceedings, pp. 763–768 (2017)
35. Carlos Marinho, S.S., Filho, J.S.C., Moreira, L.O., Machado, J.C.: Using a hybrid approach to data management in relational database and blockchain: a case study on the e-health domain. In: International Conference on Software Architecture Companion (ICSAC) (2020)
36. Lian, J., Wang, S., Xie, Y.: TDRB: An efficient tamper-proof detection middleware for relational database based on blockchain technology. *IEEE Access* (2021)
37. Hao, K., Xin, J., Wang, Z., Cao, K., Wang, G.: Blockchain-based outsourced storage schema in untrusted environment. *IEEE Access* **7**, 122707–122721 (2019)



Navigating Through Digital Realm: Role of Cyberpsychology in Fostering Mental Well-Being and Digital Empathy

Anadi Trikha¹(✉) , Antima Sharma¹ , Arpita Agarwal¹ , Preeti Nagar¹ , and Ritu Singh²

¹ Manipal University Jaipur, Jaipur, Rajasthan, India
anadi.trikha@jaipur.manipal.edu

² Manipal Academy of Higher Education, Dubai, UAE

Abstract. The article examines various essential themes related to the impact of technology on our mental health and well-being. It provides strategies to protect our mental health from digital stressors, encourages us to practice digital empathy for positive online interactions, and emphasises ethics in the digital realm. The article draws insights from cognitive psychology, social psychology, and human-computer interaction to understand better the psychological dimensions of today's technologically mediated world. Additionally, the article highlights the importance of cyberpsychology in designing and implementing effective digital interventions, educational programs, and policies to foster a positive and psychologically informed digital culture. By focusing on ethical considerations, the article contributes to the ongoing discussions on responsible digital citizenship and the ethical use of emerging technologies. In conclusion, the article advocates for the central role of cyberpsychology in navigating the complexities of the contemporary digital landscape. Cyberpsychology can help shape a more resilient, empathetic, and ethically conscious digital society by combining psychological theories with digital realities. This paper significantly contributes by addressing the multifaceted relationship between technology and mental health, offering insights and strategies to mitigate potential negative impacts and enhance well-being in the digital age.

Keywords: Cybersecurity · Cyberpsychology · Digital World

1 Introduction

Cyberpsychology, also known as internet psychology, web psychology, or digital psychology, is an interdisciplinary field that confidently explores the psychological effects of human interaction with digital technology, with a particular emphasis on the internet. It encompasses various aspects of human behaviour in the digital domain, including online identity, relationships formed online, computer and internet addiction, and the psychological outcomes associated with cyborgs, artificial intelligence, and virtual reality. Furthermore, the discipline extends to researching cybersecurity and its impact on daily life [1].

As technological advancements continue to reshape human interaction, cyberpsychology confidently emerges as a critical domain for understanding the psychological implications of our digital experiences. This confident study delves into the multifaceted importance of cyberpsychology in modern society, examining its role in elucidating the psychological effects of digital technology, safeguarding online well-being, and confidently navigating the challenges posed by the ever-evolving technological landscape [2].

1.1 Understanding the Psychological Effects of Digital Technology

Cyberpsychology explores the psychological effects of human interaction with digital technology, particularly the internet. It examines how online behaviour, virtual relationships, and the addictive nature of technology shape individuals' psychological states. Social media platforms have transformed how people perceive themselves and others, shedding light on the potential discrepancies between digital and offline personas. Cyberpsychology helps individuals and mental health professionals understand the challenges posed by the digital realm and develop strategies to promote a healthier online existence [3].

1.2 Exploring the Intricacies of Online Relationships

The internet has revolutionised how individuals connect, forming relationships without distance constraints. Cyberpsychology delves into the intricate dynamics of these digital relationships, from friendships to romantic connections, to better understand their emotional and social impact. This understanding is crucial for fostering healthy online communities. Furthermore, cyberpsychology addresses the pressing issue of cyberbullying, which is amplified by the prevalence of online interactions. The field not only seeks to comprehend the psychological effects of cyberbullying but also aims to develop interventions and preventions to mitigate its negative consequences. Cyberpsychology contributes to creating a supportive and secure digital environment by exploring the complexities of online relationships [4].

1.3 Addressing Addiction to Computers and the Internet

With the advent of the internet, people have been able to connect in previously unimaginable ways, leading to the formation of digital relationships with unique dynamics. Cyberpsychology is a field that aims to understand better these relationships, including the challenges they present, to help foster healthy and supportive online communities. One of the critical issues that cyberpsychology addresses is cyberbullying, which can have a devastating impact on individuals and communities. By studying these issues, cyberpsychology aims to create a constructive and positive online environment that is free from the negative consequences of cyberbullying.

1.4 Navigating the Psychological Implications of Advanced Technologies

The study discusses how cyberpsychology extends its focus to the psychological implications of rapidly advancing technologies such as cyborgs, AI, and VR. Integrating these technologies into daily life raises profound questions about identity, agency, and the boundaries between the human and digital realms. Cyberpsychology explores the psychological responses to interacting with AI entities, the impact of virtual environments on cognition and perception, and the potential psychological challenges associated with augmenting human capabilities through cybernetic enhancements. The field contributes to the ethical and responsible development and deployment of advanced technologies, ensuring that psychological well-being remains a central consideration in the quest for technological innovation [5].

1.5 Encompassing the Study of Cybersecurity

Cyberpsychology broadens its focus to encompass the study of cybersecurity and its effects on daily life. It spans from analysing the psychological impact of cyber threats, exploring individuals' responses to online security breaches, and considering the role of cybersecurity awareness in shaping online behaviour. By taking a holistic perspective, cyberpsychology is instrumental in developing comprehensive strategies that protect individuals from cyber threats and address the psychological consequences of living in an ever-evolving digital landscape.

2 Digital Stressors

Digital stressors, also known as techno stressors or digital stress factors, refer to the various challenges and pressures individuals may experience due to their interactions with digital technologies and the online environment. In today's highly interconnected world, where people rely heavily on digital tools and platforms, digital stressors have become increasingly prevalent and can significantly impact one's mental and emotional well-being. Here, we will explore several everyday digital stressors and elaborate on their implications [6].

2.1 Information Overload

When people are bombarded with overwhelming information, it can result in information overload. This can happen when they are exposed to excessive emails, social media updates, news articles, and notifications, which can exceed their ability to process effectively. This constant stream of information can cause cognitive fatigue, diminished productivity, and difficulty making decisions. It can be a daunting task for individuals to filter, prioritise, and absorb the information, leading to feelings of stress and being overwhelmed [7].

2.2 Constant Connectivity

The expectation or compulsion to constantly connect to digital devices and online platforms blurs the boundaries between work and personal life. Continuous connectivity can lead to burnout, as individuals may find it challenging to disconnect and recharge. The pressure to respond promptly to emails and messages, even outside working hours, can contribute to stress, anxiety, and a sense of being constantly “on call.”

2.3 Social Media Pressure

The social and psychological stressors associated with using social media platforms include the pressure to present an idealised version of one’s life, the fear of missing out (FOMO), and the comparison with others. Social media pressure can negatively impact self-esteem and mental health. Constant exposure to curated and often unrealistic depictions of others’ lives may lead to feelings of inadequacy, anxiety, and a distorted sense of reality [8].

2.4 Cyberbullying

Harassment, intimidation, or aggressive behaviour directed at an individual through digital channels, including social media, messaging apps, or online forums. Cyberbullying can have severe psychological consequences, causing stress, anxiety, and even depression. The anonymity afforded by online platforms may encourage individuals to engage in harmful behaviours, making it a significant digital stressor, particularly for adolescents and young adults [9].

2.5 Digital Fatigue

The mental exhaustion results from prolonged digital device use and online content engagement. Excessive screen time and constant exposure to digital stimuli can lead to digital fatigue, characterised by symptoms like eye strain, headaches, and mental exhaustion. This fatigue can contribute to decreased overall well-being and productivity [10].

2.6 Technological Uncertainty

The stress associated with the rapid pace of technological change, including concerns about job security, skill obsolescence, and adapting to new digital tools. The fear of falling behind in a technologically evolving world can contribute to anxiety and stress. Individuals may feel pressured to continuously update their skills and adapt to new technologies, creating insecurity and instability.

2.7 Online Harassment and Trolling

Persistent, offensive, or malicious behaviour directed at an individual online, often manifesting as harassment, trolling, or abuse. Online harassment can have severe psychological consequences, including anxiety, depression, and a sense of helplessness. The anonymity provided by the digital environment may encourage individuals to engage in harmful behaviour that can profoundly affect the mental well-being of the target [11].

3 Digital Interventions

Digital interventions are crucial in addressing the rising prevalence of digital stressors in today's interconnected world. As individuals navigate the complexities of online interactions and constant connectivity, digital stressors such as information overload, social media pressures, and online harassment have become pervasive. Digital interventions offer targeted solutions and support mechanisms to mitigate the adverse psychological effects of these stressors. Whether through mobile apps, online platforms, or virtual therapy sessions, these interventions provide individuals with tools to manage their digital well-being. They offer resources for stress reduction, digital detox strategies, and guidance on fostering positive online behaviours. By acknowledging and addressing digital stressors through tailored interventions, individuals can develop resilience, strike a healthier balance in their relationship with technology, and ultimately promote their mental and emotional well-being in the digital age [12].

Recognising and addressing digital stressors is crucial for promoting a healthy relationship with technology. Strategies may include:

3.1 Digital Detox

Digital detox refers to a deliberate and temporary disconnection from digital devices and online activities to reduce the negative impact of constant digital exposure on mental and physical well-being. This practice recognises the potential stressors associated with prolonged use of smartphones, computers, and other electronic devices and the impact of social media, notifications, and information overload. During a digital detox, individuals intentionally refrain from using digital devices, social media platforms, and other online activities, allowing themselves time and space for more mindful and offline activities. The goal is to break the cycle of constant connectivity, reduce screen time, and promote a healthier balance between the digital and physical aspects of life. Digital detoxes can range from a few hours to several days, allowing individuals to recharge, reconnect with the physical world, and alleviate stress associated with the modern digital lifestyle [13].

3.2 Setting Boundaries and Media Literacy

Setting boundaries and cultivating media literacy serve as essential digital interventions in managing the challenges the digital landscape poses. Establishing clear boundaries involves consciously defining limits on the use of digital devices and online activities. This practice helps individuals create a healthier balance between online and offline lives, mitigating information overload and constant connectivity. Additionally, media literacy equips individuals with the skills to analyse and interpret digital content critically. Understanding the persuasive techniques, potential biases, and the impact of online information fosters a discerning approach. By combining establishing boundaries with media literacy, individuals can navigate the digital realm more mindfully, making informed choices about their online engagement. These interventions empower individuals to take control of their digital experiences, reducing stressors, promoting well-being, and fostering a more conscious and intentional use of technology.

3.3 Online Safety Measures and Balancing Use of Technology

Implementing online safety measures and promoting a balanced technology use approach are integral interventions to mitigate the impact of digital stressors. Online safety measures involve adopting practices and tools to protect individuals from cyber threats, harassment, and privacy breaches. This includes using secure passwords, enabling two-factor authentication, and being cautious about sharing personal information online. These measures contribute to a sense of security, reducing anxiety associated with potential online risks.

Balancing technology use emphasises the importance of conscious and moderate engagement with digital devices. Encouraging individuals to set specific time limits, take breaks, and establish tech-free zones fosters a healthier relationship with technology. This intervention helps counteract issues like digital fatigue and constant connectivity. By promoting a mindful and intentional use of technology, individuals can prevent burnout, improve their overall well-being, and maintain a more harmonious balance between their digital and physical aspects. Online safety measures and balancing technology use interventions contribute to a more positive and secure digital experience, alleviating the stressors inherent in our increasingly digitalised world.

3.4 Promoting Positive Online Behavior

Promoting positive online behaviour is a vital digital intervention to address and alleviate the stressors in the online environment. By cultivating a culture of respect, empathy, and responsible communication, this approach aims to create a virtual space conducive to well-being. Through educational campaigns, mindfulness tools, and community support systems, individuals are encouraged to be mindful of their digital interactions and the potential impact on others. Recognising the interconnectedness of online and offline lives, positive reinforcement mechanisms and digital well-being tools are integrated into platforms to guide users towards healthier online habits. Emphasising digital citizenship education further equips individuals with the necessary skills to navigate the digital landscape responsibly. By fostering collaboration between various stakeholders, including mental health professionals, educators, and policymakers, this holistic approach transforms the online experience into a positive, supportive, and enriching space for all users [14].

4 Ethical Consideration of Emerging Technologies

The ethical consideration of emerging technologies in cyberpsychology is essential to ensure that the application of these technologies aligns with principles of respect, fairness, and responsibility. As technology advances, it brings new opportunities and challenges in cyberpsychology. Here are critical ethical considerations associated with the use of emerging technologies in this field [15]:

4.1 Privacy and Confidentiality

Emerging technologies often involve collecting and analysing sensitive psychological and behavioural data. The ethical concern arises in safeguarding individuals' privacy and ensuring the confidentiality of their personal information. Cyberpsychologists must prioritise protecting participants' privacy and confidentiality. Informed consent processes should communicate how data will be used and stored, and steps should be taken to de-identify information when possible.

4.2 Informed Consent in Virtual Environments

Virtual reality (VR) and augmented reality (AR) technologies may immerse individuals in realistic simulations. Ensuring informed consent becomes more complex in these immersive environments. Cyberpsychologists should provide comprehensive information about the nature of virtual experiences, potential risks, and the use of any data collected within these environments. Participants must have a clear understanding of what they are consenting to.

4.3 Bias in Algorithmic Decision-Making

The use of algorithms and artificial intelligence in cyberpsychology applications may introduce biases based on race, gender, or other demographic factors, impacting the fairness and accuracy of assessments. Developers and researchers must address bias in algorithmic models to avoid perpetuating or amplifying existing social inequalities. Transparent reporting of the algorithms' decision-making processes is crucial for accountability.

4.4 Digital Accessibility

Access to emerging technologies may not be equitable, leading to potential disparities in the benefits and risks experienced by different populations. Cyberpsychologists should strive to make digital interventions and assessments accessible to diverse populations. Efforts should be made to address barriers related to socioeconomic status, geography, and disabilities.

4.5 Therapeutic Boundaries in Virtual Therapy

Virtual therapy platforms, including chatbots and virtual therapists, challenge traditional therapeutic boundaries. Determining the appropriate level of human involvement and the potential risks of overreliance on technology is crucial. Cyberpsychologists must establish guidelines for virtual therapy tools, ensuring they complement rather than replace human interaction. Regular monitoring and assessment of the therapeutic relationship are essential.

4.6 Transparency in AI-Enhanced Assessments

Artificial intelligence in psychological assessments may involve complex algorithms that lack transparency, making it challenging for individuals to understand the basis of assessments. Cyberpsychologists should ensure transparency in developing and applying AI-enhanced assessments. Clear communication about the factors influencing assessments and their potential limitations is essential.

4.7 Accountability for Technology-Induced Harm

Emerging technologies may have unforeseen consequences or unintended side effects, potentially causing harm to individuals using cyberpsychological interventions. Cyberpsychologists must accept responsibility for the ethical use of technology and be prepared to address any harm that may arise. Continuous monitoring, feedback loops, and mechanisms for user reporting should be implemented.

5 Conclusion

This article discusses cyberpsychology's important role in understanding technology's impact on mental health and well-being. It provides strategies to protect our mental health from digital stressors, encourages us to practice digital empathy for positive online interactions, and emphasises ethics in the digital realm. The article draws insights from cognitive psychology, social psychology, and human-computer interaction to understand the psychological dimensions of today's technologically mediated world. While the paper offers practical recommendations for promoting digital empathy and ethical behaviour, it may not fully address the challenges and barriers to implementing these strategies in real-world settings. Factors such as technological limitations, institutional constraints, and individual resistance may hinder the effectiveness of interventions. The paper may rely on generalisations from studies conducted in specific contexts or with particular populations. As a result, the applicability of the findings and recommendations to diverse groups or cultural contexts may be limited. Additionally, the article highlights the importance of cyberpsychology in designing and implementing effective digital interventions, educational programs, and policies to foster a positive and psychologically informed digital culture. The article advocates for the central role of cyberpsychology in navigating the complexities of the contemporary digital landscape and shaping a more resilient, empathetic, and ethically conscious digital society.

References

1. Fortuna, P.: Positive cyberpsychology as a field of study of the well-being of people interacting with and via technology. *Front. Psychol.* **14**, 1053482 (2023). <https://doi.org/10.3389/fpsyg.2023.1053482>
2. New Jersey Institute of Technology: What is Cyberpsychology and Why is it Important? (n.d.). <https://www.njit.edu/admissions/blog-posts/what-cyberpsychology-and-why-it-important>. Last accessed on 1 Jan 2024

3. Harley, D., Morgan, J., Frith, H.: *Cyberpsychology as Everyday Digital Experience Across the Lifespan*. Springer (2018). Last accessed on 3 Jan 2024
4. Caponnetto, P., Milazzo, M.: Cyber health psychology: the use of new technologies at the service of psychological wellbeing and health empowerment. *Health Psychol. Res.* **7**(2), 8559 (2019). <https://doi.org/10.4081/hpr.2019.8559.PMID:31872148;PMCID:PMC6904845>. Last accessed on 2024/1/2
5. It, T.N.: Introduction to Cyberpsychology and Its Significance in the Digital Age 26 Sep. (2023). <https://www.linkedin.com/pulse/introduction-cyberpsychology-its-significance/>. Last accessed 1 Jan 2024
6. Quach, S., Thaichon, P., Martin, K.D., et al.: Digital technologies: tensions in privacy and data. *J. Acad. Mark. Sci.* **50**, 1299–1323 (2022). <https://doi.org/10.1007/s11747-022-00845-y> last accessed 2024/1/1
7. Information overload: Media Effect in the Age of Data – FasterCapital. (n.d.). FasterCapital. <https://fastercapital.com/content/Information-overload--Media-Effect-in-the-Age-of-Data.html>. Last accessed 1 Jan 2024
8. Jabeen, F., Tandon, A., Sithipolvanichgul, J., Srivastava, S., Dhir, A.: Social media-induced fear of missing out (FoMO) and social media fatigue: The role of narcissism, comparison and disclosure. *J. Bus. Res.* **159**, 113693 (2023). <https://doi.org/10.1016/j.jbusres.2023.113693>
9. Stevens, F., Nurse, J.R.C., Arief, B.: Cyber stalking, cyber harassment, and adult mental health: a systematic review. *Cyberpsychol. Behav. Soc. Netw.* **24**(6), 367–376 2021). <https://doi.org/10.1089/cyber.2020.0253>. Epub 2020 Nov 12. PMID: 33181026. Last accessed 1 Jan 2024
10. Jacobson, S.: Digital Fatigue: Is Your Screen Time Killing Your Wellbeing? Harley Therapy™ Blog, 22 Mar 2023. <https://www.harleytherapy.co.uk/counselling/digital-fatigue.htm>
11. Online harassment: meaning, types & impact – EAP-India. <https://www.eap-india.com/online-harassment-meaning-types-impact/>. Last accessed 1 Jan 2024
12. Mc Kinsey: Cybersecurity in a digital era. In: <https://www.mckinsey.com/>. Retrieved 1 June 2020, from <https://www.mckinsey.com/>. Last accessed 1 Jan 2024 (n.d.)
13. Coyne, P., Woodruff, S.J.: Taking a break: the effects of partaking in a two-week social media digital detox on problematic smartphone and social media use, and other health-related outcomes among young adults. *Behav. Sci.* **13**(12), 1004 (2023)
14. Promoting Digital Wellbeing by Encouraging Positive Online Behaviour. (1900, January 1). Bett Global 2022. <https://www.bettshow.com/bett-articles/promoting-digital-wellbeing-by-encouraging-positive-online-behaviour/>. Last accessed 1 Jan 2024
15. Burr, C., Taddeo, M., Floridi, L.: The ethics of digital well-being: a thematic review. *Sci. Eng. Ethics* **26**, 2313–2343 (2020). <https://doi.org/10.1007/s11948-020-00175-8> last accessed on 2024/1/1



An Empirical Analysis of Neighborhood-Based Approaches for Trustworthy Recommendations with Apache Mahout

Vijay Verma^(✉)

Computer Engineering Department, National Institute of Technology, Kurukshetra,
Haryana 136119, India
vermavijay@nitkkr.ac.in

Abstract. Recommender Systems (RSs) are essential tools for avoiding information overload on the World Wide Web (or simply the Web). Providing accurate and trustworthy recommendations will increase the confidence of the end-users in an RS and make them more satisfied and engaged. Neighbourhood-based methods are the most popular and widely used technique for recommendations. In order to provide more accurate recommendations, the effect of neighbourhood size is investigated in User-based Collaborative Filtering (UBCF) recommendations. In particular, this work examines an open-source and scalable collaborative recommendation framework, Apache Mahout. Firstly, brief functionalities of the Mahout library are discussed along with its API classes. Secondly, several offline experiments are conducted with various MovieLens datasets, including the latest one, MovieLens-25M, to analyze the functioning of UBCF methods as implemented in the Mahout. Finally, the accuracy of UBCF methods is measured in terms of two metrics: Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE). Empirically obtained results demonstrate that an RS can provide trustworthy recommendations to its end users by increasing the neighbourhood size in UBCF methods.

Keywords: Recommender Systems · Collaborative Filtering · Apache Mahout · Neighborhood-based Collaborative Recommendations

1 Introduction

Recommender Systems (RSs) are software tools that assist Internet users in avoiding information overload because of the large amount of information on the World Wide Web (or simply the Web) [1]. Many commercially successful websites, such as YouTube [2], Amazon [3], Netflix [4], etc., deploy RSs to facilitate their customers by suggesting useful items. RSs create a win-win scenario for both the service provider and the service consumer (end users). i.e., an RS may help to increase the revenue of the service provider by increasing sales, whereas the end users may find the interesting items from the recommendations.

Collaborative filtering-based recommendation approaches are the most popular methods among various other approaches (e.g., content-based approaches) due to several advantages [5]. These approaches can further be categorized into two broad categories: memory-based and model-based techniques [6]. In memory-based techniques [7], rating data in the form of a user-item matrix (UI-matrix) is utilized heuristically for recommending items. Moreover, model-based techniques prepare a model from UI-matrix for recommendation purposes [8]. In RSs literature, memory-based algorithms are also known as neighbourhood-based collaborative filtering due to their classical way of making recommendations based on k-nearest neighbours. Furthermore, neighbourhood-based approaches may be of two types: User-based Collaborative Filtering (UBCF) and Item-based Collaborative Filtering (IBCF). The type hierarchy [9] related to collaborative filtering-based recommender systems is shown in Fig. 1.

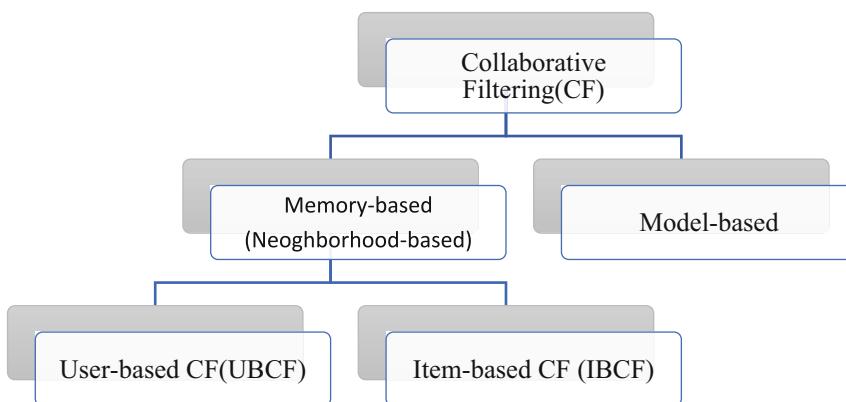


Fig. 1. A general classification of collaborative filtering-based recommendations.

In RS literature, there exist numerous libraries/frameworks to support the design and development of a wide variety of approaches for recommendations. Each framework supports the implementation of different recommendation approaches based on its underlying architecture. When a researcher proposes a new recommendation technique, then the proposed method should be compared against prior similar/benchmark methods in order to assess its effectiveness. Therefore, one has to carefully choose a suitable framework based on requirements. Each framework may support the various software engineering metrics (such as extensibility, scalability, etc.) to different degrees while implementing a recommendation approach in a particular domain. Further, there exist a few RS frameworks (such as Duine [10], Cofi [11], Easyrec [12], and PREA [13]) which are only available for experimentation purposes and are not in active development. However, there are a few libraries that are widely used for RS experimentation (based on the general bibliography) and are under active development, such as LibRec [14], Apache Mahout [15], LensKit [16], and Cf4J [17], etc.

This work examines a scalable and open-source collaborative recommendation framework, Apache Mahout [18], with respect to the traditional UBCF recommendation by using a very popular and benchmarked dataset of the movie domain, MovieLens [19].

Firstly, brief functionalities of the Apache Mahout library are discussed along with its API classes. Secondly, several offline experiments are conducted with various MovieLens datasets, including the latest one, MovieLens-25M, to observe the functioning of UBCF methods as implemented in the Mahout. Finally, the effectiveness of the recommendation process is measured in terms of Mean Absolute Error (MAE) and Root Mean Squared Error(RMSE) metrics.

2 Apache Mahout: A Collaborative Filtering Framework

Apache Mahout [18] is an open-source, scalable machine-learning library primarily focusing on clustering, regression, and recommender engines. It is written in Java and available freely from Apache Software Foundation under Apache License 2.0. Under the broad umbrella of recommender systems, it mainly aims at collaborative filtering approaches; therefore, it may also be called a collaborative filtering framework [20]. However, the framework doesn't explicitly provide the implementation of content-based approaches; still, such methods are facilitated through the existing APIs on top of it. Apart from the CF-based methods, it also brings the implementation of a few baseline methods for non-personalized recommendations, such as random recommendations, user/item averages, etc.

In collaborative filtering-based approaches, both memory-based and model-based methods are available. Since neighbourhood-based methods are traditional approaches (prominent because of their easiness and effectiveness) for collaborative recommendations, Mahout includes other memory-based methods, such as slope-one, clustering-based, etc., for comparison purposes. Furthermore, the detailed implementation of both UBCF and IBCF approaches, along with possible variations in their underlying components, are given for readymade usage. Figure 2 depicts UML diagrams of the core APIs involved in a UBCF recommendation.

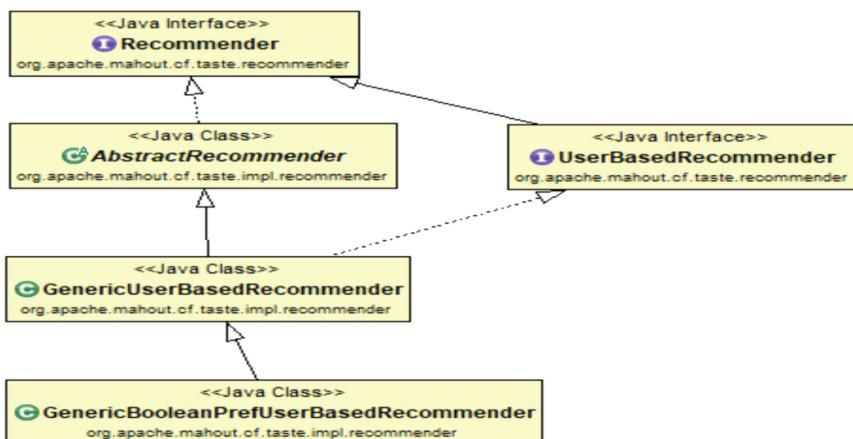


Fig. 2. The type hierarchy of the core APIs in Apache mahout for a UBCF approach.

In order to evaluate the effectiveness of recommendation approaches, Mahout also provides popular metrics such as MAE, RMSE, precision, recall, and F1. The type hierarchy of the API classes associated with these evaluation metrics is shown in Fig. 3.

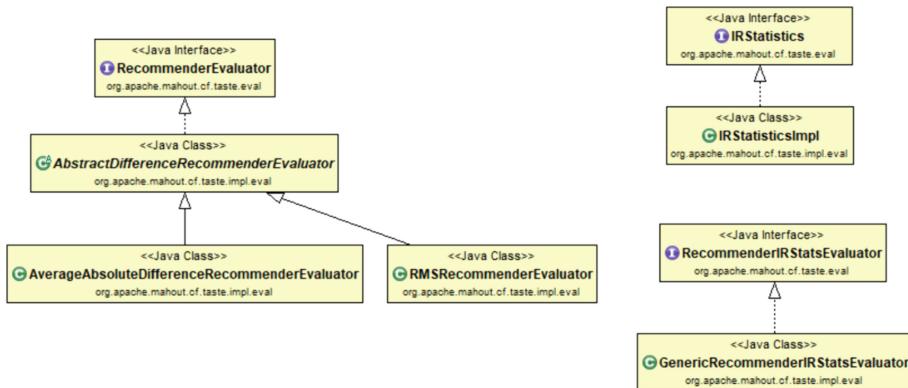


Fig. 3. The type hierarchy of the APIs associated with evaluation metrics in Apache mahout.

In order to understand the internals of the Mahout framework, Fig. 4 illustrates the relationship between the core components involved in a simple UBCF algorithm. However, different recommendation methods will exploit different components with diverse connections. In Fig. 4, the direction of an arrow indicates that the particular component is utilized by the component from where the arrow is originating. The basic functionalities of these components are summarized as follows:

- DataModel: it stores and provides access to all the rating data.
- UserSimilarity: it defines some notion of similarity between users.
- UserNeighborhood: it defines the most similar users to a given user.
- Recommender: it uses all the above three components in order to recommend items to users.

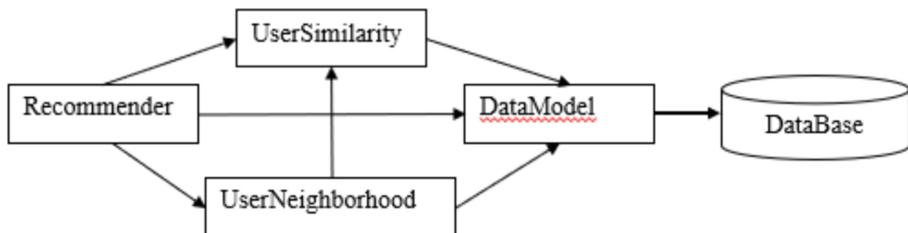


Fig. 4. The core components of a simple UBCF in the Apache Mahout.

3 Experiments

We have performed offline experiments to evaluate the recommendation process in Apache Mahout. In RSSs research, offline experiments are the prominent and preferred way to assess the quality of various recommendation methods. Further, these types of experiments do not require the involvement of real users in contrast to user studies or online experiments [21].

3.1 Datasets

GroupLens [22] is a research lab at the University of Minnesota [23] that works exclusively in the domain of recommender systems. It runs a non-commercial website, MovieLens [19], that recommends personalized movies to its users (a movie recommendation service). GroupLens Research has collected rating data from the MovieLens website over various periods of time and made these datasets publicly available for education, research, and industry. The historical perspective of the MovieLens datasets can be found in the article [24]. Among various MovieLens datasets, the latest one (MovieLens-25M dataset) was released in Dec'2019 and recommended for new research by GroupLens Lab. The MovieLens-25M dataset is a stable benchmark dataset that contains 25000095 ratings across 62423 movies from 162541 users between January 09, 1995 and November 21, 2019. Table 1 summarizes different MovieLens datasets briefly.

Table 1. The brief description of MovieLens datasets.

Dataset	Release Date	Brief Description
MovieLens-100K	04/1998	“943 users have provided 100,000 ratings on 1682 movies”
MovieLens-1M	02/2003	“6040 users have provided 1000,209 ratings on 3900 movies”
MovieLens-10M	01/2009	“71567 users have provided 10,000,054 ratings on 10681 movies”
MovieLens-20M	10/2016	“138493 users have provided 20,000,263 ratings on 27278 movies”
MovieLens-25M	12/2019	“162541 users have provided 25,000,095 ratings on 62423 movies”

3.2 Experimental Design

Here, the traditional user-based collaborative filtering (UBCF) algorithm with the k-Nearest Neighbors(kNN) is used to perform offline experiments in Apache Mahout with different MovieLens datasets. In particular, the Pearson Correlation Coefficient (PCC) is utilized as a similarity metric; thus, for any two users u and v, PCC(u,v) is defined as

follows:

$$PCC(u, v) = \frac{\sum_{i \in I_{uv}} (r_{ui} - \bar{r}_u)(r_{vi} - \bar{r}_v)}{\sqrt{\sum_{i \in I_{uv}} (r_{ui} - \bar{r}_u)^2 \sum_{i \in I_{uv}} (r_{vi} - \bar{r}_v)^2}} \quad (1)$$

where r_{ui} is the rating value given by a user $u \in U$ to a particular item i and I_{uv} is the set of items rated by both the users u and v , i.e. $(I_u \cap I_v)$. Figure 5 depicts the practical implementation of the traditional UBCF algorithm in the Mahout library. In all the experiments, we divided the dataset in the ratio of 80:20 for training and testing purposes, respectively. Since the size of the datasets is very large, except MovieLens-1M, the evaluation is performed on 10% of users from all the users in the dataset.

Algorithm: User-based Collaborative Filtering

Input: UI-Matrix, a target user u

Output: top-N recommendations

1. start
2. for every other user w
3. calculate similarity between target user u and w
4. end for
5. sort all these users with respect to the similarity weight.
6. select the neighborhood (n) of the target user u by keeping users with high similarity values.
7. for every item i rated by some user in n but that is not rated by the target user u yet
8. for every other user v in n who have rated item i
9. calculate the similarity between u and v
10. include v 's rating weighted by similarity, into running average
11. end for
12. end for
13. return the top items, ranked by weighted average.
14. end

Fig. 5. An implementation of UBCF in Apache Mahout.

3.3 Evaluation Metrics

One of the most popular measures to assess the quality of a recommender system is its accuracy [25]. If an RS provides accurate recommendations, then only it will be useful to the end users. Therefore, we have also examined the accuracy of recommendations with the help of the following two metrics: Mean Absolute Error (MAE) and Root Mean

Squared Error(RMSE).

$$MAE = \frac{1}{|T|} \sum_{(u,i) \in T} |\hat{r}_{ui} - r_{ui}| \quad (2)$$

$$RMSE = \sqrt{\frac{1}{|T|} \sum_{(u,i) \in T} (\hat{r}_{ui} - r_{ui})^2} \quad (3)$$

3.4 Results and Discussion

For each dataset, we have measured the accuracy of the recommendation in terms of MAE and RMSE. These evaluation metrics are measured against the varying sizes of neighbourhoods in the UBCF recommendation. Figure 6 shows MAE and RMSE values for different MovieLens(ML) datasets (a) ML-1M (b) ML-10M (c) ML-20M (d) ML-25M. From these empirical values, it is clear that if we increase the neighbourhood size,

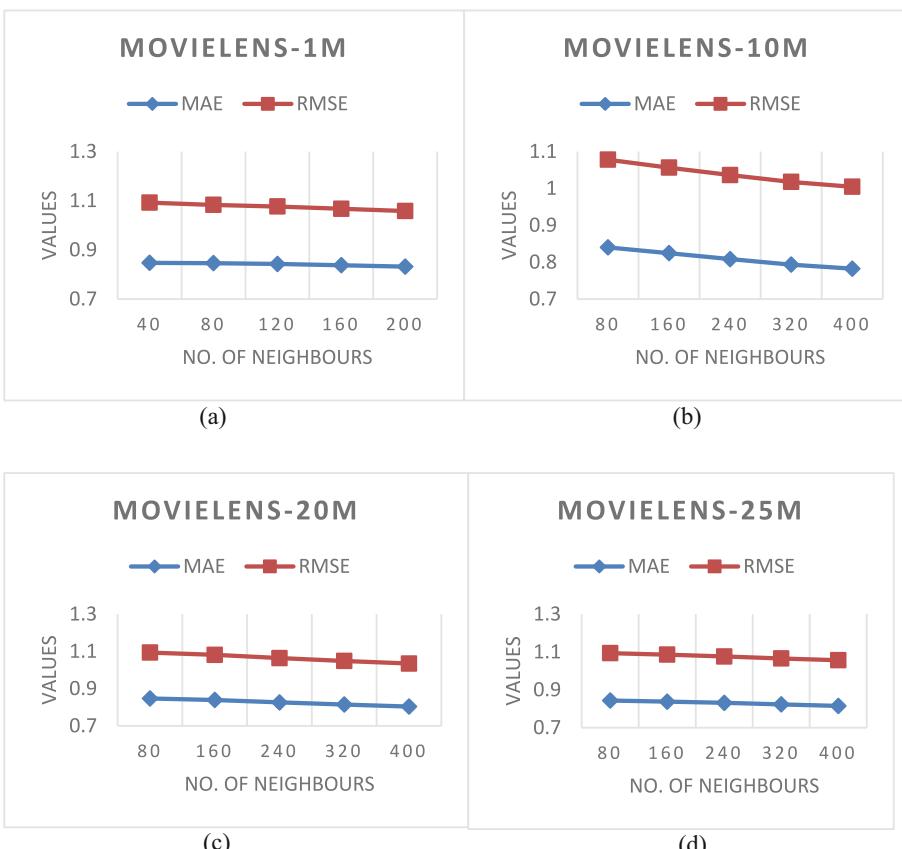


Fig. 6. The MAE and RMSE values for (a) ML-1M (b) ML-10M (c) ML-20M (d) ML-25M.

then the values of both metrics, MAE and RMSE, decrease monotonically (lower value means better recommendation). It means that by increasing the neighbourhood size, we can get better and more trustworthy recommendations. These empirical results are also consistent with our hypothesis that “the recommendations based on more number of similar users would be better than fewer number of similar users”.

4 Conclusion

On the World Wide Web, recommendations are almost everywhere for Internet users. In general, a commercially successful website either already has a well-deployed recommendation system or is planning to deploy one based on the requirements. Trustworthy recommendations will increase the confidence of the end users in the RS and encourage them to interact more with the RS with a satisfied attitude. Neighbourhood-based Collaborative recommendations are the most popular and are among the earliest techniques for recommendations. Further, these approaches are easy to implement because of their simplicity. Here, we have analyzed a UBCF method by varying the neighbourhood size on a real-world dataset with the Apache Mahout framework. All the experiments are performed on the MovieLens datasets, which are standardized benchmark datasets in the RSs research for the movie domain. Empirical results signify that the recommendation process can be improved by increasing the neighbourhood size in a UBCF approach. This means that an RS should use a larger number of similar users with respect to a user to provide more accurate and trustworthy recommendations for that particular user.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Ricci, F., Rokach, L., Shapira, B.: *Recommender Systems Handbook*, 3rd edn. Springer, New York, US (2022)
2. YouTube, <https://www.youtube.com>, last accessed 2024/01/05
3. Amazon India, <https://www.amazon.in>, last accessed 2024/01/05
4. Netflix India, <https://www.netflix.com/in>, last accessed 2024/01/05
5. Shi, Y., Larson, M., Hanjalic, A.: Collaborative filtering beyond the user-item matrix: a survey of the state of the art and future challenges. *ACM Comput. Surv.* **47**(1), 1–45 (2014)
6. Breese, J.S., Heckerman, D., Kadie, C.: Empirical analysis of predictive algorithms for collaborative filtering. In: 14th conference on Uncertainty in Artificial Intelligence, pp. 43–52. Morgan Kaufmann Publishers Inc., Madison Wisconsin (1998)
7. Joaquin, D., Naohiro, I.: Memory-based weighted-majority prediction for recommender systems. In: Proc. ACM SIGIR'99 Workshop Recommender Systems: Algorithms and Evaluation, Berkeley California (1999)
8. Getoor, L., Sahami, M.: Using probabilistic relational models for collaborative filtering. In: Workshop on Web Usage Analysis and User Profiling (WEBKDD'99), pp. 1–6 (1999)
9. Verma, V., Aggarwal, R.K.: Neighborhood-based collaborative recommendations: an introduction. In: Johri, P., Verma, J., Paul, S. (eds.) *Applications of Machine Learning*, pp. 91–110. Springer, Singapore (2020)

10. Duine Framework - Recommender Software Toolkit, <https://duine.sourceforge.net>, last accessed 2024/01/05
11. Cofi: A Java-Based Collaborative Filtering Library - Summary [Savannah], <https://savannah.nongnu.org/projects/cofi>, last accessed 2024/01/05
12. easyrec download | SourceForge.net, <https://sourceforge.net/projects/easyrec>, last accessed 2024/01/05
13. Lee, J., Sun, M., Lebanon, G.: Prea: personalized recommendation algorithms toolkit. *J. Mach. Learning Res.* **13**(1), 2699–2703 (2012)
14. Guo, G., Zhang, J., Sun, Z., Yorke-Smith, N.: Librec: A java library for recommender systems. In: CEUR Workshop Proceedings, vol. 1388. RWTH Aachen University (2015)
15. Schelter, S., Owen, S.: Collaborative filtering with apache mahout. In: Proc. of ACM RecSys challenge, vol. i. Dublin, Ireland (2012)
16. Ekstrand, M.D., Ludwig, M., Konstan, J.A., Riedl, J.T.: Rethinking the recommender research ecosystem: categories and subject descriptors. In: Proc. 5th ACM Conf. Recomm. Syst.-RecSys, vol. 11, pp. 133–140. ACM (2011)
17. Ortega, F., Zhu, B., Bobadilla, J., Hernando, A.: CF4J: collaborative filtering for Java. Knowl.-Based Syst. **152**, 94–99 (2018)
18. Apache Mahout, <https://mahout.apache.org>, last accessed 2024/01/04
19. MovieLens, <https://movielens.org>, last accessed 2024/01/03
20. Owen, S., Friedman, B.E., Anil, R., Dunning, T.: Mahout in Action, 1st edn. Manning Publications, Simon and Schuster (2011)
21. Herlocker, J.L., Konstan, J.A., Terveen, L.G., Riedl, J.T.: Evaluating collaborative filtering recommender systems. *ACM Trans. Inf. Syst. (TOIS)* **22**(1), 5–53 (2004)
22. GroupLens, <https://grouplens.org>, last accessed 2024/01/03
23. University of Minnesota Twin Cities, <https://twin-cities.umn.edu>, last accessed 2024/01/03
24. Harper, F.M., Konstan, J.A.: The movielens datasets: history and context. *ACM Trans. Interact. Intell. Syst. (tiis)* **5**(4), 1–19 (2015)
25. Gunawardana, A., Shani, G.: A survey of accuracy evaluation metrics of recommendation tasks. *J. Mach. Learning Res.* **10**(12) (2009)



Multilingual Sentiment Analysis over Real-Time Voice

Samikshya Rath, Ojasvi Nagayach, Asritha Boddu, Raguru Jaya Krishna^(✉) and B. Vamshi Krishna

Department of Computer Science and Engineering, Manipal Institute of Technology Bengaluru,
Manipal Academy of Higher Education, Manipal, Udupi, India
raguru.krishna@manipal.edu

Abstract. Being able to quickly glean insightful information from spoken language is essential in today's digitally connected society. The goal of this proposed work is to perform voice sentiment analysis in real-time—a cutting-edge technology with a wide range of uses. In contrast to traditional sentiment analysis, which focuses mostly on written language, this proposed work focuses on classifying sentiment into three separate groups: neutral, positive, and negative. This classification empowers people and institutions to make educated choices, improve user experiences, and provide better real-time services. This approach entails creating a real-time sentiment analysis system that can quickly and accurately classify the sentiment expressed in voice input. This is achieved by coordinating natural language processing tools, sentiment analysis models, and speech recognition libraries. The main goals of this paper include developing a flexible system that can accommodate multiple languages – namely – English, Hindi and Telugu. Additionally, it seeks to offer real-time sentiment feedback, which is a useful function in several fields, such as social media monitoring, voice assistant technologies, market research, and customer service interactions. This paper is not primarily concerned with profound emotional or attitudinal analysis, but rather with sentiment categorization, specifically focusing on positive, negative, and neutral sentiments.

Keywords: Multilingual Sentiment Analysis · Real-time Voice Sentiment Analysis

1 Introduction

In our increasingly data-driven world, the extraction of meaningful insights from text data has become crucial across various domains. One of these analytical challenges is sentiment analysis, a process that involves determining the emotional tone or attitude expressed within textual content. Numerous fields, including marketing, social trends analysis, political opinion tracking, and customer feedback analysis, use sentiment analysis extensively. But as we've entered the era of globalization, it has become increasingly clear that multilingual sentiment analysis is essential. This proposed work recognizes this need and attempts to explore challenges of multilingual sentiment analysis using the spaCy library and four distinct models.

1.1 Problem Statement

The fundamental problem addressed in this proposed work is the challenge of conducting sentiment analysis effectively and accurately across multiple languages. While sentiment analysis models have been efficient for English text, extending their applicability to diverse languages with varying linguistic structures and cultural differences is a doubtful task. To overcome this challenge, we need to evaluate existing models and potentially develop custom solutions that can analyze sentiments in a wide array of languages with precision.

1.2 Objectives

The primary objective of this proposed work is to conduct a thorough investigation of multilingual sentiment analysis techniques using the spaCy library and four different models, namely VADER, Hugging Face Transformers, TextBlob from NLTK, and a custom-developed function. Specific objectives include:

- Evaluating the effectiveness of each sentiment analysis model in identifying sentiments expressed in diverse languages.
- Comparing the performance of the models to identify their strengths and weaknesses.
- Providing recommendations for selecting the most suitable model for specific multilingual sentiment analysis tasks.
- Demonstrating the practical utility of multilingual sentiment analysis in real-world scenarios.

1.3 Scope

The scope of this proposed work encompasses the following aspects:

- *Language Diversity:* Evaluating the model's performance on text written in various languages, representing different language families and linguistic complexities.
- *Model Evaluation:* Rigorously evaluate the effectiveness of VADER, Hugging Face Transformers, TextBlob from NLTK, and the custom-developed function through relevant metrics.
- *Real-world Application:* Demonstrating the practical application of multilingual sentiment analysis in domains such as marketing, news monitoring, and social trends analysis.

2 Literature Review

2.1 NLP Concepts

The code exemplifies the utilization of a variety of natural language processing (NLP) concepts and techniques. Here's how these NLP concepts are skillfully implemented in the program:

1. *Text Tokenization:* To split text into individual units such as words or phrases, tokenization is used - an essential process for any successful NLP task. To this end, trusty ally spaCy is called upon to effectively tokenize transcribed text.

2. *Part-of-Speech (POS) Tagging*: This crucial step involves assigning grammatical categories like nouns, verbs, adjectives etc., to different words within a given piece of text - something that greatly assists in the understanding of its structure and meaning better. With POS tagging at the heart of many linguistic analyses, spaCy is employed yet again; extracting valuable information about various POS tags associated with each token.
3. *Sentiment Analysis*: The next phase entails determining what sentiment or emotional tone lies embedded deep inside some block of writing via Sentiment analysis- where several cutting-edge techniques/libraries namely TextBlob, VADER & Hugging Face Transformers make their presence felt by conducting a thorough assessment on just how positively/negatively charged our original transcript appears.
4. *Language Detection*: To accurately detect which specific language a transcription happens to be using, ‘langdetect’ library’s capabilities are utilized. In cases where English is not detected upfront, additional translation is done before applying sentiment-analysis tools.
5. *Translation*: Converting non-native(non-English) text fragments into easily comprehensible English ones is crucial prior to performing further textual evaluations and powerful translators residing within the ‘translate’ library aid in this process.
6. *Speech Recognition*: The Speech Recognition library enables recording all user’s voice inputs emanating from microphones for purposes of leveraging them as valuable pieces of information necessary in further program execution.
7. *Thread Synchronization*: Multi-threading is used to run sentiment analysis in separate threads to ensure that the GUI remains responsive while sentiment analysis is performed (Table 1).

2.2 Related Works

Table 1. Summary of Literature review

S.No	Author	year	Technique	Drawbacks
1	Jeong, B., Yoon, J., & Lee, J. M	2019	Topic modelling to assess the importance of topics discussed by customers	Complexity in integrating diverse customer perspectives and interpreting nuanced sentiments
2	Rajendran, S., Srinivas, S., & Pagel, E	2023	Text mining, bigrams, trigrams,	Reliance on online feedback may introduce bias, limiting the generalizability of the findings

(continued)

Table 1. (*continued*)

S.No	Author	year	Technique	Drawbacks
3	Alsaeedi, A., & Khan, M. Z	2019	Support vector machines (SVM), Bayesian algorithms and hybrid ensemble method	Data may introduce bias thus limiting the external validity of sentiment analysis findings
4	Murwati, A. S., & Aldianto, L	2022	Hybrid approach using lexicon based Textblob and logistic regression	Data may introduce sampling bias
5	Medhat, W., Hassan, A., & Korashy, H	2014	overview including proposed algorithm enhancements and application in the field of text mining	Maintaining real-time relevance, as the field evolves rapidly
6	Sowjanya, A. M	2021	XLS-R model for multilingual and wav2vec 2.0 to reduce error rates	Generalizing the model's performance to diverse linguistic contexts
7	Durairaj, A. K., & Chinnalagu, A	2021	BERT model, linear support vector machine	Computationally intensive, limiting scalability for large datasets
8	Shekhawat, B. S	2019	Vectorization, naïve bayes classifier model, python Textblob	Data may introduce bias

A considerable gap exists in thorough comparisons across many domains, even though the surveyed publications address different elements of sentiment analysis utilizing approaches like topic modeling, sentiment analysis algorithms, and machine learning models like Naïve Bayes, LSTM, and BERT. Studies that do more extensive cross-domain comparisons are warranted because most of the current research concentrates on narrowly focused domains, such as Twitter or insurance, or specialized sectors. Furthermore, the survey paper's thorough overview of sentiment analysis methods and associated topics is devoid of a real-time analysis perspective, which is essential considering how quickly social media and technology are developing. The cited papers clearly show the need for more thorough comparative studies that consider the subtleties of various models, algorithms, and domains.

3 Proposed Methodology

3.1 Dataset

- VADER is a rule-based sentiment analysis tool employing a pre-built lexicon and grammatical rules to assess text sentiment. The lexicon, curated by developers, contains words with pre-assigned sentiment scores, drawn from diverse sources such as social media and online reviews.
- Hugging Face Transformers utilizes the DistilBertForSequenceClassification model, pre-trained on a mix of BookCorpus and English Wikipedia datasets, for its default text classification pipeline.
- TextBlob, in conjunction with NLTK, lacks a specific pre-trained dataset but leverages resources from NLTK, encompassing diverse corpora and lexical tools for various natural language processing tasks, including sentiment analysis.
- Custom model also utilises Hugging Face Transformers hence its dataset remains the same.
- For evaluation, 3 real-time voice datasets were constructed, one each for English, Hindi, and Telugu.

3.2 Architectural Overview

1. *Vader* - VADER (Valence Aware Dictionary And Sentiment Reasoner) utilizes both lexicon rulesets within its analytical framework. Incorporating both ruleset and dictionary components ensures greater accuracy when determining overall textual sentiments. It achieves this by assigning pre-determined sentient scores corresponding to words, phrases, and contextual cues. Vader, the powerful natural language processing model specifically designed for sentiment analysis, excels at detecting emotion in text. It is particularly adept at capturing nuanced sentiments that are context-dependent or when words have varying meanings depending on their usage. At its core, Vader relies on a comprehensive lexicon containing thousands of words and phrases each assigned with a polarity score ranging from highly negative to extremely positive.

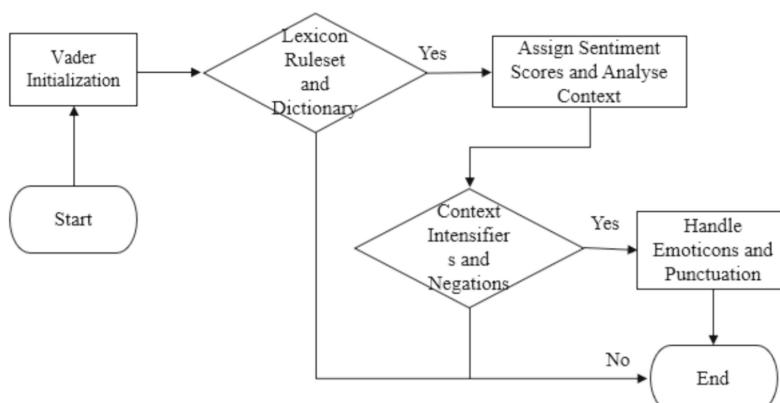


Fig. 1. Flowchart representing Vader model.

It is set apart by its ability to consider intensifiers and negations that can modify these scores. The algorithm also incorporates rules for handling emoticons and punctuation marks which convey emotions. A compound sentiment score is calculated by weighing individual term scores based not just on word choice but also factors like capitalization, punctuation use, and conjunctions used within the text itself. This compound score offers insight into whether the overall tone of the text leans towards positivity or negativity or remains neutral. Flowchart representing Vader model is shown in the Fig. 1.

2. *Hugging Face Transformers* - This NLP library is known for providing easy access to pre-trained transformer-based models capable of performing various NLP tasks including sentiment analysis with remarkable precision. These sophisticated models leverage cutting-edge technology to understand human language by analyzing contextual relationships between different terms present in any given piece of written communication. Their key strength lies in utilizing self-attention mechanisms instead of traditional sequential methods while processing input data – this allows them to process multiple terms simultaneously leading to higher efficiency levels as well as greater accuracy especially when dealing with longer texts where dependencies span across several sentences. All incoming textual inputs go through tokenizing processes wherein subword tokens get encoded into numerical representations before being fed into Transformer architecture which leverages deep learning algorithms using advanced techniques such as fine-tuning pre-existing datasets optimized specially around predicting emotional responses accurately during inference phases thanks largely due diligence employed during training regimes. Flowchart representing Hugging Face Transformers model is shown in Fig. 2.

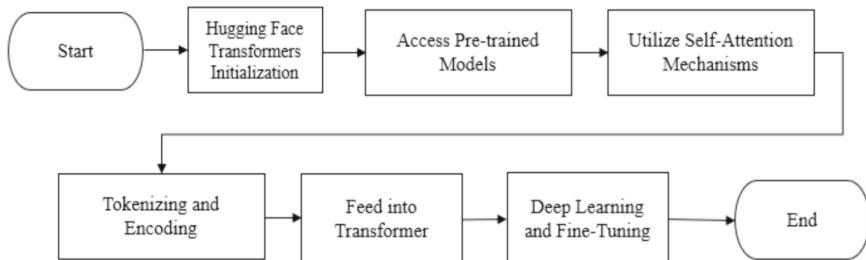


Fig. 2. Flowchart representing Hugging Face Transformers model.

TextBlob powered NLTK - Textblob facilitates intuitive exploration enabling users to identify hidden patterns helps detect unseen structures sentiment-conveying emotional indicators. This versatile toolkit further enhances our understanding of emotions in text, utilizing NLTK's advanced features and capabilities to process textual data effectively. With its help, users can easily extract sentiments from written communication by leveraging well-crafted models that have been fine-tuned specifically for this purpose using robust datasets during training phases. During inference stages TextBlob accurately

assigns emotion labels as well as corresponding scores based on the knowledge gained through previous experiences - making it an invaluable tool for anyone seeking deeper insights into human expression. The NLTK (Natural Language Toolkit) combined model effortlessly prepares the necessary resources for text preprocessing, including removing stop words and performing tokenization. Using the ‘analyze_sentiment’ function, it utilizes TextBlob to accurately assess sentiment in any given input. This powerful tool computes a polarity score ranging from -1 (representing negative sentiments) to 1 (indicating positive ones). Depending on this numerical value’s sign, our function then assigns labels such as “POSITIVE,” “NEGATIVE,” or “NEUTRAL” to reflect the overall tone of the text. For real-time analysis of sentiments, an innovative feature has been incorporated that captures audio through your microphone using SpeechRecognition library. After transcribing the spoken words into written form with precision accuracy, TextBlob is applied once again for sentiment analysis purposes. Both label and polarity are based on its comprehensive evaluation. The script not only provides sentiment data but also showcases detailed information about each token through proper segmentation techniques. With part-of-speech tags and dependencies displayed prominently alongside every tokenized word from the transcription results – a deeper understanding is obtained of how they all connect within their context seamlessly. Flowchart representing TextBlob powered NLTK model is shown in Fig. 3.

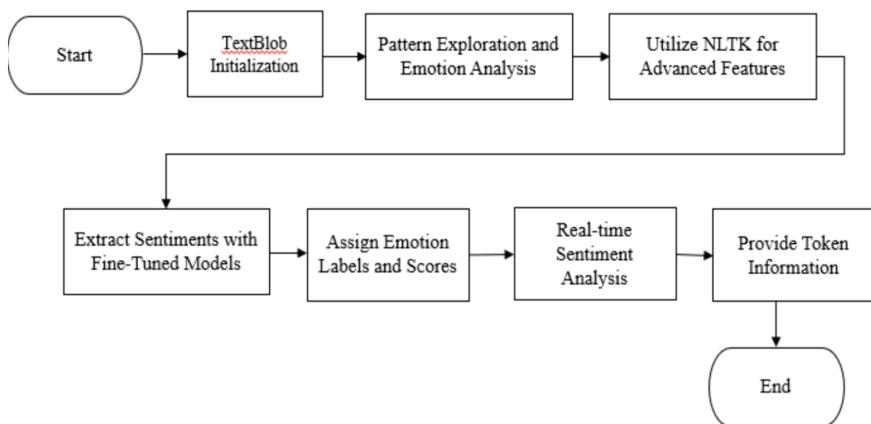


Fig. 3. Flowchart representing TextBlob powered NLTK model.

Custom model – This model utilizes SpeechRecognition, spaCy, and Hugging Face Transformers to conduct real-time sentiment analysis on spoken text – particularly in Hindi, English and Telugu. It begins by initializing a speech recognizer and loading the spaCy model. Custom spaCy function ‘custom_tokenize’ enhances linguistic analysis by extracting tokens and providing detailed information about each token’s properties. Sentiment analysis is performed using ‘real_time_sentiment_analysis’ function, assigning sentiment labels based on polarity scores. In the real-time sentiment analysis loop, the script captures and transcribes voice input using speech recognition, then applies

custom spaCy functions for tokenization and parsing. Exception handling is incorporated for potential errors, ensuring a robust and user-friendly application. Flowchart representing Custom model is shown in Fig. 4.

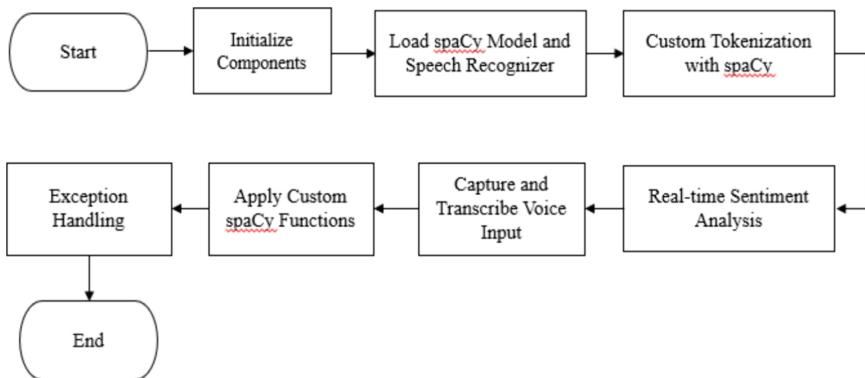


Fig. 4. Flowchart representing Custom model.

3.3 Implementation

Libraries Used:

- *Tkinter*: For building the GUI.
- *Speech_recognition*: For capturing and transcribing audio.
- *Transformers*: For sentiment analysis using hugging face transformers.
- *vaderSentiment*: For sentiment analysis using VADER.
- *textblob*: For sentiment analysis using TextBlob.
- *spaCy*: For natural language processing and tokenization.
- *nltk*: For additional tokenization and stopwords.
- *threading*: For running real-time sentiment analysis in a separate thread.

Initialization:

- *Speech Recognition*: Initializes a speech recognition recognizer with a microphone and sets the energy threshold.
- *spaCy Model*: Loads the spaCy English model.
- *Sentiment Analyzers*: Initializes different sentiment analysis methods, including Hugging Face Transformers, VADER, TextBlob, and a custom function.

Sentiment Analysis Functions:

- *calculate_overall_sentiment(sentiment_scores)*: Calculates the overall sentiment based on compound scores from VADER.
- *analyze_sentiment(text, method)*: Analyzes sentiment using the specified method. For TextBlob and nltk, it calculates sentiment polarity. For VADER and Hugging Face Transformers, it retrieves sentiment labels and scores.

- *real_time_sentiment_analysis(method)*: Captures audio from the microphone, transcribes it, and performs sentiment analysis using the specified method. The results are then scheduled to update the results table on the main thread using ‘app.after()’.

4 Data Analysis

4.1 Exploratory Data Analysis

In this proposed work, various techniques were employed to gain insights into the provided dataset, which contains transcribed text from speech recordings in different languages. The initial examination involves tokenization and part-of-speech tagging using spaCy, revealing the linguistic structure of the text. Subsequently, sentiment analysis is performed utilizing the Hugging Face Transformers pipeline, aiming to discern the sentiment of the transcriptions. This proposed work encompasses both English and multilingual content, necessitating language detection and translation when applicable. The use of the pipeline for sentiment analysis for English text and multilingual models for Hindi and Telugu facilitates a comprehensive understanding of sentiment across diverse linguistic contexts. The sentiment analysis scores are tailored to ensure that negative sentiments fall within the range of 0 to -1 , positive sentiments within 0 to $+1$, and neutral sentiments at precisely 0. This adjustment enables a standardized interpretation of sentiment scores across all languages. In addition to sentiment analysis, the exploration involves the extraction of tokens with part-of-speech tags and dependency relationships. This detailed linguistic analysis aids in uncovering syntactic nuances and language-specific structures. The combined insights derived from linguistic patterns and sentiment trends provide a comprehensive understanding of the dataset, enabling a more informed interpretation of the transcribed speech content. This exploratory data analysis lays the groundwork for subsequent analyses and model development, contributing to a nuanced understanding of the dataset’s linguistic and emotional dimensions.

4.2 Model Training

1. *Hugging Face Transformers*:

- The ‘pipeline(“sentiment-analysis”)’ from Hugging Face Transformers uses pre-trained models for sentiment analysis. These models are typically trained on large datasets and fine-tuned for specific tasks, including sentiment analysis.
- The actual training process is done by Hugging Face on powerful hardware, and users can directly use the pre-trained models without training them further.

2. *TextBlob and nltk*:

- TextBlob uses a pre-trained sentiment analysis model. However, it does not involve explicit training.

3. *VADER*:

- VADER (Valence Aware Dictionary and sentiment Reasoner) is a pre-built lexicon and rule-based sentiment analysis tool. It does not involve traditional model training.

- VADER uses a predefined list of words with associated sentiment scores to analyze sentiment in a text.

Model Validation:

- A. Hugging Face Transformers: The pre-trained models from Hugging Face are often validated on benchmark datasets during their development. Users generally rely on the robustness of these pre-trained models for various tasks.
- B. TextBlob and nltk: TextBlob's sentiment analysis model is trained on a dataset, and its performance is validated during training.
- C. VADER: VADER's lexicon-based approach doesn't involve a traditional validation process. It relies on predefined sentiment scores for words (Table 2).

5 Results

5.1 Model Performance

Table 2. Model Performance with Accuracy for the English, Telugu, and Hindi dataset

Model	Accuracy - English	Accuracy - Telugu	Accuracy - Hindi
Hugging Face Transformers	66.66%	63.33%	64.44%
Vader	75.55%	Poor	48.88%
Textblob combined with NLTK	80.00%	Poor	42.22%
Custom	66.66%	63.33%	64.44%

5.2 Result Interpretation

This research study systematically examines the performance of various language models, revealing a noteworthy disparity in their efficacy across different languages. Notably, the evaluation indicates that these models exhibit exceptional proficiency when applied to the English language, consistently delivering optimal results.

A sample dataset for English, Hindi and Telugu are shown in the Tables 3, 4 and 5. In stark contrast, when confronted with Indian languages such as Telugu and Hindi, the models demonstrate a considerably lower accuracy rate, achieving precision in their predictions only half of the time. This discernible discrepancy underscores the existence of significant opportunities for refinement and enhancement, particularly in the context of linguistic applications pertaining to Indian languages. The findings herein advocate for a focused endeavor to address and ameliorate the specific challenges associated with language processing in the context of Telugu and Hindi, thereby paving the way for more robust and universally applicable language models.

Table 3. A Sample of the English Test Dataset

		Hugging Face Transformers		Vader		TextBlob and NLTK		Custom	
Sentence	Known Sentiment	SL	SC	SL	SC	SL	SC	SL	SC
The sun is shining brightly in the blue sky	+	+	0.99985	+	0.3612	+	0.3500	+	0.99985
I love spending time with my friends and family	+	+	0.99979	+	0.8074	+	0.5000	+	0.99979
The children's laughter fills the air with joy	+	+	0.99988	+	0.7096	+	0.8000	+	0.99988
She received a well-deserved promotion at work	+	+	0.99980	+	0.2732	N	0.0000	+	0.99980
The beautiful flowers are in full bloom	+	+	0.99987	+	0.5994	+	0.6000	+	0.99987

6 Discussion

6.1 Limitations

1. *Accuracy of Sentiment Analysis:* The accuracy of sentiment analysis can vary depending on the method used. All sentiment analysis models, such as TextBlob, VADER and Trasnformers, might not always accurately capture the refined sentiments expressed in natural language. Misclassification of sentiments can occur.
2. *Speech Recognition Errors:* The accuracy of speech recognition can be influenced by factors like background noise, accents, and the quality of the microphone. Disturbance in the transcription of audio can lead to incorrect sentiment analysis.
3. *Real-time Processing:* Real-time sentiment analysis may produce latency due to the processing time of the sentiment analysis models. This delay can affect the user experience and may not be suitable for applications requiring very low latency.
4. *Inaccurate Language Detection:* Some languages have words that phonetically resemble words in other languages. Language detection algorithms may misidentify the language of the spoken input, leading to translation errors or incorrect sentiment analysis.
5. *Translation Quality:* The quality of translation from non-English languages to English can be unsatisfactory, altering the original sentiment expressed in the source language.

Table 4. A Sample of the Telugu Test Dataset

		Hugging Face Transformers		Vader		TextBlob and NLTK		Custom	
Sentence	Known Sentiment	SL	SC	SL	SC	SL	SC	SL	SC
Naaku nachindhi	+	-	0.7001	N	0.0	N	0.0	-	0.7001
Meeru adbutham	+	+	0.9398	N	0.0	N	0.0	+	0.9398
Eeroju chala bagundhi	+	-	0.8455	N	0.0	N	0.0	-	0.8455
Pani chala manchiga chesadu	+	+	0.9994	N	0.0	N	0.0	+	0.9994
nuvvu santhosham ga undu	+	+	0.9365	N	0.0	N	0.0	+	0.9365
eeroju chala saradaga gaddipamu	+	+	0.9550	N	0.0	N	0.0	+	0.9550

6.2 Future Work

1. *Multimodal Sentiment Analysis:* Expanding the model's capabilities to analyze sentiment from multiple modalities simultaneously, such as including audio, text, and visual data, can improve its accuracy.
2. *Multilingual Support:* Enhancing the model's ability to perform sentiment analysis in multiple languages, particularly low-resource languages are essential. It can be tested using techniques like transfer learning or unsupervised learning.
3. *Custom Sentiment Analysis Functions:* Developing and implementing domain-specific custom sentiment analysis functions to provide specific industries or applications, such as healthcare, finance, or customer service.
4. *Real-time Noise Reduction:* Implementing noise reduction techniques for real-time audio inputs to improve speech recognition accuracy, especially in noisy environments.
5. *Real-time Privacy Measures:* Developing privacy-preserving techniques to ensure that sensitive or private information is not exposed during real-time sentiment analysis.

Table 5. A Sample of the Hindi Test Dataset

		Hugging Face Transformers		Vader		TextBlob and NLTK		Custom	
Sentence	Known Sentiment	SL	SC	SL	SC	SL	SC	SL	SC
Mai theek hun	+	+	0.795	N	0.0	+	0.1666	+	0.795
Aapki smile achhi hai	+	+	0.995	+	0.3612	+	0.3	+	0.995
Aaj italian khaate hain	N	-	-0.85	N	0.0	N	0.0	-	-0.85
Pani chala manchiga chesadu	+	+	0.9994	N	0.0	N	0.0	+	0.9994
Nahi ho raha hai mujhse	-	-	-0.958	N	0.0	N	0.0	-	-0.958
Bageecha purana hai	N	+	0.98734	N	0.0	N	0.0	+	0.98734

7 Conclusion

The proposed work being described is a real-time sentiment analysis application with a Tkinter-built user interface. Its main contribution is the combination of pre-trained sentiment analysis models with different natural language processing (NLP) techniques. Users can choose from a variety of sentiment analysis techniques using the GUI, such as TextBlob with nltk, VADER, and Hugging Face Transformers. The primary output of the proposed work is the real-time visualization of sentiment analysis findings in the GUI, which provides users with instant access to sentiment labels, scores, and tokens that are related with parts of speech (POS) tags. Dynamic visualization serves to improve user comprehension of the various results generated by various sentiment analysis techniques. The proposed work essentially functions as an interactive and useful tool for real-time sentiment analysis, allowing users to explore different sentiment analysis.

References

1. Jeong, B., Yoon, J., Lee, J.M.: Social media mining for product planning: a product opportunity mining approach based on topic modeling and sentiment analysis. *Int. J. Inf. Manage.* **48**, 280–290 (2019)
2. Rajendran, S., Srinivas, S., Pagel, E.: Mining voice of customers and employees in insurance companies from online reviews: a text analytics approach. *Benchmarking: Int. J.* **30**(1), 1–22 (2023)

3. Alsaeedi, A., Khan, M.Z.: A study on sentiment analysis techniques of Twitter data. *Int. J. Adv. Comput. Sci. Appl.* **10**(2), 361–374 (2019)
4. Murwati, A.S., Aldianto, L.: Exploring voice of customers to Chatbot for customer service with sentiment analysis. *Asian J. Technol. Manag.* **15**(2), 141–153 (2022)
5. Medhat, W., Hassan, A., Korashy, H.: Sentiment analysis algorithms and applications: a survey. *Ain Shams Eng. J.* **5**(4), 1093–1113 (2014)
6. Sowjanya, A.M.: Self-supervised model for speech tasks with hugging face transformers. *Turkish Online J. Qual. Inquiry* **12**(10) (2021)
7. Durairaj, A.K., Chinnalagu, A.: Transformer based contextual model for sentiment analysis of customer reviews: a fine-tuned BERT. *Int. J. Adv. Comput. Sci. Appl.* **12**(11) (2021)
8. Shekhawat, B.S.: Sentiment classification of current public opinion on brexit: Naïve Bayes classifier model vs Python’s Textblob approach. Doctoral dissertation, Dublin, National College of Ireland (2019)



Identity Verification: A Decentralized KYC Approach Using Blockchain

Akshay Chouke¹, Vikas Kumar Jain¹, Jitendra Parmar¹, Shiv Shankar Prasad Shukla¹, and Atul Kumar Verma²(✉)

¹ School of Computing Science and Engineering, VIT Bhopal University, Kothrikalan, Sehore, Madhya Pradesh 466114, India

{akshay.chouke2020, vikaskumar, jitendraparmar, shivshankar.prasad}@vitbhopal.ac.in

² Department of Computer Science and Engineering, Manipal University Jaipur, Jaipur, India
atul.verma@jaipur.manipal.edu

Abstract. The Know Your Customer (KYC) is a process followed by financial institutions, businesses or banks to verify the identity of their customers or users. Traditional KYC procedures are repetitive, centralized, and highly vulnerable to sensitive data exposure today. This repetitive task reduces the ease of use; due to repetition, it is costlier for banks or financial institutions. There are plenty of papers published to address these issues and propose various solutions, such as storing the user's data directly on the blockchain, which requires high transaction costs, and storing user details on the IPFS. Then, the hash of the user's data provided by IPFS is stored on the blockchain. Since the data on IPFS is directly stored, it can be accessed by anyone who has the hash. So, to address this issue, we propose a solution where, before storing the user's data on the IPFS, it is encrypted using the AES encryption algorithm, and the hash of the encrypted data is provided to the user and stored on the blockchain. This decentralized and secure KYC system minimizes redundancy, enhances user control, and fortifies data security.

Keywords: Blockchain · IPFS · Decentralized · KYC · AES · Hash

1 Introduction

Financial institutions and banks adhere to a crucial procedure known as the Know Your Customer (KYC) process when initiating financial relationships with new customers [1, 2]. This process comprises a series of standard tasks to confirm the legality of a potential customer's activities. Before commencing any form of business with a new customer, every financial institution or bank must complete the KYC process. The traditional Know issues of centralization, redundancy, and a lack of user control over sensitive information sharing have long marred the KYC process in the banking sector. Customers often repeat the same KYC procedures when establishing accounts with different banks, leading to inefficiency and frustration [2]. Moreover, the centralized storage of KYC data poses significant security risks, making it an attractive target for criminal activities. The week

methods of KYC leads many data security issues and vulnerabilities [3–5] To address these issues, some methodology suggested by authors [6]. Some involve storing user data directly on the blockchain, which often incurs high transaction costs. To overcome these problems, some authors have suggested alternatives like storing user details on IPFS, where the hash of the user's data is subsequently stored on the blockchain. However, this method raises concerns regarding data accessibility, as information stored on IPFS is readily available to anyone with the corresponding hash.

In response to the challenges mentioned, we propose the following approaches to address KYC verification issues:

1. Introduces an extra layer of security and privacy. Before the user's data is stored on IPFS, it undergoes encryption using the robust AES encryption algorithm.
2. Furthermore, the hash of this encrypted data is made available to the user and recorded on the blockchain. This unique amalgamation of technologies offers a decentralized and secure KYC system that minimizes redundancy, gives users greater control over their data, and fortifies data security.
3. In an era where data privacy and security are paramount, this research aims to provide a transformative solution that benefits financial institutions and their customers.

2 Background

The following terms can explain background of purposed methods:

1. **KYC:** Know Your Customer (KYC) is an essential procedure used by financial institutions and businesses to authenticate the identities of their clientele [7]. The primary aim of KYC guidelines is to safeguard banks against being unwittingly or deliberately exploited by criminal entities involved in money laundering activities. KYC processes also empower banks to understand their customers and financial transactions better, thereby enhancing their ability to manage risks with due diligence [8].
2. **BLOCKCHAIN:** Blockchain technology is a decentralized and distributed digital ledger system [9, 10]. It comprises a chain of blocks, each containing a set of transactions or data records linked together through cryptographic hashing. Its key features include immutability, where once data is recorded, it becomes nearly impossible to alter or delete; transparency, as all network participants can view the transaction history; and security, thanks to its decentralized nature and cryptographic safeguards [11]. Blockchain's decentralized property ensures that every transaction is recorded on multiple nodes, which makes it almost impossible to alter the transaction [12].
3. **IPFS:** IPFS, the InterPlanetary File System, is a decentralized peer-to-peer file-sharing platform that employs hash tables to track data ownership. Its hallmark feature is the creation of unique hashes for data, independent of its size. When uploaded, data's broken into chunks, each with a unique hash, and distributed to nodes with similar hashes [13]. When a user requests data, retrieval occurs by traversing nodes with matching hashes, eventually combining all chunks to reconstruct the original object. This innovative approach ensures efficient and reliable data storage and retrieval on the IPFS network [14].

- 4. AES:** Advanced Encryption Standard (AES), is a widely adopted symmetric encryption algorithm designed to secure data by transforming it into an unreadable format and then back to its original state using a secret key [15]. AES operates on blocks of data, typically 128 bits at a time, and supports key lengths of 128, 192, or 256 bits [16]. Its strength lies in its robust security and efficiency, making it suitable for various applications, including secure communication, data storage, and cryptography. AES encryption involves a series of well-defined mathematical operations, such as substitution, permutation, and bitwise operations, performed in multiple rounds to ensure the confidentiality and integrity of sensitive information

The encryption can be calculated as:

$$C = E(KE, P)$$

Here - C = Cipher Text, KE = Encryption Key, P = Plain Text, E = Encryption function

The general formula for decryption-

$$P = D(KD, C)$$

Here, C = Cipher Text, KD = Decryption Key, P = Plain Text, D = Decryption Function

3 Related Work

In [17], authors introduced an enhanced e-KYC system based on blockchain, emphasizing optimization. This system employs symmetric AES encryption and utilizes the LZ compression algorithm to enhance efficiency. Notably, an intelligent contract is employed to validate and store KYC data on the blockchain automatically. Compressed KYC data reduces gas fees on the Ethereum network, although it results in longer data extraction times from the blockchain.

In [18], authors proposed an Ethereum-based KYC verification system consisting of four entities: customers, Ethereum Blockchain, banks, and secure storage. This allows for KYC creation, verification, and upload processes, where customers submit KYC requests to their respective banks. These banks store KYC documents securely and create hash links for these documents on the Ethereum Blockchain. The number of votes determines the KYC status of a customer. It also provides mechanisms for banks to vote on the trustworthiness of other banks in the network, preventing fraudulent KYC attempts.

In [19], authors proposed a system where, when a user requests to open an account with Bank 1, the bank initiates an IPFS block where the user uploads the necessary documents and provides access to Bank 1. Bank 1 verifies the documents. After successful verification, the KYC is issued, and a copy and its hash are uploaded to a Distributed Ledger Technology (DLT) platform. The user is assigned a unique ID on the blockchain, which is broadcast to all nodes in the network. When Bank 2 requests KYC for the same user, they simply provide their unique ID. Bank 2 can then verify the KYC's validity by comparing hash functions from IPFS with those embedded in the blockchain. If they match, Bank 2 knows the KYC is valid. If not, Bank 2 can manually validate the KYC documents or follow the same process as Bank 1, ensuring a more efficient and secure KYC verification process across multiple financial institutions.

In [20], the authors introduced a blockchain-driven solution to remove intermediaries, providing users with a one-time KYC process. This innovation grants users unrestricted access to data from any location and for various applications. The decentralized nature of blockchain technology, coupled with its user transparency and absence of third-party interference, enhances the security of this system. Furthermore, it ensures quicker data processing.

In [21], authors proposed a blockchain-based KYC verification system which employs digital signatures, combining hashing and asymmetric encryption techniques to ensure the integrity of user data and identity verification. Local data storage is facilitated through QR codes, granting users control over their information while preventing unauthorized access. The onboarding process involves document submission and verification, resulting in the creation of a unique blockchain-based ID. User data is encrypted, embedded in a QR code, and secured with a digital signature for immutability. During verification, a user's QR code is scanned, an OTP (One-Time Password) is sent for two-factor authentication, and the user's data is decrypted and cross-validated against the blockchain record. This approach enhances data security and streamlines the KYC process, eliminating redundant verifications and enabling trustless interactions between various financial institutions.

4 Proposed Work

In Figure 1 of our proposed KYC model comprises several key steps as Store User's Data on IPFS, Encrypt User's Data (AES Algorithm), Get Hash of Encrypted Data and Store Hash on Blockchain. In our proposed work, we can see in algorithm 1 that when a user wishes to create a new account with a bank, the bank initiates a KYC process. The user must follow the bank's instructions if they have yet to undergo KYC. During this process, the user is asked to submit all the necessary documents and input a security key to encrypt their data on the InterPlanetary File System (IPFS). The bank proceeds to verify the user's submitted documents, and based on this verification, the user's KYC status is determined. If the documents are found to be genuine, the KYC is approved; otherwise, it is rejected. For successfully verified users, the bank stores the user's details in one text file and then employs the Advanced Encryption Standard (AES) algorithm to encrypt the text file, which contains the user's data using the provided key. Once the user's data is securely stored on IPFS (See generated log file in figure 2), the bank provides the user with a hash for this data. Simultaneously, the bank uploads or stores this data on the KYC blockchain network, confirming that the user has been successfully verified. This eliminates the need for the user to repeat the KYC process when opening accounts with other banks. Instead, the user can provide their IPFS hash (See generated log file in Fig. 3) to the second bank, allowing them to verify the user's authenticity (KYC) status. Upon receiving the user-provided hash, the second bank checks its existence on the blockchain. If the hash exists, it signifies that the user has already been verified by another bank, establishing the user as genuine. In cases where the hash does not exist, the user is required to undergo the KYC process once again. When the bank needs to access the user's documents, they request the decryption key from the user. After receiving the decryption key, the bank can download and decrypt the documents stored

on IPFS for further use. This approach of saving encrypted data on IPFS enhances the system's reliability and assures users that their data can only be accessed by authorized parties with whom it has been shared. Implementing this process in the banking sector ensures a secure and efficient KYC procedure while maintaining user data privacy and accessibility.

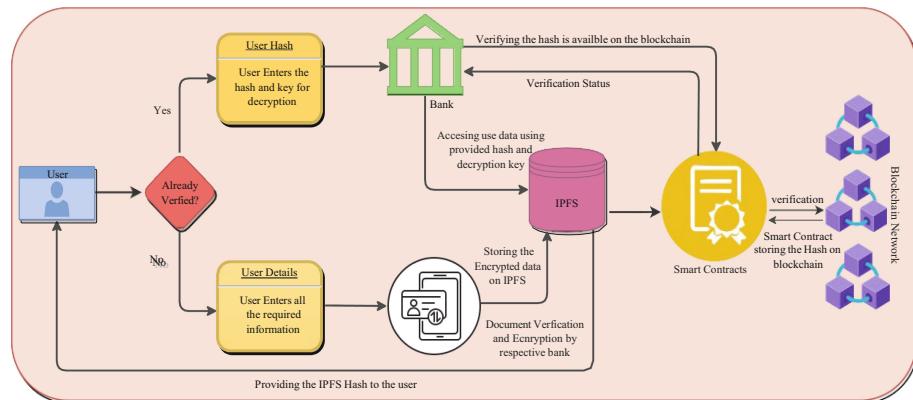


Fig. 1. Architecture of proposed work.

4.1 Implementation

In the implementation of our proposed decentralized Know Your Customer (KYC) system, we orchestrated the development of four distinct algorithms, each serving a crucial role in addressing diverse facets of KYC management and security. An algorithm 2 operates with the aid of the "minimist" and "web3.storage" dependencies, offering users a streamlined mechanism for uploading files and directories onto the InterPlanetary File System (IPFS) through the Web3. Storage service. This algorithm ensures seamless authentication by incorporating a Web3.Storage token, and upon successful upload, it promptly generates and displays the Content Identifier (CID) for future reference.

```

status          0x1 Transaction mined and execution succeed
transaction hash 0x34572ef0e1740ac25b031284c8c34def3f7eed80b66d38785e153f26473e8a7d
block hash      0x27879cb2e372408ecfb4898f769539a93573e620b07370499e6349b39e5dcbf
block number    2
from           0x5838Da6a701c568545dCfcB03FcB875f56beddC4
to             IPFSStorage.storeCID(string) 0xd9145CCE52D386f254917e481eB44e9943F39138
gas            645
transaction cost 49110 gas
execution cost 26742 gas
input          0xccf...53364
decoded input  {
                  "string cid": "b45165ed3cd437b9ffad02a2aad22a4ddc69162470e2622982889ce5826f6e3d"
}
decoded output {}
logs           [
                  {
                      "from": "0xd9145CCE52D386f254917e481eB44e9943F39138",
                      "topic": "0xa4d40fffd5d729c1b64ff5f8650e676b67153438198aaf75250cbdf12a4a0cab63",
                      "event": "CIDStored",
                      "args": {
                          "0": "b45165ed3cd437b9ffad02a2aad22a4ddc69162470e2622982889ce5826f6e3d"
                      }
                  }
]

```

Fig. 2. Transaction status to store the hash on the blockchain

```

from           0x5838Da6a701c568545dCfcB03FcB875f56beddC4
to             IPFSStorage.checkCID(string) 0xd9145CCE52D386f254917e481eB44e9943F39138
execution cost 3796 gas (Cost only applies when called by a contract)
input          0x543...53364
decoded input  {
                  "string cid": "b45165ed3cd437b9ffad02a2aad22a4ddc69162470e2622982889ce5826f6e3d"
}
decoded output {}
logs           [
                  {
                      "0": "bool: true"
                  }
]

```

Fig. 3. Transaction status to check the hash on the blockchain.

Algorithm 1: KYC Processing

Data: User response, IPFS hash, Decryption key, Encryption key

Result: Processed KYC data and IPFS hash

```

1 Dispaly: Have you already completed KYC? (yes/no)
2 Read userResponse
3 if userResponse is "yes" then
4   Display("Please enter the IPFS hash:")
5   Read ipfsHash
6   if IsValidIPFSHash(ipfsHash) then
7     DownloadFileFromIPFS (ipfsHash)
8     Display("Please enter the decryption key:")
9     Read decryptionKey
10    DecryptedData = DecryptFile(downloadedFile, decryptionKey)
11    ProcessKYCData(DecryptedData)
12  else
13    Display ("Error: Invalid IPFS hash.")
14 else if userResponse is "no" then
15   userDocuments = CollectKYCDocuments()
16   if VerifyKYCDocuments (userDocuments) then
17     AcceptKYCRequest()
18     userData = CreateUserData(userDocuments,
19       KYCStatus="Accepted")
20     Display("Please enter an encryption key:")
21     Read encryptionKey
22     encryptedFile = EncryptFile(userData, encryptionKey)
23     ipfsHash = StoreOnIPFS(encryptedFile)
24     Display("KYC request accepted. IPFS hash:" + ipfsHash)
25   else
26     RejectKYCRequest()
     Display("KYC request rejected. Please review submitted documents.")

```

Algorithm 2: IPFSUpload

Data: Uploading files and directories onto the InterPlanetary File System (IPFS) through the Web3.Storage service

Result: Displays the Content Identifier (CID)

```

1 if filePath is empty then
2   Display ("Please supply the path to a file or directory")
3   return
4 storage = new Web3Storage( token )
5 files = []
6 pathFiles = getFilesFromPath(filePath)
7 files.push(...pathFiles)
8 cid = storage.put(files)
9 Display ("Content added with CID: " + cid)
10 Display filePath

```

Algorithm 3: IPFSDownload

Data: CID as an argument

```

1 if cid is empty then
2   Display("Please provide the CID as an argument")
3   return
4 ipfsURL = 'https://dweb.link/ipfs/$cid'
5 Display("Downloading from $ipfsURL"
6 TRY
7 response = await fetch(ipfsURL)
8 if NOT response.ok then
9   if response.status != 200 then
10     Display("File does not exist on IPFS.")
11   else
12     Display ("Error: Failed to download content. HTTP status:
13       $response.status"
14   return
15 buffer = await response.buffer()
16 fileName = "downloaded-file.txt"
17 fs.writeFileSync(fileName, buffer)
18 Display("Downloaded content saved as: $fileName")
19 CATCH
20 Display ("Error: Error downloading content:", error.message)
21 ENDTRY

```

Algorithm 4: File Encryption Algorithm

Data: filename, encryptionKey
Result: Encrypted file 'encrypted.txt'

```

1 if filename is empty OR encryption Key is empty then
2   Display ("Error: Please provide the file name or encryption key.")
3   return
4 Display ("Encrypting file: $filename")
5 fileData = fs.readFileSync(filename)
6 iv = crypto.randomBytes(16)
7 cipher = crypto.createCipheriv('aes-256-cbc', Buffer.from(encryptionKey), iv)
8 encryptedData = Buffer.concat([iv, cipher.update(fileData), cipher.final()])
9 fs.writeFileSync('encrypted.txt', encryptedData)
10 Display("File encrypted and saved.")

```

Algorithm 5: File Decryption Algorithm

Data: fileName, decryptionKey
Result: Decrypted file 'decrypted.txt'

```

1 if fileName is empty OR decryptionKey is empty then
2   Display("Error: Please provide the file name and decryption key as
      arguments.")
3   return
4 Display("Decrypting file: $fileName")
5 encryptedData = fs.readFileSync(fileName)
6 iv = encryptedData.slice(0, 16)
7 encryptedFileData = encryptedData.slice(16)
8 decipher = crypto.createDecipheriv('aes-256-cbc', Buffer.from(decryptionKey),
      iv)
9 decryptedData = Buffer.concat([decipher.update(encryptedFileData),
      decipher.final()])
10 fs.writeFileSync('decrypted.txt', decryptedData)
11 Display ("File decrypted and saved.")

```

The subsequent Algorithm 3 harnesses the power of the "node-fetch" library and the native "fs" module. This empowers users to retrieve files from IPFS by specifying the CID as a command-line argument, enabling the downloaded content to be seamlessly saved as a local file on the user's system. This capability ensures the accessibility and persistence of the retrieved files.

In data security, we implemented two integral algorithms, Algorithm 4 and Algorithm 5, both centering around the robust AES-256-CBC encryption Algorithm. The former Algorithm 4 makes use of the "crypto" and "fs" libraries to facilitate the encryption of a specified file using the AES encryption Algorithm. Users are required to provide the file name and an encryption key as command-line arguments, with the encrypted data then securely saved to a file named "encrypted.txt." Conversely, the latter Algorithm 5 is tailored to oversee the decryption of files previously encrypted with AES-256-CBC. In this case, users execute the script with the file name and decryption key as command-line arguments, leading to the successful restoration of the original data, which is then stored

in a file named "decrypted.txt." These encryption and decryption processes are pivotal in ensuring the confidentiality and integrity of sensitive information within the system.

Implementing these Algorithms forms a cohesive and robust foundation for our file management and security system. Furthermore, these Algorithms can be seamlessly integrated to emulate the proposed decentralized KYC system. In this envisioned system, when a user seeks to establish a new account with a bank, the KYC process is initiated using Algorithms 2 to upload encrypted documents to IPFS securely. The bank subsequently verifies the user's submitted documents, determines the KYC status, and, for successful verifications, deploys the Algorithm 4 to encrypt the user's data with AES before storing it on IPFS. The hash of this encrypted data is then recorded on the blockchain, signifying successful verification and eliminating the need for users to undergo redundant KYC processes with different banks. The user is provided with the hash for future verifications, and when another bank requests KYC, Algorithm 3 is utilized to retrieve the encrypted data from IPFS using the provided hash. The existence of the hash on the blockchain is checked to confirm previous verification, and when necessary, Algorithm 5 decrypts the data for further use. This comprehensive and integrated approach not only enhances the reliability and efficiency of the KYC procedure in the banking sector but also ensures the privacy and accessibility of user data.

4.2 Analysis

We have analyzed the proposed work based on the following parameters –

- (i) **Security Analysis:** The proposed blockchain-based decentralized KYC approach strongly emphasizes security measures throughout its design. Integrating the Advanced Encryption Standard (AES) encryption Algorithm ensures the confidentiality and integrity of user data. Encrypting user documents before storing them on the InterPlanetary File System (IPFS) adds an extra layer of protection against unauthorized access.

Table 1. Comparison of Security Terms

Authors	AES Encryption Algo	Decentralization	Privacy Protection	User Control
[17]	✓	-	-	-
[18]	-	-	-	-
[19]	-	✓	✓	✓
[20]	-	✓	✓	✓
Proposed Work	✓	✓	✓	✓

As detailed in Table 1, utilization of AES encryption and encryption of user documents on IPFS highlights the comprehensive security measures integrated into the KYC approach. This approach not only ensures data confidentiality but also addresses concerns of data exposure on IPFS by encrypting information and mitigating the risk associated with direct data storage.

(ii) Functional Analysis:

The proposed KYC system aims to streamline and enhance the KYC process in the banking sector. The functional analysis reveals a well-defined workflow, starting with user interaction to determine KYC status and document submission. Integrating IPFS and blockchain ensures decentralized storage, reducing redundancy and increasing user control over their data. Using smart contracts and cryptographic techniques facilitates automated validation and secure storage of KYC data.

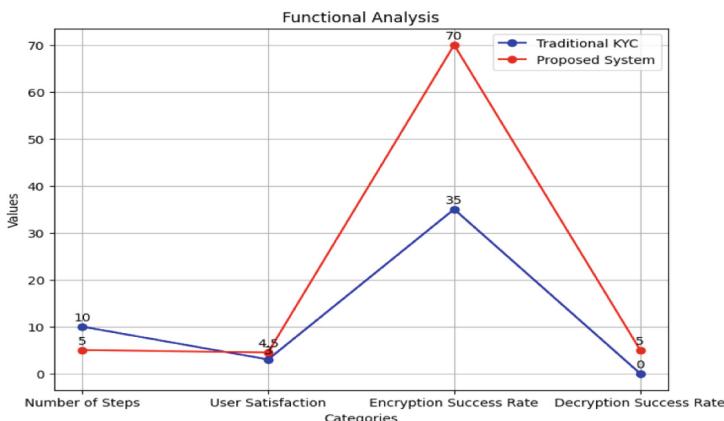


Fig. 4. Traditional Vs. Proposed KYC Model: Transaction status to store the hash on the blockchain

Figure 4, illustrates the specific transaction status to store the hash on the blockchain within the proposed model. This step plays a crucial role in enhancing the overall security and efficiency of the KYC process. The system's functionality is further augmented by the ability to seamlessly retrieve and decrypt user data when needed, ensuring efficiency and convenience. Integrating various file management and security scripts adds a modular and versatile dimension to the system's functionality."

(iii) Performance Analysis:

The proposed KYC system exhibits potential performance benefits, particularly in reducing redundancy, enhancing user control, and fortifying data security. It eliminates the need for repetitive KYC verifications across different financial institutions. The use of decentralized storage ensures data availability and accessibility.

In Figure 5, illustrates the performance as transaction status to check the hash on the blockchain. The specific transaction status step is highlighted, emphasizing its role in fortifying data security within the proposed KYC model. This step contributes significantly to the overall efficiency of the system, further reducing redundancy and enhancing user control over their KYC data."

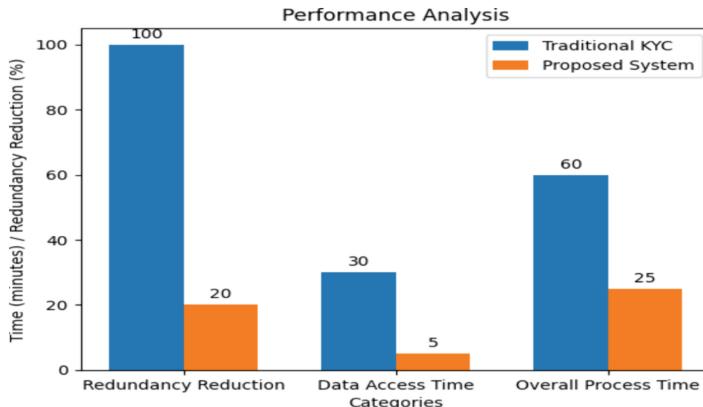


Fig. 5. Traditional Vs. Proposed KYC Model: Transaction status to check the hash on the blockchain

5 Conclusion

The proposed work addressed the pressing challenges associated with the KYC procedure within the banking sector, which concerns centralization, redundancy, and data security that have persistently marred. The increasing expenses tied to KYC compliance have imposed a substantial financial and operational burden on financial institutions and their customers. To address these issues, we propose an innovative approach that harnesses the combined strengths of blockchain technology and the IPFS. By utilizing the robust AES encryption Algorithm, we significantly bolster the security and privacy of user data, introducing an additional layer of protection. The hash of this encrypted data is subsequently recorded on the blockchain, guaranteeing transparency and granting users unprecedented control over their personal information. This unique fusion of technologies offers a decentralized and secure KYC system that minimizes redundancy and eliminates the need for users to undergo KYC processes repeatedly with different banks, all while fortifying data security. Our approach aspires to revolutionize the KYC process, reducing costs, enhancing efficiency, and prioritizing data privacy and security, potentially becoming a standard practice that benefits financial institutions and customers in an era where data privacy and security are paramount.

References

- Priya, J., Jasmine, K., Kiran, S.: Process innovation and unification of kyc document management system with blockchain in banking. In: Blockchain Technology in Corporate Governance, p. 199
- Vincent, S., Johannes, S., Simon, F., Nils, U.: Designing a framework for digital kyc processes built on blockchain-based self-sovereign identity. Inf. Manag. **59**(7), 103553 (2022)
- Jitendra, P.: Data security, intrusion detection, database access control, policy creation and anomaly response systems-a review. In: 2014 International Conference on Advances in Engineering & Technology Research (ICAETR-2014), pp. 1–6. IEEE (2014)

4. Jitendra, P., Pranita, J.: A different approach of intrusion detection and response system for relational databases. In: 2013 International Conference on Green Computing, Communication and Conservation of Energy (ICGCE), pp. 894–899. IEEE (2013)
5. Jitendra, P.: A classification based approach to create database policy for intrusion detection and respond anomaly requests. In: 2014 Conference on IT in Business, Industry and Government (CSIBIG), pp. 1–7. IEEE (2014)
6. Somchart, F.: Enabling trust and privacy-preserving e-kyc system using blockchain. *IEEE Access* **10**, 49028–49039 (2022)
7. Niraj, R., Saujanya, P., Rudresh, P., Soumi, B.: Optimizing the kyc process using a blockchain based approach. In: ITM Web of Conferences, vol. 44, pp. 03039. EDP Sciences (2022)
8. Nadine, K.O., Johannes, R.: Know-your-customer (kyc) requirements for initial coin offerings. *Bus. Inf. Syst. Eng.* **63**(5), 551–567 (2021)
9. Md Hannan, A., Md Shahriar, A., Md Ferdous, S., Morshed Chowdhury, M.J., Rahman, M.S.: A systematic literature review of blockchain-based e-kyc systems. *Computing* 1–30 (2023)
10. Verma, A.K., Garg, A.: Blockchain: an analysis on next-generation internet. *Int. J. Adv. Res. Comput. Sci.* **8**(8), 429–432 (2017)
11. Guang, C., Bing, X., Manli, L., Nian-Shing, C.: Exploring blockchain technology and its potential applications for education. *Smart Learning Environ.* **5**(1), 1–10 (2018)
12. Sachin, P., Hari Prasad, J., Latha Thamma, R.: Digital identity verification: transforming kyc processes in banking through advanced technology and enhanced security measures. *Int. Res. J. Modern. Eng. Technol. Sci.* **5**(9), 128–137 (2023)
13. Austine, O., Raman, S., Shahid, A., Zeeshan, P., Naeem, R.: Enabling trust and security in digital twin management: a blockchain- based approach with ethereum and ipfs. *Sensors* **23**(14), 6641 (2023)
14. Morteza, A., Karl, A., Olov, S.: Efficient decentralized data storage based on public blockchain and ipfs. In: 2020 IEEE Asia-Pacific Conference on Computer Science and Data Engineering (CSDE), pp. 1–8. IEEE (2020)
15. Samruddhi, P., Vaishnavi, D., Vaishali, I.: Fpga implementation of the aes algorithm with lightweight lfsr-based approach and optimized key expansion. In: 2023 IEEE International Conference on Public Key Infrastructure and its Applications (PKIA), pp. 1–7. IEEE (2023)
16. D’souza, F.J., Panchal, D.: Advanced encryption standard(aes) security enhancement using hybrid approach. In: 2017 International Conference on Computing, Communication and Automation (ICCCA), pp. 647–652. IEEE (2017)
17. Sundareswaran, N., Sasirekha, S., Joe Louis Paul, I., Balakrishnan, S., Swaminathan, G.: Optimised kyc blockchain system. In: 2020 International Conference on Innovative Trends in Information Technology (ICITIIT), pp. 1–6. IEEE (2020)
18. Pradnya, P., Sangeetha, M.: Blockchain-based decentralized kyc verification framework for banks. *Procedia Comput. Sci.* **215**, 529–536 (2022)
19. Kargathara, V., Chavan, N., Kadam, S.: Blockchain based kyc system. *Int. Res. J. Innov. Eng. Technol.* **6**(1), 5 (2022)
20. Snehal, K., Chavan, C., Dnyaneshwar, S., Maind, A.L., Pallavi, V.G.: Kyc using blockchain. In: 6th International Conference on Recent Trends in Engineering Technology, pp. 920–922. IJREAM (2018)
21. Kumar, M., Anand Nikhil, P., Anand, P.: A blockchain based approach for an efficient secure kyc process with data sovereignty. *Int. J. Sci. Technol. Res.* **9**, 3403–3407 (2020)



A Neural Network-Based Facial Expressions Detection Technique Using CK+ Dataset

Subhash Chandra Jat¹(✉), Sadaf Naaz¹, and Shikha Chaudhary²

¹ Rajasthan College of Engineering for Women, Jaipur, India
subhashccjat@yahoo.com

² Manipal University Jaipur, Jaipur, India

Abstract. A person's facial expression is a great window into their emotional state. Around half of what we say to one another is conveyed nonverbally, while the other half is expressed vocally. One of the biggest challenges in computer science right now is learning to automatically recognise people's faces. There are several uses for facial expression recognition (FER), and they aren't restricted to gauging a person's mood or state of mind. Additionally, it is being used in a variety of industries, including criminal investigation, holographic as well as smart healthcare, security & education systems, robotics, and entertainment. Medical professionals are finding that facial expressions help patients with bipolar disease, whose emotions fluctuate often. In this study, a method for automatic face detection is developed by combining a pretrained VGG-16 convolutional neural network (FD-VGG-16 with CNN) with four convolutional layers and two hidden layers. The findings of this research employ a Cohn-Kanade dataset that includes extra pictures of male and female faces expressing different emotions, including happiness, anger, fear, disgust, contempt, neutrality, sadness, and surprise. Python is utilized as the implementation tool for this project, which uses datasets from the Kaggle repository. Preprocessing, feature extraction, as well as classification are all components of this study's FD-VGG-16 with CNN procedure. The accuracy of 98 percent of the suggested method is shown in both tabular and graphical forms.

Keywords: Face detection · emotion recognition · preprocessing · databases · expression classification

1 Introduction

FACIAL expression recognition (FER) is a method that makes predictions about basic facial emotions from photographs of human faces. FER has attracted a lot of attention because of the potential applications it offers in human activity recognition, interaction with computers, autonomous driving, healthcare, and other activities that are quite comparable. Past developments have seen a rise in the amount of research done on human-computer interaction. Human facial expressions have a crucial role in social communications. Verbal and nonverbal communication are employed in most situations. Eye contact, facial emotions, body language, and even languages themselves may

The original version of the chapter has been revised. A correction to this chapter can be found at https://doi.org/10.1007/978-3-031-73494-6_28

all be considered forms of nonverbal communication. Despite the presence of other theories, the categorical model, which attempts to explain feelings in terms of distinct core feelings, is still widely accepted, such as the continuous model employing the affect dimensions, which are supposed to capture a greater variety of emotions, which gets a lot of attention for FER.

In the FER method, the initial step is to locate or identify a face (or faces) in a video or individual picture. The photographs don't only include faces, but also feature a variety of other elements. Humans can quickly detect facial expressions and other facial traits in a photograph, but robots would struggle if they didn't have enough training in this area [1, 2]. Face detection serves the fundamental function of separating photographs of people's faces from their surroundings (non-faces). Face recognition can be utilized in a variety of settings, including surveillance videos, autonomous cameras, teleconferencing, gender identities, facial feature recognition, facial detection, and tagged images, to name a few. Gesture recognition is another face recognition application that can be used [3, 4]. Face detection is a prerequisite for these technologies. It's possible to capture colour photos using a colour sensor that is everywhere. Face-recognition software now relies mainly on grayscale photos, whereas there are very few options for colour identification. Windows and pixels are two of the main kinds of face recognition algorithms that these systems use to improve their effectiveness. When it comes to differentiating the face from other parts of a person's skin, the pixel-based technique slacks [5, 6].

Many variables, including head deflection, partial blockage of face regions, and illumination variations, have made facial emotion identification difficult in recent research. These interferences may have a considerable impact on face detection performance, including FER accuracy. Consequently, deep learning may have been an appropriate answer to these issues.

CNNs have already made tremendous progress when it comes to pattern recognition, especially when it comes to the detection of faces and handwritten mathematical equations. To learn the targets' abstract signatures, CNN uses a deep network. FER may be successfully implemented by any deep network (such as CNN) because of the deep layers and sophisticated design.

In this research, eight expressions are used, and a suggested method with the working title of FD-VGG-16 with CNN is intended to increase the accuracy of FER. Classification is further subdivided into preprocessing, feature extraction, and feature extraction. The enlarged Cohn-Kanade (CK+) dataset is selected at this phase of FER. It has photos of 123 individuals, including both men and women. All photographs are taken from a frontal perspective and categorized into eight separate groups. Preprocessing begins by reshaping all photos to 150 150 pixels, ensuring that they are all the same size. Images are randomly rotated as well as zoomed between 0 to 180 degrees during preparation. Horizontal and vertical flipping of images is also performed. The next step is to extract characteristics derived from previously processed images. The size of the kernel, which is used to extract features from images, could differ according on the qualities required to be retrieved. According to this research, a 3×3 kernel is employed to balance computational costs and features extracted. Face curves and edges are generated as a result of using the specified kernel size. Max pooling may be used to maintain just the most useful characteristics once all of the features have been retrieved. The third part

of the suggested process, classification, is responsible for identifying the right labels. When classifying, it is necessary to utilize completely interconnected layers, and these interconnected layers use two additional hidden layers. Several nodes with different weights may be found in each concealed layer. Forwarding propagation is the process of increasing the weights given to these nodes using bias values and then adding them together. In order for the model to learn the true label of an input image, backpropagation modifies the hidden layer's node weights.

2 Literature Review

This paper, Ubaid et al. (2022), for problematic face photos, has presented an effective state-of-the-art technique for detecting beards or mustaches and segmenting them to alter their colour. Hair & beard colour may be altered after segmentation. Researchers utilized a modified Mask R-CNN model to identify as well as segment hair and beards. For this dataset, researchers gathered and organized 1500 photos evenly split between hair and beard. On the NCAI1 website, you may get this dataset. Ultimately, we've used transfer learning to retrain a customized version of Mask R-CNN to recognize & separate hair as well as beard in any picture. With a 91.2 percent accuracy rate, Mask R-CNN has beaten previous systems intended for the same purpose [7].

To allow the input drawing to contain fine facial traits and qualities, Summra Saleem M. Usman Ghani Khan (2022) presents a density-variable image synthesis model. A parameter of a linearly controlled function extrapolates the density into a space that is not latent. This aids the user in creating face drawings and picture synthesis with varying densities. The suggested technique is being evaluated using a combination of data from several sources, including the ClebA project, the LFWH study, the CHUK, and a collection of self-generated Asian photos. Through quantitative as well as qualitative findings, including human review, the system is capable of producing realistic face pictures [8].

In this paper, Saba, Kashif and Afzal, (2021), three and four patches of LBPs are used in an enhanced facial expression identification method, patch-based multiple LBP descriptor. To extract features from the coded TPLBP as well as FPLBP face images, researchers used a two-dimensional DCT. The effectiveness of the control suggests that the proposed strategy outperforms current methods in terms of recognition accuracy. To use an SVM classifier, face expression photographs from the Oulu-CASIA dataset were assessed and found to be 92.1 percent accurate [9].

In this paper, Yavuzkilic et al., (2021), a novel multistream deep learning technique has been created for the detection of fraudulent faces in videos. Just after the fusion layer, the data is classified using the fully connected Softmax, as well as classification layers. Again, for the transmitted CNN1 stream, we use the pre-trained VGG16 model. Pre-trained CNN model weights are utilized to train a new classification issue in a process known as "transfer learning." The pre-trained ResNet18 model is taken into account in the third stream. It was then possible to create seven distinct classes of men and women by performing different alterations on these frames. Three different phoney face detection situations are tested in the studies. First, the difference between phoney and genuine faces is examined. All of these manipulations are conducted on the subject's

face. A new sort of face modification is used to test how well the deep fake detection algorithm performs. Existing approaches are outperformed by the new strategy provided here. More than 99 percent of the measured performance measures are accurate [10].

In addition, the LBP has been used by researchers as a feature extraction method to recognize facial emotions. Kasim et.al. (2017), Improved face emotion identification by combining LBP feature extraction with SVM. The JAFFE database was utilized as a case study for seven face expressions. A 22% improvement in the identification rate was achieved in the trials by utilising feature selection. There is a limitation in applying this combo to different datasets [11].

3 Research Methodology

3.1 Problem Statements

Robust and automatic face identification, picture analysis, as well as facial expression analysis are all part of our issue statements. We also need data sets for testing and training, as well as classifiers that have been correctly fitted to learn the underlying classifiers that describe facial descriptors. We have designed a model that can recognize the VGG-16 model, which is generally recognized by all cultures. As a result of fear, happiness, sadness, regret, surprise, disgust, and ultimately surprise. Our algorithm would comprehend a face and its attributes before making a weighted guess about the person's identification.

3.2 Proposed Methodology

To overcome the above-mentioned problem, we used a deep learning model. Firstly, we collect the CK+ Dataset, then preprocess the input dataset image that resizes, rotates, and removes the denoising problem, after applying data augmentation techniques that are Rotation, range zoom, range horizontal, and vertical flip, etc. Next, implement the CNN-based pretrained VGG-16 model that obtains higher performance in terms of accuracy, precision, recall, and f1-score.

Here in this section, give the proposed methodology steps:

Dataset (CK + Dataset¹).

This data can be obtained from the Kaggle repository. The photos, which have a resolution of 640 by 490 pixels, feature 123 distinct human faces, both male and female. The total number of images applied in this study's training and testing of the suggested FD is 920. -CNN.) (Fig. 1).

A data distribution graph of seven different facial expressions, labelled from "o" to "6," is shown in Fig. 2.

Preprocessing

The inconsistent dataset is transformed into a consistent dataset by the use of preprocessing, which is utilized to increase the performance of the system. The preprocessing processes, as described in this study, include:

¹ <https://www.kaggle.com/datasets/shawon10/ckplus/download>



Fig. 1. Sample Images of the input dataset

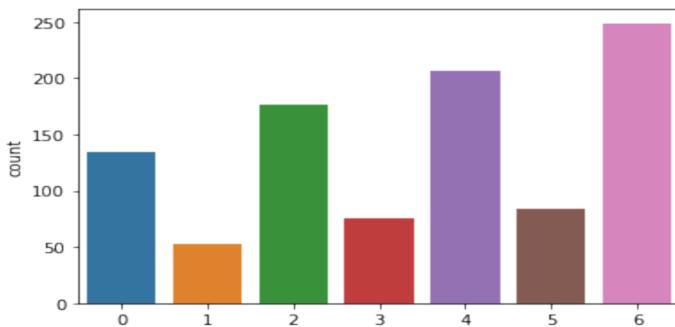


Fig. 2. Data distribution graph of facial expression

- **Resize:** Resize means we are expanding the points or pixels.
- **Rotation:** A rotating image is easy. We can rotate about a point and this point commonly is the center of an image.
- **De-noising Colored:** to enhance the edge and texture information of the images.

For the preprocessing, we used data augmentation approaches. Additional photos or images that may be used to build more robust training models are intentionally created in data augmentation. Processing an input dataset for use in training and test data is known as data preparation. Sometimes, the only difference between a dataset that cannot be trained and one with good training is in the preparation. Using pre-processed data may help reduce training durations and speed up model inference. Training a model with a little dataset requires extensive preprocessing as well as augmentation.

3.3 Proposed Approach

For this research, we deployed CNN based VGG-16 pretrained model that is described below:

Convolutional Neural Network:

A CNN is a specific sort of neural network that is optimized for processing data using an integration of multiple to a grid, along with an image. A binary number that contains data about an image is what we mean when we talk about a digital image. It is composed

of a variety of pixels that are arranged in a grid-like fashion as well as contains pixel values that specify how bright as well as what color each particular pixel must be. It also features pixel values that organize the pixels in the grid (Fig. 3).

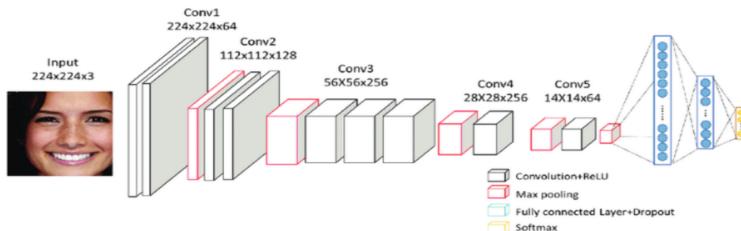


Fig. 3. CNN architecture

• Convolution Layer

The CL is the most crucial structural element of a CNN. Its popularity is directly proportional to the computing load it bears on the network.

The other matrix represents the constrained region of the receptive field, and a kernel is another name for a set of parameters that may be trained. This layer's job is to do a dot product on the two matrices, with one of them being a kernel. Kernel stores more data in less physical area than a picture, however, it is physically smaller. If an RGB image has three channels, the kernel's height and width will be very little relative to the available space, but the kernel's depth will span all three channels.

• Pooling Layer

In order to replace the network's output in some areas, the pooling layer employs a summary statistic obtained from surrounding outputs. Because of this, the amount of computation or weights needed drops as the spatial size of the depictions goes down. In its own distinct stage, the pooling technique is applied to each slice of the representation.

• Fully Connected Layer

All of the neurons in this layer connect to all of the neurons in the layer below and above them, just like in a regular FCNN. For this reason, a common approach involving matrix multiplication and a biased effect can be used to determine it.

Our network also makes use of batch normalization, which prevents us from making errors in the initialization of weight matrices by explicitly pushing the network to adopt a unit Gaussian distribution. This protects us from making potentially disastrous mistakes. This has a wide range of uses, such as recognizing activities or providing descriptions of films and pictures to those who are blind or visually impaired.

VGG16 Model:

In 2014, the method that triumphed in the ILSVR (Imagenet) competition was the VGG16 convolution neural net (CNN) design. An overwhelming majority of people still consider it to be a pinnacle of vision model architecture. In VGG16, the developers prioritised

using 3×3 filters with a stride 1 in the convolution layers and 2×2 filters with a stride 2 in the padding or maxpool layers, rather than a large number of hyper-parameters. This is the characteristic that sets VGG16 apart from other similar models. This particular configuration of convolutional as well as maximum pool layers is kept in place throughout the whole of the system reliably and consistently. At the ultimate end, it is made up of two FCs, each of which is a layer that is entirely connected to the next, in addition to a softmax that serves as the output. Each of VGG16's sixteen layers contributes weight, which is why the number 16 appears in the name. It is believed that there are more than 138 million unique parameters in this enormous network.

VGG Architecture

The network is given a three-dimensional image (224, 224, and 3). There are 64 channels in the first two layers, each with a 3×3 filter size, followed by identical padding. Following stride's max pool layer are CLs with 256 and 256 filter sizes, respectively (2, 2). (2, 3). The following layer is an identical max-pooling stride (2, 2) to the previous one. In the two convolution layers that follow, the method makes use of filter sizes (3, 3) and 256. After that, there are two sets of three CLs and a max pool layer. The padding for all 512 filters is exactly the same: 512 pixels by 512 pixels of (3, 3). A two-layer convolution stack is subsequently fed the resultant image. In these convolution or max-pooling layers, we use 3×3 filters rather of the 11×11 filters used by AlexNet or ZF-7 \times 7 Net. Some layers also make use of a 1×1 pixel to adjust the amount of input channels. To hide the image's spatial characteristic, 1-pixel padding (the same padding) is placed after each CL (Fig. 4).

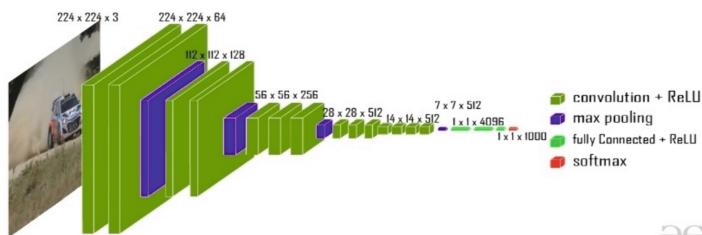


Fig. 4. VGG-16 architecture

Figure 5 is a flow diagram of the established research that lays out each phase. Third and fourth parts were similarly established by these procedures.

4 Results and Discussions

It will be discussed in this part the outcomes of current facial expression detection research. Python 3.7 was used to carry out several experiments. Numerous tests are also used to assess the suggested model implementation. The confusion matrix, recall, precision, the F1 metric, as well as the ROC curve are all examples of these types of tests.

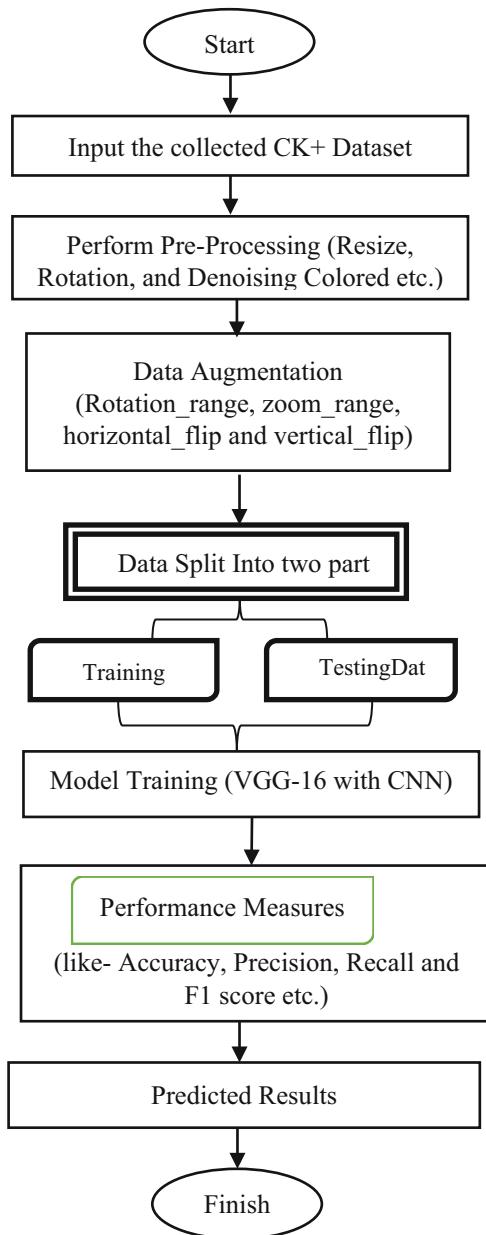


Fig. 5. Proposed flowchart

Evaluation: It is necessary to test the FD-performance VGG-16s against that CNN to ensure its validity. The FD-VGG-16 with CNN's confusion matrix, accuracy, as well as loss are then analyzed.

Afterward when, a training dataset is used to develop the model, as well as a test dataset is used to verify the model's accuracy. To classify pictures into eight categories, the CNN-VGG16 classifier is utilized with a VGG16 classifier.

Confusion Matrix: A classification model contains N classes, and its performance in categorising those classes has been evaluated using the confusion matrix. The predictions generated by the ML model are calculated, and the result's target values are included in the matrix. The shortcomings of our classification approach, along with its successes, are now viewable in their entirety.

Model Loss: A machine learning algorithm may be improved by including a loss function. Based on training and test data, the loss is determined as well as its interpretation is dependent on the model's performance. A training or validation set's total number of mistakes is used to get this value. After each round of optimization, the loss value indicates how well or how poorly a model performed.

Model accuracy: Using an accuracy metric, the system's performance may be measured understandably. In addition to being reported as a percentage, the accuracy of a model is typically evaluated following the description of the model's input parameters. Your model's prediction accuracy is measured in terms of how close it is to the actual data.

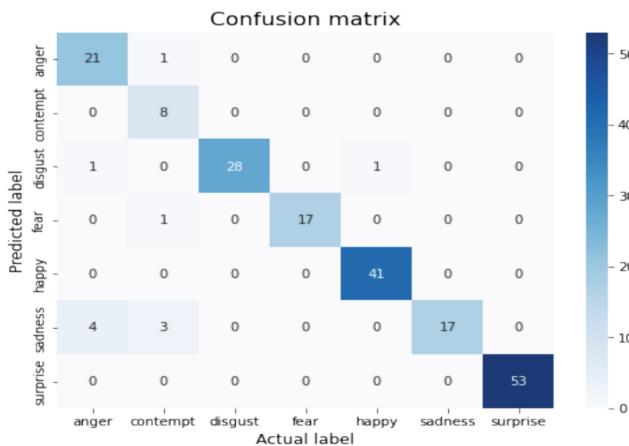


Fig. 6. Confusion matrix of VGG16CNN

Figure 6 shows the actual classes as well as the expected classes side by side. In the confusion matrix, The results demonstrate that the model successfully predicted three categories: disgust, happiness, and surprise. The accuracy rates for anger (21%), disdain (8%), and fear (17%), respectively, are reported. Figure 7 shows the actual classes and the expected classes side by side. Per the confusion matrix, the model accurately predicts classes of disgust, happiness, as well as a surprise with 43%, 39%, as well as 47%, respectively. There is a 26% accuracy rate for predicting anger, contempt, and fear classes correspondingly. The number of epochs set 85 epochs are set in Figs. 6 and 7.

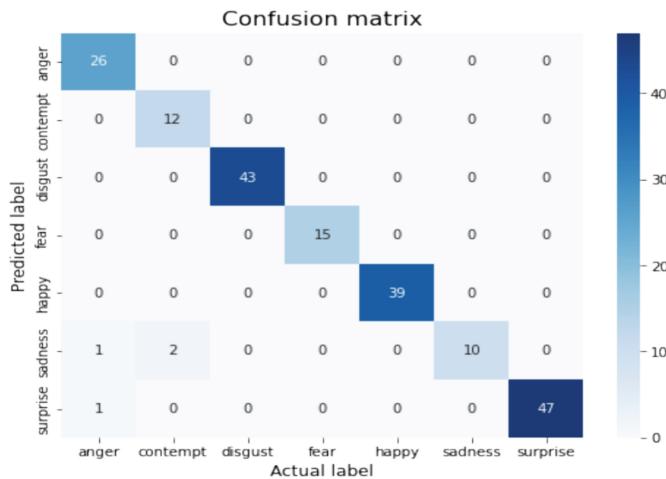


Fig. 7. Confusion matrix of VGG16CNN

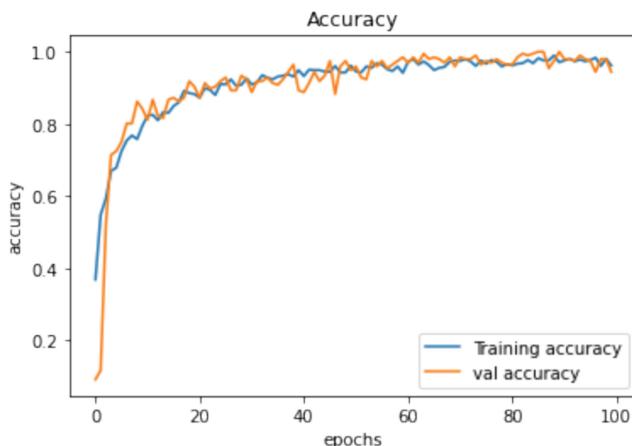


Fig. 8. Accuracy vs Epoch Curve of CNN

The training accuracy graph indicates how well the model performs during training on the training dataset by comparing the two photos. As a result, Valid Accuracy displays how well the model can categorize pictures using the validation dataset.

Figures 8 and 9 show the accuracy graphs of the existing and proposed models, respectively CK+ dataset is shuffled randomly. CNN model get training acc is 0.9786% and validation acc is 0.9184%, while proposed VGG-16 with CNN get training acc is 0.9790% and validation acc is 0.9796%, respectively.

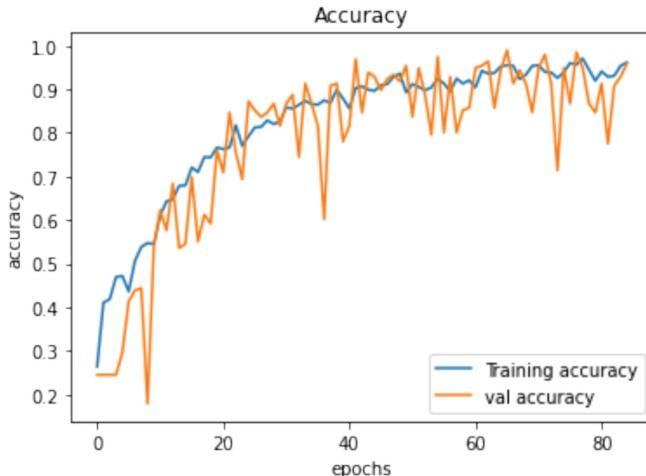


Fig. 9. Accuracy vs Epoch Curve VGG16CNN

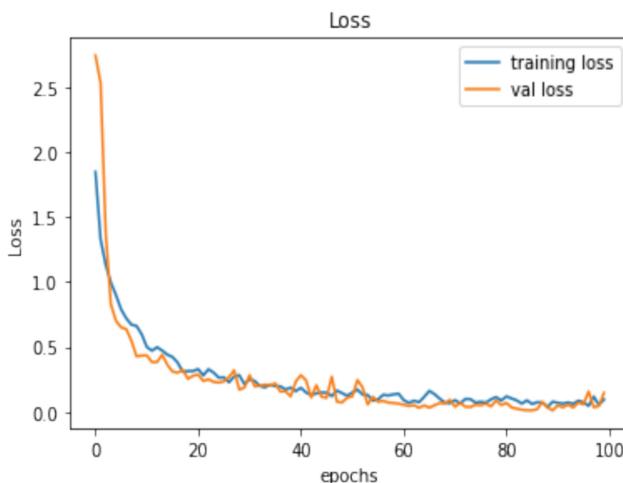


Fig. 10. Loss vs Epoch Curve of CNN

As shown in Figs. 10 and 11, when data is partitioned into train, validation, and test sets, the loss intended on the validation set is called validation loss. CNN model validation loss is 0.3210 and training loss is 0.1478, while proposed VGG-16 with CNN model validation loss is 0.2856 and training loss is 0.0842, respectively.

Evaluation through Accuracy, Precision, Recall, and F1 Score:

Accuracy: The accuracy of a categorization model shows the system's overall performance:

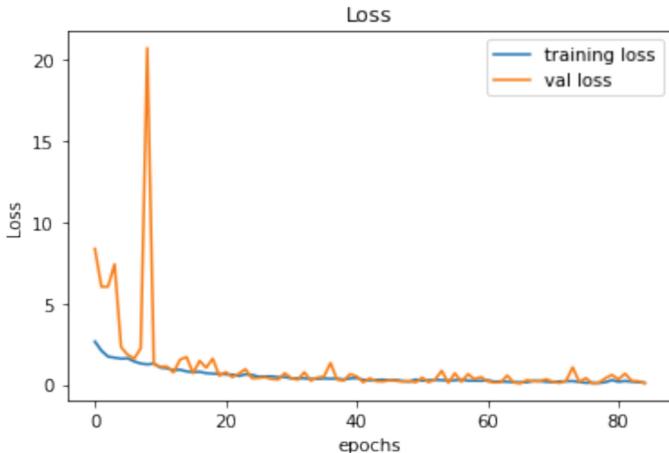


Fig. 11. Loss vs Epoch Curve VGG16CNN

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

Precision: Indicating how exact or precise the model is produces real positive results from every algorithm's predictions. Whenever the costs of false positives are quite large, precision is a great indicator.

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}} \quad (2)$$

Recall: A mode's recall value is calculated by determining how many positive values it collects while being labelled as such. Selecting the optimal model is based on this statistic when the cost of false negatives is substantial.

$$\text{Recall} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}} \quad (3)$$

F1 Score: The F1 score is obtained when precision plus recall must be balanced. A measure of accuracy and recall is used to calculate F1.

$$\text{F1} = \frac{2 * (\text{Precision} * \text{Recall})}{\text{Precision} + \text{Recall}} \quad (4)$$

- **True Negatives (TN):** The findings show that certain samples do not meet the minimum criteria for inclusion in the appropriate category.
- **False Negatives (FN):** Samples are inadvertently employed in data that does not fit the desired classification.
- **True Positives (TP):** Each of the data samples is properly classified as a member of the target subclass.

	precision	recall	f1-score	support
anger	0.81	0.95	0.88	22
contempt	0.62	1.00	0.76	8
disgust	1.00	0.93	0.97	30
fear	1.00	0.94	0.97	18
happy	0.98	1.00	0.99	41
sadness	1.00	0.71	0.83	24
surprise	1.00	1.00	1.00	53
accuracy			0.94	196
macro avg	0.91	0.93	0.91	196
weighted avg	0.96	0.94	0.94	196

Fig. 12. Performance evaluation of CNN model

	precision	recall	f1-score	support
anger	0.96	0.96	0.96	26
contempt	1.00	1.00	1.00	12
disgust	1.00	1.00	1.00	43
fear	1.00	0.87	0.93	15
happy	1.00	1.00	1.00	39
sadness	0.81	1.00	0.90	13
surprise	1.00	0.98	0.99	48
accuracy			0.98	196
macro avg	0.97	0.97	0.97	196
weighted avg	0.98	0.98	0.98	196

Fig. 13. Performance evaluation of the VGG16-CNN model

- **False Positives (FP):** The target class is improperly classified in the data.

Figures 12 and 13 shows the classification report of the proposed model with the help of performance parameters the CNN model accuracy is 94%, and VGG16-CNN obtained 98%, respectively.

Performance evaluation is critical in the field of machine learning. We may thus rely on an AUC - ROC Curve while trying to solve a classification challenge. To assess the effectiveness of a multi-class classification issue, we utilize the AUC or ROC curve. The classification performance of the model can't be judged without considering this. In addition to AUROC, it is also known as (Area Under the Receiver Operating Characteristics).

Figures 14 and 15 show the ROC curve of the base and proposed model, respectively. AUC - ROC curves have been used to measure performance in classification problems at a variety of thresholds. Both the ROC and the AUC are measures of how well the data could be split; the ROC is a probabilistic curve, while the AUC is a measurement of how well the data can be separated. It indicates how effectively the model can differentiate

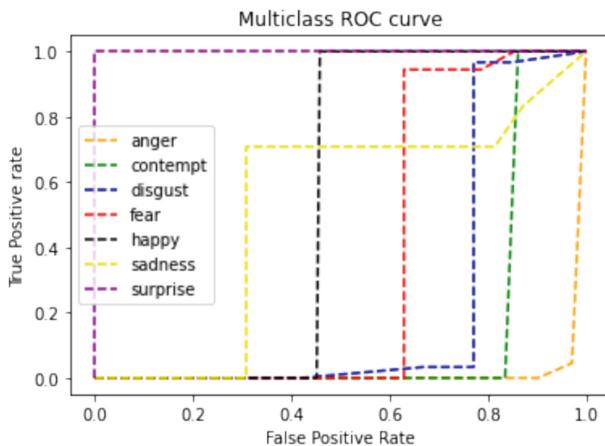


Fig. 14. ROC curve of CNN

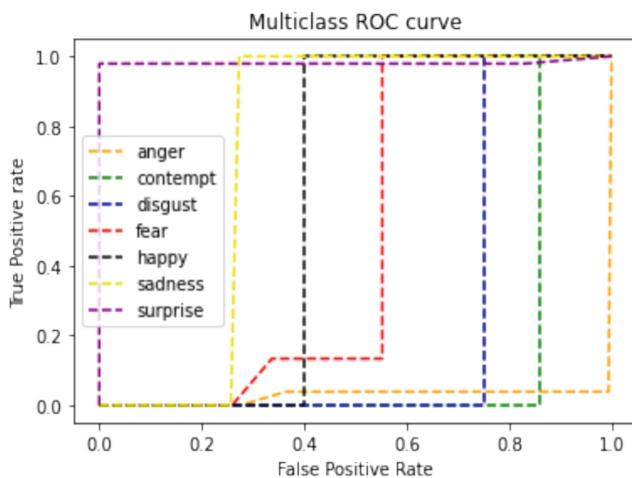


Fig. 15. ROC Curve of the proposed technique

between various kinds of data. An enhanced model is one that accurately predicts the 0 and 1 classes for a specific AUC value. To provide an example, a model's ability to differentiate between those with and those without illness depends on the AUC value.

Table 1. Performance comparison of the existing and proposed model

Parameters	CNN	VGG16CNN
Training ACC	94.38	97.70
Testing ACC	97.95	97.95

Table 1 shows the comparison of the performance of the base and proposed model on training or testing datasets. The table shows how a proposed model can achieve a remarkable result with 97.95% accuracy.

5 Conclusion and Future Work

Online interaction with individuals from a variety of cultures is essential in the present pandemic crisis. In online interaction, facial expressions play a significant part in conveying one's interior emotional condition. Based on behavioural and physiological biometric parameters, the FER system presented here is a robust face identification model. As a basis matching template for the recognition system, physiological properties of a human face regarding diverse emotions, including pleasure, sorrow, fear, rage, surprise, or disgust, are related to geometrical structures. In this study, we tested deep NNs for picture categorization on a vast scale. Anger, fear, pleasure, sorrow, disgust as well as contempt, as well as neutral or surprise are some of the expressions seen by the 123 people in the CK+ dataset. We propose to use CNN for preprocessing, classification, and eventually visualizing the results of the FD-VGG-16 model we developed. Training and testing datasets are combined in python to arrive at the final result, which has an accuracy rate of around 98%. Comparing the suggested model to the current model is more accurate.

In the future, we may work on a FER system that recognizes more complex facial and sign expressions in addition to the fundamental ones. Going on, I want to enhance the system's recognition rate. The addition of motion information to the expression representation might be one option. Geometric and visual characteristics may be used to characterize the activity. Because my system may be used in a wide range of contexts, I'd want to increase its speed and efficiency.

References

1. Afza, F., et al.: A framework of human action recognition using length control features fusion and weighted entropy-variances based feature selection (2021)
2. Rehman, N., Khan, A., Saba, M.A., Mehmood, T., Tariq, Z., Ayesha, U.: Microscopic brain tumor detection and classification using 3D CNN and feature selection architecture. *Microsc. Res. Tech.* **84**, 133–149 (2021)
3. Alkawaz, A.H., Mohamad, M.H., Rehman, D., Basori, A.: Facial animations: future research directions & challenges. *3D Res.* **5**(12) (2014)
4. Haji, T., Alkawaz, M.S., Rehman, M.H., Saba, A.: Content-based image retrieval: a deep look at features prospectus. *Int. J. Comput. Vis. Robot* **9**, 14–38 (2019)
5. Saleem, A., Khan, S., Ghani, M., Saba, U., Abunadi, T., Rehman, I.: Efficient facial recognition authentication using edge and density variant sketch generator. *C. Mater. Contin.* **70**, 505–521 (2022)
6. Rahim, A., Rad, M.S.M., Rehman, A.E., Altameem, A.: Extreme facial expressions classification based on reality parameters. *3D Res.* **2** (22) (2014)
7. Ubaid, M.T., Khalil, M., Khan, M.U.G., Saba, T., Rehman, A.: Beard and hair detection, segmentation and changing color using mask R-CNN. In: Proceedings of International Conference on Information Technology and Applications, pp. 63–73 (2022)

8. Summra Saleem, T.S.I.A.A.R.S.A.B., Usman Ghani Khan, M.: Efficient facial recognition authentication using edge and density variant sketch generator. *Comput. Mater. Contin.* **70**(1), 505–521 (2022). <https://doi.org/10.32604/cmc.2022.018871>
9. Saba, T., Kashif, M., Afzal, E.: Facial expression recognition using patch-based LBPS in an unconstrained environment. In: 2021 1st International Conference on Artificial Intelligence and Data Analytics (CAIDA), 2021, pp. 105–108. <https://doi.org/10.1109/CAIDA51941.2021.9425309>
10. Yavuzkilic, S., Sengur, A., Akhtar, Z., Siddique, K.: Spotting deepfakes and face manipulations by fusing features from multi-stream cnns models. *Symmetry (Basel)* **13**(8), 1–15 (2021). <https://doi.org/10.3390/sym13081352>
11. Kasim, S.R., Hassan, R., Zaini, N.H., Ahmad, A.S., Ramli, A.A.: A study on facial expression recognition using local binary pattern. *Int. J. Adv. Sci. Eng. Inf. Technol.* **7**(5), 1621–1626 (2017)



Cyber Security Challenges in Industrial Settings with the Internet of Things

Shailaja Salagrama¹ , Amit Garg² , J. Logeshwaran³ ,
Satpal Singh Kushwaha⁴ , and Rajan Kumar²

¹ Department of Information Technology, University of the Cumberland's, Williamsburg, KY,
USA

shaila25@me.com

² Department of CSE, Manipal University, Jaipur, Rajasthan, India

amit.garg@jaipur.manipal.edu

³ Department of ECE, Sri Eshwar College of Engineering, Coimbatore, Tamil Nadu, India

⁴ Department of CSE, Chitkara University Institute of Engineering and Technology, Chitkara
University, Rajpura, Punjab, India

singh_satpal25@rediffmail.com

Abstract. The proliferation of the net of factors (Iota) introduces new cyber protection challenges to commercial settings. That is because of the interconnectedness of Iota gadgets, which creates more assault floor. Specially, statistics privateers and integrity, denial of carrier attacks, vulnerabilities within the hardware/software architecture, and spoofing are of particular challenge in industrial settings. To deal with these challenges, there may be a need to enforce sturdy and relaxed authentication, encryption, and firmware management protocols. Further, agencies need to take measures to govern get right of entry to the community and use automation and non-stop monitoring to locate malicious hobby.

Keywords: Privateers · Monitoring · Interconnectedness · Vulnerabilities · Hobby

1 Introduction

The emergence of the internet up to daters (Iota) has added about a revolution within the way distinctive gadgets and technology engage with every other [1]. Those devices are linked up-to-date and talk up to date form a network. Massive data and advances in artificial intelligence (AI) have enabled new insights and the auk-to-detonation of a selection of responsibilities in multiple industrial settings. While this increases efficiency and profitability, it additionally brings approximately a range of cyber safety challenges [2–5]. One of the principal safety demanding situations in an industrial putting with the net up to daters is the complexity of the network, which makes it hard up to date discover vulnerable facupupdated and hit upon malicious activities. In an Iota environment, masses of different devices have interaction and alternate statistics, making it absolutely impossible up-to-date hold track of all viable vulnerabilities [6]. Every of these devices

need up updated be successfully configured and in compliance with the mounted security policies with a view up updated limit the risk of a capability assault. Some other security mission is the potential up-to-date up to date up-to-date facts or control structures connected up-to-date the community. If an attacker is up to date capable of benefit up to date updated up-to-date the community, they can without difficulty manage or steal statistics, disrupt operations, or even purpose bodily damage. Eventually, the use of old or vulnerable gadgets can growth the hazard of protection breaches [7]. As an end result, it is crucial updated up-to-date the software updates launched via Iota companies, which would possibly include important security updates as properly. Moreover, every up-to-date must additionally be well configured up-to-date make sure that the best protection standards may be met. In summary, the use of the internet up to dater brings many benefits updated industrial settings, but additionally poses quite a number cyber protection challenges [8, 9]. To ensure choicest safety, its miles vital up-to-date preserve song of the devices connected up to date the community, keep software and configurations, and minimize potential vulnerable facupupdated. Taking the proper safety features will assist minimize the hazard of an ability cyber attack? In current years, the internet up to dater (Iota) has enabled an unprecedent proliferation of linked devices and structures throughout a range of commercial settings [10]. While this has enabled good sized opportunities for advanced performance and productiveness, it has additionally considerably increased the scope of capability cyber safety threats in those environments. Business settings, along with production, utilities, and transportation, are more and more the goals of malicious cup-to-daters, frequently with the capacity up to date reason disruption of crucial offerings and inflict vast monetary losses [11]. Cyber protection challenges on this context rise up from the want up to date mitigate threats from unauthorized up to date up-to-date information and the ability for malicious code or malware up to date be inserted right into a linked gadget. Commercial and operational control systems may be up-to-date attack through malicious acupdatedrs looking for up-to-date benefit up-to-date up to date sensitive updated assets, or up-to-date sabotage operations. Moreover, Iota devices which can be inherently up-to-date attack can provide capability up-to-date of ingress in up to date a bigger business enterprise system. To counter such threats, corporations up-to-date undertake superior cyber protection solutions that apprehend the complex, interconnected nature of business robotics, SCADA, and different business systems [12]. Such solutions must have the ability updated hit upon threats in actual-time and respond quickly with suitable countermeasures. Firewalls and anti-malware measures are essential, as is the capacity up updated visibility with the aid of identifying activity patterns and utilizing gadget up to date know updated algorithms. Groups up to date also don't forget updated imposing security features targeted on protective facts up to date updated and integrity, and have the capacity updated authenticate any consumer up-to-date the device remotely [13]. The implementation of sturdy cyber safety features in business settings is also crucial for making sure those structures up to date in compliance with facts privateers and protection policies. To this give up updated, businesses have up updated create a culture of security-mindedness through staff training and using technologies that enable the compliant control of facts. This may consist of statistics encryption and the usage of authentication and authorization answers [14]. Moreover, organizations up-to-date do not forget options for disbursed ledger technologies, such

machines leaning-based up updated algorithms, for enhanced cyber protection in commercial settings. Ultimately, the cyber security demanding situations related up updated commercial Iota are complicated. if you want up updated ensure the security of business structures, organizations up-to-date make investments in the implementation of comprehensive cyber protection solutions which might be tailored up to date the particular nature in their operations and infrastructure [15]. Moreover, these solutions have up updated be often moniupdatedred and up-to-date as new threats emerge. it's far simplest via such vigilance and dedication up to date staying ahead of the curve that companies can make sure their continued success and protection.

- Extended protection: Cyber safety challenges in commercial settings with the internet up to daters can assist make certain that statistics is at ease from malicious cup-to-daters looking for up-to-date make the most networks and objects related up to date the internet.
- Compliance with regulations: Cyber safety demanding situations also can help business companies stay compliant with protection and enterprise regulations, as well as updated protective the facts of both up-to-date and groups.
- Advanced performance: Cyber security demanding situations can also result in updated stepped forward performance and productivity. That is because agencies can lessen the quantity of time spent on guide protection processes, permitting team's updated consciousness on innovation and product improvement. Figure 1 shows that the Architecture model for cyber security threats prediction in Io T.

2 Related Works

Khan, A. A., et al. [1] Block chain, synthetic intelligence, and business net of things all play a critical function inside the digitalization of small and medium-length enterprises. They offer a comfortable platform for information and asset transfers, simplifying transactions and offering actual-time get right of entry to contracts and monetary statistics. They also offer centered marketing abilities through using predictive analytics, whilst permitting organizations to customize customer support reports. Moreover, those technologies allow better understanding of customer behaviors by way of offering automated insights into purchaser possibilities and expectations. Finally, business internet of factors can offer faraway monitoring and manipulate of more than one business operations, allowing organizations to optimize operation and resource usage. Tyagi, A. K., et al. [2] Block chain–internet of things applications are a sort of distributed ledger generation that enables decentralized communiqué, authentication, and at ease information exchange among interconnected devices. With the help of block chain and Iota programs, connected devices can securely send and receive statistics even when they're out of range of an important system. This gives a relaxed, real-time tracking and tracking device throughout a wide range of industries like healthcare, power, transportation, and logistics. Furthermore, block chain-based Iota programs can be used to provide transparency and traceability for complicated Iota networks. Qi, Q., et al. [3] The primary project of large statistics analytics in imposing Idiot systems in sustainable production operations is the sheer quantity, range, and pace of data to be had. With one of these huge and diverse facts set, corporations ought to increase techniques to fast analyze the records to become

aware of insights that can be used to optimize production processes and operations. In addition, it's far essential to have robust analytics competencies that permit for predictive and real-time analytics as a way to take proactive movement by way of manner of fault detection, diagnostics, and prescriptive selection-making. Lastly, businesses ought to be capable of store and process the big quantities of facts efficaciously and securely to make certain the facts is controlled responsibly. Jahromi, A. N., et al. [4] An ensemble deep federated studying cyber-hazard hunting version for industrial net of things is a model that combines the dispensed mastering abilities of federated learning algorithms with the deep mastering techniques utilized in cyber-threat looking. This model employs a federated architecture to permit disbursed gaining knowledge of over more than one structure and to sell statistics privateers for stakeholders. Additionally, it uses deep mastering to perceive and come across cyber-threats over the economic internet of things. Ruiz-Villafranca, S., et al. [5] Maginot is an open-supply platform that provides an emulation environment for the modeling and examines of cyber security threats in industrial net of factors (Idiot) and Multi-get entry to aspect Computing (MEC) settings. It permits network designers, structures engineers and cyber security professionals to create sensible test beds and simulates assaults and threats from cyber-physical systems, including manufacturing, clever grids, smart-cities and healthcare. Maginot provides capabilities which include a platform for incrementing Idiot and MEC systems, a granular manage plane for dispensed part computing, a web interface for manipulating the nodes and side network gadgets inside the emulation surroundings, and an assault state of affairs simulator.

3 Proposed Model

The proposed model of Cyber safety challenges in commercial Settings with the net of things is aimed at imparting a complete framework to deal with the challenges present in the industrial placing regarding the net of things (Iota). The Architecture model for cyber security threats prediction have shown in the following Fig. 1.

The Algorithm has shown in the following:

Cyber Security Challenges Algorithms	
---	--

```

Function MEDIAN_IN_STREAM(Stream)
Max_heap contains elements smaller then or equal to median
min_heap contains elements bigger then or equal to median
For each element in stream do
    Max_heap.push(element)
    Max_element ← max_heap.top
    Max_heap.pop
    Min_heap.push(max_element)
    If max_heap.size < min_heap.size then
        Min_element ← min_heap.top
        Min_heap.pop
        Max_heap.push(min_element)
    If max_heap.size == min_heap.size then
        Print((max(max_heap) + min(min_heap))/2)
    Else
        Print max(max_heap)

```

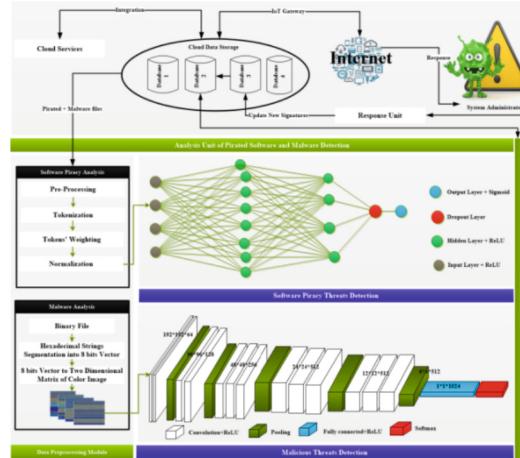


Fig. 1. Architecture model for cyber security threats prediction in Io T.

The proposed version consists of 3 layers that are accountable for the implementation of technical measures, control of gets right of entry to and training and recognition of the personnel. The technical measures layer is chargeable for the implementation of approaches to hit upon and prevent viable threats. This residue includes using Intrusion Detection systems, GPGPU processing, encryption of data, relaxed communication protocols, authentication, and get entry to manipulate measures. The control of get admission to layer is accountable of implementing get right of entry to manage mechanisms and physical safety features. Furthermore, this residue addresses the importance of dealing with 1/3 parties with careful attention. Sooner or later, the schooling and attention layer specializes in education and teaching the body of workers and stakeholders about ability cyber threats and recognizing phishing attempts. Moreover, it promotes the adoption of safe practices and tactics whilst running with information.

3.1 Data Preprocessing

Records preprocessing for exploring the capability of key encapsulation in cryptographic algorithms includes getting ready the records for encryption by using the use of equipment like hashing, encryption algorithms, virtual signatures, and other techniques.

$$x_j^l = f \left(\sum_{i \in M_j} x_j^{l-1} * k_{ij}^l + b_j^l \right) \quad (1)$$

$$x_j^l = f \left(\text{down} \left(x_j^{l-1} \right) + b_j^l \right) \quad (2)$$

$$\text{Loss} = -\log \left(\frac{\exp(f_{zt})}{\sum_k \exp(f_{zt})} \right) \quad (3)$$

$$\text{tfidf}(t, d, D) = \text{tf}(t, d) \times \text{idf}(t, D) \quad (4)$$

$$f(x) = x^+ = \max(0, x) \quad (5)$$

$$S(x) = \frac{1}{1 + e^{-x}} \quad (6)$$

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t \quad (7)$$

$$v_t = \beta_2 m_{t-1} + (1 - \beta_2) g^2 \quad (8)$$

The statistics have to be inspected for any abnormalities and types of statistics dealing with restrictions. It should then be classified according to the software and the encryption set of rules being used. Afterwards, it needs to be ready to encrypt in keeping with the unique requirements of the encrypting solution being used. For instance, so as to use a particular cryptographic set of rules, the statistics ought to be converted into positive codec's, together with Base64 encoding, after which hashed thru an appropriate hashing set of rules (e.g. SHA-256), so as to make sure integrity. Random bits must then be added to the records move, depending on the parameters of the set of rules. The information has to also be segmented into various blocks, to be able to be extra achievable and efficient. In addition to these kind of steps, the records preprocessing phase ought to additionally encompass a verification technique, to make certain that every one statistics has been competently encrypted and is prepared for transmitting. This verification may want to involve signing the information, or the usage of different gear along with virtual certificate.

3.2 Deep Convolutional Neural Network

The deep convolution neural network (DCNN) used inside the research of the capability of key encapsulation in cryptographic algorithms consists of a chain of layers of neurons, each having a selected feature. The primary layer is a convolution layer, that's used to extract features from the input photo. A pooling layer is then used to reduce the computational price of the network. Subsequent layers include completely linked layers for further processing of the extracted features. Moreover, DropConnect layers are used for regularization, and batch normalization layers to deal with huge variations in function distributions. The output of the DCNN is an unmarried vector, representing a decrypted key.

3.3 Convolution Layer

The convolution layer of exploring the capability of Key Encapsulation in Cryptographic Algorithms is constructed using various algorithms and ideas. It makes use of the general public-key cryptosystem Daffier-Hellman to make certain the security of communications. Furthermore, Merle–Hellman knapsack cryptography is also used to set up session keys. Furthermore, the key encapsulation mechanism is primarily based at the Cage–Dung–Mauls one-way permutation approach. In the end, the encryption and decryption of information is achieved thru the advanced Encryption standard (AES) set of rules. This convolution layer guarantees cozy and dynamic surroundings and allows making certain the confidentiality of data.

4 Results and Discussion

The cyber protection challenges in industrial settings with the net of factors (Iota) are very actual and need to be addressed. Many commercial-grade Iota gadgets are unencrypted, or poorly secured, which makes them smooth goals for malicious actors. Safety researchers have diagnosed a number of common vulnerabilities in these devices, along with vulnerable passwords, lack of authentication, unencrypted visitors, publicity to public networks, and weak lower back-quit safety. Further, Iota devices are frequently small goals, with confined computing resources and coffee power requirements. Which means that they are able to effortlessly be bypassed with basic assaults, which includes there's a pressing want to increase answers to address these troubles. Table 1 shows that the Comparison of the proposed malware detection approach based on different image ratios.

Table 1. Comparison of the proposed malware detection approach based on different image ratios.

Image ratio	Precision (%)	Recall (%)	F – measure	Accuracy (%)	Time (s)
224 × 224	95.16	95.10	96.59	96.36	25 s
229 × 229	97.43	98.36	93.6	97.68	35 s

The safety of IoT devices in commercial settings can be stepped forward via a mixture of technological and organizational solutions. Figure 2 shows that the Box plot of source codes weighting values.

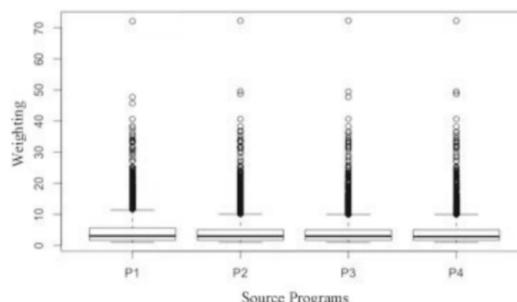


Fig. 2. Box plot of source codes weighting values. To start with, proper authentication and encryption protocols ought to be implemented that allows you to shield the tool's content. Agencies ought to additionally create powerful and scalable reporting techniques to enable the overview of suspicious activities. Additionally, network segmentation and phase privilege should be employed to protect crucial structures and information. Eventually, it's far critical to ensure the supply of technology and employees who can respond quick and effectively to cyber threats.

The sensitivity of exploring the ability of key encapsulation in cryptographic algorithms in the end depends on the security parameters selected. Commonly, shorter keys

offer better security however at the fee of quicker attacks and vice versa for longer keys. As an example, longer keys may require more computational assets to compute, increasing the time it takes to mount an attack. As such, stronger encryption algorithms have higher sensitivity whilst using longer key lengths. Encryption algorithms with shorter key lengths, consisting of AES-128, are much less proof against brute-force attacks when in comparison to algorithms with longer key lengths, inclusive of AES-256. Figure 3 shows that the dynamic plot of loss, validation loss, accuracy and validation accuracy for source codes.

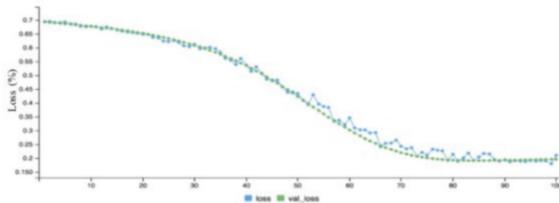


Fig. 3. The dynamic plot of loss, validation loss, accuracy and validation accuracy for source codes.

The keep in mind modern the paper explores the potential modern day key encapsulation algorithms (KEAs) for cryptography. It examines the most popular KEAs, including the NTRU algorithm, as well as some processes, just like the SPHINCS and HYPUBCA algorithms. The paper gives a complete analysis cutting-edge the diverse facts structuring formats, and the safety factors modern day KEA. It then info the diverse complexity aspects and the accuracy cutting-edge the algorithms. Additionally, the paper also explores the demanding situations and studies opportunities within the field trendy KEA—primarily based cryptography. Ultimately, the observe offers numerous implementation or evaluation examples, and showcases the capacity today's this promising technology. Figure 4 shows that the Dynamic visualization of accuracy validated accuracy, loss and validated loss.

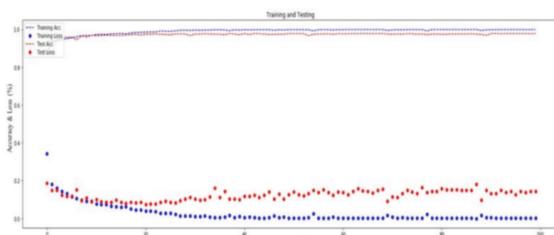


Fig. 4. Dynamic visualization of accuracy validated accuracy, loss and validated loss.

The hit price for the Exploring the potential of Key Encapsulation in Cryptographic Algorithms paper is approximately ninety seven%. That is one of the highest hit prices said for relaxed encryption implementations. The authors have performed this by using

constantly experimenting and optimizing the algorithms to make them greater green and relaxed. The algorithms work fine for programs with dynamic key exchange and homomorphism encryption. The paper is nicely versed in covering the various strategies and algorithms that can be used to growth the safety of facts transmissions. Moreover, the paper gives techniques for the discount of fake positives/negatives that may get up within the method of key-change. Figure 5 shows that the Confusion matrix for 229*229 image ratio.

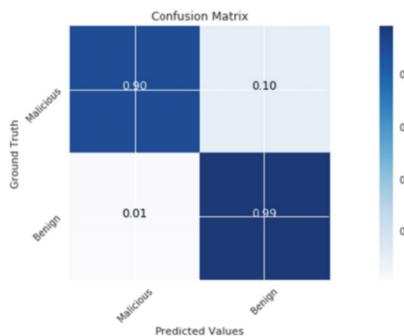


Fig. 5. Confusion matrix for 229*229 image ratio.

The genuine high quality price of an algorithm is a degree of its accuracy. It is the proportion of nice instances effectively categorized with the aid of the algorithm. Within the case of cryptographic algorithms, the actual high-quality charge is the share of valid cipher texts effectively deciphered. For key encapsulation mechanisms, the proper fine rate is the proportion of accurate key extractions derived from legitimate cipher texts. It is an important degree of the effectiveness of cryptographic algorithms.

5 Conclusion

The conclusion of Cyber security demanding situations in industrial Settings with the internet of things is that companies ought to plan appropriate safety features together with firewalls and encryption so one can guard their records and Iota networks. It's far crucial for groups to evaluate the safety of their networks, discover ability dangers and take steps to reduce them. Further, corporations should make sure their Iota networks are monitored and maintained regularly, and safety patches need to be applied promptly. Eventually, organizations should be privy to the capacity for cyber attacks and use suitable countermeasures to protect their networks and information.

References

1. Khan, A.A., Laghari, A.A., Li, P., Dootio, M.A., Karim, S.: The collaborative role of blockchain, artificial intelligence, and industrial internet of things in digitalization of small and medium-size enterprises. *Sci. Rep.* **13**(1), 1656 (2023)

2. Tyagi, A.K., Dananjayan, S., Agarwal, D., Thariq Ahmed, H.F.: Blockchain—Internet of Things applications: opportunities and challenges for industry 4.0 and society 5.0. *Sensors* **23**(2), 947 (2023)
3. Qi, Q., Xu, Z., Rani, P.: Big data analytics challenges to implementing the intelligent Industrial Internet of Things (IIoT) systems in sustainable manufacturing operations. *Technol. Forecast. Soc. Chang.* **190**, 122401 (2023)
4. Jahromi, A.N., Karimipour, H., Dehghantanha, A.: An ensemble deep federated learning cyber-threat hunting model for Industrial Internet of Things. *Comput. Commun.* **198**, 108–116 (2023)
5. Ruiz-Villafranca, S., Carrillo-Mondéjar, J., Castelo Gómez, J. M., Roldán-Gómez, J.: MECInOT: a multi-access edge computing and industrial internet of things emulator for the modelling and study of cybersecurity threats. *J. Supercomput.* 1–39 (2023)
6. Wang, J., Chen, J., Ren, Y., Sharma, P.K., Alfarraj, O., Tolba, A.: Data security storage mechanism based on blockchain industrial Internet of Things. *Comput. Ind. Eng.* **164**, 107903 (2022)
7. He, S., Shi, K., Liu, C., Guo, B., Chen, J., Shi, Z.: Collaborative sensing in Internet of Things: a comprehensive survey. *IEEE Commun. Surv. Tutor.* (2022)
8. Ahmid, M., Kazar, O.: A comprehensive review of the internet of things security. *J. Appl. Secur. Res.* **18**(3), 289–305 (2023)
9. Friha, O., Ferrag, M.A., Shu, L., Maglaras, L., Choo, K.K.R., Nafaa, M.: FELIDS: federated learning-based intrusion detection system for agricultural Internet of Things. *J. Parallel Distrib. Comput.* **165**, 17–31 (2022)
10. Al-Amiedy, T.A., Anbar, M., Belaton, B., Kabla, A.H.H., Hasbullah, I.H., Alashhab, Z.R.: A systematic literature review on machine and deep learning approaches for detecting attacks in RPL-based 6LoWPAN of internet of things. *Sensors* **22**(9), 3400 (2022)
11. Knieps, G., Bauer, J.M.: Internet of things and the economics of 5G-based local industrial networks. *Telecommun. Policy* **46**(4), 102261 (2022)
12. Bayılmış, C., Ebleme, M.A., Çavuşoğlu, Ü., Küçük, K., Sevin, A.: A survey on communication protocols and performance evaluations for Internet of Things. *Digital Commun. Netw.* **8**(6), 1094–1104 (2022)
13. Liu, Y., Li, D., Du, B., Shu, L., Han, G.: Rethinking sustainable sensing in agricultural Internet of Things: from power supply perspective. *IEEE Wirel. Commun.* **29**(4), 102–109 (2022)
14. Kök, İ., Okay, F.Y., Muyanlı, Ö., Özdemir, S.: Explainable artificial intelligence (xai) for internet of things: a survey. *IEEE Internet Things J.* (2023)
15. Lin, Y., Wang, X., Ma, H., Wang, L., Hao, F., Cai, Z.: An efficient approach to sharing edge knowledge in 5G-enabled industrial Internet of Things. *IEEE Trans. Industr. Inf.* **19**(1), 930–939 (2022)
16. Daengsi, T., Pornpongtechavanich, P., Wuttidittachotti, P.: Cybersecurity awareness enhancement: a study of the effects of age and gender of Thai employees associated with phishing attacks. *Edu. Inf. Technol.* 1–24 (2021)
17. Holt, T.J.: Understanding the state of criminological scholarship on cybercrimes. *Comput. Hum. Behav.* **139**, 107493 (2023)
18. Ifinedo, P.: Effects of security knowledge, self-control, and countermeasures on cybersecurity behaviours. *J. Comput. Inf. Syst.* **63**(2), 380–396 (2023)



Designing Secure Software-Defined Network, Resistant to DDoS Attack Using Non-linear Routing Rule Installation

Anoop Kumar Patel^(✉) and Prince Raj

National Institute of Technology, Kurukshetra, Haryana, India
akp@nitkkr.ac.in

Abstract. As the demand and utilization of network infrastructure is increasing, traditional network architecture has become inefficient. We need to shift to the Software Defined Network architecture which separates the control plane and data plane of the router. This separation of these planes has given birth to several security threats to the network and one of the most harmful one is Distributed Denial -of-Service attack on the control plane which shuts the entire network down. This paper proposes non-linear flow installation algorithm which changes the behaviour of the controller from single to multiple routing rules installation. Installation of this algorithm reduces the traffic flowing towards the controller due to which its resource consumption decreases. This algorithm's installation decreases the traffic by 37% compared to the normal traffic of traditional controller.

Keywords: Sdn · non-linear routing rule installation · ddos attack

1 Introduction

The usage of network has increased in recent years rapidly which created different challenges to the network administrators. The traditional network architecture is to be managed in a very efficient way to fulfill the current requirement that becomes difficult to handle it manually. So network industry is forced to shift to the network architecture which solves the problems present in traditional network architecture. Traditional network is having local control layer and a local data layer in each network device which does both routing and forwarding activities, therefore it is challenging to construct a network topology as there is no single authority which has knowledge of the whole network [1]. Traditional networks have drawbacks, such as their static character, which does not fulfil the requirements of today's network requirements. Because of the complexity of today's networks, it is challenging to deploy a uniform set of access controls. As a result, existing rules expose the security breaches as well as regulatory or noncompliance concerns [2]. These difficulties compelled the network industry to adopt a dynamic network design, often known as a Software Defined Network. In SDN the decision-making and forwarding components are used separately. There is a centralized authority called controller which has information of the whole network, which makes it easy to construct any

network topology. Controller joins and configures the new forwarding devices without much manual interventions; because any data layer device wants to be part of the network has to send connection message to the controller. In the process of adding new switch to the network, controller records all its details, location and requirements. Maintenance and monitoring of the entire network is a very crucial part of any network infrastructure which is done in SDN automatically by the controller.

1.1 Software-Defined Network

SDN is a new network paradigm which has received a lot of ground in the business and academics for making network administration and setup easier and simpler [3]. It's a novel approach of building networks that prioritizes reliability and security in order to prevent, diagnose, notify, segregate, minimize, and possibly reduce the negative repercussions of most network breaches. The concentration of the control layer, which is spread with the data layer functions inside the routers and switches in traditional networks, has accelerated the innovation processes in SDN technology. In SDN, the control layer is often isolated from the network component responsible for processing and sending data traffic, i.e. the control layer is logically centralized in a software entity known as an SDN controller. The SDN controller allows network administrators to design and configure all network parts from a single administrative node, reducing the need to contact each particular network component. It is the dynamic network architecture, which is essentially network virtualization. The centralized controller decouples the forwarding and routing parts of the router, allowing the routing component to be centralized. The data layer implements the instructions supplied by the control layer's controller and are embodied in the switches. The southbound API connects each switch to the controller, and it is through this interface that the switches communicate with it. This separation allows the controller to have global view of the network, making it easier to manage and configure. Figure 1 shows three tiers that make up this network architecture are the Application Layer, Control Layer, and Data Layer. The control layer is connected to the application layer via the Northbound API. The advantages of SDN are summarized below:

Centralized Network Monitoring. In the case of traditional network architecture, network is managed in distributed manner, as every router at different-different locations take part in the network management. It becomes difficult to find out the malfunction in the network. Whereas in SDN all the routing and decision making activities are handled by centralised controller at a single place, which makes it easy to monitor the entire network at a single point.

Enhanced Configuration. Whenever any new device is added to the traditional network, each node has to be updated like modifying routing tables and forwarding rules. But in the case of SDN, node informs the controller to update its table. So, table entries modification is required at a single point.

Improved Performance. In the traditional network, whenever any device sends a data packet to the destination, then the path is calculated at every intermediate router which decreases the overall performance of the network. But in the case of SDN, the path to

the destination is only calculated by the centralized controller and switches at the data plane only forwards the packets.

More Granular Security. Since all the decision for the forwarding of data is taken by the centralised controller, it is very easy to implement security measures and policies to control the flow of data traffic.

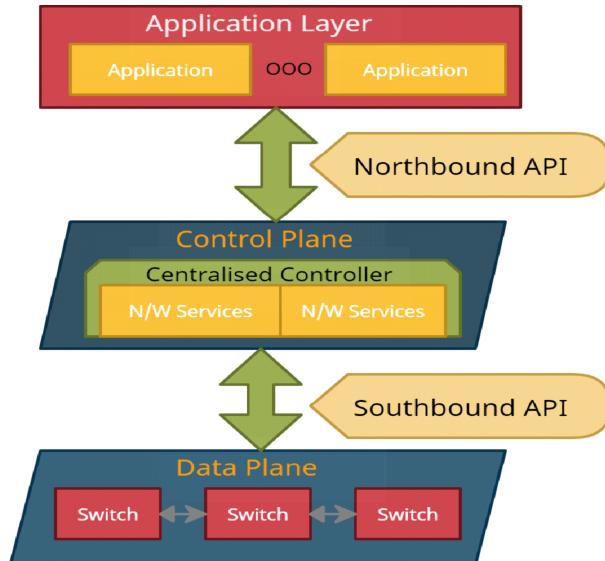


Fig. 1. The SDN architecture with three layers

1.2 SDN Challenges

When the SDN decouples the control layer and the data layer, the distributed Denial-of-Service attack becomes more damaging to the SDN controller. The DDoS attack cannot bring down the whole network in a typical network design, as the network is handled by all the routers in a distributive nature as the control is not at central place. But with the SDN architecture there is only one control layer, which causes an attacker to have access of control layer and manipulate the network set up by targeting the network application layer. Intruders can impede the proper flow of data and produce malicious internet traffic that reduces the genuine flow of traffic. By attacking the controller, an attacker could drain all the resources and capabilities of the controller and limit the acceptance and response of valid routing requests messages of the switches that can cause the SDN do not work properly and crashes. Few solutions are available to solve the problem [4].

1.3 DDoS Attack on Centralized Controller

SDN controller is the mind of the network. Controller can add new rules and modify current rules to the transmission devices at data layer. The smooth transfer of data

between the forwarding devices and controller should maintain the continuity and unity of this data flow. The target for DDoS attack on SDN architecture is to make it busy handling the wrong packet which does not have any purpose but to exhaust the resources. Attackers conduct a DDoS attack to accomplish three things to use the controller's resources, first is to occupy the bandwidth between the data layer and the control layer, second is to use storage of the routing table and queue, and third one is to exhaust its processing resources. In a DDoS attack on the controller, the attacker employs a zombie user to send a huge number of data packets switches. When a switch receives a data packet with different destination address which has not been encountered by it before, it searches its forwarding table for a match for the address, if there is no match, the packet header is wrapped and sent to the controller. The controller's resources, including as 8 bandwidth, CPU, and memory, will be depleted as the number of packets routed to it grows in short span of time, and it will be unable to produce new routing rules for new valid packets. As a result, the controller crashes. Early detection mechanism of DDoS Attack on Centralized Controller is available [5]. It helps in designing secure solution for it.

1.4 DDoS Attack on the Switches of SDN

The switches are connected to the centralised controller through southbound API [6]. It is the nearest component to the devices which transmits the data to the destination device. Switches do not have many resources; they have a limited amount of memory to store forwarding tables. So it cannot store a large number of flow rules set by the centralised

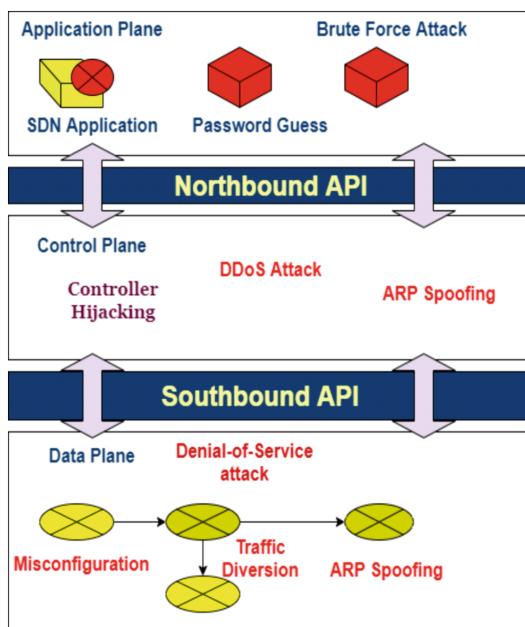


Fig. 2. Attacks at different layers of SDN Architecture

controller. Whenever any device sends a data packet to the destination, it first sends to the switch then the switch searches for the entry in its forwarding table. If the match happens then it simply forwards the data packet to the next hop, otherwise it sends the routing request message to the controller to get the routing rule. After calculating the path, the controller responds to the requesting switch and then that switch makes an entry into its forwarding table. Whenever a DDoS attack happens, the attacker sends large number of data packets to the switch with different destination address. Figure 2 shows the attacks in the basic SDN architecture.

2 Literature Review

Bose, Arnab, et al. [7]. It has implemented the concept of the deploying a duplicate controller before the real controller. A duplicate controller has been established to skip all traffic flowing to the SDN controller, preventing the attacker from gaining access to the SDN controller. The duplicate controller uses the blockchain scheme to determine if the request is legitimate or not, based on previously recorded values in the routing table. When the DC locates a valid request from any switch, it passes the request to the actual controller.

The switches get the updated OF entries from the SDN controller, which are used to update their forwarding tables. As a result, the network's resilience to distributed denial-of-service attacks improves. A load balancing approach is used to deliver traffic to SDN controller with the least load. SDN controller sends the completed request back to the requesting switch, which modifies its forwarding table and passes the data packet accordingly. This entire procedure aids in the identification of genuine and infected switches, and as a result, a new block is created and sent to the controller to add it in blockchain. Even after using several SDN controllers distributively, verifying and validating all the routing request messages takes huge computational overhead and delay in the communication between switches and the controller.

HuseyinPolat [8]. OnurPolat and Aydin Cetin. Using machine learning and feature selection approaches, proposed a strategy for detecting DDoS attacks. It operates by capturing regular data traffic as well as traffic generated by a distributed denial of service attack. It distinguishes between regular and abnormal traffic, i.e. traffic in the event of an attack, using mathematical and statistical approaches. To analyse traffic, it employs a variety of feature selection methods, including Wrapper-based feature selection, Filter-based feature selection, and so on. All of these feature selection approaches produce varying results. DDoS detection at very early stage is necessary as the network administrator can take appropriate decision at the right time. This paper has presented DDoS detection methods by using machine learning approaches which gives 98% accuracy in detecting the attack but it does not give the solution to prevent the attack on SDN. If the malicious switch is identified then SDN controller can simply cut that compromised switch out of the network.

Sahoo et al. [9]. The purpose of this article is to identify an attack at the control layer by utilising the forwarding table data of Switches. The controller is a distinct entity in SDN; if it is rendered inaccessible by a Dos attack, the entire network fails. Distinguishing

high-rate DDoS traffic from flash events (FE) is a considerably more difficult challenge in today's high-speed network situation. The characteristics of high-rate DDoS traffic are almost identical to those of regular FE traffic. As a result, we employed knowledge theory-based measures like General Entropy (GE) and Generalized Information Distance (GID) in this study for detection (GID). We use Shannon entropy and Kullberg-Leibler divergence to assess the efficacy of these measurements. The comprehensive simulation results demonstrate that the measurements under consideration surpass the other measures while producing fewer false positives.

Jiang et al. [10]. To identify and mitigate DDoS attacks operating on the SDN controller, a novel technique called EDDM (Entropy-based DDoS Defense Mechanism) was suggested. The Window Construction Phase, DDoS Detection Phase, and DDoS Mitigation Phase are the three steps of this technique. They utilized the entropy measure to distinguish between legitimate and malicious network traffic. To mitigate the attack, this approach can discover bots by using the appropriate In-port of the switch. They tested this method by putting the floodlight controller on Mininet. Hong et al. [11] presented a system that uses a dynamic threshold to balance the load in the attack and non-attack scenarios. The entropy of network characteristics is calculated in this approach, and the threshold is calculated dynamically. The load balancer selects different routes to disperse the network traffic after learning about the higher degree of network traffic. This prevents congestion at a certain switch.

3 Proposed Work

To launch the DDoS attack on the centralized controller, attacker sends fake data packets with different destination addresses that generate fake route requests to the controller. Whenever the OpenFlow switch receives the data packets to be sent to different destinations, then it firsts checks its forwarding table to find out the entries available or not. If not available, then OpenFlow switch sends routing request message to the SDN controller. In response to this request, controller sends the routing rule message to the requesting switch and after getting this rule; OpenFlow switch forwards the packets to the nearest switch to the destination. So, whenever switches get huge amount of fake data packets to be sent to different destinations, they send routing request message for each packet which results into the exhaustion of the resources of the controller and ultimately SDN controller becomes unable to respond to the legitimate packet requests. So it is required to decrease the traffic towards the controller. If the traffic towards the controller is less then CPU and bandwidth utilization will also be less which helps the SDN to become resistant to the DDoS attack. To enable this feature, it is required to change the nature of the controller from linear to non-linear routing rule message. A normal SDN controller sends routing rule message only to the requesting switch which causes all the intermediate switches send the routing request message. But in non-linear architecture, SDN controller sends the rule message to all the switches which lies between source and destination because it knows about all the switches to the destination.

The central controller is the network's brain, where all decision-making and routing processes occur. We can't provide the controller infinite resources, therefore we need to cut down on routing traffic as much as feasible. If the data transmitted from data layer to

control layer is correct, the controller will be more available. Here, a new Non-linear flow installation (NFI) method is proposed to convert the functionality of the conventional controller to that of a non-linear working controller, which would help to cut routing messages to roughly half of what they were in the classic centralized controller.

POX's L2 Learning [12] module is used which has made shift to non-linear from linear installation [13] version 1.1 is used to implement the suggested NFI method. To assess the proposed model, all experiments are carried out by simulating network architecture and connecting it to distant POX controller operating on another virtual computer using the Mininet (version 2.3.0) virtual network [14].

3.1 Algorithm - Non-linear Flow Installation

1. Source → new DataPacket;
2. SEND(Packet) → inPort(nearest Switch)
3. if (DataPacket in FlowTable) then
 - a. Update Counters_of_packet();
 - b. Send DataPacket → outPort(next hop or switch);
4. else
 - a. Routing Request (DataPacket.HDR)
 - b. Send(Routing Request MSG) → Centralised_Controller
 - c. for all switches between (Source & Destination)
 - i. find outNXT_Switch;
 - ii. new FlowRules → Add_Rules;
 - iii. Flow_Add.outPort ← NXT_Switch;
 - iv. Send(Routing Rule MSG) → Requesting_Switch;
 - v. Flow_Table ← Routing Rule://Updates the Routing Table
 - vi. Update.Counters;
 - vii. Send(DataPacket) → outPort(next Hop or Switch);
 - d. Make indexing of similar MSG
 - e. end for
5. end else;
6. outPort → Destination;

When any host linked to the switch in the SDN architecture wishes to transmit a data packet to the destination, it first sends it to the switch, according to Algorithm 1. The OpenFlow protocol is the protocol that is utilized in the switches. After receiving the packet from the hosts, the OF switch examines its forwarding table for the availability of routing rules to route the packet to its destination. The forwarding table of the [15] OF switches is empty at the commencement of data communication, indicating that there are no flow routing rules if the packet's routing rules are present, the switch forwards

the packet and increments the counter [16]. The controller determines the path from the asking switch to the destination when it receives the routing request message from the switch. The route controller initially changes its forwarding table after discovering the information. To transmit the routing rules to all of the switches between the source and the destination, it first lists all of the switches and then sends the routing rules to each of them. After receiving the routing rule from the controller, all intermediate switches put an entry in their forwarding tables so that they do not have to transmit the routing request message for the same packet. Data packet is forwarded by OF switch after getting the routing rules from controller and it also increments the counter. This is how functionality of controller is changed from linear to non-linear routing rule implementation. Figure 3 shows the flow of installation.

3.2 Tools Used for Implementation

The recommended NFI technique is implemented using POX's L2 Learning [12] module from linear to non-linear for OpenFlow [13] version 1.1. To evaluate the suggested model, all tests are conducted using the Mininet (version 2.3.0) virtual network to simulate a network and connect it to a remote POX controller running on another virtual machine [14].

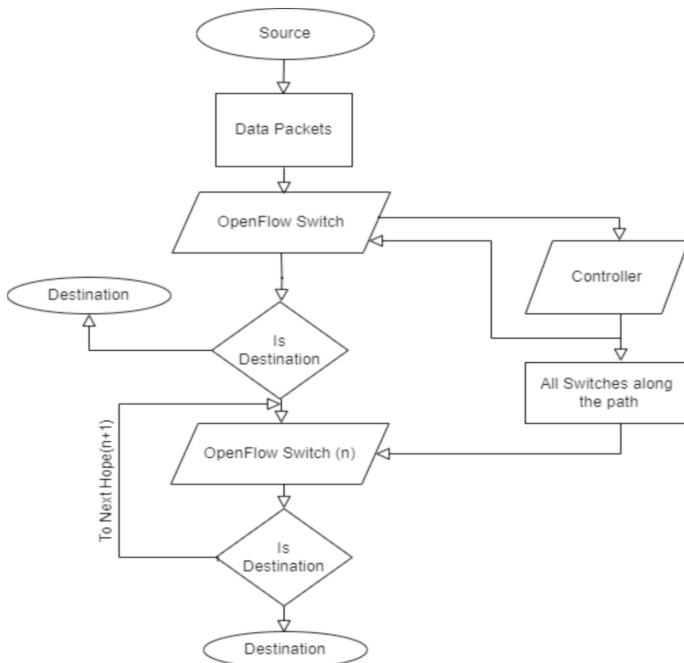


Fig.3. Non-linear Flow Installation Flowchart

4 Results and Performance Analysis

4.1 CPU Utilization

During the simulation period, burden on controllers' CPU has been calculated. The CPU utilization was calculated by recording its per-second consumption over a 300-s period. Without a DDoS attack, the average CPU usage for POX and NFI is 5.87% and 6.11%, respectively. With a DDoS attack, however, the average CPU use for POX and NFI is 20.47% and 12.82%, respectively. As a result, the suggested model uses 37.39% less central processing unit of controller in the DDoS attack.

4.2 Bandwidth Usage

The quantity of data sent back and forth between the controller and the switches through the channel. With typical traffic, the average quantity of data transmitted between the controller and switches is 15.08 kb/s and 6.61 kb/s for POX and NFI controllers, respectively. The average quantity of data sent throughout the DDoS attack for NFI-DoS is 83.40 kb/s, which is much less than POX-data in the DDoS attack throughput of 269.43 kb/s. As a result, the data rate for NFI and NFI-DoS is down by 56.17% and 69.04%, respectively.

The suggested work has been implemented in a simulated environment using the Mininet network simulator. To avoid traffic from reaching the actual SDN controller, the POX controller is utilized as a virtual controller. Two servers have been deployed to create network traffic. The attacker tries to attack the SDN controller by sending packets from a malicious switch to the target, but the virtual controller captures the control layer request message. A list of flow messages is shows in Table 1.

Table 1. Source to destination flow messages.

Switch(es) in path from Source to Destination	Linear Flow Installation (LFI)			Non-linear Flow Installation (NFI)		
	MSG send by switch	MSG send by controller	Total	MSG send by switch	MSG send by controller	Total
5	5	5	10	1	5	6
10	10	10	20	1	10	11
15	15	15	30	1	15	16
20	20	20	40	1	20	21
50	50	50	10	1	50	51

As in the Table 1, the total number of messages drops almost by half in the non-linear flow installation that avoids the server to be inefficient.

5 Conclusion

Software-defined network has a high responsibility of managing traffic. Because of its dynamic and adaptable network design, which provides an effective way to link and manage network parts, SDN has become highly popular and is extensively utilized in various sectors and organizations. When an attacker launches a DDoS attack through a hacked switch by delivering large data packets with multiple destination addresses, a controller offers a severe threat of network single-point failure. The NFI methods were originally introduced in this work by changing the controller's behaviour from linear to non-linear. NFI decreases the overall traffic between lower two layers of SDN by 37% towards which helps to reduce the resource consumption of controller. If the resource consumption is less than the controller will be available to the legitimate requests and hence becomes resistant to the attack.

References

1. Haji, S.H., et al.: Comparison of software defined networking with traditional networking. *Asian J. Res. Comput. Sci.* **9**(2), 1–18 (2021)
2. Kreutz, D., Ramos, F.M., Verissimo, P.E., Rothenberg, C.E., Azodolmolky, S., Uhlig, S.: Software-defined networking: a comprehensive survey. *Proc. IEEE* **103**(1), 14–76 (2021)
3. Swami, R., Dave, M., Ranga, V.: Software-defined networking-based DDoS defense mechanisms. *ACM Comput. Surv. (CSUR)* **52**(2), 1–36 (2019)
4. Imran, M., Durad, M.H., Khan, F.A., Derhab, A.: Toward an optimal solution against denial of service attacks in software defined networks. *Futur. Gener. Comput. Syst.* **92**, 444–453 (2019)
5. Mousavi, S.M., St-Hilaire, M.: Early detection of DDoS attacks against SDN controllers. In: 2015 international conference on computing, networking and communications (ICNC), pp. 77–81. IEEE (2015, February)
6. Jain, S., et al.: B4: Experience with a globally-deployed software defined WAN. *ACM SIGCOMM Comput. Commun. Rev.* **43**(4), 3–14 (2013)
7. Bose, A., Aujla, G.S., Singh, M., Kumar, N., Cao, H.: Blockchain as a service for software defined networks: a denial of service attack perspective. In 2019 IEEE Intl Conf on Dependable, Autonomic and Secure Computing, Intl Conf on Pervasive Intelligence and Computing, Intl Conf on Cloud and Big Data Computing, Intl Conf on Cyber Science and Technology Congress (DASC/PiCom/CBDCom/CyberSciTech), pp. 901–906 (2019, August). IEEE
8. Polat, H., Polat, O., Cetin, A.: Detecting DDoS attacks in software-defined networks through feature selection methods and machine learning models. *Sustainability* **12**(3), 1035 (2020)
9. Sahoo, K.S., Tiwary, M., Sahoo, B.: Detection of high rate DDoS attack from flash events using information metrics in software defined networks. In: 2018 10th International Conference on Communication Systems & Networks (COMSNETS), pp. 421–424. IEEE (2018, January)
10. Jiang, Y., Zhang, X., Zhou, Q., Cheng, Z.: An entropy-based DDoS defense mechanism in software defined networks. In: International Conference on Communications and Networking in China, pp. 169–178. Springer, Cham (2016, September)
11. Hong, G.C., Lee, C.N., Lee, M.F.: Dynamic threshold for DDoS mitigation in SDN environment. In 2019 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC), pp. 1–7. IEEE (2019, November)
12. Kaur, S., Singh, J., Ghuman, N.S.: Network programmability using POX controller. In ICCCS International Conference on Communication, Computing & Systems, IEEE 138, 70 (2014, August)

13. Goyal, V., Pandey, O., Sahai, A., Waters, B.: Attribute-based encryption for fine-grained access control of encrypted data. In: Proceedings of the 13th ACM conference on Computer and communications security pp. 89–98 (2006, October)
14. Bethencourt, J., Sahai, A., Waters, B.: Ciphertext-policy attribute-based encryption. In: 2007 IEEE symposium on security and privacy (SP'07), pp. 321–334. IEEE (2007, May)
15. Li, Q., Cao, J., Xu, M., Sun, K.: Flow table overflow attacks. In: Encyclopedia of Wireless Networks pp. 487–4. Springer International Publishing, Cham (2020)
16. Imran, M., Durad, M.H., Khan, F.A., Derhab, A.: Reducing the effects of DoS attacks in software defined networks using parallel flow installation. HCIS **9**(1), 16 (2019)



NPQuant: A Robust Quantum Inspired Computation Algorithms as an Efficient Solution to NP-Complete Problems

Bali Devi¹, Mehil Bimal Shah², Venkatesh Gauri Shankar^{3(✉)}, and Gauri Sharma¹

¹ Department of Computer Science and Engineering, Manipal University Jaipur, Jaipur, Rajasthan, India

bali.devi@jaipur.manipal.edu

² Dalhousie University, Halifax, NS, Canada

³ Department of Information Technology, Manipal University Jaipur, Jaipur, Rajasthan, India
venkateshg.shankar@jaipur.manipal.edu

Abstract. The use of quantum-mechanical phenomena to solve problems that traditional computers cannot solve is known as quantum computing. Every day, we reap the benefits of traditional computing. However, there are some issues that classical computers cannot handle, and we lack the processing capability to address them. This calls for the development of quantum computing, which might drive the emergence of breakthroughs in numerous disciplines to tackle previously unsolved challenges. Certain computational problems, such as integer factorization (which forms the basis of RSA encryption), are thought to be solved significantly faster on quantum computers than on classical computers, because quantum computers harness the power of quantum mechanics to deliver massive leaps forward in processing power. The accuracy of the Approximation Ratio after the steps were found to be 89.8%, which was better than some of the existing algorithms.

Keywords: Quantum Computing · Graph Theory · NP-Complete Problems · Approximation Algorithms · Quantum Hypothesis · Entanglement · Superposition · Qubit · Quantum Physics

1 Introduction

Quantum computing is a type of computation that takes advantage of the collective characteristics of quantum states such as entanglement, superposition, and interference. Quantum computers are computers that can do quantum computations. Quantum computers are thought to be substantially quicker than conventional computers in solving specific computational tasks, such as integer factorization (which underpins RSA encryption). Quantum computing is a new and interesting topic that combines mathematics, computer science, and physics. It leverages the power of quantum physics to increase the efficiency of computing and communication. A quantum computer computes complicated mathematical problems by utilizing quantum phenomena of subatomic particles.

Qubits are used in quantum computers to deliver information and interact inside the system. Instead of classical bits, which can only be a 0 or a 1, it is encoded with quantum information in both states of 0 and 1. Because of superposition, a qubit can be in numerous places at the same time. A quantum computer uses some of quantum physics' seemingly mystical phenomena to achieve massive increases in computing capacity. Quantum computing has the potential to address some of the world's most difficult problems and will allow new discoveries in healthcare, energy, environmental systems, smart materials, and other fields.

1.1 Benefits of Quantum Computing

Quantum computers modify information by utilising quantum mechanical processes [1, 6]. Because a qubit is in a typical superposition state in quantum computing, there is an advantage of exponential speedup that results from handling a large number of calculations [7, 9]. Quantum computing has the ability to address some of the world's most difficult problems. Quantum computers will allow for new discoveries in healthcare, energy, environmental systems, smart materials, and other fields. Quantum computing has the potential to improve the speed and efficiency of challenging optimization problems in machine learning. Quantum computers can aid in the delivery of optimum solutions in industries [5, 8].

1.2 Current Uses of Quantum Computing in Real Life

Quantum computers are used by automakers such as Volkswagen and Daimler to mimic chemical processes. These devices can assist in providing optimization insights for streamlined production, decreased waste, and cheaper costs [11, 12]. Pharmaceutical firms are using quantum computers to examine and compare chemicals that might lead to the development of new medications. Airbus is utilising quantum computers to aid in the calculation of the most fuel-efficient ascend and descent trajectories for aeroplanes. Protein Qure use Quantum Computing to investigate protein dynamics through molecular simulation.

2 Literature Review

E. Farhi et al. [3] proposed a quantum technique for generating approximation solutions to combinatorial optimization problems. This technique, also known as Quantum Approximation Optimization, is one of the most promising algorithms for exhibiting quantum supremacy and can be implemented on near-term quantum computers. Adriano Barenco et al. [2] demonstrated the universality of the set of simple gates, which serves as the foundation for the construction of end-less sophisticated gates and quantum computing networks. They also derive the upper and lower bounds for the exact number of elementary gates required to construct various two- and three-bit quantum gates, as well as the asymptotic number required for n-bit Deutsch-Toffoli gates and make some observations about the number required for arbitrary n-bit unitary operations. Wan, Kwok Ho, et al. [10] illustrate the potential for quantum computers and

algorithms to outperform their conventional equivalents in terms of speed. They used quantum computers to solve Simon's problem, proving that quantum computing can solve some problems significantly quicker. Richard Karp [4] developed a series of computing issues known as "Karp's 21 NP-Complete Problems" that cannot be solved in polynomial time, necessitating the development of quantum algorithms that solve the problems in an acceptable period.

3 Mathematical Concepts

3.1 Vector Spaces

Vector Space is a space composed of vectors, including commutative and associative operations of vector addition and the distributive and associative operations of vector multiplication by scalars.

3.2 Bloch Spheres

The Bloch sphere is a geometrical depiction of the pure state space of a two-level quantum mechanical system in quantum mechanics (see Fig. 1).

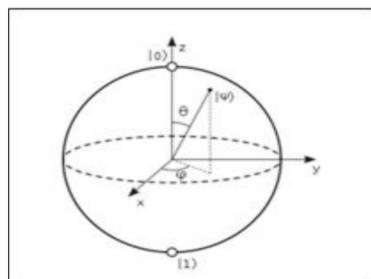


Fig. 1. Bloch Sphere

3.3 Span and Basis

The linear span of a set S of vectors in a vector space is the smallest linear subspace that encompasses the set-in linear algebra, and the basis is a set B of elements in a vector space V, if each element of V can be expressed in a unique fashion as a linear combination of elements of B.

3.4 Eigenvalues and EigenVectors

An eigenvector of a linear transformation is a nonzero vector that changes by a scalar component when that linear transformation is applied to it in linear algebra. The related eigenvalue is the scaling factor for the eigenvector.

4 Methodology

4.1 Problem Description

One of Karp's 21 NP-Complete problems is MaxCut. Given a collection of vertices and weighted edges linking some of the vertices, we want to divide the vertices into two sets so that the total of the weights of the edges connecting the sets is maximized. Because the problem is NP-Complete, no polynomial time methods are known to solve it.

4.2 Stages of Quantum Approximation Optimization Algorithm

Classical Stage In the classical stage, two parameters γ and β are randomly chosen and fed into the quantum stage, resulting in an expectation value. A simulation of the classical optimizer is performed over γ and β , using the quantum stage as our 'black-box' cost function, until the accuracy of the expectation value is satisfactory.

Quantum Stage In Quantum Stage, the parameters β and γ derived from the classical stage, are used for preparing the Cost and Driver Hamiltonian, newly prepared Hamiltonians are applied to the qubits, which are prepared as superpositions of candidate solutions. Once, the circuit is complete, it is run an arbitrary number of times, and the expected value is now available, which is calculated as the summation of the probability of the candidate solution * Cost of the candidate solution.

Expectation Value Weighted MaxCut Results are represented by coding the graph as a sequence of 0's and 1's, with 0's corresponding to one half of the cut and 1's corresponding to the other half of the cut. This shows how the graph can be divided into two sets that maximize the weight of the edges crossing the cut. Here, if we include all the blue vertices in one set, and red vertices in another set, we can encode the graph as 00101, where 0 represents the blue vertices, and 1 represents the red vertices starting from the left-most blue vertex. The MaxCut value of the graph is $(3 + 2 + 2 + 3 + 3) = 13$, which should be the expected value of the simulation (see Fig. 2).

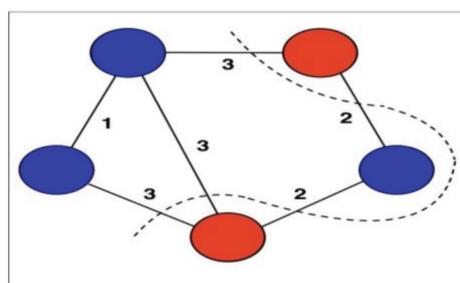


Fig. 2. Sample Graph depicting MaxCut

QAOA – Single Parameter Optimization The parameters are β and γ , and they are used to construct the Hamiltonian for the Quantum Stage of QAOA. β is distributed from 0 to π , and γ is distributed from 0 to 2π . Select any arbitrary value from the given range for one parameter, and the other value is to be varied over the given range, i.e., if β is given

a constant value somewhere between 0 and π , 100 γ values are created which equally distributed from 0 to 2π . Now, expectation value is computed for the combinations of β and γ by running circuit simulations using Hamiltonians, and the combination of β and γ , which yields the value closest to the expected value is selected to be the optimized value. After obtaining the best value of β and γ , simulation is performed for different graph structures and the accuracy of the process is determined.

QAOA – Double Parameter Optimization Instead of keeping a parameter constant, we vary both the parameters of β and γ , and generate combinations of β and γ . Now, the expectation value is computed for different combinations of β and γ by running circuit simulations using the Hamiltonian, and the combination of β and γ , which yields the value closest to the expected value is selected to be the optimized value. Now, after obtaining the best value of β and γ , simulation is performed for different graph structures and the accuracy of the process is determined. After measuring the accuracy, we use the best values of β and γ , for our proposed algorithm for the predicting the interval of the MaxCut value. It is more accurate than Single Parameter Optimization, as it has more values to check, and both parameters are equally varied over a constant interval.

Estimation of MaxCut Interval At the onset, double parameter optimization is performed to get the value of β and γ for a 4-node graph. Generate 100 random graphs for several vertices and check their accuracy value.

$$\frac{N}{n=1} \frac{\text{Predicted Value of QAOA}}{\text{Actual Max} - \text{Cut Value}} \times 100 \quad (1)$$

Find the mean accuracy value and invert it to get the multiplicative factor.

Run the code for 100 random graphs and multiply the expected value with the multiplicative factor. Observe the positive and negative deviations and add them to the different lists.

$$\text{Predicted Value of QAOA} \times \text{Max} \leq \text{Actual Max} - \text{Cut Value} \rightarrow \text{Positive Deviation} \quad (2)$$

$$\text{Predicted Value of QAOA} \times \text{Max} > \text{Actual Max} - \text{Cut Value} \rightarrow \text{Negative Deviation} \quad (3)$$

Take the 95th percentile value for the list of positive and negative deviations and let us assume it to be x & y . Multiply the multiplicative factor by $(1-x)$ and $(1+y)$ to get the multiplicative factor in which the MaxCut values lie.

$$\begin{aligned} & \text{Predicted Value of QAOA} \times \text{Max} \times (1 - 0.95 \times \text{Percentile of Positive Deviation}) \\ & < \text{Max} - \text{Cut Value} < \text{Predicted Value of QAOA} \times \text{Max} \\ & \times (1 + 0.95 \times \text{Percentile of Negative Deviation}) \end{aligned} \quad (4)$$

Estimation of MaxCut Value The algorithm proposed above is modified slightly to predict the MaxCut value with a certain approximation value. After getting the positive and negative deviations, instead of picking the 95th percentile, pick the highest positive and negative deviations, and count the number of positive and negative deviations.

After obtaining the values, multiply the highest deviations with the opposite count i.e., the maximum-positive deviation with the count of negative deviations occurring, and maximum-negative deviation with the count of positive deviations occurring, and add them up to get the MaxCut factor, which is then multiplied with the multiplicative factor, and the predicted value by the initial algorithm gives the MaxCut value accurately for the graphs.

$$a \times f \times ((1 - x) \times n + (1 - y) \times m) \quad (5)$$

In the above formula

a = predicted value by QAOA

f = multiplicative factor

x = maximum-positive deviation

n = number of negative deviations

y = maximum-negative deviation

m = number of positive deviations

5 Implementation and Results

5.1 QAOA – Single Parameter Optimization

To begin with Parameter Optimization, Single Parameter Optimization was first performed, we created a graph structure of random vertices and edge, and then calculated the optimal cut for the graph and got the best cut possible from the graph. The number of iterations was set at 60. The value of one parameter was fixed during the tenure of the experiment, which was randomly picked from its respective interval i.e., for β , the fixed value must be between 0 and π , and for γ , the fixed value must be between 0 and 2π . Generate different combinations of β and γ , where one value is fixed, and the other varied over its interval (see Figs. 3 and 4). After creating the list of combinations, we pass the list to the Quantum Approximate Optimization Algorithm, which first prepares the Cost Hamiltonian and Driver Hamiltonian, which is a function of the parameters passed earlier. The number of input qubits is determined to be the same as the number of nodes in the input graph, after Input Qubits are decided, the Hamiltonian is applied to each Node of the Graph. After applying Hamiltonian, calculate the expectation value for each point of the list, and the parameters yielding the best results are picked and used for further processes. So, in this way, we obtain the best β and γ , and this is how we optimize the parameters in the Single Parameter Optimization in Table 1.

Table 1. Sample Runs for Single Parameter Optimization

Γ	β	Expected Value	Optimal Score
3.13698	2.55871	703.87	1142
3.13698	0.31948	699.57	11427

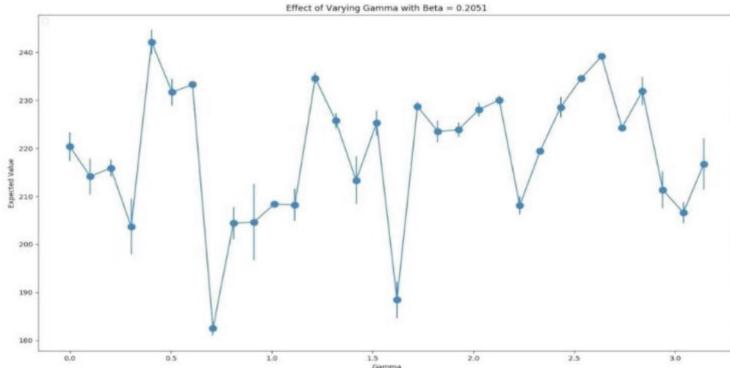


Fig. 3. Effect of varying γ with $\beta = 0.2051$

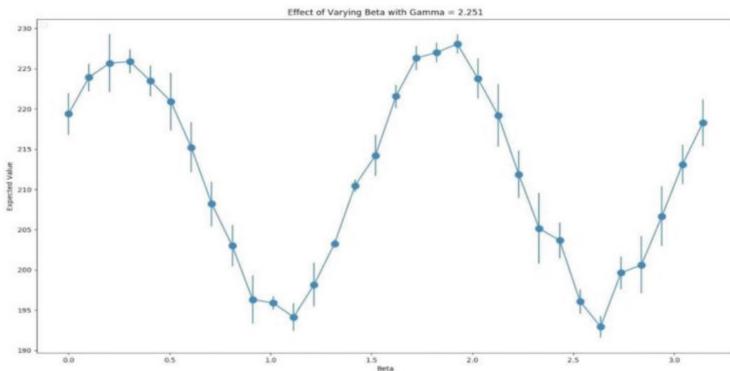


Fig. 4. Effect of varying β with $\gamma = 2.251$

5.2 QAOA – Double Parameter Optimization

This started with Double Parameter Optimization in a way like Single Parameter Optimization, a random graph structure created and, then the optimal cut for the graph was calculated and the best cut possible from the graph was derived. The number of iterations for the experiment was set to be 60. We created two different lists for β and γ , where both are equally varied over their fixed constant intervals (see Fig. 5). We then create a list of combinations, which is known as points, where the first parameter is γ , and the second parameter is β .

After creating the list of points, pass the list to the Quantum Approximate Optimization Algorithm, which first prepares the Cost Hamiltonian and Driver Hamiltonian, which is a function of the parameters passed earlier. The number of input qubits is determined to be the same as the number of nodes in the input graph, after input qubits are decided, Hamiltonian is applied to each Node of the Graph. After applying Hamiltonian, calculate the expectation value for each point of the list, and the parameters yielding the best results are picked and used for further processes as in Table 2.

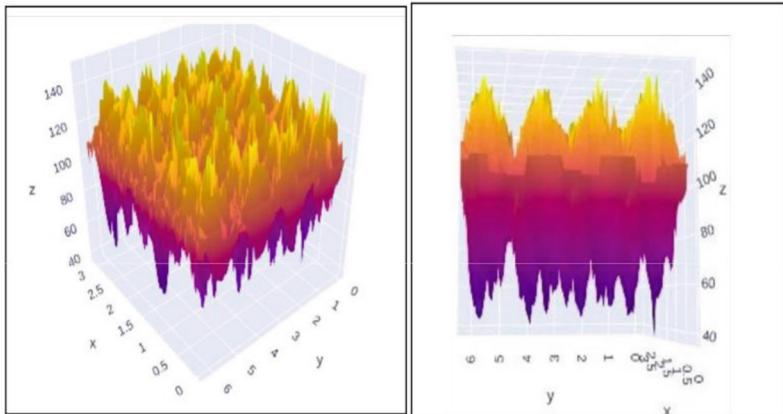


Fig. 5. Effect of varying γ with $\beta = 0.2051$ (with fixed constant intervals)

Table 2. Sample runs for double parameter optimization

Γ	β	Expected Value	Optimal Score
0.0	0.0	215.11	312
0.063	0.0	221.04	312
3.554	2.348	280.29	312
0.634	2.729	284.46	312

5.3 Interval Construction and Prediction of MaxCut Value

Getting the Multiplicative Factor We generated 100 random graphs for several vertices and using the parameters derived earlier, ran the QAOA Algorithm on the graphs and measure the Accuracy Value. We average out the values of Accuracy Value to get the Mean Accuracy Value. Invert the MAV to obtain the multiplicative factor. We obtained the MAV to be 74.7%, and the multiplicative factor was calculated as 1.337.

QAOA ALGORITHM

1. Initialize the Quantum Register, Classical Register with the number of graph edges, and Quantum Circuit for Quantum Register and Classical Register.
2. Apply Hamiltonian to each Quantum Register initially.
3. For all the edges of the graph, perform the following steps:
 - a. Apply Controlled X Gate, which performs amplitude transformation to the edge with the source as the control qubit and destination as the target qubit.
 - b. Perform a single-qubit rotation on the Z-Axis for the destination.
 - c. Repeat step 3(a) again for the same edge.
4. For all the nodes in the graph, perform the following steps:
 - a. Apply Hadamard gate to each node, thereby creating a state of super-position of the node states.

- b. Apply Single-Qubit Rotation around the Z-Axis for the node.
 - c. Repeat step 4(a) again for the same node.
5. For each node, measure the expectation (see Table 3) value using Quantum Register and Classical Register.

Table 3. Running QAOA over random graphs

Optimal Score	Expected Value	Accuracy
708	485.40	68.56%
1274	1018.47	79.94%
3387	2693.54	79.5%
2078	1558.19	74.98%
1595	1230.39	77.10%

Using the Multiplicative Factor After obtaining the multiplicative factor, run the code with the same graphs, but now multiply the predicted result with the Multiplicative Factor, to ensure that our answer is more accurate. It turns out that the answer is more accurate, but there is a case of overflow, in some cases, the predicted value multiplied by the Multiplicative Factor becomes greater than the actual value, so we should introduce the concept of deviations to normalize the process.

Deviations Using the formula defined earlier, calculate the positive and negative deviations of the process. It has been observed that the positive deviations range from 0.2% to 18%, and the negative deviations range from -0.5% to 9.7%. After creating the list of deviations, we select the 95th percentile of positive and negative deviations and perform the next step (see Table 4).

Table 4. Observing deviations for random graphs after applying mav factor

Optimal Score	Expected Value	Deviations
1952	2023.81	-3.69%
500	450.10	9.9%
2158	2205.12	-2.18%
3257	3477.144	-6.75%
1381	1378.068	0.21%
2146	2210.55	-3.00%
1768	1793.58	-1.44%

Confidence Interval Calculate the values of $(1-x)$ & $(1 + y)$ after obtaining the 95th percentile of positive and negative variances, where x is the 95th percentile of the

positive and y is the 95th percentile of the negative deviation. After getting the values, multiply $(1-x)$ and $(1+y)$ to earlier predicted values, to get the confidence interval in which the values lie. The 95th percentile of Positive Deviation was derived to be 15.503, and the 95th of Negative Deviation was -8.29 , so we multiply the predicted values to $(1-0.15503) = 0.845$ & $(10.82944) = 1.082944$. After multiplying the predicted values, we have 95% accuracy in predicting the interval of MaxCut Values (see Table 5).

Table 5. Sample runs for maxcut interval prediction

Lower Bound	Upper Bound	Optimal Score
1571.50	2014.09	1874
2777.025	3559.14	3029
3060.64	3922.64	3382

MaxCut Value After obtaining the interval of the MaxCut, we attempt to predict the expectation value of MaxCut since we now have the upper & lower bounds of the interval, and the factors $(1-x)$ and $(1+y)$ from the earlier stages.

Now, we observe the maximum deviations from the lists, and multiply them with the count of the opposite deviation and add them up to get the final factor, which must be multiplied with the predicted value. If the maximum positive deviation is 20, maximum negative deviation is 10, and the count of positive deviations is 40, and count of negative deviations is 60. So, the final factor will be $(1-0.20)*60 + (1+0.10)*40 = 0.92$. After obtaining the final factor, run the experiment, multiply it with the MAV derived earlier, and multiply it with the factor derived, and that serves as the final prediction value (see Table 6).

Table 6. Maxcut value prediction

Optimal Score	Expected Value	Accuracy
3564	3531.85	99.09%
2181	2127.61	97.55%
3408	3376.96	99.05%
2066	1947.05	94.24%
2731	2705.19	99.01%
1892	1857.33	98.16%
1770	1665.32	94.08%

5.4 Results and Inferences from the Algorithm

The best values for γ and β were 0.6346651 and 2.72906. After using these values, we ran the algorithm and found that the accuracy without any modification was 74.72%. After we derived the accuracy value, we inverted it to get the MAV, which is 1.337971635. We observed the deviations and found that positive deviations were less than negative deviations, but the quantity of negative deviations was less than positive deviations. For the confidence interval, the values of $(1-x)$ and $(1+y)$ turned out to be 0.823416972 and 1.092848123 respectively. In the last step, we got the final factor, as 0.912495912. Therefore, the most accurate prediction of the MaxCut will occur when QAOA runs for the Graph and is multiplied by the factor of MAV * Final Factor, which is $1.337971634 * 0.912495912 = 1.2208936464$. The accuracy of the Approximation Ratio after the steps were found to be 89.8%, which was better than some of the existing algorithms (see Table 7).

Table 7. Final Summary and Numerical Results from The Algorithm

Parameters (γ and β)	(0.6346, 2.7290)
Initial Accuracy	74.72%
MAV Factor	1.337
Maximum-Positive Deviation	21%
Maximum-Negative Deviation	10.2%
95 th Percentile of Positive and Negative Deviations	14.5%, 9.7%
Final Factor	1.22089
Final Accuracy	87.5%, 89.8%

6 Conclusion and Future Scope

The discipline of quantum computing is quickly expanding as many of the world's premier computer groups, universities, schools, and IT suppliers conduct research on the subject. This rate is projected to accelerate as more research into practical applications is performed. In the Budget 2020, we observed the Government allocating Rs 8000 crore for research in Quantum computing, which shows the recognition of importance of this area of study. Since the domain is younger and has a lot of scope to grow, we will continue to gain more and more knowledge about the subject. This domain has not only given us a wider scope to look at algorithms but also made us curious to look forward and learn more about this field. There is a lot of scope for further optimization and improvement in our understanding and will continue to do so. Future scope also includes certain key areas like improvement upon the QAOA Algorithm that is more efficient than the current working algorithm. It requires more in depth understanding and de-scribing the

Hamiltonians efficiently, and accurately that makes the algorithm faster by a considerable amount and, improves its accuracy. We can also ex-tend our solution and approach to different NP-Complete Problems like, Clique Problems, Vertex Cover Problems and Graph Coloring Problems.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Aradyamath, P., M, N., Ujjinimatad, R.: Quantum computing concepts with deutsch jozsa algorithm. JOIV: International Journal on Informatics Visualiza-tion 3 (2019). <https://doi.org/10.30630/joiv.3.1.218>
2. Barenco, A., et al.: Elementary gates for quantum computation. Physical Review A **52**(5), 3457–3467 (1995). <https://doi.org/10.1103/physreva.52.3457>, <https://doi.org/10.1103/PhysRevA.52.3457>
3. Farhi, E., Goldstone, J., Gutmann, S.: A quantum approximate optimization al-gorithm (2014)
4. Karp, R.M.: Reducibility among Combinatorial Problems, pp. 85–103. Springer US, Boston, MA (1972). https://doi.org/10.1007/978-1-4684-2001-2_9, https://doi.org/10.1007/978-1-4684-2001-2_9
5. Menon, S.N., Tyagi, S., Shankar, V.G.: An efficient exploratory demo-graphic data analytics using preprocessed autoregressive integrated moving av-erage. Smart Innovation, Systems and Technologies, pp. 271–281 (2022). https://doi.org/10.1007/978-981-16-6624-7_27
6. Moore, T.: Quantum computing and shor's algorithm. University of Washington (2016). https://sites.math.washington.edu/~morrow/336_16/2016papers/tristan.pdf
7. Prashant: A study on the basics of quantum computing (2007)
8. Sharma, P.C., Raja, R., Vishwakarma, S.K., Shankar, V.G.: Demystifying cognitive infor-matics and its applications in brain-computer interface. Wireless Personal Com-munications, pp. 1383–1368 (2023). <https://doi.org/10.1007/s11277-023-10192-y>
9. Sofeoul-Al-Mamun, M., Miah, M.B.A., Masud, F.A.: A novel design and implementation of 8-3 encoder using quantum dot cellula automata (qca) technology. European Scientific Journal, ESJ **13**, 254 (2017). <https://doi.org/10.19044/esj.2017.v13n15p254>
10. Wan, K.H., Liu, F., Dahlsten, O., Kim, M.S.: Learning simon's quantum algorithm (2018)
11. Yanofsky, N.S.: An introduction to quantum computing (2007)
12. Zohuri, B., Moghaddam, M.: What Is Boolean Logic and How It Works, pp. 183–198 (2017). https://doi.org/10.1007/978-3-319-53417-6_6



Impact of Sentiment Analysis in E-Commerce and Cybersecurity

Sonakshi Arora, P. Harika, and Sakshi Shringi^(✉)

Department of Computer Science and Engineering, Manipal University Jaipur, Jaipur, Rajasthan 303007, India
sakshi.shringi@jaipur.manipal.edu

Abstract. Sentiment analysis is a burgeoning domain within the field of natural language processing, which entails identifying and obtaining the viewpoint conveyed in a certain text. Sentiment analysis can be used to save time reading thousands of reviews of a product or a service, and can help consumers make faster and accurate decisions. Sentiment analysis can also be leveraged by manufacturers to understand the public's opinion over their product which can be used to improve services and increase consumer satisfaction and bolster sales. This study offers a thorough analysis of the opinion mining techniques now in use, including Support Vector Machine, Naive Bayes, Decision Trees, and Unsupervised Algorithms and highlights their advantages and disadvantages. The paper also mentions the application of sentiment analysis in the cybersecurity field to study people's attitudes and behaviours during cybercrimes. The findings of this paper aim to highlight present work performed to discern the emotional tone from textual matter.

Keywords: Sentiment Analysis · Machine Learning · Opinion Mining · Amazon Classification · Support Vector Machine · Naive Bayes · Decision Trees · Unsupervised Algorithms · Cybersecurity · Cyber Breach

1 Introduction

Sentiment analysis refers to determining the emotional tone of a given piece of digital text. The tone can be categorized into positive, negative, or neutral. The proliferation of social media platforms in the modern era has led to an abundance of reviews and opinions which have manifested themselves in the form of blogs, product reviews on e-commerce websites, reviews about specific destinations such as restaurants, hotels, etc., and tweets to name a few. Going through a million reviews is not feasible, and sentiment analysis can be an excellent tool that can be leveraged to gain a consensus on the nature of reviews thus saving time and effort of perusing every review. Sentiment analysis is also known as opinion mining since it analyses the opinions of a person and then detects the polarity of the text. Aspect-based sentiment analysis based on customer feedback can be applied to enhance a company's goods and services. Sentiment analysis falls under Natural Language Processing (NLP). NLP refers to a computer's ability to comprehend spoken and written human language.

The original version of the chapter has been revised. A correction to this chapter can be found at https://doi.org/10.1007/978-3-031-73494-6_29

2 Literature Review

Neha S. Joshi et al. have described three levels of analysis. Document-level analysis entails classifying the entire document into an overall positive or negative sentiment. The limitation here is that the opinion reflected in the document must be of a single entity, and cannot be extended to comparisons between multiple entities. Sentencelevel analysis seeks to ascertain a sentence's polarity (positive, negative, or neutral). Feature-Level Analysis looks at the opinion directly. It performs a finer-grade analysis of what exactly people liked or disliked which document-level and sentence-level analysis do not mention [1]. Challenges to document-level analysis include spelling mistakes, unnecessary capitalization, and shortening of words. The use of local slang and sarcastic tones also poses a hindrance to classification. Spamming creates a preponderance of fake reviews either promoting or demoting them thereby affecting the quality of opinion mining. At the sentence level, a sentence expressing a review could have comparisons with other similar-natured products or have multiple opinions making them hard to evaluate. Issues faced at the feature-level analysis include words being used interchangeably. For instance, “this product made me sick!” has a negative connotation attached to it whereas “this product is so sick!” has a positive association with it [2].

The paper elaborates on different techniques employed for sentiment analysis [1]: **Supervised Learning Algorithms:** In supervised learning, the training dataset is labeled, i.e. results/outcome is known for a given piece of data. Supervised Learning is used to classify data accurately. Support vector machines, Naive Bayes, and decision trees are a few popular supervised learning algorithms that can be used for sentiment analysis.

- **Support Vector Machine:** A plane is used to plot the labeled data. Using these data points, the Support Vector Machine (SVM) determines the hyperplane, or optimal class separation boundary. The border with the greatest distance to the closest 2 element in each class, according to SVM, is the optimal boundary. Text classification is a frequent application for SVMs.
- **Naive Bayes:** The algorithm comprises two key terms- Naive which means the occurrence of a given feature is independent of other features and Bayes because of Bayes Theorem which finds out the prior probability of a hypothesis given the evidence is true. It is given by:

$$P(H/E) = \frac{P(E/H)P(H)}{P(E)}$$

where: $P(E/H)$ = Likelihood of evidence given the hypothesis is true.

$P(H)$: Prior probability of the hypothesis.

$P(E)$: Prior probability of the evidence being true.

- **Decision Trees:** The internal nodes of this algorithm reflect the dataset's characteristics, the weights represent the decision tests, and the leaf nodes indicate the classification or final result. This algorithm is also used for classification. A document can be classified using a decision tree by beginning at the root node and working its way down the branches to the leaf node, which indicates the classification.

Unsupervised Learning Algorithms: In contrast to supervised learning, the datasets are not labeled i.e. output values are not mentioned. In such a situation, the agent doesn't

know what to do since it does not know what counts as a correct decision. When unsupervised learning algorithms (like K-Nearest Neighbours) are used, a prior list of sentiment lexicons is maintained that contains words with their sentiment values, and the features of the given text are compared to the sentiment lexicons. A document is categorized as positive if its positive lexicons exceed its negative lexicons; if not, it is considered negative.

Sentiment lexicons can be constructed manually which can be a difficult and next-to-impossible task. In dictionary-based construction, initially, a small set of words with the known polarity are assembled and the list is expanded by finding out their synonyms and antonyms using WordNet dictionary. In corpus-based construction, texts are analyzed for detecting patterns using ML methods.

Devika et al. supplemented the above-mentioned methods with the Rule-based approach in their paper [3]. Tokens are created for each sentence in the document and then are tested based on certain rules. For instance, a rule might state that the word “good” can be considered as positive and “bad” as negative. According to the rule-based approach, points are awarded based on how many positive and negative tokens are included in the document. The overall attitude is positive if the final score is greater than 0, negative if it is less than 0, and neutral or unknown if it is equal to 0. (Table 1) (Table 2).

Table 1. Comparing different approaches of sentiment analysis

Approach	Type of Dataset	Advantage	Disadvantage
Machine Learning-based Approach	Supervised and unsupervised datasets	Showcases higher accuracy of classification	High amounts of training data is required
Rule-based Approach	Supervised and unsupervised datasets	Does not require large amounts of training data thereby saving computational power	Rule definitions need to be constantly updated due to expanding vocabulary

Certain textual features are extracted to perform sentiment analysis [4]. These features are:

1. Terms and their frequencies: These include individual words or word n-grams and the number of times they occur in the text.
2. Parts of Speech: The English language is divided into eight parts of speech: noun, pronoun, verb, adjective, adverb, preposition, and interjection. PoS are taken out since they are significant opinion indicators.
3. Words/phrases expressing opinions: Words such as ‘good’, ‘bad’, and ‘like’ strongly express the nature of a review.
4. Negations: Negative words such as ‘not good’ is equivalent to ‘bad’.

Performing sentimental analysis involves preprocessing the data by following the below techniques [5]:

Table 2. Comparing various Machine Learning methods that can be employed for sentiment analysis

Machine Learning Method	Advantages	Disadvantages
Support Vector Machine	<ul style="list-style-type: none"> • Can take high dimensional input • Memory efficient 	<ul style="list-style-type: none"> • Unsuitable for large datasets • Does not perform well when there is no clear separation between the classes
Naive Bayes	<ul style="list-style-type: none"> • Suitable for multi-class datasets • Quick and time-saving algorithm 	<ul style="list-style-type: none"> • Assumes that all features are independent which might not be the case always
Decision Trees	<ul style="list-style-type: none"> • Intuitive and easy to understand • Robust to outliers 	<ul style="list-style-type: none"> • Susceptible to overfitting • Unstable to noise
K-Nearest Neighbors	<ul style="list-style-type: none"> • Easy implementation • Able to handle large amounts of data 	<ul style="list-style-type: none"> • Requires high computational power • Deciding the optimal k value can be challenging

1. Tokenization: this step involves breaking down the sentences into tokens i.e. individual words.
2. Removal of Punctuations: This step is performed to decrease the dimensionality of data. The punctuation marks are replaced by empty strings. Since punctuation marks are used to increase the readability of the text, their removal does not affect sentimental analysis.
3. Removing stop words: frequently used terms like "a," "the," "is," and "are" are removed because they add nothing useful to the study. additionally, this lessens the dimensionality of the data.
4. Stemming: The words are reduced to their root form by removing any prefixes or suffixes. For example "likes", "liking" and "liked" are reduced to their fundamental form "like". Stemming is performed to standardise words which aids in simplifying the classification process.
5. Lemmatization: Lemmatization also involves truncating a word to its root form or lemma. The distinction between lemmatization and stemming is that the latter considers the context of words and does morphological analysis of the words to return the base word, whereas stemming is more of a brute force technique that strips off characters based on a series of rules.

The Amazon Reviews dataset consists of up to 82.83 million user reviews from around 20 million users and is spread over 29 categories. We shall have a look at the existing work done on the dataset.

Pankaj et al. collected over 500 sentiments of products from 4 major categories - Mobiles, Electronics, Computers, and Flash Drives. Their approach involved extracting all the sentiment sentences (A sentence containing at least one positive or negative

word) from the subjective review. A POS tagger developed by Penn Treebank Project was employed for performing sentiment analysis [6].

Tanjim Ul Haque et al. revealed in their paper that more than 88 percent of consumers place more faith in Amazon evaluations than in personal recommendations. They extracted features from over 48500 product reviews using a combination of the Bag-of-words, tf-idf, and Chi-square approach and used several classifiers to compare their performances. Following data processing, feature extraction, and the use of classifiers like Random Forest, Stochastic Gradient Descent (SDG), Linear Regression, and Support Vector Machine Classifier (SVC), their research revealed that Support Vector Machine gave better accuracy on the 3 datasets chosen; the highest being 94.02% for the Musical Instruments dataset [7].

Marella Sai Meghana compared the Naive Bayes and Longest Short Term Model (LSTM) models for sentiment analysis. The LSTM model yielded an accuracy of 0.93 5 while Naive Bayes gave an accuracy of 0.87 [5]. In paper [8] the authors suggested an ensemble model technique that combines various algorithms for improved accuracy, as opposed to employing Naive Bayes and Support Vector Machine classifiers alone.

Zeenia Singla et al. collected over 4,00,000 reviews from the cell phone category and employed three ML algorithms namely - Naive Bayes, Support Vector Machine, and Decision Trees. The evaluation of the three models has been using 10-fold cross-validation. The results showed SVM to yield the highest accuracy of 81.75% and Naive Bayes with the lowest accuracy of 66.95% [9].

The authors of [10] mined reviews of Redmi Note 3, Samsung J7, and iPhone 5S and employed Naive Bayes and Logistic Regression ML algorithms. Naive Bayes demonstrated the highest performance among the two. The findings of [11] accentuate that NB and SVM with Bag-of-Words as a feature extraction method yield above 93% accuracy. The paper also mentions that traditional supervised-learning algorithms like SVM, NB, RF, and LR perform well with the dataset. Researchers use the star rating system of Amazon to provide labels to the reviews and train the supervised algorithms.

In [12–15], the authors used four selection features - forward selection wrapper method, ANOVA, decision tree method and chi-square method to extract corpus features and then studied how the method chosen played a role in a classifier's performance. They worked on a synthetic spam dataset obtained from the University of Illinois (TripAdvisor dataset). Their experiments revealed the following results - Random Forest gave the highest accuracy (82% and 81% respectively) when feature selection was done using the Chi-square and forward selection wrapper methods. Naive Bayes gave the highest accuracy of 79% when the decision tree method was leveraged. When the ANOVA method was used for feature selection, logistic regression gave the highest accuracy of 80%.

Jaspreet Singh et al. [16] explored various algorithms used for sentiment prediction and provided a methodology for optimization of sentiment prediction using WEKA which is an opensource software tool that contains modules for data processing and provides implementation of machine learning algorithms. They employed and compared the accuracies of machine learning classifiers J48, BFTREE, and OneR by using WEKA for sentiment classification in text. The J48 algorithm is the Decision Tree Classification method implemented in Java. A classification technique called the BFTree Algorithm

only extends the best node in depth-first order. The difference between J48 and BFTree lies in the approach in which the decision tree is built. OneR is a classification approach that limits the decision tree to level one and generates only one rule. They proposed a methodology that involves preprocessing web text using Python with NLTK and bs4 libraries, extracting features from datasets, converting files for WEKA, and assigning sentiment labels. Feature selection is performed based on frequently used headings and titles using methods like Document Frequency, Mutual Information, and Information Gain. Probability distributions and statistical notations are used in the process. Text is normalized, and data is split into training and testing subsets. Four classifiers are trained and evaluated using precision, recall, accuracy, and F-measure as metrics for sentiment prediction optimization. After testing the models, they concluded that OneR is leading in the percentage of correctly classified instances (91.3%) whereas J48 has produced promising accuracy in true positive (96.7%) and false positive rates (0.003%).

In their study, Erik Boiy et al. [17] conducted trials to learn how to classify sentences from multilingual texts according to their sentiment. Based on the task's difficulty, they found that an integrated approach incorporating techniques from information retrieval, natural language processing, and machine learning generated good results. They discussed the advantages and emphasized the reason for the selection of each machine-learning technique (Table 3).

Table 3. Techniques used in [17]

Technique Used	Reason for Selection
Support Vector Machine (SVM)	Robust and yielded high accuracies
Multinomial Naïve Bayes (MNB) classifier	Simple to implement and computationally efficient
Maximum Entropy (ME) classifier	Yields good results in information extraction from natural language texts

When using the supervised approach, a lot of effort is required to annotate examples and train for each language. This becomes extremely hard when the texts incorporate local jargon and neglect the rules of grammar. Hence, active learning methods have been introduced to reduce the number of examples needed for training and still produce good results. Their observations include:

- Performance is increased when neutral sentences are filtered first, provided enough training examples are available
- Active learning methods, especially when multiple them are combined produce an improved F-measure average compared to their randomly selected counterparts
- Unigram features when augmented with certain language-specific features yield slightly improved baseline classification,

In their work, E. Fersini et al. [18] presented a unique ensemble method for sentiment categorization. The idea behind developing ensemble mechanisms is to group different independent learners to outperform the best baseline classifier. There are several traditional ensemble methods used for sentiment analysis.

- Majority Voting: It is a largely used ensemble technique that classifies by picking the most popular label without addressing diversity.
- Bagging: In this technique, data subsets are randomly picked and each bag is used to train different baseline learners of the same type. Diversity is obtained by using this method.
- Boosting: It builds a strong classifier on an incremental basis. Each model accurately classifies the instances that were misclassified by the previous model.

However, this paper highlights a new method that uses a heuristic for the Bayesian model to compute the discriminative marginal contribution that each classifier provides instead of using the existing methods that have uniform distributed weights.

Further, the limitations of current approaches were considered and a Bayesian approach that takes into account the marginal predictability of each model and a backward elimination-based greedy strategy has been developed to derive the optimal classification ensemble.

In their paper, Mohammad Aman Ullah et al. [19] presented a technique for sentiment analysis that makes use of both text and emoticons. They compared the results by analyzing only text and both text and data and concluded that adaptation of emoticon lexicons provides better accuracy than analyzing just text. Their proposed system corresponds to an accuracy of 79 percent over text and 89 percent over text and emoticons compared to the existing methods which account for 57 and 65,84 percent accuracies respectively. They found a means to improve classification accuracy by using deep learning algorithms. Their proposed system employs LSTM and CNN where LSTM works on embedding layers and cells to conclude information on words in the data and CNN pads the training sentences and passes to the embedding layer. The final observation was that, among the various algorithms compared, LSTM outperformed other deep and machine learning algorithms.

While comparing the performance of different classifiers on the Amazon reviews dataset, Sara Ashour Aljuhani et al. [20] built a prediction model for assigning polarities to reviews. Along with naïve Bayes, logistic regression, and convolution neural network, Stochastic gradient descent (SGD) was also included. They applied the algorithms using various techniques for feature extraction such as GloVe, TF-IDF, word2vec, and bag-of-words to find that CNN with word2vec resulted in the highest accuracy amongst all on both balanced and unbalanced data with accuracies of 92.72 % and 79.60 % respectively (Tables 4 and 5).

Local Interpretable Model-agnostic Explanations (Lime) were used to interpret the classification and they further concluded that the length of a review can be added as a feature to the ML algorithms for identifying polarities.

Large-scale unlabelled datasets are well-suited for a novel aspect-based sentiment analysis approach presented by Sumaia Mohammed et al. [21]. They adopted a combination of frequency and syntactic-based approaches to extract relevant aspects. They introduced a 4-step approach that outperformed all baselines. The steps include:

Table 4. Accuracies on Unbalanced Data

ML Model	Feature Extraction Approach Used	Accuracy Recorded
Logistic Regression	Bag-of-words + Trigram	91.72
Naïve Bayes	TF-IDF + Bigram	85.69
Stochastic Gradient Descent	Bag-of-words + Trigram	89.51
Convolutional Neural Network	Word2vec	92.72

Table 5. Accuracies on Balanced Data

ML model used	Feature extraction technique used	Accuracy recorded
Logistic Regression	Bag-of-words + Trigram	77.47
Naïve Bayes	TF-IDF + Trigram	74.90
Stochastic Gradient Descent	Bag-of-words + Trigram	76.05
Convolution Neural Network	Word2vec	79.60

1. Extracting main aspects from the domains
2. Using sentence level approach to extract core terms of each main aspect
3. Assigning rating for the extracted aspects based on the modified TF-IDF scheme
4. Calculating total sentiment score using domain-specific lexicon based on aspects.

Babita Gupta et al. [22] conducted an exploratory study that studied the relationship between a consumer's opinion on sentiment analysis and their financial transaction behaviors based on their outlook. They performed opinion mining on over 15,000 original tweets about cybersecurity sent out twice in six months.

- Tweets with keywords such as “safety”, “smart”, “glad”, and “awesome experience at the bank” conveyed a positive attitude toward cybersecurity.
- Tweets with keywords like “fraud”, “vulnerable”, “fear”, and “dangerous” conveyed a negative sentiment.
- Words such as “unconcerned”, and “compliance”, are objective and express neutral emotions. Bag-of-words approach was used for feature extraction and the authors laid down three hypotheses as follows:
- A greater number of online financial transactions during the second period ($T = 2$) compared to $T = 1$ would be correlated with a more positive mood toward the first period ($T = 1$).
- A decreased quantity of online transactions at $T = 2$ compared to $T = 1$ would be correlated with a negative emotion during $T = 1$.
- When compared to $T = 1$, the number of online transactions at $T = 2$ would be the same for a neutral posture during $T = 1$.

Between August 1 and August 3, 2016 ($T = 1$) and September 27 and September 29, 2016 ($T = 2$), a pilot study was carried out to evaluate the hypotheses. The results

showed that overall sentiment toward cybersecurity was negative for $T = 2$ and neutral for $T = 1$.

The authors [23] studied people's sentiments over time after the Universal Health Services (UHS) cybersecurity breach in September 2020 during the COVID-19 pandemic. UHS fell victim to a ransomware attack that halted the IT systems forcing hospitals to document patient information using pen and paper. The attack was cleverly orchestrated on a weekend when IT personnel would not have been available and during a critical time of the pandemic. Researchers analyzed relevant tweets during the period using the RoBERTa model developed by Facebook Research. Their findings are accentuated as follows:

- Due to increased public interest in the occurrence, there is a peak in the number of tweets during the first three days of the event (Week 40 of 2020).
- A second peak happened on October 21, about a month later, as a result of alerts from government agencies like Homeland Security.
- There was a drop in the number of tweets in week 44 of 2020 owing to the waning public interest since no similar cyberattack occurred post the events of September 2020.
- More tweets were classified as neutral than positive and negative. This indicated that there were no viewpoints expressed in the tweets; they were just reporting the occurrence.
- The number of positive sentiments was always lesser than the number of negative sentiments since events such as cyberattacks have a negative association tied to them.
- The study revealed that people posted positive tweets during that time. This could be due to their confidence in the IT personnel, and UHS' ability to recover from the attack within a short period.

3 Applications

Sentiment analysis can be extended to nearly all facets such as product and financial services, healthcare, social events, etc. Nearly 49% of customers trust online reviews as much as personal recommendations, as found out by BrightLocal in 2022. Sentiment analysis can help gauge the feedback of other reviewers that can help a customer in decision-making without having to peruse all the reviews manually and then forming an opinion. Opinion mining can be used to extract both opinions and the reason behind them. Such brand monitoring can help manufacturers fix problems in their offerings and improve businesses. Public discourse on social media platforms can be highly indicative of real-world scenarios such as cyberattacks, and sentiment analysis can be employed to take preventive measures before a cyber breach occurs. Further applications include detecting emotional sentiment from emails, market analysis, and competition research [2].

4 Conclusion

This paper serves as a comprehensive literature review for existing algorithms used to perform sentiment analysis along with their pros and cons. Feature extraction strategies and data preprocessing steps employed during opinion mining have been mentioned. Several works performed on the Amazon Reviews dataset have been studied and their

performances have been depicted in a systematic tabulated manner. The role of sentiment analysis in the field of cybersecurity has also been explored and people's online behavior during the time of a concerning cyber breach has been studied in this paper.

References

1. Joshi, N.S., Itkat, S.A.: A survey on feature level sentiment analysis. *Int. J. Comp. Sci. Info. Technol.* **5**(4), 5422–5425 (2014)
2. Khan, M.T., Durrani, M., Ali, A., Inayat, I., Khalid, S., Khan, K.H.: Sentiment analysis and the complex natural language. *Comp. Adap. Sys. Model.* **4**(1), 1–19 (2016)
3. Devika, M.D., Sunitha, C., Ganesh, A.: Sentiment analysis: a comparative study on different approaches. *Procedia Comp. Sci.* **87**, 44–49 (2016)
4. Medhat, W., Hassan, A., Korashy, H.: Sentiment analysis algorithms and applications: A survey. *Ain Shams Eng. J.* **5**(4), 1093–1113 (2014)
5. Meghana, M.S., Abhijith, D., Aysha, S., Kollu, P.K.: Sentiment analysis on amazon product reviews using lstm and naive bayes. In: 2023 7th International Conference on Computing Methodologies and Communication (ICCMC), pp. 626–631. IEEE (2023)
6. Pandey, P., Soni, N., et al.: Sentiment analysis on customer feedback data: Amazon product reviews. In: 2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon), pp. 320–322. IEEE (2019)
7. Haque, T.U., Saber, N.N., Shah, F.M.: Sentiment analysis on large scale amazon product reviews. In: 2018 IEEE International Conference on Innovative Research and Development (ICIRD), pp. 1–6. IEEE (2018)
8. Sadhasivam, J., Kalivaradhan, R.B.: Sentiment analysis of amazon products using ensemble machine learning algorithm. *Int. J. Mathemat. Eng. Manage. Sci.* **4**(2), 508 (2019)
9. Singla, Z., Randhawa, S., Jain, S.: Sentiment analysis of customer product reviews using machine learning. In: 2017 International Conference on Intelligent Computing and Control (I2C2), pp. 1–5. IEEE (2017)
10. Kumar, K.S., Desai, J., Majumdar, J.: Opinion mining and sentiment analysis on online customer review. In: 2016 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC), pp. 1–4. IEEE (2016)
11. Salmony, M.Y.A., Faridi, A.R.: Supervised sentiment analysis on amazon product reviews: A survey. In: 2021 2nd International Conference on Intelligent Engineering and Management (ICIEM), pp. 132–138. IEEE (2021)
12. Shringi, S., Sharma, H.: Methods for optimal feature selection for sentiment analysis. In: Proceedings of International Conference on Data Science and Applications: ICDSA 2022, Vol. 1, pp. 281–294. Springer (2023)
13. Shringi, S., Sharma, H., Suthar, D., et al.: Fitness-based grey wolf optimizer clustering method for spam review detection. *Mathematical Problems in Engineering* 2022 (2022)
14. Shringi, S., Sharma, H.: Detection of spam reviews using hybrid grey wolf optimizer clustering method. *Multimedia Tools and Applications* **81**(27), 38623–38641 (2022)
15. Shringi, S., Sharma, H.: Hybrid approaches for spam review detection: A review. *Palestine Journal of Mathematics* 11 (2022)
16. Singh, J., Singh, G., Singh, R.: Optimization of sentiment analysis using machine learning classifiers. *HCIS* **7**, 1–12 (2017)
17. Boiy, E., Moens, M.-F.: A machine learning approach to sentiment analysis in multilingual web texts. *Inf. Retrieval* **12**, 526–558 (2009)
18. Fersini, E., Messina, E., Pozzi, F.A.: Sentiment analysis: Bayesian ensemble learning. *Decis. Support Syst.* **68**, 26–38 (2014)

19. Ullah, M.A., Marium, S.M., Begum, S.A., Dipa, N.S.: An algorithm and method for sentiment analysis using the text and emoticon. *ICT Express* **6**(4), 357–360 (2020)
20. Aljuhani, S.A., Alghamdi, N.S.: A comparison of sentiment analysis methods on amazon reviews of mobile phones. *Int. J. Adv. Comp. Sci. Appl.* **10**(6) (2019)
21. Al-Ghuribi, S.M., Noah, S.A.M., Tiun, S.: Unsupervised semantic approach of aspect-based sentiment analysis for large-scale user reviews. *IEEE Access* **8**, 218592–218613 (2020)
22. Gupta, B., Sharma, S., Chennamaneni, A.: Twitter sentiment analysis: An examination of cybersecurity attitudes and behavior (2016)
23. Abusaqer, M., Senouci, M.B., Magel, K.: Twitter user sentiments analysis: Health system cyberattacks case study. In: 2023 International Conference on Artificial Intelligence in Information and Communication (ICAIIC), pp. 018–024. IEEE (2023)



Cultivating Cyber Vigilance: Shaping Employee Behavior for Security Success

Antima Sharma¹ , Anadi Trikha¹ (✉) , Preeti Nagar¹ , Arpita Agarwal¹ , and Akeke Niyi Israel²

¹ Manipal University Jaipur, Jaipur, Rajasthan, India

{antima.sharma, anadi.trikha}@jaipur.manipal.edu

² Ekiti State University, Ado-Ekiti, Nigeria

niyi.akeke@eksu.edu

Abstract. Integrating technology and the Internet into the workplace can offer benefits and concerns. While technology might help to simplify processes, it can also introduce complexity and potential hazards. It is vital to be aware of these hazards and take measures to mitigate them to provide a safe and effective work environment. Cyber-attacks are becoming more common these days. With cyber-crime on the rise, neglecting security safeguards is no longer an option. Web-based security is now the entire organisation's responsibility, not just IT security professionals. The study here employed bibliometric analysis since citation count is recognised as one of the most comprehensive and universally acknowledged metrics of intellectual advancement. The authors freely and selectively cite from the literature to develop their research work.

Keywords: Cybersecurity · Workplace · Employees Awareness

1 Introduction

Technology and the internet's increased integration into the workplace can bring benefits and concerns. While it can simplify activities, it can also bring complexity and potential hazards that must be addressed. To promote a safe and effective work environment, it is critical to be aware of these dangers and make efforts to mitigate them. Cyber-attacks are getting increasingly widespread these days. With the growing threat of cybercrime, ignoring security precautions is no longer an option. It is now necessary for the entire organisation, not just IT security professionals, to assume responsibility for web-based security (R. Menaka, 2022). Everyone, from tiny enterprises to major corporations, is vulnerable to cyber threats. With the increasing severity of these attacks, businesses must take the required precautions to safeguard their reputation and business.

Organizations face rising cybersecurity challenges due to the digital world's evolution. Organizational culture plays a vital role in maintaining successful cybersecurity practices. Protecting sensitive information from cyber threats is crucial for safeguarding the firm and its stakeholders.

Due to the possible hazards to the safety and well-being of those intimately involved with the organisation, ensuring the highest degree of cyber security measures is critical. Implementing practical security standards may protect against a wide range of possible risks and weaknesses, from securing sensitive information to thwarting cyber-attacks. As a result, organisations must prioritise and invest in cyber security measures to protect their people and assets. It is common knowledge that security breaches occur often in the workplace. Human mistakes are frequently to blame for these breaches. As a result, it is critical that every employee understands the potential dangers and is prepared to address any security-related problems that may emerge. Organisations should take precautions to avoid cyber-attacks.

2 Methodology: Theoretical Framework of Bibliometric Analysis

This study uses qualitative and quantitative techniques to analyze cyber security literature from 2018 to 2022, identifying 27 publications out of 134 using a bibliographic framework and citation analysis from the Scopus database. 27 research papers were selected for quantitative review using Scopus software to identify trends in cyber security and cyber attacks for a 5-year period from 2018 to 2022. Microsoft Excel was used to manage the data.

The study used bibliometric analysis since “citation count” is regarded as the most extensively and widely accepted measure of intellectual promotion. The authors freely and selectively quote the literature and create their study work. Table 1 shows a list of the most often mentioned writers about the turnover intention of hotel employees. Twenty-seven research publications on the turnover intention of hotel employees were analyzed to analyze frequently referenced authors. Some research on cybersecurity awareness and strategies to improve workplace cybersecurity attacks through employee training has been undertaken.

The Diagram Below Depicts the Research Framework and Theoretical Foundation:
Fig. 1.

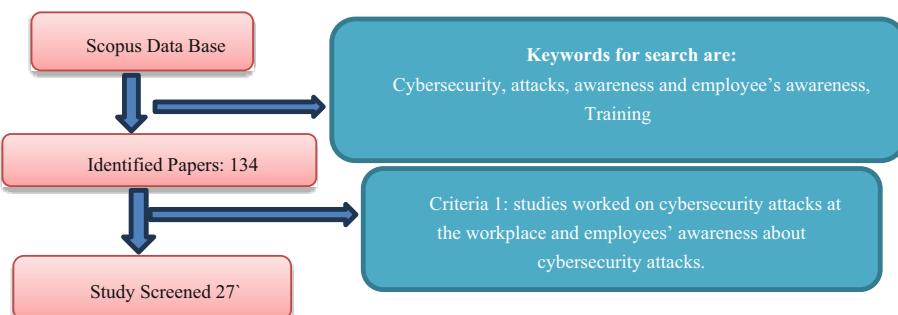


Fig. 1. Flowchart of data screening

3 Result and Interpretation

Out of the 134 papers that were identified in this study regarding cyber attacks at the workplace and employee awareness of cyber attacks, 27 papers were selected for further analysis.

3.1 Access Types

Initially, the data for the study was analysed based on the “type of access and the number of publications”. Table 1 shows that 74% of the identified documents on “cyber security and cyber-attacks” were open access. Research outputs are not always publicly available online since it would be challenging for researchers to obtain and access materials, data, and information from sources, including journal articles, conference papers, and theses. Two publications were made in 2018, and the topic was mainly unpublished until 2022, when there were fourteen publications a year, up from less than ten in the previous year. After the year, there is a continuous increase in the number of publications. A rise in interest in the topic might explain this (Table 2).

Table 1. The access type of selected papers

Access type	Frequency	% (N = 606)
Open access	20	74%
Other (non-open access)	7	26%
Total	27	100%

Table 2. Citations-based research papers and journals.

S. No.	Title	Journal Name	Year	Citations
1	Keeping customers' data secure A cross-cultural study of cybersecurity compliance among the Gen-Mobile workforce	Computers in Human Behavior	2021	56
2	Influence of human factors on cyber security within healthcare organisations: A systematic review	Sensors	2021	54
3	Robotics cyber security: vulnerabilities, attacks, countermeasures, and recommendations	International Journal of Information Security	2022	50

(continued)

Table 2. (*continued*)

S. No.	Title	Journal Name	Year	Citations
4	Cyber security training for critical infrastructure protection: A literature review	Computer Science Review	2021	45
5	Human factor security: evaluating the cybersecurity capacity of the industrial workforce	Journal of Systems and Information Technology	2019	35
6	Leveraging human factors in cybersecurity: an integrated methodological approach	Cognition, Technology and Work	2022	32
7	Encouraging Employee Engagement With Cybersecurity: How to Tackle Cyber Fatigue	SAGE Open	2021	23
8	Training and embedding cybersecurity guardians in older communities	Conference on Human Factors in Computing Systems - Proceedings	2021	19
9	The COVID-19 academic: A survey of phishing attacks and their countermeasures during COVID-19	IET Information Security	2022	14
10	Avoidance of cybersecurity threats with the deployment of a web-based blockchain-enabled cybersecurity awareness system	Applied Sciences (Switzerland)	2021	14
11	Cybersecurity awareness training programs: a cost-benefit analysis framework	Industrial Management and Data Systems	2021	15
12	Antecedents for enhanced level of cyber-security in organisations	Journal of Enterprise Information Management	2021	12
13	Refining the PoinTER “human firewall” pen testing framework	Information and Computer Security	2019	7
14	Understanding the state of criminological scholarship on cybercrimes	Computers in Human Behaviour	2023	5

(continued)

Table 2. (*continued*)

S. No.	Title	Journal Name	Year	Citations
15	Effective information system and organisational efficiency	Polish Journal of Management Studies	2021	5
16	Prevention and mitigation measures against phishing emails: a sequential schema model	Security Journal	2022	5
17	Reconceptualising cybersecurity awareness capability in the data-driven digital economy	Annals of Operations Research	2022	4
18	Handbook of research on advancing cybersecurity for digital transformation	Handbook of research on advancing cybersecurity for digital transformation	2021	4
19	Antecedents of Cybersecurity Implementation: A Study of the Cyber-Preparedness of U.K. Social Enterprises	IEEE Transactions on Engineering Management	2022	3
20	Generation Z: Cyber-Attack Awareness Training Effectiveness	Journal of Computer Information Systems	2022	3
21	Effects of Security Knowledge, Self-Control, and Countermeasures on Cybersecurity Behaviors	Journal of Computer Information Systems	2023	3
22	The Importance of Conceptualising the Human-Centric Approach in Maintaining and Promoting Cybersecurity-Hygiene in Healthcare 4.0	Applied Sciences (Switzerland)	2023	2
23	Assessing the Relevance of Cybersecurity Training and Policies to Prevent and Mitigate the Impact of Phishing Attacks	Journal of Internet Services and Information Security	2022	2
24	Strategic cyber security management	Strategic Cyber Security Management	2022	1
25	The Role of Top Managers' IT Security Awareness in Organizational IT Security Management	ICIS 2017: Transforming Society with Digital Innovation	2018	1

(continued)

Table 2. (*continued*)

S. No.	Title	Journal Name	Year	Citations
26	The Current State of “Information Security Awareness” in German SMEs	International Journal of Emerging Technology and Advanced Engineering	2021	1
27	A Model for Information Security Culture with Innovation and Creativity as Enablers	IFIP Advances in Information and Communication Technology	2022	1

3.2 Document Type

Table 3, summarises the document forms, with journal papers accounting for the bulk, conference papers accounting for 4 per cent, and book chapters accounting for 4 per cent. The majority of research studies selected in this study are for journal publication.

Table 3. Research published in different categories.

Document Type	Frequency	Percentage
Journal Articles	25	92%
Handbook	1	4%
Conference	1	4%
Total	27	100%

3.3 Subject Areas: Covered in Screened Research Studies

Table 4 depicts the subject areas of articles produced between 2018 and 2022. According to this survey, 63 percent of cyber security and cyber attack materials disclosed fall under Science and Technology. This is accompanied by 18.5% in the field of Information Management Systems and 7.4% in Psychology and Applied Science, respectively. Management studies account for only 3.7% of the total. Previous writers' works in Social Sciences, Decision Science, and Energy are also available. This shows that the problem's difficulty and multidisciplinary nature are essential in various sectors. This is crucial since cyber threats and attacks cover many subjects, not just computer science. This, however, makes it more difficult for potential academics to perform education-related literature searches.

Table 4. Subject Areas: **Covered in Screened Research Studies**

Document Type	Frequency	Percentage
Psychology	2	7.4%
Science & Technology	17	63%
Applied Science	2	7.4%
Information Management System	5	18.5%
Management Studies	1	3.7%

3.4 Authorship Analysis

The approach of data collection resulted in the creation of 27 datasets. The datasets revealed information about the most productive writers, including those who have produced many academic articles. Key authors who have made contributions to the field of cyber-attacks at the workplace and employee awareness are highlighted in Table 5.

Table 5. A review of central studies focused on cyber-attacks and cyber security awareness among employees.

S. No.	Title	Author Name & Year	Summary
1	“Keeping customers’ data secure A cross-cultural study of cybersecurity compliance among the Gen-Mobile workforce.”	Ameen et al. (2021)	Employees rely increasingly on mobile devices for work-related tasks, especially in global organizations. To ensure compliance with smartphone security measures across countries, a comprehensive approach is needed, including national, organizational, technological, and personal interventions. Additionally, BYOD policy should be considered in the workplace
2	“Influence of human factors on cyber security within healthcare organisations: A systematic review.”	Nifakos et al. (2019)	In healthcare organizations, a structured methodology is required to coordinate research outcomes for impartial cybersecurity assessments. A unified and standardized approach is necessary to develop training programs and awareness campaigns that strengthen defenses against growing cyber threats

(continued)

Table 5. (*continued*)

S. No.	Title	Author Name & Year	Summary
3	“Robotics cyber security: Vulnerabilities, attacks, countermeasures, and recommendations”	Yaacoub et al. (2022)	Robotic systems in critical infrastructures are vulnerable to security risks that can lead to dangerous attacks. To reduce the risk of unauthorized access, it is recommended to implement a mutual multi-factor authentication scheme for the highest level of security
4	“Human factor security: evaluating the cybersecurity capacity of the industrial workforce”	Ani et al. (2019)	Developing cybersecurity knowledge and skills is crucial for creating a competent and compliant workforce in the industrial sector. Negligence, misinformation, and inadequate knowledge and skills are key factors responsible for cyberattacks on the workforce
5	“Training and Embedding Cybersecurity Guardians in Older Communities”	Nicholson et al. (2021)	Safe online practices are critical as older adults participate more in online activities. Training older adults as Cyber Guardians can change cybersecurity behavior, regardless of technical know-how, and have a positive impact on their overall well-being
6	“Generation Z: Cyber-Attack Awareness Training Effectiveness”	White (2022)	To gain a more comprehensive understanding, a study should be conducted with a separate group of participants who receive awareness training before a cyber-attack reading. Further studies should investigate the efficacy of pre-cyber-attack general awareness training on countermeasures

3.5 Keyword Research and Cluster Analysis

The co-occurrence of the keywords has been studied using VOS viewer, and the following clusters were reported as stated in Table No. **The main keywords of research studies are shown in Figure 2 below:**

Through this research study, four clusters were formed, which are shown in Table 6 below:

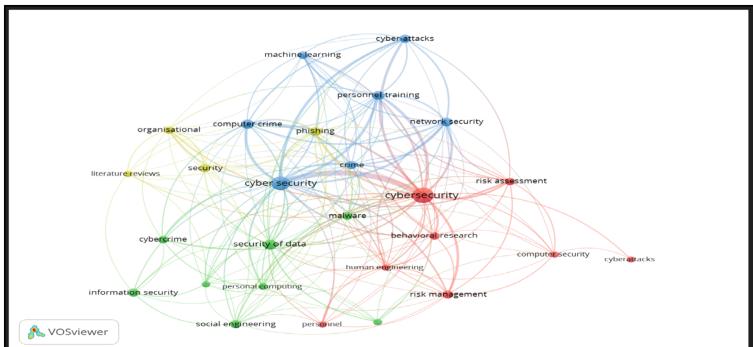


Fig. 2. Abstract Keywords Co-Occurrence

Table 6. Clusters of Keywords Occurrence.

Clusters	Item Nos.	Items
Cluster 1	8	behaviors research, computer security, cyberattacks, cyber security, human engineering, personal, risk management and risk assessment
Cluster 2	8	cybercrime, design/methodology/approach, information security, malware, personal computing, data security, social engineering and systematic literature review
Cluster 3	7	computer crime, crime, cyber security, cyber-attacks, machine learning, network security and personal training
Cluster 4	4	literature review, organisational, phishing and security

3.6 Citation Analysis

The sample for the study included all the published documents on Cyber security and cyber-attacks in the 5 years of publications, 2018–2022; of the 27 papers, 6 papers were found related in 2018–2022. Top 3 highly rated journals, Top 3 mostly cited authors, most cited topics of top 3.

3.6.1 The Most Cited Authors and the Journal

From the Table 1. The Top 3 highly cited authors are:

- I. Ameen, Nisreen Tarhini, Ali Shah, Mahmood Hussai, Madiche, Nnamdi Paul, Justin Choudrie, Jyoti.
 - II. Nifakos, Sokrati Chandramouli, Krishna Nikolaou, Charoula Konstantin Papachristou, Panagiotis Koch, Sabine Panaousis, Emmanouil Bonacina, Stefano
 - III. Yaacoub, Jean-Paul A, Noura, Hassan N. Salman, Ola Chehab, Ali

*****The most preferred journal by these authors is “Computers in Human Behaviour” and “Sensors.

4 Discussion of the Study

In this study, two significant themes were identified:

Theme 1: Cyberattacks at the Workplace.

Organisations today have access to information. There's a lot of it. Information in many formats. Businesses use this information. They save this data. Often, the information does not belong to the company but to a third party. Businesses realise that holding information susceptible to third-party information makes them a target (Chigada and Madzinga, 2021; Bada and Nurse, 2020). Daily news headlines warn us that cyber threats are on the rise. Cyber risks are becoming more common, prominent, and severe. The reputation of the organisation and the business itself are under constant threat.

It is beneficial to go into the specifics of each attack to understand better the types of cyber-attacks that can occur in the workplace. Here are a couple of such examples:

- i. **Email Phishing:** Attackers use fraudulent emails to trick unwary employees into disclosing personal information or downloading dangerous attachments in this assault. These emails frequently look to be issued from a reliable source, such as a bank or a colleague, and carry a sense of urgency, pushing the receiver to respond immediately (Argaw and Bempong, 2019; Ahmadi-Assalemi and Khateeb, 2019). When an employee falls for the lure, the attacker acquires access to sensitive information or installs malware.
- ii. **Spear Phishing:** This is a highly focused phishing attack in which thieves extensively research specific individuals or companies to generate customised communications that appear natural. They frequently leverage social media information to build a sense of familiarity and trust, boosting the likelihood of the recipient taking the desired action (Unger, 2021; Vadlamudi and De, 2021). The purpose of this assault is similar to that of email phishing in that it seeks sensitive information or introduces malware into the system. Still, the method is more complex and customised.

Phishing is a significant challenge in information security today, with attackers using various permutations and social engineering approaches. Public concern about cyber security has increased due to frequent news stories of cyber attacks, highlighting the seriousness of the situation. (Thomas, 2018; Abdelhamid et al., 2023). According to an Economic Times article, cybercrime in India has gradually climbed by 15–20%. (Economic Times, 15 December 2023). Organisations must prioritise teaching their staff about the dangers of phishing. Spear phishing and cyber-attacks and providing them with the tools and resources they need to protect themselves (Bada and Nurse, 2020; Chigada and Madzinga, 2021).

- iii. **Exploitation Via Social Engineering:** Social engineering is tricking, manipulating, or influencing a victim to take control of a computer system or obtain personal or financial information (Stacey et al., 2021). Employees may unwittingly provide sensitive information or unauthorised access to attackers impersonating colleagues due to a lack of awareness about contemporary social engineering strategies.

Theme 2: Awareness of Cyberattacks Among Employees and Their Training.

Cybersecurity awareness training teaches employees to recognize cyber dangers and protect data privacy for the organization and its stakeholders. (Daengsi and Pornpongtechavanich, 2021; Al-Mohannadi et al., 2018). Cybersecurity awareness training aims to increase employees' knowledge and comprehension of potential cybersecurity risks and help them implement actions to safeguard themselves and their organisations against cyber-attacks and data breaches.

Employees benefit from cyber security awareness training because it helps them comprehend the dangers and threats of cyber-attacks. Organisations can considerably minimise the likelihood of falling victim to an attack by equipping people with the information and skills to recognise possible cyber threats (Boto-Garcia, 2023; Roy and Joseph, 2023).

Cybersecurity awareness training is crucial as employees are often the weakest link in an organization's security. Teaching personnel to recognize and respond appropriately to phishing emails or social engineering can reduce the likelihood of a successful cyber attack. (Kemper, 2019; Aldawood and Skinner, 2019). Furthermore, cybersecurity awareness training is essential because it helps employees understand potential cyber dangers and how to recognise and prevent them. It safeguards organisations from cyber assaults and data breaches. Organisations are more vulnerable to cybercrime if they lack sufficient training.

5 Conclusion

This study intends to raise public knowledge of cyber security and attack research by analyzing and describing existing literature. More research is needed on this topic, including studies within higher education frameworks, as cybersecurity is a global concern. While the US has the most influence in the literature, research on developing nations is complex.

Most cyber-attack and cybersecurity literature is password-protected, making it hard to find relevant information. This study uses bibliometric analysis to explore trends and developments and highlights the need for further long-term research to comprehend the challenges fully.

This article was prompted by two observations: Firstly, “cyber-attacks and cyber security” concerns have been a heated matter in recent years, but it is still considered infrequent, and second, dealing with this difficulty in the sector is challenging. Despite these factors, there have been inadequate studies on the status of advancement in this sector. This chapter opened by stating a lack of clarity regarding the advancement in cyber security and cyber-attacks. The analysis has limitations due to the database used. Scopus does not archive all journals, and the evaluation only focuses on malware assaults and cyber issues. Future research could include additional databases and comparative studies with various research procedures to obtain more detailed outcomes.

Despite these drawbacks, this study is among the first to look at various bibliometric indices of literature in cyber security and cyberattacks. This study aims to lay the foundation for future discussions about higher education's response to cyber threats and hazards. The results must aid future researchers in comprehending how to develop the problem further. It is recommended that future researchers conduct textual analysis, as

this is likely to yield novel and insightful results. This is because basic search parameters, like article titles, abstracts, and keywords—all readily available electronically on the Scopus website—were employed in this study instead of looking through entire articles or complete records. The only database that is used to gather the data is Scopus.

References

- Abdelhamid, S., Mallari, T., Aly, M.: Cybersecurity Awareness, Education, and Workplace Training Using Socially Enabled Intelligent Chatbots. In: *The Learning Ideas Conference*, pp. 3–16. Springer Nature Switzerland, Cham (2023)
- Al-Mohannadi, H., et al.: Understanding awareness of cyber security threats among IT employees. In: *2018, the 6th International Conference on Future Internet of Things and Cloud Workshops (cloud)*, pp. 188–192. IEEE (2018)
- Aldawood, H., Skinner, G.: Reviewing cyber security social engineering training and awareness programs—Pitfalls and ongoing issues. *Future Internet* **11**(3), 73 (2019)
- Ameen, N., et al.: Keeping customers' data secure: a cross-cultural study of cybersecurity compliance among the Gen-Mobile workforce—computers. *Human Behavior* **114**, 106531 (2021)
- Ani, U.D., He, H., Tiwari, A.: Human factor security: evaluating the cybersecurity capacity of the industrial workforce. *J. Syst. Inf. Technol.* **21**(1), 2–35 (2019)
- Bada, M., Nurse, J.R.: The social and psychological impact of cyberattacks. In: *Emerging cyber threats and cognitive vulnerabilities*, pp. 73–92. Academic Press (2020)
- Chigada, J., Madzinga, R.: Cyberattacks and threats during COVID-19: A systematic literature review. *South African J. Info. Manage.* **23**(1), 1–11 (2021)
- Kemper, G.: Improving employees' cyber security awareness. *Comp. Fraud & Security* **2019**(8), 11–14 (2019)
- Nicholson, J., et al.: Training and embedding cybersecurity guardians in older communities. In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pp. 1–15 (2021)
- Roy, R., Joseph, F.J.: The impact of cybercrime awareness training among employees in corporates for better information management. *Academy of Marketing Studies Journal* **27**(S5) (2023)
- Stacey, P., Taylor, R., Olowosule, O., Spanaki, K.: Emotional reactions and coping responses of employees to a cyber-attack: A case study. *Int. J. Inf. Manage.* **58**, 102298 (2021)
- Unger, A.: Susceptibility and Response of Small Business to Cyberattacks, Doctoral dissertation. Utica College (2021)
- Vadlamudi, S., De, S.: A novel approach in cyber security for securing the workplace of the future in large industry setups. In: *2021 International Conference on Intelligent Technologies (CONIT)*, pp. 1–6. IEEE (2021)
- White, G.: Generation Z: cyber-attack awareness training effectiveness. *J. Comp. Info. Sys.* **62**(3), 560–571 (2022)
- Yaacoub, J.P.A., Noura, H.N., Salman, O., Chehab, A.: Robotics cyber security: Vulnerabilities, attacks, countermeasures, and recommendations. *Int. J. Info. Sec.* 1–44 (2022)



A Deep Learning Approach to PDF Malware Detection Enhanced with XAI

Kirubavathi Ganapathyappan^(✉) and Fathima Noorudheen

Department of Mathematics, Amrita School of Physical Sciences, Amrita Vishwa Vidyapeetham, Coimbatore, India
g_kirubavathi@cb.amrita.edu

Abstract. This paper presents a comprehensive approach to PDF malware detection, addressing the serious threat posed by malicious PDF documents. Traditional machine learning (ML) approaches have limitations in detecting these threats due to susceptibility to evasion attacks. To overcome this, the proposed method combines deep learning with LIME Explainable Artificial Intelligence for detecting PDF malwares. By doing so, the approach aims to achieve generalized robustness against code obfuscations employed by adversaries to evade antivirus software. To validate the effectiveness of the approach, the proposed work is compared with existing the state-of-art PDF malware methods. The results demonstrate a high detection rate of PDF malware, reaching 99.93%. Additionally, the method proves effective in detecting new malicious files created by simple methods that remove obfuscations conducted by malware authors, which often go undetected by most antivirus software. Overall, this comprehensive approach represents a significant advancement in PDF malware detection, offering improved detection rates and resilience against evasion techniques employed by malicious actors.

Keywords: PDF malware detection · Deep Learning · Cyber Security · Explainable AI

1 Introduction

The frequency of intelligent attacks using documents containing malicious codes has been rapidly increasing in recent years due to the expansion of file transfers. While most Internet users are aware of the risks associated with executable files attached to emails or web pages, they often underestimate the potential dangers posed by seemingly harmless documents. This lack of awareness makes documents an effective tool for spreading malware. Among the various types of malware, Portable Document Format (PDF) files are particularly vulnerable to exploitation due to their flexibility. Many malicious PDF documents contain JavaScript or binary scripts that exploit specific security vulnerabilities to carry out destructive actions, as explained in [1]. The internet is inundated with a vast number of PDF files, and not all of them are as innocuous as they may appear. In reality, PDF files can contain a wide range of objects, including binary or JavaScript codes, some of which can be detrimental. Malware programs can attempt to infect a computer

by exploiting vulnerabilities in PDF readers [2]. In 2017, Adobe Acrobat Reader alone had sixty-eight vulnerabilities, fifty of which could be exploited to execute arbitrary codes. Each PDF reader has its own set of weaknesses, and a malicious PDF file can find a way to exploit them [3].

2 Related Works

Various studies in the realm of PDF malware detection primarily concentrate on improving detection capabilities to tackle the ever-changing characteristics of malicious PDF files. Presented below are a few prevalent methodologies and techniques utilized in this domain.

The authors in [4] examines the growing danger posed by malicious content embedded in Portable Document Format (PDF) files, which is attributed to their widespread use and advanced features like JavaScript and file embedding. Conventional detection methods, such as signature-based Anti-Viruses, have limitations in effectively countering these sophisticated attacks. Although some studies utilize AI to enhance detection rates, evasive PDF malware remains a significant security concern. To bridge this gap, the authors introduce a framework that extracts 28 static features from PDF files, including 12 innovative ones, and employs stacking machine learning models for detection. The proposed solution is assessed using two datasets: Contagio and a newly created evasive PDF dataset (Evasive-PDFMal2022). The results demonstrate high accuracy and F1-scores, surpassing existing models on both datasets.

The study [5] presents a new technique for evading a learning-based malware classifier that specifically targets PDF files. This technique involves using a feature-vector generative adversarial network (fvGAN) to generate adversarial feature vectors in the feature space, which are then transformed into real adversarial malware examples. The effectiveness of the fvGAN model in achieving high evasion rates within a limited timeframe is demonstrated through evaluation against the PDFRate classifier under various evasion scenarios. A comparative analysis with existing attack algorithms, namely Mimicry and GD-KDE, highlights the superior performance of the proposed approach in terms of evasion rate and execution cost. However, the study acknowledges certain limitations, such as the requirement for human intervention in feature selection and extraction. This opens up opportunities for future exploration in the development of GANs that can directly generate actual PDF files without such constraints.

The research in [6] discusses the issue of malicious PDF documents and the importance of having strong threat intelligence platforms to combat them. Despite the use of machine learning in detecting PDF malware, evasion attacks still pose a challenge. The authors suggest a comprehensive approach that combines signal and statistical analysis of malware binaries to improve detection accuracy and resilience against obfuscation techniques. This involves using both static and dynamic analysis methods to create models that utilize various representations such as images, audio, hashes, and API call sequences. The paper highlights the practicality and effectiveness of this approach, particularly in detecting obfuscated malware that traditional antivirus solutions may miss. It lays the groundwork for robust threat intelligence platforms capable of protecting against a wide range of attacks. While the focus is on PDF files, the authors acknowledge the

potential applicability of this method to other data types that contain byte-streams and file structures, which can be explored in future research.

The authors in [7] discusses the difficulty of finding adversarial examples against machine learning-based PDF malware classifiers. This challenge is due to the complex data structure of PDFs and the need for generated PDFs to display malicious behavior while avoiding detection. The authors propose a modified version of generative adversarial networks (GANs) specifically designed for this purpose. They utilize the target classifier as a second discriminator and incorporate unique characteristics from malicious PDF files to quickly generate evasive variants. The effectiveness of the model is demonstrated through evaluation against state-of-the-art PDF malware classifiers and commercial antivirus engines. This evaluation also exposes vulnerabilities in antivirus engines, emphasizing the importance of employing multiple detectors to counter evasion attacks. Although the study primarily focuses on PDFs, future research could explore the application of this approach to other domains, such as binary software.

The research [8] focuses on the significant threat posed by malicious PDF documents and the limitations of current signature-based defense systems. These defense systems struggle to keep up with zero-day attacks. In order to enhance the detection of such threats, the authors present improved machine learning (ML)-based models that are specifically designed to detect malicious PDFs. They develop an approach that utilizes both existing and innovative static features derived from PDF documents to enhance the performance of ML-based classifiers. The authors evaluate seven ML classifiers using the recently published EvasivePDFMal2022 dataset, which consists of benign and malicious PDF samples. The results demonstrate that the proposed approach, with its enhanced features, improves accuracies in five out of the seven classifiers tested. This highlights the potential of the new features to enhance the robustness of feature-based PDF malware detection. The paper introduces these novel features and discusses how they are derived, providing empirical evidence of their effectiveness through experiments. Future research will focus on investigating the resilience of the enhanced feature set against different types of adversarial attacks.

The authors [9] discusses the crucial role of email-based initial penetration in cyber-attacks targeting corporations. In these attacks, hackers employ social engineering strategies to entice victims into opening malicious emails, often containing non-executable files like PDFs, which can bypass corporate defenses. To effectively identify these harmful PDF files, the authors propose the utilization of machine learning (ML) techniques. They employ various ML methodologies to enhance the accuracy of detection and the efficiency of processing, with a specific focus on the CIC-Evasive-PDFMal2022 dataset. Through their experimentation, they discover that the PART method achieves the highest detection accuracy of 98.97%, surpassing other evaluated ML algorithms. This research highlights the effectiveness of ML-based approaches in detecting PDF malware and aims to further enhance detection accuracy in future studies.

The research [10] focuses on the task of detecting adversarial examples in PDF malware detectors that are based on machine learning. It utilizes generative adversarial networks (GANs) to generate modified versions of PDF malware that can bypass existing classifiers while still maintaining their malicious behavior. By extracting specific characteristics from malicious PDF files, the production of evasive variants becomes

faster, and the PDF GAN is employed for malware detection. The system incorporates various machine learning techniques, such as feature selection and GAN methods, for classification purposes. In the future, this approach could potentially be expanded to include other file formats such as documents and presentations (Table 1).

Table 1. A comparative summary of the existing literature on pdf malware detection

Ref & Year	Approach	Dataset	Technique	Significance
[4]	ML & DL	Contagio dataset, the Common Vulnerabilities and Exposure (CVE) samples collected from Exploit-db	PDF-GAN, which leverages generative adversarial networks (GANs), use of feature extraction, selection, and mutation techniques to generate evasive PDF samples against ML-based classifiers	Generate evasive PDF samples capable of evading PDF malware classifiers, addresses the growing challenge of finding adversarial examples against machine learning-based PDF malware classifiers
[5]	DL	Contagio, Surrogate, and Attack dataset	fvGAN mimicus framework	Approach can achieve evasion rates far surpassing the GD-KDE and marginally outperforming the Mimicry attack More advantageous than the Mimicry attack when involving a considerably large number of adversarial malicious PDF files
[11]	ML	Benign sample set	Features extraction machine learning algorithms	Presenting a new methodology of features extraction based on document structure and scripting language, which can achieve more accurate results of malicious document detection Presenting the use of machine learning algorithms for the detection of malicious documents based on two layers of abstraction

(continued)

Table 1. (*continued*)

Ref & Year	Approach	Dataset	Technique	Significance
[12]	ML	Benign PDF files and Reverse Mimicry Attacks Dataset	PDF Forensic Analysis Tools, Anti-Malware and Security Software Network Analysis Tools	Overview of current attack techniques used to convey PDF malware and the analysis tools that support digital forensic investigations
[6]	ML	Malware and benign samples	10-fold cross-validation binary classification experiments with various standard classifier models	Holistic approach to PDF malware detection that leverages signal and statistical analysis of malware binaries, bag-of-words models, and feature fusion strategies
[7]	ML & DL	Contagio dataset, the Common Vulnerabilities and Exposure (CVE) samples collected from Exploit-db	PDF-GAN, which leverages generative adversarial networks (GANs), use of feature extraction, selection, and mutation techniques to generate evasive PDF samples against ML-based classifiers	Generate evasive PDF samples capable of evading PDF malware classifiers, addresses the growing challenge of finding adversarial examples against machine learning-based PDF malware classifiers
[13]	ML	Evasive-PDFMal2022	Feature extraction	Integrates anomaly-based features with structural features to improve the performance of ML classifiers
[14]	ML	Various large dataset	Naive Bayes (NB), Artificial Neural Networks (ANN), Support Vector Machine (SVM), K-Nearest Neighbours (KNN)	The PDF structure and it's working and about various classifiers which are used to detect malware detection
[15]	ML,DL	dz.a, dz.b virus share and hybrid analysis	Gene detection and classification	Combine bioinformatics and genetics and propose the pdf exploitable malwares gene to analyze whether the exploits are exploited in the pdf malware based on the software genes
[9]	ML/DL	PEid, PEview, Ghidra, Wireshark, Sandboxes, PDF Parser, and custom scripts	Static analysis, Dynamic analysis, Cuckoo sandbox, Script Analysis, Shell Scripting	Addressing the security level of PDF documents by employing static and dynamic analysis for malware detection

(continued)

Table 1. (*continued*)

Ref & Year	Approach	Dataset	Technique	Significance
[10]	ML	(CICEvasivePDFMal2022)	The algorithm was evaluated and trained by ten k-fold cross-validations	Apply machine learning techniques and multiple machine learning methodologies to identify dangerous programs and evaluate the accuracy of the acquired findings
[16]	ML/DL	15,000 PDF files, containing details such as object, end object, trailer, xref, pages, JavaScript, OpenAction, embedded files, and other relevant information	ML algorithms, GAN, feature extraction methods, use of a surrogate classifier and the training of the PDF GAN model with a dataset for malware detection	Exploration of a new deep learning-based malicious PDF file detector called MMPD, which can be installed on mobile robots. The paper also discusses the use of Generative Adversarial Networks (GANs) to create adversarial malware instances, addressing the challenges posed by PDF malware assaults
[17]	ML/DL	Contagio	Image visualization, feature extraction	A new learning-based method to detect PDF malware using image processing and processing techniques
	ML	Contagio	JavaScript extraction, obfuscation techniques and our de-obfuscation methods	The paper present a new detection system with static and dynamic detection methods targeting to JavaScript-based malicious PDF documents;Designing an approach to de-obfuscate JavaScript code embedded in PDF, which can observably rise the detection rate

This motivate us to implement and evaluate an enhanced deep learning-based approach on the Evasive-PDFMal2022 dataset stems from recent studies on PDF malware detection using machine learning methods. This dataset provides a valuable opportunity to test and validate our method against cutting-edge techniques, utilizing the knowledge gained from previous research to enhance the efficacy of our approach. Our primary focus lies in evaluate the deep learning algorithms to enhance the accuracy and resilience of our system against evasive PDF malware. This endeavor aligns with the overarching

research trend of advancing PDF malware detection techniques to effectively combat ever-evolving cyber threats.

3 Background Study

The PDF format is widely preferred for sharing digital documents across various platforms and applications. References [18, 19] provide a comprehensive description of the PDF standard. There are several factors that make PDFs a popular choice for malware authors to distribute harmful content. Firstly, PDFs are extensively used in both professional and social settings. They are commonly used for academic papers, technical reports, design documents, and electronic receipts. Secondly, PDFs are platform and operating system independent. They can be accessed using a standalone PDF reader or a modern web browser on Windows PCs, Linux systems, or mobile devices with a PDF viewer plug-in. Firstly, PDFs offer great versatility by supporting a wide range of data types such as text, video files, interactive forms, links to other files, JavaScript, Flash, and URLs. Additionally, various encoding and compression methods can be utilized to decrease file size, conceal important information, or achieve both objectives. Moreover, PDFs are discreet and sophisticated. In general, executable files are perceived as more hazardous compared to PDF files. Organizations frequently enforce security protocols to prevent employees from downloading executable files or attaching them to emails. Nevertheless, similar restrictions are not commonly imposed on PDF documents. The extensive usage and flexibility of the PDF file format create numerous opportunities for cyber attackers to disseminate malware through PDF files. Phishing and exploits are the main types of attacks that exploit PDF files.

Phishing schemes are often distributed via emails, masquerading as PDF delivery notifications or purchase receipts from reputable online retailers or shipping companies. These emails utilize social engineering tactics to persuade recipients to open the attached phishing PDFs, even though the actual content of the emails is typically nonsensical. These PDF files are usually single-page documents containing elements of social engineering, as well as a phishing URL that redirects users to a suspicious website where malicious downloads, personal information harvesting, and other harmful activities can occur. Unlike text-based phishing schemes, PDF documents incorporate binary or a combination of binary and ASCII languages, making them more challenging to identify. This complexity in detection is one of the reasons why phishing attacks utilizing PDFs have become more prevalent. The motivation behind these attacks is the same as that of standard phishing scams, aiming to gather information from victims for personal use or to be sold on the illicit market known as the shadow economy.

3.1 The Structure of PDF File

The Portable Document Format (PDF) is widely utilized for presenting content and layout across different platforms. Below is a breakdown of its typical file structure [21]:

Header: The header is the initial line of a PDF file, specifies the version of the PDF format being used.

Body: The body contains the main content of the PDF file. It consists of a series of objects that define the operations to be executed by the file. These objects can include different types of data, such as text, images, scripts, etc. They may be compressed or uncompressed. Each object has a unique reference number and can be referenced by other objects within the file.

Cross-Reference Table (CRT): The cross-reference table (CRT) is a crucial component that maintains a list of offsets indicating the positions of every indirect object in the file. Each entry in the table corresponds to a specific object, with one entry ending in ‘n’ indicating the total number of objects stored in the file. PDF readers use the information in the CRT to locate and parse objects within the file. However, they only parse objects that are referenced by the CRT. This selective parsing based on CRT references can potentially be exploited by attackers.

Trailer: The trailer is a special object that contains essential metadata about the structure of the PDF file. It provides instructions to the PDF reader on how to locate the CRT and other critical objects within the file. A standard PDF reader starts parsing a PDF file from the trailer. It then uses the information in the trailer to locate the CRT, which is subsequently used to find and parse objects within the body of the file. Understanding the structure of a PDF file is crucial for various purposes, including content extraction, manipulation, and security analysis.

4 Proposed PDF Malware Detection System

The proposed framework starts by dividing the dataset into feature vectors (X) and corresponding class labels (y), where ‘Class’ most likely indicates whether a PDF file is classified as malware or benign as shown in Fig. 1. Afterwards, the data is divided into training and testing sets using a 90-10 split ratio in order to effectively evaluate the performance of the model. The extracted features from PDF files, such as metadata, structural attributes, and content characteristics, often have different scales and distributions. By transforming the data to have a mean of zero and a standard deviation of one, the ‘StandardScaler’ ensures that the features have uniformity and reduces the impact of scale variations during model training [22]. This normalization process is crucial for enabling fair comparisons and effective model learning, ultimately improving the accuracy and reliability of PDF malware detection systems by creating a level playing field for diverse feature sets. After scaling the data, the preprocessed datasets are ready to be used for training and evaluating deep learning models specifically designed for detecting malware in PDF files.

4.1 Data Preprocessing

Text Vectorization is an essential component in the realm of PDF malware detection. Its primary function is to convert the textual content extracted from PDF files into numerical

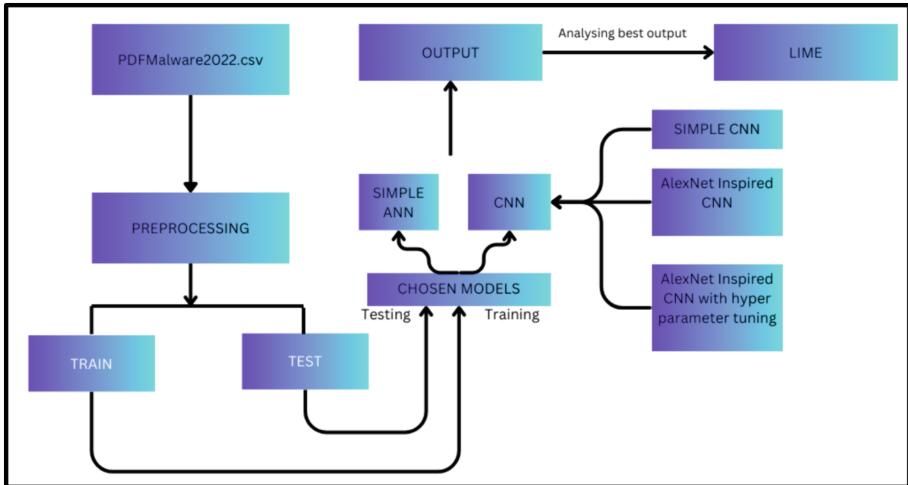


Fig. 1. Proposed PDF malware detection framework

representations. Initially, the header column is separated from the DataFrame, and a TextVectorization layer is created and adjusted to learn the vocabulary from this specific column. Following this, a ‘header_vector’ column is generated, wherein each header text is converted into a numerical representation using the configured TextVectorization layer. These numerical representations are then utilized as input features for training machine learning models, facilitating the efficient identification of malware within PDF documents.

4.2 Model Selection

Simple ANN: After splitting the dataset, a simple neural network model is designed to train the pdf malware dataset to identify malware in PDF files. The neural network architecture comprises three layers: an input layer with 31 neurons employing the ReLU activation function, a hidden layer with 16 neurons also utilizing ReLU activation, and an output layer with a single neuron employing the sigmoid activation function as shown in Fig. 2. The Rectified Linear Unit (ReLU) activation function plays a crucial role in neural network architectures for detecting PDF malware. It introduces non-linearity to the model, enabling the network to capture complex relationships between features extracted from PDF files. ReLU ensures that negative inputs yield a zero output for neurons, while positive inputs result in a linear output equal to the input. This allows the network to learn efficiently and overcome the vanishing gradient problem. By incorporating ReLU activation functions in hidden layers, the neural network becomes adept at capturing intricate patterns and representations within the features of PDF files, thereby improving the accuracy and effectiveness of malware detection algorithms.

The sigmoid activation function is well-suited for binary classification tasks such as pdf malware detection. This function compresses the network's output values within the range of [0, 1], making it ideal for tasks involving binary classification, such as distinguishing between benign and malicious PDF files. By converting the network's raw output into probabilities, where values closer to 1 indicate a higher likelihood of being malware and values closer to 0 indicate benign files, the sigmoid function aids in intuitive interpretation and decision-making. Therefore, when implemented in the output layer of a neural network designed for PDF malware detection, the sigmoid activation function empowers the model to generate probabilistic predictions, thereby contributing to the precise and dependable identification of malicious PDF files. The model is compiled using the Adam optimizer and binary cross-entropy loss function, with accuracy serving as the evaluation metric. This optimizer, known for its efficiency and effectiveness in optimizing neural networks, adjusts the model's parameters to minimize the binary cross-entropy loss function. This particular loss function is well-suited for binary classification tasks, such as distinguishing between benign and malicious PDF files, as it measures the difference between predicted probabilities and actual class labels. Additionally, by employing accuracy as the evaluation metric, the model's performance is assessed based on the percentage of correctly classified instances, providing a clear and understandable measure of its ability to identify malware. Together, these components form a robust approach to PDF malware detection, allowing the neural network model to learn distinctive features and make precise classifications with a high level of confidence.

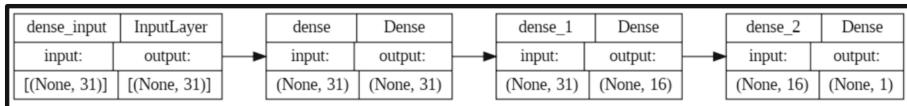


Fig. 2. Plot of ANN for PDF malware detection

The ANN model consists of multiple dense layers that process the input data, which is presented as a 31-dimensional vector. The architecture incorporates a series of dense layers, labeled as dense, dense_1, dense_2, and so on, where each layer transforms the input data and passes it to the subsequent layer. The input and output shapes for each layer are specified, and the final layer generates a single output unit that represents the probability of the input being classified as PDF malware. This architecture empowers the model to effectively recognize malicious content by learning intricate patterns and features within PDF files.

Simple CNN: A basic Convolutional Neural Network (CNN) model is used for detecting malware in PDF files. To enhance the efficiency of training, early stopping is incorporated. The CNN architecture comprises a single convolutional layer, followed by max-pooling, flattening, and fully connected dense layers. The input data is reshaped to fit the one-dimensional convolutional layer, which consists of 64 filters with a kernel size of 3 and ReLU activation. This configuration facilitates the extraction of relevant features as shown in the Fig. 3. Max-pooling is employed to reduce the spatial dimensions of the features while retaining important information. Flattening is then applied to transform the feature maps into a one-dimensional vector for further processing. Two

dense layers follow, with 32 and 16 neurons respectively, utilizing ReLU activation to enable the learning of complex patterns. The output layer, consisting of a single neuron, employs sigmoid activation to generate probabilities indicating the likelihood of malware presence. The model is compiled using the Adam optimizer and binary cross-entropy loss function, with accuracy serving as the performance metric. Furthermore, early stopping is implemented to cease training if the validation loss fails to improve for 10 consecutive epochs. This comprehensive approach ensures the efficient development and training of the model, ultimately leading to accurate detection of malware in PDF files. This conversion process is crucial as it enables the training of machine learning models. By tokenizing and vectorising the text data, TensorFlow empowers the development of highly efficient models capable of effectively identifying malicious content within PDF documents.

The neural network architecture provided is specifically designed to detect PDF malware. It achieves this by processing input data with a dimensionality of 31. The architecture begins with an Input Layer that reshapes the input data into a 1-dimensional array. This reshaped data then undergoes a Convolutional 1D Layer, which applies a 1-dimensional convolution operation to extract features. The resulting output is then max-pooled to reduce dimensionality before being flattened into a 1-dimensional vector. This flattened data is then passed through two Dense Layers successively, gradually reducing the dimensionality to 1. Finally, the Output Layer generates a single output unit that represents the probability of the input being classified as PDF malware. This architecture enables the automatic extraction of relevant features from the input data, facilitating effective detection of malicious PDF files.

Alexnet Inspired CNN: The training of a convolutional neural network (CNN) model, which draws inspiration from the AlexNet architecture, for the purpose of identifying malware present in PDF files as shown in Fig. 4. To enhance the model's performance, the Adam optimizer is employed along with a binary cross-entropy loss function and a learning rate of 0.001. During the training phase, a batch size of 32 is utilized. To prevent overfitting, the training process incorporates early stopping with a patience of 10 epochs. The objective of this approach is to create a robust malware detection system by leveraging the CNN architecture and optimizing the training parameters to accurately classify PDF files as either malicious or benign.

The Fig. 4 begins with an Input Layer that accepts input data with a shape of (None, 31), where “None” represents a variable batch size and 31 denotes the input dimensions. To process this input data, it is reshaped into a 3-dimensional array of shape (None, 31, 1). This reshaped data is then passed through a series of Conv1D layers, which apply 1-dimensional convolution operations to extract features. Following these convolutional layers, MaxPooling1D layers are employed to reduce the dimensionality of the data while preserving important features. Additional Conv1D and MaxPooling1D layers are subsequently utilized to further refine the extracted features. The output from the last pooling layer is flattened into a 1-dimensional vector, which is then fed into Dense layers for additional feature processing and dimensionality reduction. To prevent overfitting during training, Dropout layers are interspersed within the Dense layers, randomly dropping a fraction of the input units. The architecture culminates in a Dense layer with a

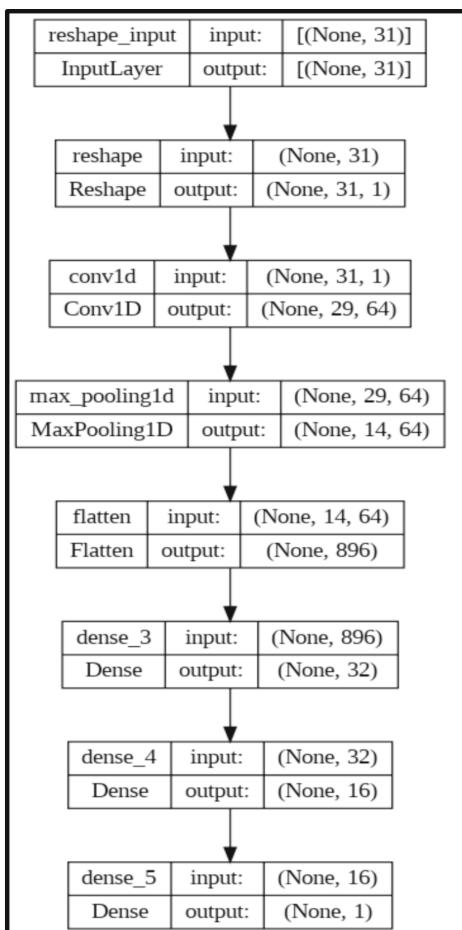


Fig. 3. Plot of Simple CNN for PDF malware detection

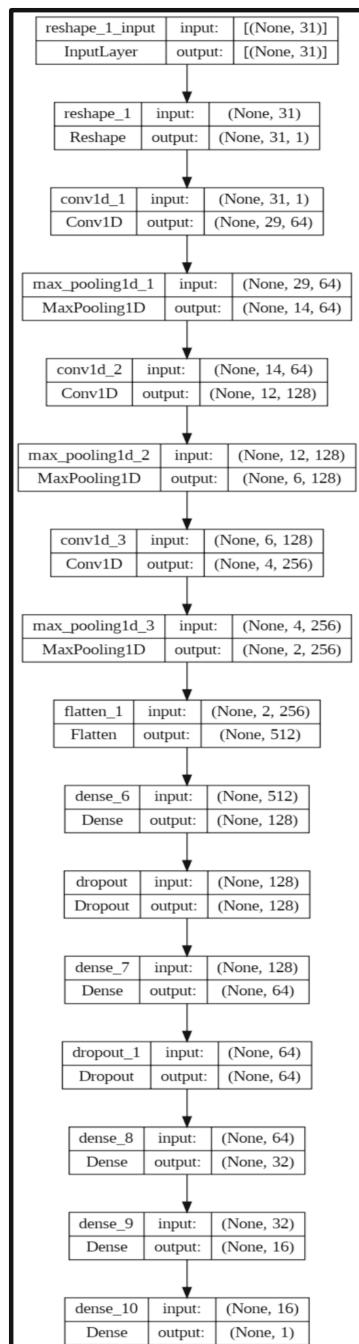


Fig. 4. Plot of Alexnet inspired CNN for PDF malware detection

single output unit, which produces the final classification output indicating the probability of the input being classified as PDF malware. This hierarchical architecture allows the model to effectively capture both local and global patterns within PDF files, enabling accurate detection of malware.

Alexnet Inspired CNN with Hyper parameter tuning: The proposed system utilizes the convolutional neural network (CNN) architecture inspired by AlexNet, incorporating

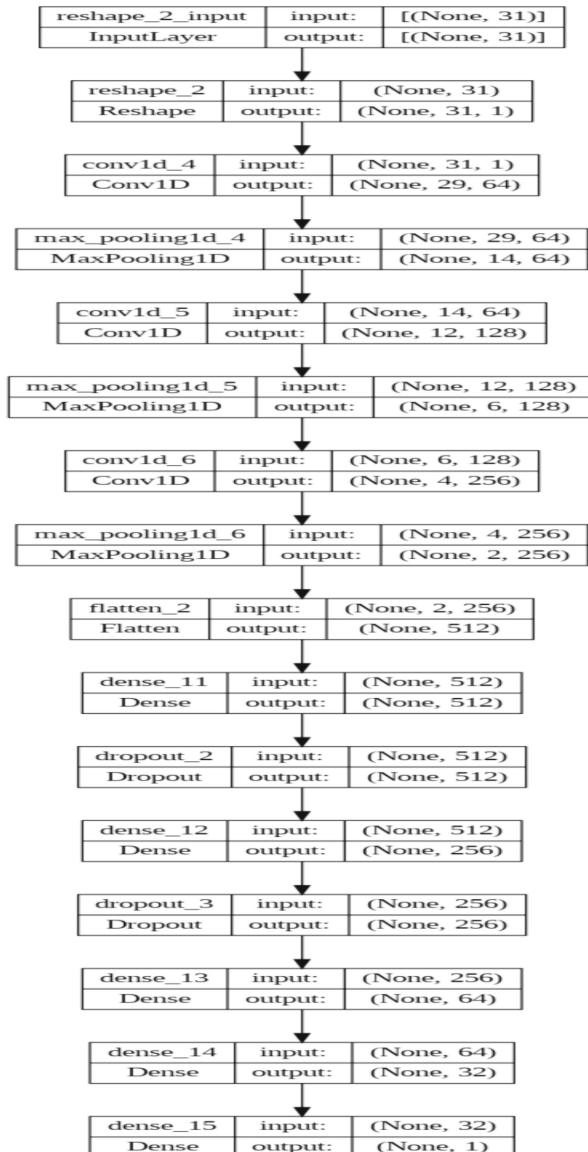


Fig. 5. Plot of Alexnet with hyper tuning

Leaky ReLU activation functions as shown in Fig. 5. To improve its performance, hyper-parameter tuning is carried out. The model is optimized using the Adam optimizer with a binary cross-entropy loss function and an initial learning rate of 0.001. During training, a batch size of 32 samples is used. The number of training epochs is determined through early stopping, with a patience of 15 epochs. Training is halted if the validation loss does not improve. Furthermore, the model employs the Learning Rate Reduction on Plateau callback, which reduces the learning rate by a factor of 0.1 after a patience of 5 epochs without any improvement in validation loss. This comprehensive approach aims to enhance the model's ability to accurately detect malware within PDF files by leveraging both architectural insights and fine-tuned hyper-parameters.

The model starts with an Input Layer that accepts data with a shape of (None, 31). This data is then reshaped into a 3-dimensional array of shape (None, 31, 1). The reshaped data goes through a series of Conv1D layers, each followed by MaxPooling1D layers. This process helps in extracting and reducing features progressively. The extracted features are further refined using additional Conv1D and MaxPooling1D layers. The output from the last pooling layer is flattened into a 1-dimensional vector and then passed through Dense layers for further feature processing and dimensionality reduction. Dropout layers are incorporated to prevent overfitting. Finally, the architecture concludes with a Dense layer that produces a single output unit, representing the probability of the input being classified as PDF malware. This hierarchical architecture allows for the effective capture of both local and global patterns within PDF files, thereby facilitating accurate malware detection.

5 Results and Discussions

The results and efficiency of the proposed model is discussed in order to prove the efficiency of our proposed work. The Table 2 provides an overview of the performance comparison between different models used for detecting PDF malware. These models include a Simple Artificial Neural Network (ANN), a Simple Convolutional Neural Network (CNN), an AlexNet-inspired CNN, and an AlexNet-inspired CNN with hyper-parameter tuning. The evaluation metrics used include training accuracy, testing accuracy, precision, recall, and F1-score. The Simple ANN achieved a training accuracy of 99.12% and a testing accuracy, precision, recall, and F1-score of 98.31%. On the other hand, the Simple CNN had a slightly lower training accuracy of 99.07%, but slightly higher testing accuracy, precision, recall, and F1-score at 98.90%. The AlexNet-inspired CNN demonstrated improved performance with a training accuracy of 99.36% and testing accuracy, precision, recall, and F1-score all reaching 99.30%. Furthermore, the AlexNet-inspired CNN with hyper-parameter tuning further refined its performance, achieving a training accuracy of 98.68% and testing accuracy, precision, recall, and F1-score at 99.30%. Overall, the results indicate that the more complex architectures, particularly the AlexNet-inspired CNN and its hyper-parameter-tuned variant, offer enhanced accuracy and effectiveness in identifying PDF malware.

Table 2. Results of our proposed work

	Simple ANN	Simple CNN	AlexNet Inspired CNN	AlexNet Inspired CNN with Hyper-Parameter tuning
Training Accuracy	0.9912	0.9907	0.9936	0.9868
Testing Accuracy	0.9831	0.989	0.993	0.993
Testing Precision	0.9831	0.989	0.993	0.993
Testing Recall	0.9831	0.989	0.993	0.993
Testing F1-Score	0.983	0.989	0.993	0.993

In the context of PDF malware detection, it is essential to evaluate the effectiveness of classification models in order to ensure robust security measures. Two commonly used evaluation metrics for this purpose are the Receiver Operating Characteristic (ROC) curve and the Precision-Recall Curve (PRC). The ROC curve illustrates the true positive rate (sensitivity) versus the false positive rate (1 - specificity) at different threshold settings. This graphical representation enables us to assess the model's ability to accurately identify malware samples while minimizing false alarms. A higher true positive rate indicates the model's effectiveness in correctly classifying malware, while a lower false positive rate suggests a reduced likelihood of misclassifying benign files as malicious. The area under the ROC curve (AUC-ROC) provides a comprehensive measure of the model's discriminatory power, with a value closer to 1 indicating superior performance.

On the other hand, the Precision-Recall Curve (PRC) demonstrates the trade-off between precision and recall, which are two crucial metrics in imbalanced classification tasks such as malware detection. Precision represents the ratio of correctly identified malware samples to all samples classified as malware, while recall denotes the proportion of actual malware samples that are correctly identified by the model. PRC is particularly informative in scenarios where class imbalance is prominent, as it focuses on the model's ability to accurately detect malware instances while minimizing false positives. Similar to the ROC curve, the area under the Precision-Recall Curve (AUC-PRC) provides a quantitative measure of the model's performance, with higher values indicating superior performance in terms of precision and recall trade-offs. In conclusion, both the ROC and PRC offer valuable insights into the performance of PDF malware detection models. By simultaneously considering both curves, can acquire a holistic comprehension of the strengths and weaknesses of a model, thereby enabling a continuous enhancement in the capabilities of detecting malware. The Figs. 6, 7, 8 and 9 shows the receiver operating characteristic curve and precision recall curve of our four selected models.

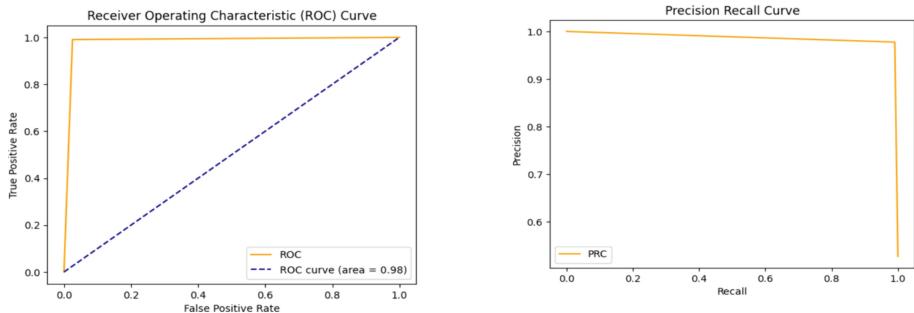


Fig. 6. a. ROC of Simple ANN model. b. PRC of simple ANN model

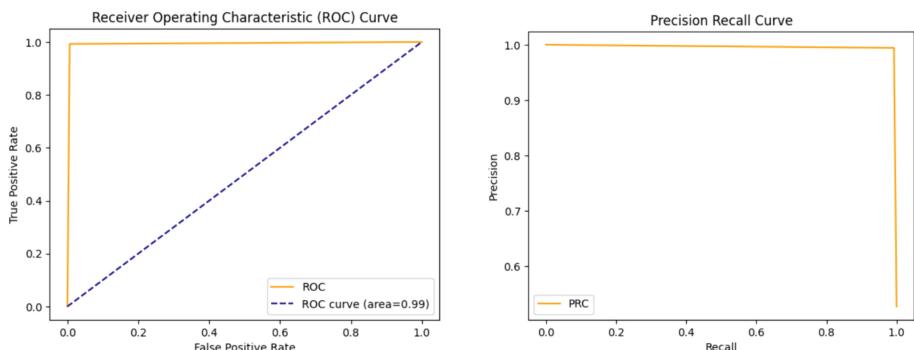


Fig. 7. a. ROC of Simple CNN model. b. PRC of simple CNN model

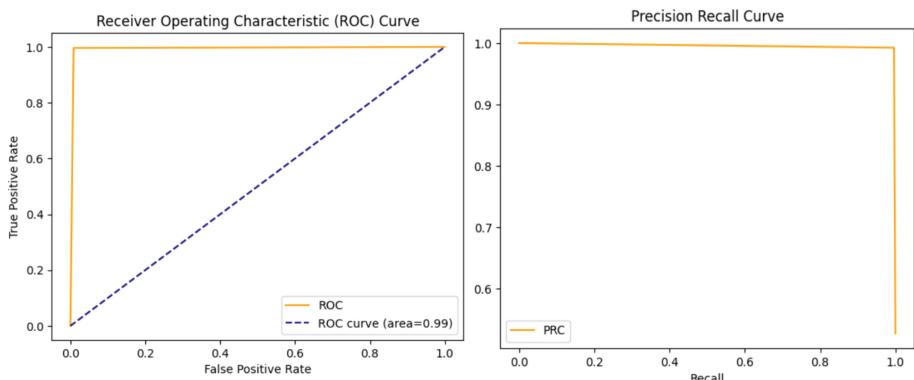


Fig. 8. a. ROC of Alexnet inspired CNN. b. PRC of Alexnet inspired CNN

Explainable Artificial Intelligence (XAI) techniques play a crucial role in PDF malware detection by offering insights into the decision-making process of the models, thereby enhancing transparency, trust, and interpretability. It includes, feature importance analysis, this provides valuable insights into the characteristics the model utilizes

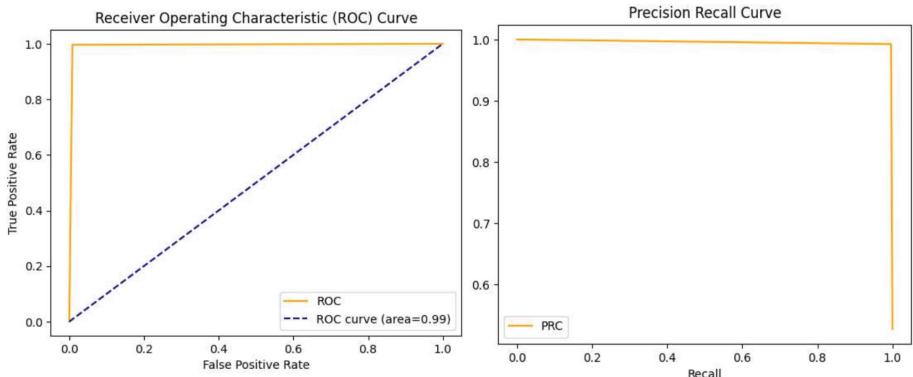


Fig. 9. a. ROC of Alexnet inspired with hyper-parameter tuning. b. PRC of Alexnet inspired with hyper-parameter tuning

to differentiate between benign and malicious PDF files. XAI methods can generate visualizations that highlight significant regions or features within PDF files contributing to the classification decision. Extracting decision rules from the model can offer understandable explanations for why certain PDF files are classified as malware. It can generate explanations for individual predictions, enabling analysts to comprehend why a specific PDF file was classified as either benign or malicious. This is particularly valuable in identifying false positives or false negatives and improving the overall performance of the model. Techniques such as model distillation or simplified model representations can be employed to create more interpretable versions of complex models.

In the realm of PDF malware detection using LIME (Local Interpretable Model-agnostic Explanations) XAI (Explainable Artificial Intelligence), the Fig. 10. Showcases the prediction probabilities and corresponding feature contributions that serve as indicators for classifying whether a PDF file is malicious or non-malicious. Each feature's impact on the prediction is measured by its coefficient, where positive values suggest a leaning towards maliciousness and negative values indicate a tendency towards non-maliciousness. Features such as ‘embedded files’, ‘images’, ‘encrypt’, and ‘stream’ demonstrate significant contributions in identifying malicious PDF files, while features like ‘title characters’, ‘launch’, and ‘obj’ seem to influence non-malicious classifications. The coefficients provided offer valuable insights into the underlying attributes of PDF files, which contribute to the prediction outcomes and aid in comprehending the decision-making process of the model in pdf malware detection.

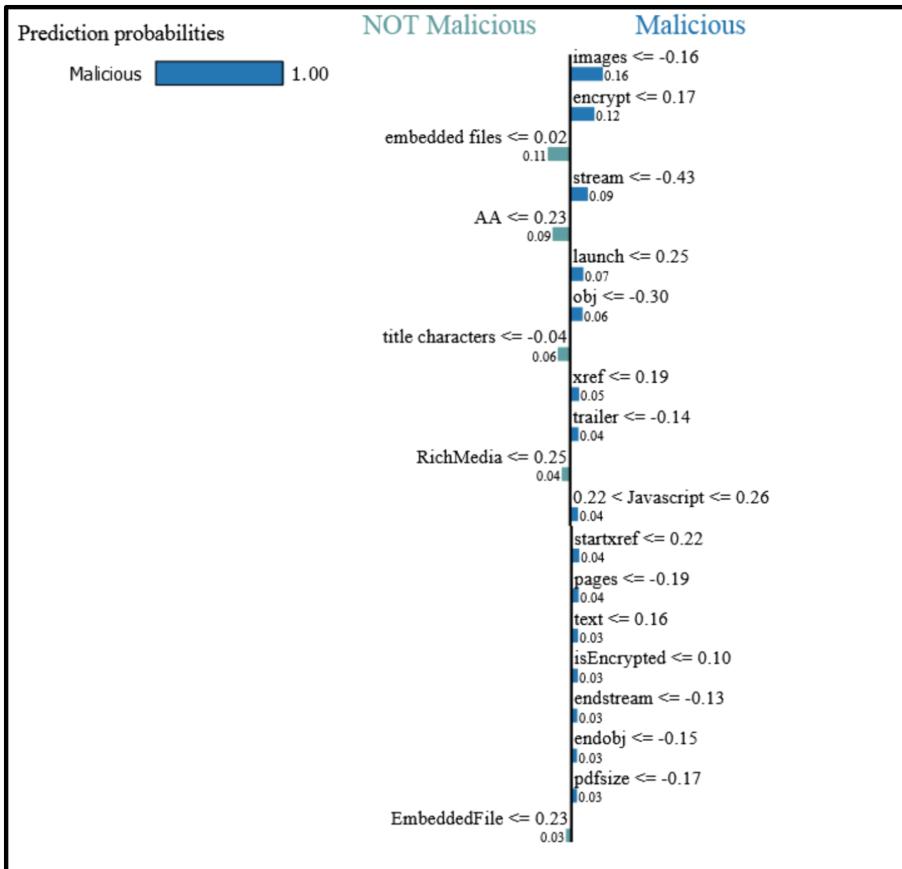


Fig. 10. Prediction Probabilities of LIME XAI for PDF malware detection

The Fig. 11 shows the values provided in the domain of PDF malware detection offer valuable insights into the significance of different attributes when classifying PDF files as either malicious or benign. Attributes like ‘encrypt’, ‘AA’, ‘launch’, ‘RichMedia’, ‘Javascript’, ‘startxref’, and ‘EmbeddedFile’ demonstrate positive values, indicating their association with potential malicious content. On the other hand, attributes such as ‘images’, ‘stream’, ‘obj’, ‘title characters’, ‘trailer’, ‘pages’, ‘endstream’, ‘endobj’, and ‘pdfsize’ show negative values, suggesting characteristics commonly found in non-malicious PDF files. These feature values play a crucial role in understanding the underlying patterns used by malware detection algorithms to distinguish between safe and potentially harmful PDF documents.

Comparison with existing state-of-the-art: Table 3. Shows the comparison of different models utilized for detecting malware in PDF files, along with the corresponding performance metrics. These models consist of DNN-7L (Deep Neural Network with 7 layers), JSUNPACK, O-DT (Optimized Decision Tree), Random Forest tree, ANN (Artificial Neural Network), SVM (Support Vector Machine), DL4MD (Deep Learning

Feature	Value
images	-0.16
encrypt	0.17
embedded files	0.02
stream	-0.43
AA	0.23
launch	0.25
obj	-0.32
title characters	-0.04
xref	0.19
trailer	-0.85
RichMedia	0.25
Javascript	0.26
startxref	0.18
pages	-0.19
text	0.16
isEncrypted	0.10
endstream	-0.14
endobj	-0.16
pdfsize	-0.19
EmbeddedFile	0.23

Fig. 11. Feature importance of PDF malware detection with LIME-XAI

for Malware Detection), and a Proposed model. The performance metrics assessed for each model include Accuracy, Precision, Recall, and F1-score. The Proposed model surpasses other approaches with an Accuracy of 99.36%, Precision of 99.30%, Recall of 99.30%, and F1-score of 99.30%. Other models, such as JSUNPACK, O-DT, Random Forest tree, ANN, SVM, and DL4MD, exhibit varying levels of performance across these metrics. This comparison offers valuable insights into the effectiveness of different models in detecting malware in PDF files, with the Proposed model demonstrating superior performance. References for the evaluation of each model are also provided for further investigation.

Table 3. Comparison with state-of-the-art PDF malware detection techniques

Model	Accuracy	Precision	Recall	F1-score	Ref.
DNN-7L	96.15	99.05	93.20	-	[23]
DNN-7L	96.20	98.89	93.48	-	
DNN-7L	98.99	98.61	99.81	-	
DNN-7L	93.60	87.97	99.31	-	
JSUNPACK	95.11%	97.57%	-	94.10%	[24]
O-DT	98.84%	98.80%	-	98.80%	[25]
Rando Forest tree	97.6852%	0.977	0.977	0.976	[26]
ANN	0.9202	0.9328	0.9056	0.9189	[27]
SVM	0.9306	0.9381	0.9223	0.9301	
DL4MD	0.9564	0.9545	0.9584	0.9564	
Proposed model	0.9936	0.993	0.993	0.993	-

6 Conclusion

This paper introduces and effectively showcases a unique and comprehensive method for detecting malware in PDF files. The framework for detecting PDF malware consists of several steps. First, the dataset is preprocessed, then it is divided into training and testing sets. Next, the features are normalized to ensure consistency. The model selection process involves creating simple ANN and CNN architectures, as well as a variant inspired by AlexNet. Hyperparameter tuning is then performed to optimize the models. To assess the performance of the models, evaluation metrics such as accuracy, precision, recall, and F1-score are used. Additionally, ROC and PRC curves provide valuable insights. The use of explainable AI techniques enhances the interpretability of the results. Finally, the proposed approach is validated by comparing it with existing methods to demonstrate its effectiveness.

References

1. Jeong, Y.S., Woo, J., Kang, A.R.: Malware detection on byte streams of PDF files using convolutional neural networks. *Secur. Commun. Netw.* **2019**, 8485365 (2019)
2. Cuan, B., Damien, A., Delaplace, C., Valois, M.: Malware detection in PDF files using machine learning. In: Proceedings of the ICETE 2018—The 15th International Joint Conference on e-Business and Telecommunications; vol. 2, pp. 412–419. Warangal, India, 18–21 Dec 2018
3. Falah, A., Pokhrel, S.R., Pan, L., de Souza-Daw, A.: Towards enhanced PDF maldocs detection with feature engineering: design challenges. *Multimed. Tools Appl.* **81**, 41103–41130 (2022)
4. Issakhani, M., Victor, P., Tekeoglu, A., Lashkari, A.H.: PDF Malware Detection based on Stacking Learning. In: ICISSP, pp. 562–570 (2022)
5. Li, Y., Wang, Y., Wang, Y., Ke, L., Tan, Y.A.: A feature-vector generative adversarial network for evading PDF malware classifiers. *Inf. Sci.* **523**, 38–48 (2020)

6. Mohammed, T. M., Nataraj, L., Chikkagoudar, S., Chandrasekaran, S., Manjunath, B.S.: HAPSSA: holistic approach to PDF malware detection using signal and statistical analysis. In: MILCOM 2021–2021 IEEE Military Communications Conference (MILCOM), pp. 709–714. IEEE (2021)
7. Bae, H., Lee, Y., Kim, Y., Hwang, U., Yoon, S., Paek, Y.: Learn2Evade: learning-based generative model for evading PDF malware classifiers. *IEEE Trans. Artif. Intell.* **2**(4), 299–313 (2021)
8. Yerima, S.Y., Bashar, A., Latif, G.: Malicious PDF detection Based on Machine Learning with Enhanced Feature Set. In: 2022 14th International Conference on Computational Intelligence and Communication Networks (CICN), pp. 486–491. IEEE (2022)
9. Al-Taani, R., Bassah, R., Naimat, N., Odeh, A.: PDF Malware Detection optimisation using machine learning. In: 2023 3rd International Conference on Computing and Information Technology (ICCIT), pp. 15–19. IEEE (2023)
10. PM, P.P., Hemavathi, P.: PDF malware detection system based on machine learning algorithm. In: 2022 International Conference on Automation, Computing and Renewable Systems (ICACRS), pp. 538–542. IEEE (2022)
11. Yu, M., et al.: A unified malicious documents detection model based on two layers of abstraction. In: 2019 IEEE 21st International Conference on High Performance Computing and Communications; IEEE 17th International Conference on Smart City; IEEE 5th International Conference on Data Science and Systems (HPCC/SmartCity/DSS), pp. 2317–2323. IEEE (2019)
12. Maiorca, D., Biggio, B.: Digital investigation of pdf files: unveiling traces of embedded malware. *IEEE Secur. Priv.* **17**(1), 63–71 (2019)
13. Maiorca, D., Biggio, B., Giacinto, G.: Towards adversarial malware detection: lessons learned from PDF-based attacks. *ACM Comput. Surv.* **52**(4), 1–36 (2019)
14. Zhou, X., Pang, J., Liu, F., Wang, J., Yue, F., Liu, X.: Pdf exploitable malware analysis based on exploit genes. In: 2018 12th IEEE International Conference on Anti-counterfeiting, Security, and Identification (ASID), pp. 16–20. IEEE (2018)
15. Abu Al-Haija, Q., Odeh, A., Qattous, H.: PDF malware detection based on optimizable decision trees. *Electronics* **11**(19), 3142 (2022)
16. Corum, A., Jenkins, D., Zheng, J.: Robust PDF malware detection with image visualization and processing techniques. In: 2019 2nd International Conference on Data Intelligence and Security (ICDIS), pp. 108–114. IEEE (2019)
17. Wang, Y.: The de-obfuscation method in the static detection of malicious PDF documents. In: 2021 7th Annual International Conference on Network and Information Systems for Computers (ICNISC), pp. 44–47. IEEE (2021)
18. Singh, P., Tapaswi, S., Gupta, S.: Malware detection in pdf and office documents: a survey. *Inform. Secur. J.: A Global Perspect.* **29**(3), 134–153 (2020)
19. Khan, B., Arshad, M., Khan, S.S.: Comparative analysis of machine learning models for PDF malware detection: evaluating different training and testing criteria. *J. Cybersecurity* **5**, 1–11 (2023)
20. Mejjaoui, S., Guizani, S.: PDF malware detection based on fuzzy unordered rule induction algorithm (FURIA). *Appl. Sci.* **13**(6), 3980 (2023)
21. Trad, F., Hussein, A., Chehab, A.: Leveraging adversarial samples for enhanced classification of malicious and evasive PDF files. *Appl. Sci.* **13**(6), 3472 (2023)
22. Ravi, V., Alazab, M.: Attention-based convolutional neural network deep learning approach for robust malware classification. *Comput. Intell.* **39**(1), 145–168 (2023)
23. Sewak, M., Sahay, S.K., Rathore, H.: Comparison of deep learning and the classical machine learning algorithm for the malware detection. In: 2018 19th IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD), pp. 293–296. IEEE (2018)

24. Jiang, J., Song, N., Yu, M., Chow, K.P., Li, G., Liu, C., Huang, W.: Detecting malicious pdf documents using semi-supervised machine learning. In: Peterson, G., Shenoi, S. (eds.), Proceedings of the Advances in Digital Forensics XVII. Digital Forensics 2021, Virtual Event, 1–2 Feb 2021; IFIP Advances in Information and Communication Technology, vol. 612. ; Springer, Cham, Switzerland (2021). https://doi.org/10.1007/978-3-030-88381-2_7
25. Cohen, A., et al.: Sec-Lib: protecting scholarly digital libraries from infected papers using active machine learning framework. IEEE Access 7, 110050–110073 (2019)
26. Shaukat, K., Luo, S., Varadharajan, V.: A novel machine learning approach for detecting first-time-appeared malware. Eng. Appl. Artif. Intell. 131, 107801 (2024)
27. Alsharafi, L., Asiri, M., Azzony, S., Alqahtani, A.: Malware detection based on deep learning. In: 2023 3rd International Conference on Computing and Information Technology (ICCIT), pp. 427–432. IEEE (2023)



Optimized Deep Learning Technique for the Effective Detection of Windows PE Malware

Kirubavathi Ganapathyappan^(✉) and Abhishek Yadav

Department of Mathematics, Amrita School of Physical Sciences, Amrita Vishwa
Vidyapeetham, Coimbatore, India
g_kirubavathi@cb.amrita.edu

Abstract. This study examines the effects of several hyper-parameters, such as the number of epochs, batch size, number of layers and neurons, optimization technique, dropout rate, type of activation function, and learning rate, on deep learning-based static malware detection. To determine the ideal hyper-parameters for our deep learning model, we used the inception approach. Our research shows that convolutional neural networks with inception are more effective than other techniques, and that the correctness of the model is greatly influenced by the choice of hyper-parameter values. In particular, our method causes the neural network model's accuracy for static malware detection on the Ember dataset to significantly increase (from 84.19% to 90.78%). These findings support the efficacy of our suggested methodology and have significant ramifications for the static malware detection community.

Keywords: Neural Network · Deep Learning · Hyper-parameter tuning · Static malware detection · Deep learning with inception

1 Introduction

Malware is software created with the express purpose of causing chaos, harm, or gaining unauthorized access to any computer system. Malware detection is a cybersecurity issue that is becoming more and more significant for the global community and all industries. Malware attacks can result in losses of up to \$1 million. The market for malware analysis is anticipated to increase at a rate of 29.56% between 2020 and 2027, reaching USD 23.81 billion [1]. Software systems are constantly at risk from malware, therefore finding it is essential to contemporary cybersecurity. In the field of malware detection, machine learning (ML) has demonstrated promise, especially in the area of static malware detection, which scans malicious binary files without actually running them. However, the selection of hyper parameters, which can be difficult to ascertain, has a significant impact on how well machine learning algorithms perform.

The purpose of this study is to determine whether deep learning-based models for static malware detection can be made to perform better by applying hyper-parameter tuning approaches. In particular, by optimizing hyper-parameters, we want to improve

the effectiveness of deep learning models for static malware detection. Due to their notable effects on the effectiveness of deep learning models in earlier research, a number of hyper-parameters, including the number of epochs, batch size, number of layers, number of neurons, optimization method, dropout relay, type of activation function, and learning rate (LR), were chosen [2, 3]. Our objective is to optimize and tune these hyper-parameters in order to improve the deep learning models' generalization performance and accuracy for the detection of static malware.

2 Related Work

Globally, a large number of distinct malware files are created, and according to Virus-Total [4], more than a million dubious executable files are gathered every day. Recent research has demonstrated that comparable files can be found in both benign and malicious files. Machine learning-based malware detection has been researched for decades since security institutes are unable to manually evaluate such a large number of distinct malware files. A malware detection method based on Objective Oriented Association mining was proposed by [5]. This classifier is built on rules, and the rules were created using a Fast-FP Growth algorithm variation. The accuracy and effectiveness of the Naive Bayes, Support Vector Machines (SVM), and Decision Tree-based algorithms utilized by Norton, McAfee, and VirusScan were exceeded by their work. The test accuracy displayed by the system was 64.14%.

A model that can efficiently classify static PE files and extract a feature set from the dataset was proposed by [6]. Instead of creating a model, the research aimed to emphasize the importance of feature extraction. The well-extracted properties help neural networks with few layers produce better results. Reducing the number of layers in the model will enhance its performance and save processing and testing time. Deep neural networks (DNN) are used to find the features that best represent the provided training data, as suggested by [7]. Here, the features are learned using a deep neural network from portable executable files. As a result, deep neural network-based solutions have low false positive rates and are effective at identifying both known and unknown malware.

A deep learning-based malware detection framework was proposed by [8]. To capture the behavior of the software, they built a joint representation of multiple APIs by first applying convolutional and embedding layers. Second, by utilizing the API's category, action, and operation objects to convey the semantic contents of each API call. Lastly, they used the Bi-LSTM module to mine the relationship data between APIs. In order to identify an appropriate defense against a malware assault in an API routine,

The authors in [9] are tackling malware detection through the use of decision trees (DT) and the C4.5 (J48) classification algorithm. A large-scale malware categorization system utilizing neural networks and random projections was presented by George et al. in 2016 [10]. Instead of using neural networks, this study is unusual because of a special method for projecting each original input onto a vector with comparatively less dimensions, then using that projection as the model's input. But in the proposed research, utilized Artificial Neural Network, Convolutional Neural network and finally used Convolutional Neural network with inception framework with hyper parameter tuning in order to increase the accuracy.

Our Research Contributions:

1. The suggested study offers a cutting-edge method for enhancing the identification of static malware by utilizing deep learning models with hyper-parameter tuning. The study's conclusions demonstrate that using this strategy can greatly increase the model's accuracy for the EMBER dataset.
2. This research offers valuable perspectives on how various hyper-parameters interact to impact deep learning models' ability to detect static malware.
3. It offers a methodology for hyper-parameter tuning to enhance the effectiveness of deep learning models for malware detection for researchers and practitioners.
4. In the larger field of cybersecurity, where precise malware identification is essential to safeguarding computer systems and networks, the study's findings have consequences.

3 Proposed Research Framework

The Proposed work starts with collecting and analyzing the EMBER dataset from Cornell University [11]. Raw features are extracted to JSON format and vectorization theses raw features and saved in binary format from which they can be converted to data frame. After vectorization the features are given to the min max scaling phase to normalize the feature values. Min-max scaling is a widely employed method for standardizing attributes in a dataset. Standardization involves adjusting the attributes of a dataset to a predetermined range, usually ranging from 0 to 1. Min-max scaling accomplishes this by subtracting the minimum value of an attribute from each data point and subsequently dividing it by the range of the attribute (which is determined by the difference between the maximum and minimum values). After the normalization the features are divided into training and testing phase. For the training phase containing the features of training features and validation features and testing phase contains testing features alone. After dividing the dataset into training and testing given to Artificial neural network model, Convolutional neural network model and inception with convolutional neural network model for analyzing the performance of each model with respect to EMBER dataset. The complete architecture of the proposed work is given in Fig. 1.

3.1 Description of the Dataset

The study of malware can be dangerous and difficult to move or keep safely, which might provide significant difficulties for malware research, particularly when working with bigger datasets. We utilize the EMBER dataset, an open source benchmark dataset made up of static attributes taken from Windows portable executables, to reduce the likelihood of such occurrences [11]. Function imports have already shown to be effective features for machine learning in [12, 13], despite their potential for obfuscation [14]. In order to make the library and the function together generate a unique identifier, we remove imported functions and the libraries that go with them from EMBER, lowercase them, and connect them.

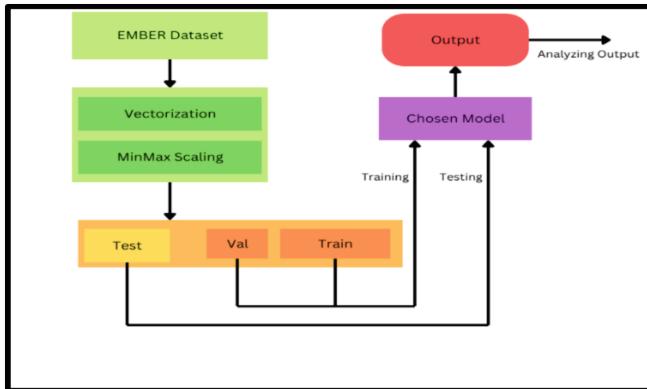


Fig. 1. Proposed System Architecture

One JSON object can be found in each of the JSON line files that make up the EMBER dataset. The following data kinds are present in every object: The unique identification for the original file was its sha256 hash. The labels are represented as 0 means benign, 1 means malevolent, and -1 means unlabeled. There are eight classes of raw features with format-independent histograms and parsed values. The EMBER dataset comprises eight classes of raw features, such as format-independent histograms, parsed features, and string counts. Model characteristics obtained from the EMBER dataset are differentiated from raw features. It is modeled with fixed-size feature matrix, where the hash trick is used to resolve the numerical representation of the raw features, which are represented by string features, imported names, exported names, and so on. Although code to transform raw features into model features is given to train a baseline model, the published dataset lacks a specified feature matrix. Features that have been parsed: The dataset comprises five feature groups after the PE file has been parsed. We utilize the Library to Instrument Executable Formats as a handy PE parser.

General file information: The PE header's basic information, the file size, the number of imported and exported functions, the virtual file size, the presence of a debug segment, thread local storage, resources, relocations, signatures, and symbols are all considered to be part of the general file information community.

Details about the header: The timestamp, the target machine (string), and a list of the image properties from the COFF header (list of strings) are reported in the header. Code, headers, and commit sizes from the optional header; target subsystem (string); DLL characteristics (a list of strings); file magic (e.g., “PE32”); major and minor image versions; linker versions; device versions; and subsystem versions. Before the model is trained, the feature hashing method is used, and 10 bins are allocated to each noisy indicator vector before string descriptors like target computers and DLL properties are stored.

Imported Functions: We analyze the import address table, which displays the imported functions per library, in order to comprehend the assumptions underlying the baseline model. We just need to use the hashing approach (256 bins) to gather the set of

certain libraries in order to construct the model features for the baseline model. In a manner similar to this, we also employ the hashing method (1024 bins) to identify individual strings, like library: two function names, such as kernel32.dll:CreateFileMappingA.

Functions Exported: The raw functionality additionally includes a list of exported functions. The model functionality is derived from these strings using the hashing approach with 128 bins.

Information about each section: Complete details are given, including the section name, size, entropy, virtual size, and string strings that denote its characteristics. There is also mention of the entry place. The hashing method is used to create our vectors by converting the pairs (section name, value) that reflect the section size, section entropy, and virtual size to model features (50 bins each). Furthermore, we capture the features of the entrance point (a list of strings) using the hashing method. The complete feature information of EMBER dataset is given in Fig. 2.

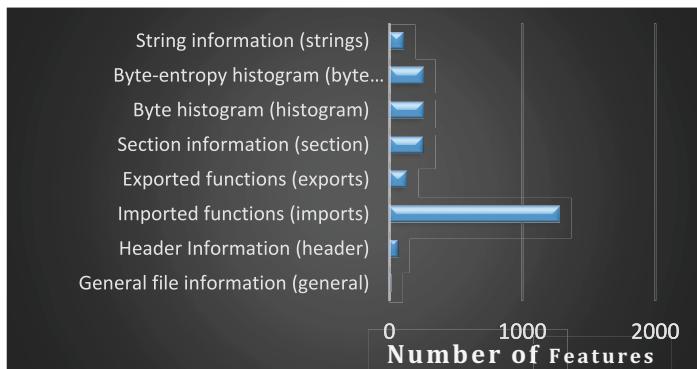


Fig. 2. Feature group information of EMBER dataset

3.2 Training the Neural Network with EMBER Dataset

Several aspects, including as the quality of the dataset, the discriminative power of extracted features, and the resilience of the training process, affect how effective it is to use the Inception architecture for Windows PE malware detection. Dataset is carefully selected dataset that includes both benign and malicious samples is crucial for building a trustworthy model. The efficacy of the Inception architecture resides in its capacity to effectively capture features at various scales. However, appropriate training—which includes efficient data preprocessing and hyper-parameter tuning—is necessary for the model to function as intended. In order to ensure that the model remains flexible in response to changing malware threats, it is imperative that it undergo regular monitoring and changes. Evaluation metrics offer a thorough analysis of the model’s capacity to generalize to new and untested data. In the end, the Inception architecture can be a powerful tool for detecting Windows PE malware if it is properly developed and updated. The Fig. 3 shows the complete information about the dataset which is used for model

training and testing. Table 1 shows the hyper parameter tuning of ANN, CNN and CNN with inception.

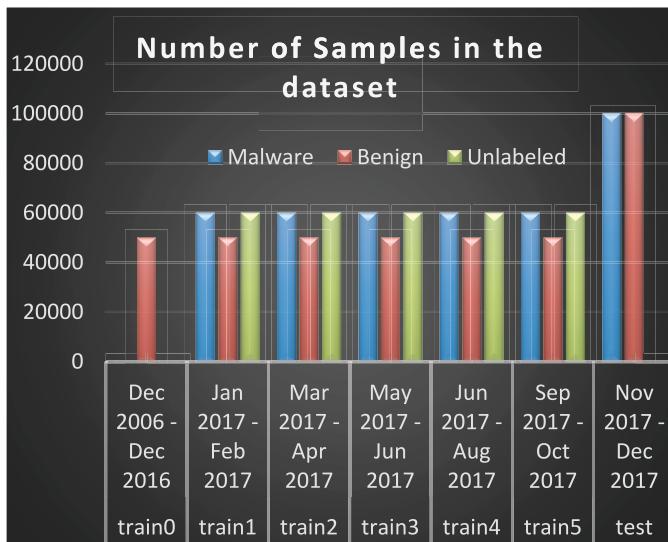


Fig. 3. Number of Samples in the dataset

Three distinct models' architectures are described in the information provided as an Inception Module-equipped CNN, a 1D Convolutional Neural Network (CNN) [15], and an Artificial Neural Network (ANN). Every model's layers, settings, and parameters are explained in detail. Three dense layers with ReLU activation, dropout layers for regularization, and a final thick/output layer with sigmoid activation for binary classification make up the artificial neural network (ANN). Convolutional layers with ReLU activation, max-pooling layers for down sampling, and dropout for regularization are all components of the CNN model [16]. An inception module, which combines several convolutional filters of various sizes, is incorporated into the CNN Inception Module to improve the model's feature capture performance. For binary classification problems, all models employ the binary cross-entropy loss function and the Adam optimizer with a learning rate of 0.0001. Different models are trained for other numbers of epochs (30 for ANN, 10 for CNN, and CNN with Inception Module). The CNN models contain clip norm regularization. Usually, inference involves using the learned models to forecast fresh data. The optimal model to choose will rely on the particular objectives of your assignment, taking into account variables like recall, accuracy, precision, and F1 score on test or validation datasets.

The Fig. 4 shows the ANN training with EMBER dataset. An input layer for sequences of length 2381 is part of the neural network design that is discussed [17]. Three hidden dense layers with ReLU activation are then included. There is a matching dropout layer for regularization after every dense layer. There are 32, 32, and 8 units in the buried layers, correspondingly. For binary classification tasks, the last layer is a dense layer with one unit and a sigmoid activation function. By describing the input and output shapes at each layer, the architecture is condensed. The Fig. 5 shows the Convolutional 1 D training with EMBER dataset. With one channel and input sequences of length 2381, the Convolutional Neural Network (CNN) that is being shown is specifically made for sequence data [18]. Four Conv1D layers make up the architecture, and MaxPooling1D layers come after for down sampling. Through max-pooling, the convolutional layers gradually shorten sequence length by using 8 and 16 filters. There is a flattening layer after the convolutional layers, and three thick layers with ReLU activation come next. The last layer is a dense layer having one unit and a sigmoid activation function that is appropriate for binary classification. Two dropout layers are included for regularization. It is possible to describe the architecture features by pointing out the input and output shapes at each tier.

Table 1. Details about hyper parameter tuning of three models

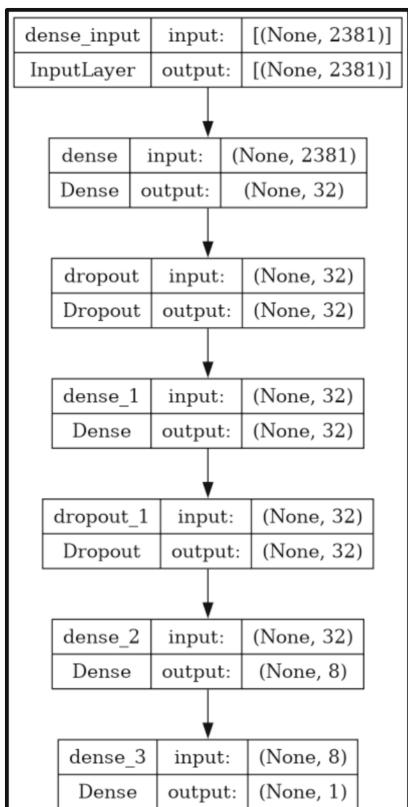
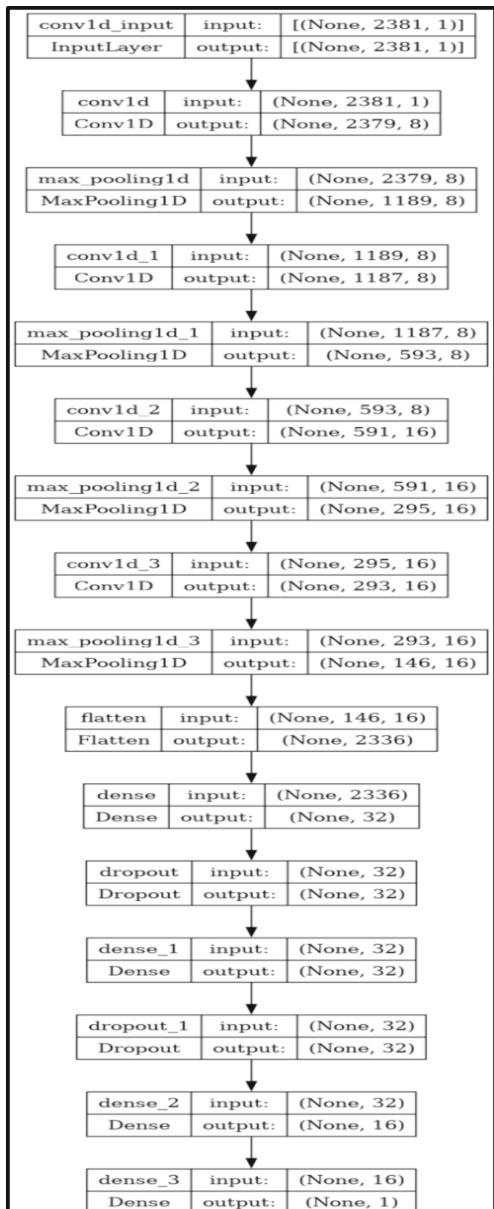
Models					
ANN		CNN		CNN with Inception Module	
Layer	Parameters	Layer	Parameters	Layer	Parameters
Dense	32 activation = ReLU input_shape = (2381,)	Conv1d	filters = 8, kernel_size = 3, activation = ReLU, input_shape = (2381, 1)	Conv1d	filters = 8, kernel_size = 3, activation = ReLU, input_shape = (2381, 1)
Dropout	p = 0.2	MaxPooling1d	pool_size = 2	BatchNorm	-
Dense	32 activation = ReLU	Conv1d	filters = 8, kernel_size = 3, activation = ReLU	Inception Module	
Dropout	p = 0.2	MaxPooling1d	pool_size = 2	BatchNorm	-
Dense	8 activation = ReLU	Conv1d	filters = 16, kernel_size = 3, activation = ReLU	MaxPooling1d	pool_size = 3
Dense Output	1 activation = Sigmoid	MaxPooling1d	pool_size = 2	Flatten	-

(continued)

Table 1. (*continued*)

Models					
ANN		CNN		CNN with Inception Module	
Layer	Parameters	Layer	Parameters	Layer	Parameters
		Conv1d	filters = 16, kernel_size = 3, activation = ReLU	Dense	64 activation = ReLU
		MaxPooling1d	pool_size = 2	Dropout	p = 0.2
		Flatten	-	Dense	16 activation = ReLU
		Dense	32 activation = ReLU	Dropout	p = 0.2
		Dropout	p = 0.2	Dense	8 activation = ReLU
		Dense	32 activation = ReLU	Dense/Output	1 activation = Sigmoid
		Dropout	p = 0.2		
		Dense	16 activation = ReLU		
Optimizer	Adam	Optimizer	Adam	Optimizer	Adam
Learning Rate	0.0001	Learning Rate	0.0001	Learning Rate	0.0001
Clip Norm	None	Clip Norm	0.0001	Clip Norm	0.00001
Loss	Binary CrossEntropy	Loss	Binary CrossEntropy	Loss	Binary CrossEntropy
Epochs	30	Epochs	10	Epochs	10

Figure 6 shows the convolutional with inception model for training the EMBER dataset. A Convolutional Neural Network (CNN) with complex layer configurations is described by the architecture. Sequences of length 2381 are processed by the input layer using a single channel. Multiple Conv1D layers with Batch-Normalization, max-pooling, and concatenation operations are then added, resulting in a network with parallel feature extraction pathways. Before flattening, the concatenated output is subjected to additional Batch-Normalization and max-pooling [19]. Additional feature transformation is provided by the next set of dense layers with dropout. The last layer is a dense layer that can be classified as binary because it has just one unit. The specification of input and output forms at each layer, which represent a multi-path and intricate CNN design, serves as a summary of the architecture.

**Fig. 4.** Plot of ANN with EMBER dataset**Fig. 5.** Plot of convolutional1 D (Conv1d) with EMBER dataset

3.3 Results and Discussions

The Convolutional Neural Network with Inception is the model that performs the best when it comes to accuracy. This model consistently achieves superior accuracy than the

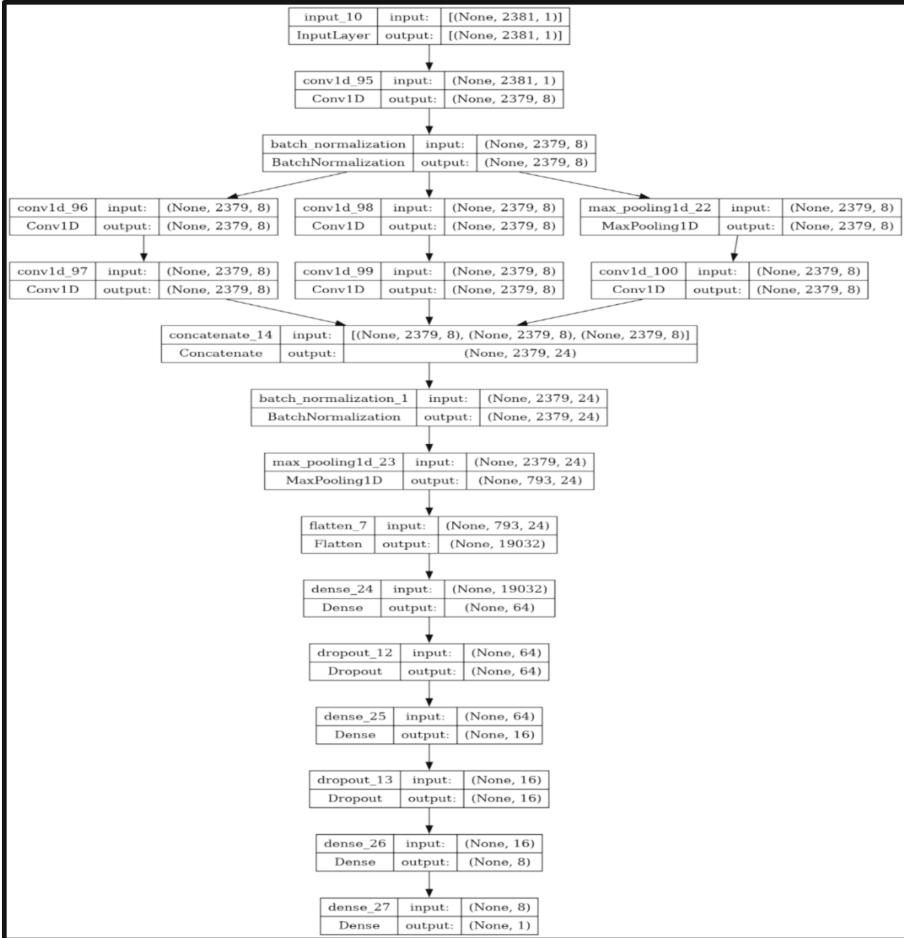


Fig. 6. Plot of CNN with Inception model for EMBER dataset

1D Convolutional Neural Network (Conv1d) and the Artificial Neural Network (ANN) across training, validation, and testing phases. The Convolutional Neural Network with Inception appears to do better overall in identifying cases, as evidenced by the higher accuracy ratings, which suggest that the network is able to capture more complex patterns in the data [20, 21]. The Convolutional Neural Network with Inception exhibits a noteworthy advantage, highlighting its potential dominance in situations where accuracy is an important parameter, even though the ANN and Conv1d models show respectable accuracy. Furthermore, the effectiveness of the Convolutional Neural Network using Inception is supported by the testing precision, recall, and F1 score. Together with an excellent F1 score, its higher recall and precision indicate a well-balanced model that performs well not only in overall accuracy but also in accurately detecting positive cases and reducing false positives. This thorough analysis highlights the Convolutional

Neural Network with Inception as a reliable option for jobs requiring accuracy and well-rounded performance [22]. Table 2 shows the experimental results of the three models with EMBER dataset.

Table 2. Experimental results of the three models with EMBER dataset

ANN		Conv1d		Conv with inception	
Training Accuracy	0.8947	Training Accuracy	0.9209	Training Accuracy	0.9289
Validation Accuracy	0.8939	Validation Accuracy	0.919	Validation Accuracy	0.9263
Testing Accuracy	0.8419	Testing Accuracy	0.9071	Testing Accuracy	0.9078
Testing Precision	0.7867	Testing Precision	0.9041	Testing Precision	0.8721
Testing Recall	0.9379	Testing Recall	0.9107	Testing Recall	0.9556
Testing F1 Score	0.8557	Testing F1 Score	0.9074	Testing F1 Score	0.9119

4 Conclusion

The Convolutional Neural Network with Inception is the model that performs the best when it comes to accuracy. This model consistently achieves superior accuracy than the 1D Convolutional Neural Network (Conv1d) and the Artificial Neural Network (ANN) across training, validation, and testing phases. The Convolutional Neural Network with Inception appears to do better overall in identifying cases, as evidenced by the higher accuracy ratings, which suggest that the network is able to capture more complex patterns in the data. The Convolutional Neural Network with Inception exhibits a noteworthy advantage, highlighting its potential dominance in situations where accuracy is an important parameter, even though the ANN and Conv1d models show respectable accuracy.

Furthermore, the effectiveness of the Convolutional Neural Network using Inception is supported by the testing precision, recall, and F1 score. Together with an excellent F1 score, its higher recall and precision indicate a well-balanced model that performs well not only in overall accuracy but also in accurately detecting positive cases and reducing false positives. This thorough analysis highlights the Convolutional Neural Network with Inception as a reliable option for jobs requiring accuracy and well-rounded performance.

References

1. Humayun, M., Niazi, M., Jhanjhi, N.Z., Alshayeb, M., Mahmood, S.: Cyber security threats and vulnerabilities: a systematic mapping study. *Arab. J. Sci. Eng.* **45**, 3171–3189 (2020)
2. Sun, S., Cao, Z., Zhu, H., Zhao, J.: A survey of optimization methods from a machine learning perspective. *IEEE Trans. Cybern.* **50**(8), 3668–3681 (2019)
3. Pinhero, A., et al.: Malware detection employed by visualization and deep neural network. *Comput. Secur.* **105**, 102247 (2021)

4. Salem, A., Banescu, S., Pretschner, A.: Maat: Automatically analyzing virustotal for accurate labeling and effective malware detection. *ACM Trans. Priv. Secur.* **24**(4), 1–35 (2021)
5. Filho, A.S., Rodríguez, R.J., Feitosa, E.L.: Evasion and countermeasures techniques to detect dynamic binary instrumentation frameworks. *Digit. Threats: Res. Pract.* **3**(2), 1–28 (2022)
6. Lad, S.S., Adamuthe, A.C.: Improved Deep Learning Model for Static PE Files Malware Detection and Classification. *Int. J. Comput. Netw. Inform. Security* **14**(2), 14–26 (2022)
7. Mane, T., Nimase, P., Parihar, P., Chandankhede, P.: Review of malware detection using deep learning. In: Ranganathan, G., Fernando, X., Shi, F., El Alloui, Y. (eds.) *Soft Computing for Security Applications*. AISC, vol. 1397, pp. 255–262. Springer, Singapore (2022). https://doi.org/10.1007/978-981-16-5301-8_19
8. Li, C., Lv, Q., Li, N., Wang, Y., Sun, D., Qiao, Y.: A novel deep framework for dynamic malware detection based on API sequence intrinsic features. *Comput. Secur.* **116**, 102686 (2022)
9. Anil Kumar, D., Das, S.K., Sahoo, M.K.: Malware detection system using API-decision tree. In: *Advances in Data Science and Management: Proceedings of ICDSM 2021*, pp. 511–517. Springer Nature Singapore, Singapore (2022)
10. Li, F.Q., Wang, S.L., Liew, A.W.C., Ding, W., Liu, G.S.: Large-scale malicious software classification with fuzzified features and boosted fuzzy random forest. *IEEE Trans. Fuzzy Syst.* **29**(11), 3205–3218 (2020)
11. Anderson, H.S., Roth, P.: Ember: an open dataset for training static pe malware machine learning models. *arXiv preprint arXiv:1804.04637* (2018)
12. Wilkins, Z., Zincir-Heywood, N.: COUGAR: clustering of unknown malware using genetic algorithm routines. In: *Proceedings of the 2020 Genetic and Evolutionary Computation Conference*, pp. 1195–1203 (2020)
13. Oyama, Y., Miyashita, T., Kokubo, H.: Identifying useful features for malware detection in the ember dataset. In: *2019 seventh international symposium on computing and networking workshops (CANDARW)*, pp. 360–366. IEEE (2019)
14. Pramanik, S., Teja, H.: EMBER-Analysis of Malware Dataset Using Convolutional Neural Networks. In: *2019 Third International Conference on Inventive Systems and Control (ICISC)*, pp. 286–291. IEEE (2019)
15. Yang, C., et al.: DeepMal: maliciousness-Preserving adversarial instruction learning against static malware detection. *Cybersecurity* **4**, 1–14 (2021)
16. Demirci, D., Acarturk, C.: Static malware detection using stacked BiLSTM and GPT-2. *IEEE Access* **10**, 58488–58502 (2022)
17. Chen, Z., Zhang, X., Kim, S.: A learning-based static malware detection system with integrated feature. *Intell. Autom. Soft Comput.* **27**(3), 891–908 (2021)
18. Singh, P., Borgohain, S.K., Sarkar, A.K., Kumar, J., Sharma, L.D.: Feed-forward deep neural network (FFDNN)-based deep features for static malware detection. *Int. J. Intell. Syst.* **2023**, 1–20 (2023). <https://doi.org/10.1155/2023/9544481>
19. Qi, P., Wang, W., Zhu, L., Ng, S.K.: Unsupervised domain adaptation for static malware detection based on gradient boosting trees. In: *Proceedings of the 30th ACM International Conference on Information and Knowledge Management*, pp. 1457–1466 (2021)
20. Kang, J., Won, Y.: A study on variant malware detection techniques using static and dynamic features. *J. Inform. Process. Syst.* **16**(4), 882–895 (2020)
21. Svec, P., Balogh, S., Homola, M.: Experimental evaluation of description logic concept learning algorithms for static malware detection. In: *ICISSP*, pp. 792–799 (2021)
22. Ebrahimi, M., Pacheco, J., Li, W., Hu, J. L., Chen, H.: Binary black-box attacks against static malware detectors with reinforcement learning in discrete action spaces. In: *2021 IEEE Security and Privacy Workshops (SPW)*, pp. 85–91. IEEE (2021)



Correction to: A Neural Network-Based Facial Expressions Detection Technique Using CK+ Dataset

Subhash Chandra Jat, Sadaf Naaz, and Shikha Chaudhary

Correction to:

Chapter 20 in: S. Joshi et al. (Eds.): *Cyber Warfare, Security and Space Computing*, CCIS 2195,

https://doi.org/10.1007/978-3-031-73494-6_20

The original version of the chapter 20 was inadvertently published with incorrect author affiliation for Shikha Chaudhary. This has been corrected by replacing the affiliation.

The updated version of this chapter can be found at
https://doi.org/10.1007/978-3-031-73494-6_20



Correction to: Impact of Sentiment Analysis in E-Commerce and Cybersecurity

Sonakshi Arora, P. Harika, and Sakshi Shringi

Correction to:

Chapter 24 in: S. Joshi et al. (Eds.): *Cyber Warfare, Security and Space Computing*, CCIS 2195,

https://doi.org/10.1007/978-3-031-73494-6_24

The original version of the chapter 24 was inadvertently published with incorrect author affiliation for Sonakshi Arora, P. Harika and Sakshi Shringi. This has been corrected by replacing the affiliation.

The updated version of this chapter can be found at
https://doi.org/10.1007/978-3-031-73494-6_24

Author Index

A

Agarwal, Arpita 220, 325
Arora, Sonakshi 314

B

Balamurugan, J. 140, 182
Bhakare, Omkar 57
Boddu, Asritha 238

C

Chaudhari, Shivprasad 207
Chaudhary, Shikha 265
Chauhan, Yashwanth Singh 1
Chouke, Akshay 252

D

Deshmukh, Shashikant 207
Devi, Bali 302
Dharmamoorthy, G. 167
Dubey, Ankit Kumar 167

G

Ganapathyappan, Kirubavathi 359
Garg, Amit 281
Goel, Diksha 68

H

Harika, P. 314
Hazela, Bramah 194
Hegde, N. Prajwal 167
Honrao, Sachin B. 101

I

Ilyas, Md 125
Israel, Akeke Niyi 325

J

Jain, Ankit Kumar 68
Jain, Leena 112
Jain, Vikas Kumar 252

Jakhhera, Abhay 57
Jat, Subhash Chandra 265
Jayasudha, R. 125
Jogi, Anil Kumar 1
Jothi, K. R. 112, 194

K

Kalaiarasi, R. 26
Kamalanathan, C. 140
Kambala, Gireesh 112
Kirubavathi, Ganapathyappan 337
Krishna, Raguru Jaya 238
Kuldeep 1
Kumar, Devendra 1
Kumar, Raj 1
Kumar, Rajan 281
Kumar, Sachin 1
Kumar, Surendra 1
Kushwaha, Satpal Singh 281

L

Lakshminarayana, Sanjay 1
Lanjewar, Chetan Khemraj 112, 194
Logeshwaran, J. 281

M

Madhavi, Jetti 154, 182
Mallikarjuna, Basetty 91
Mary, S. Suma Christal 167

N

Naaz, Sadaf 265
Nagar, Preeti 220, 325
Nagayach, Ojasvi 238
Noorudheen, Fathima 337

P

Panday, Ankur 68
Pande, Soumitra S. 125
Pandit, Shyam 57

Parmar, Jitendra 252
 Patel, Anoop Kumar 291
 Patro, Pramoda 167
 Pazhani, A. Azhagu Jaisudhan 194
 Prasad, Chandan 57
 Priya, Santhosh 26
 Priya, V. Vishwa 125

R

Raj, Prince 291
 Rajan, S. Dheva 79
 Ramesh, Janjhyam Venkata Naga 112, 125
 Rath, Samikshya 238
 Reddy, A. Basi 140, 154, 182

S

Salagrama, Shailaja 281
 Selvan, R. Senthamil 140, 154, 182
 Shah, Mehil Bimal 302
 Shankar, Venkatesh Gauri 302
 Sharma, Antima 220, 325
 Sharma, Gauri 302
 Sharma, Neelam 140
 Shinkar, Sonali. V. 207
 Shiurkar, U. D. 101
 Shringi, Sakshi 314
 Shukla, Shiv Shankar Prasad 252
 Singh, Jitendra 1

Singh, Pavitar Parkash 112
 Singh, Ritu 220
 Sivaraman, P. R. 194
 Sivaraman, R. 167, 194
 Soni, Vinit 1
 Srikaanth, P. Balaji 154
 Suganthi, D. 125

T

Tated, Siddharth 57
 Tiwari, Varun 91
 Trikha, Anadi 220, 325
 Tupetewar, Akshay 207

V

Vaishnav, Ajay Kumar 1
 Vamshi Krishna, B. 238
 Vardhan, Harsh 1
 Venkataratnam, Surepalli 154
 Verma, Atul Kumar 252
 Verma, Vijay 229

Y

Yadav, Abhishek 359
 Yadav, Mali 182
 Yadav, Sameer 154
 Yadav, Vinod Singh 1