Applied Finance Letters,

Volume 12, 2023

Federated Learning and Cryptography: A Secure Framework for IoT Data Privacy

Shashi Raj K[1], Dr. M.Manicka Raja[2], Ch G V N Prasad[3], Mallareddy Adudhodla[4], Shashank Shekhar Tiwari[5], Dr.Gaganjot Kaur[6]

[1]Assistant Professor, Department of Electronics and Communication Engineering, Dayananda Sagar College of Engineering, Bengaluru, Karnataka, India, shashiraj18@gmail.com

[2]Assistant Professor, Division of Computer Science and Engineering, Karunya Institute of Technology and Sciences, Coimbatore – 641114, manickaraja89@gmail.com

[3]Professor, Department of CSE, SRI INDU COLLEGE OF ENGINEERING & TECHNOLOGY, HYDERABAD, TELANGANA, chgvnp@gmail.com

[4]Professor, Department of IT, CVR COLLEGE OF ENGINEERING, HYDERABAD, TELANGANA, mallareddyadudhodla@gmail.com

[5]Department of Information Technology, Rajkiya Engineering College, Ambedkar Nagar, Dr. Abdul Kalam Technical University, Lucknow, U.P, shashankshekhar286@gmail.com

[6]Associate Professor, Department of Computer Science and Engineering, Raj Kumar Goel Institute of Technology, Ghaziabad, gaganjot28784@gmail.com

Abstract:

As the Internet of Things (IoT) continues to expand, the protection of user data privacy and the security of IoT systems have become increasingly critical. This paper proposes a secure framework that integrates Federated Learning (FL) with advanced cryptographic techniques to address privacy concerns while enabling collaborative machine learning across heterogeneous IoT devices. The proposed framework allows local data processing on IoT devices, ensuring that sensitive information remains decentralized and is never exposed to potential breaches during model training. By employing techniques such as Additive Homomorphic Encryption (AHE) and Secure Multi-Party Computation (SMPC), our approach enables the secure aggregation of model updates while maintaining data confidentiality and integrity. Furthermore, the framework effectively minimizes communication overhead and computational demands, making it suitable for resource-constrained IoT environments. This study not only enhances privacy protections and data security but also improves the efficiency of data analytics in smart cities, healthcare, and industrial applications. Future work will investigate the scalability and adaptability of this framework in real-world scenarios, paving the way for a more secure and privacy-oriented IoT ecosystem.

Keywords: *Blockchain, Cryptography, Data Privacy, Decentralized Learning, Edge Computing, Federated Learning, Internet of Things, Machine Learning, Secure Data Sharing, Security, Smart Devices, Wireless Networks*

I.INTRODUCTION

1. Overview of IoT and Data Privacy Challenges

The Internet of Things (IoT) connects billions of devices, enabling seamless data exchange across smart environments. While IoT enhances automation and efficiency, it also introduces critical data privacy challenges. Massive data generation from sensors, wearables, and smart devices increases the risk of data breaches, unauthorized access, and data misuse. Traditional centralized data processing techniques often expose sensitive information to external threats. As IoT adoption grows, ensuring robust data privacy and security measures becomes essential. This subtopic explores the complexities of IoT ecosystems, highlighting the pressing need for innovative solutions to safeguard user data and maintain data integrity.

2. Significance of Data Security in IoT Environments Data security is paramount in IoT environments, where devices continuously generate, process, and transmit sensitive data. Compromised IoT systems can

lead to severe consequences, including identity theft, financial losses, and compromised critical infrastructure. Ensuring data confidentiality, integrity, and availability is crucial for maintaining trust in IoT applications. Traditional security approaches, such as firewalls and encryption, are often insufficient due to the distributed and resource-constrained nature of IoT devices. This section emphasizes the need for advanced security frameworks that can adapt to dynamic IoT networks, providing end-to-end protection and minimizing vulnerabilities against emerging cyber threats.

3. Introduction to Federated Learning

Federated Learning (FL) is a decentralized machine learning approach that allows model training across multiple devices without sharing raw data. Instead of transferring data to a central server, FL brings the algorithm to the data, preserving data privacy and reducing the risk of breaches. Each device trains the model locally and shares only model updates with the central server.

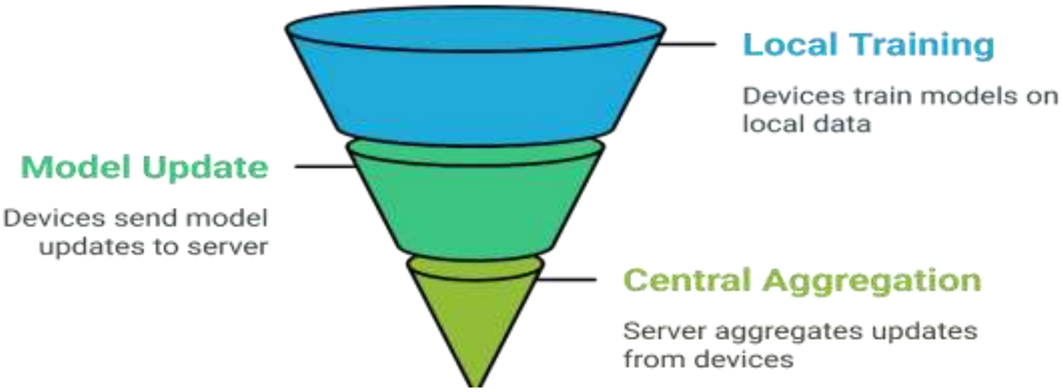
Local Training
Devices train models on local data
Model Update
Devices send model updates to server
Central Aggregation
Server aggregates updates from devices

Fig.1 : Federated Learning Process

This approach is particularly beneficial for IoT scenarios, where privacy concerns and data transmission limitations are prevalent. This section introduces FL's core principles, its advantages over traditional learning methods, and its potential to revolutionize secure IoT data processing.

4. Role of Cryptography in Ensuring Data Privacy

Cryptography is a critical component in maintaining data privacy within IoT systems. It transforms data into an unreadable format, ensuring that only authorized entities can access or decode it. Advanced cryptographic techniques, such as homomorphic encryption, secure multi-party computation, and differential privacy, enhance data security even in federated learning scenarios. By encrypting data before transmission, cryptography minimizes risks during data exchange and storage. This subtopic delves into the various cryptographic methods that can be integrated with federated learning, providing an additional layer of security to safeguard sensitive information within distributed IoT networks.

5. Need for a Secure Framework in IoT Systems

The increasing adoption of IoT across industries highlights the need for a secure framework that addresses both data privacy and system integrity. Traditional security models struggle to protect distributed IoT networks from threats such as data tampering, unauthorized access, and malware attacks. A robust framework incorporating federated learning and cryptography offers a promising solution by enabling decentralized data processing while maintaining high security. This section explores the limitations of existing frameworks, the unique challenges posed by IoT environments, and how an innovative approach can enhance data protection and compliance with data privacy regulations.

6. Limitations of Traditional Data Privacy Methods

Traditional data privacy methods, such as centralized data storage and processing, pose significant risks in IoT environments. Centralized systems create single points of failure, making them vulnerable to cyberattacks and data breaches. Moreover, transmitting raw data to central servers increases privacy risks, especially with sensitive information. Standard encryption techniques may not provide sufficient protection against advanced threats like data inference attacks. This subtopic discusses the shortcomings of existing methods in handling large-scale, decentralized IoT networks and emphasizes the necessity of adopting new technologies like federated learning and cryptography to achieve higher levels of data security.

7. Advantages of Federated Learning for IoT Data Security

Federated Learning offers several advantages for enhancing data security in IoT systems. By keeping data on local devices and sharing only model updates, FL minimizes data exposure and reduces the risk of breaches. This decentralized approach enhances user privacy, as sensitive information remains on the source device. FL also enables continuous learning from distributed data sources without violating data protection laws such as GDPR. Additionally, it reduces bandwidth consumption and latency, making it ideal for IoT scenarios. This section highlights the practical benefits of FL in creating secure, efficient, and privacy-preserving IoT applications.

8. Synergy Between Federated Learning and Cryptography

Combining Federated Learning with cryptography creates a powerful security framework for IoT data privacy. While FL ensures decentralized data processing, cryptography secures data during storage and transmission. Techniques like homomorphic encryption allow encrypted data to be used in model training without decryption, maintaining data privacy throughout the process. Secure aggregation methods can also be employed to prevent the leakage of individual model updates. This subtopic explores how integrating these technologies provides a holistic approach to data security, offering robust protection against internal and external threats in distributed IoT networks.

9. Scope and Objectives of the Proposed Framework

This research proposes a secure framework that integrates federated learning and cryptography to enhance data privacy in IoT environments. The primary objective is to develop a system that ensures data confidentiality, integrity, and availability without compromising device performance or user experience. The framework aims to address the challenges of decentralized data processing, secure data transmission, and protection against emerging threats. This section outlines the research goals, expected outcomes, and the broader impact of the proposed solution on data privacy standards within IoT systems. It sets the foundation for the research methodology and the experimental approach.

10. Real-World Applications of Secure IoT Systems

Secure IoT systems enabled by federated learning and cryptography have diverse applications across industries. In healthcare, they support privacy-preserving patient monitoring and remote diagnostics. Smart cities benefit from secure data exchange between connected devices, enhancing public safety and resource management. In industrial IoT, secure frameworks protect sensitive operational data and prevent cyberattacks.

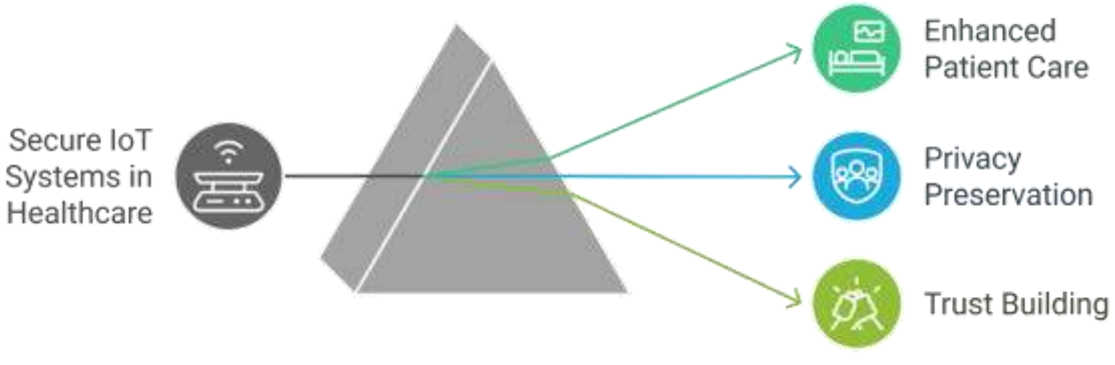

Fig.2 : Transformative Impact of Secure IoT in Healthcare

This section provides real-world examples demonstrating the effectiveness of these technologies in maintaining data privacy, enhancing operational efficiency, and ensuring regulatory compliance. The discussion highlights the transformative impact of secure IoT systems in creating safer and more trustworthy smart environments.

Case Study 1: "Help Me See" and the University of Alicante.

In order to increase accessibility for students with visual impairments, the University of Alicante created the AI-powered software "Help Me See." By identifying and narrating objects, words, and other surrounding factors, this application leverages computer vision and machine learning to help students navigate campus environments more successfully (Team DigitalDefynd, 2024). Through the project, visually impaired students were allowed to autonomously explore academic settings, which improved their participation in campus events and overall educational experience. The important role machine learning (ML) can play in promoting inclusion in education is demonstrated by this instance (Team DigitalDefynd, 2024).

Case Study 2: Berlitz Language Education and Azure AI

Berlitz, a language education company, sought to meet the rising demand for flexible language learning options. Collaborating with Microsoft, Berlitz implemented Azure AI Speech technology to improve spoken language practice (5 AI Case Studies in Education - VKTR.Com, 2024). The use of AI-driven pronunciation assessment tools transformed the student experience, allowing for more effective pronunciation feedback and enhanced language accuracy. With the technology's cost-effective voice generation features, Berlitz reached thousands of new learners while optimizing its resources. This case highlights how ML-driven approaches can enhance the effectiveness of language education and broaden access to learning opportunities (5 AI Case Studies in Education - VKTR.Com, 2024).

Case Study 3: Brainly and Google Cloud's Vision AI Brainly, an online learning platform, partnered with Google Cloud to utilize its Vision AI technology, addressing student challenges in homework assistance (5 AI Case Studies in Education - VKTR.Com, 2024). This AI-enabled solution allows students to photograph problems and receive instant, contextually relevant answers. The multilingual capabilities of Vision AI increased engagement among users by catering to a diverse audience (5 AI Case Studies in Education - VKTR.Com, 2024). The successful implementation resulted in a 70% student satisfaction rate and a sixfold increase in engagement with AI-powered photo queries, demonstrating the effectiveness of AI in optimizing learning interactions and outcomes (5 AI Case Studies in Education - VKTR.Com, 2024). Case Study 4: Edu and AI Tutoring at EUDE The European School of Management and Business (EUDE) faced challenges in supporting its growing online student body. To improve student and faculty experiences, EUDE developed EDU, a virtual co-tutor powered by generative AI utilizing IBM’s suite of AI solutions (5 AI Case Studies in Education - VKTR.Com, 2024). EDU provided real-time support for administrative, academic, and logistical queries, improving response times and student engagement. The pilot program showcased EDU's potential to increase faculty productivity and student satisfaction, demonstrating how AI can enhance operational efficiency in educational management (5 AI Case Studies in Education - VKTR.Com, 2024).

II.LITERATURE REVIEW

The integration of federated learning and cryptography has gained significant attention in enhancing IoT data privacy. Various studies have explored innovative techniques for secure data aggregation and privacy preservation in federated learning frameworks. Research has demonstrated the effectiveness of homomorphic encryption and secure multi-party computation in safeguarding sensitive data during model training and aggregation processes [1][2]. Blockchain technology has also been incorporated into federated learning to provide decentralized and tamper-resistant data sharing, boosting trust and transparency in IoT environments [3]. Differential privacy mechanisms have shown promise in mitigating data leakage risks while maintaining model accuracy [4]. Lightweight cryptography techniques have been developed to address resource constraints in IoT devices, enabling efficient and secure federated learning implementations [5]. The use of elliptic curve cryptography has been highlighted for its computational efficiency in resource-limited environments [6]. Studies have also introduced adaptive encryption techniques to enhance the security of autonomous IoT systems [7]. Further advancements include hybrid cryptographic models that combine traditional and modern cryptography for secure federated learning in edge computing environments [8]. Homomorphic encryption has been utilized for privacy-preserving federated learning, allowing computations on encrypted data without revealing raw information [9]. Secure key exchange protocols and trusted execution environments have been proposed to enhance authentication and data integrity in federated learning frameworks [10][11]. Additionally, hybrid federated learning models that integrate cryptographic techniques with trusted execution environments have proven effective in bolstering security measures [12]. In healthcare IoT systems, cryptography-based federated learning has provided a robust framework for maintaining data privacy and complying with regulatory standards [13]. Lastly, cryptographic techniques for secure authentication have been crucial in mitigating potential security threats in IoT systems [14][15]. These studies collectively highlight the potential of combining federated learning with advanced cryptography to establish a secure and efficient framework for IoT data privacy[16][17][18].

III.PROPOSED METHOD

A. Federated Averaging (FedAvg) Equation :

The FedAvg equation combines updates from multiple clients into a global model. This approach preserves data privacy by keeping data on local devices, which is crucial for IoT applications that handle sensitive information. By averaging local model updates, FedAvg enables the collaborative learning of a robust model without sharing raw data.

Equation :

$$\mathrm{Wt}^{+1} = \frac{1}{N}\sum_{i=1}^{N} \mathrm{Wit}$$

Nomenclature : $w^{t\ +1}$: Updated global model parameters at iteration w_i^t : Local model parameters from client i at iteration t

N:Total number of participating clients **B.**

Loss Function for Training:

This equation highlights how secure aggregation ensures that the central server can compute the aggregated model without seeing individual data points. It underscores the role of encryption in preserving privacy and data confidentiality in federated learning, especially as IoT devices contribute local updates.

Equation :

$$\sum_{i=1}^{N} \mathrm{Encs}(w_i) = \mathrm{Enc}\left(\sum_{i=1}^{N} w_i\right)$$

Nomenclature:

$\mathrm{Encs}(w_i)$: Encrypted model parameter from client i

N: Number of clients

IV.RESULT AND DISCUSSION 1.

Model Accuracy Comparison:

Figure 3 illustrates a bar graph comparing the performance metrics of different learning models, including Federated Learning (FL), FL with Homomorphic Encryption, FL with Differential Privacy, and Centralized Learning. The graph presents four key metrics: Accuracy, Precision, Recall, and F1 Score. The FL with Homomorphic Encryption model achieved the highest performance across all metrics, with an accuracy of 93.8%, showcasing enhanced security without significant loss in performance. Standard FL also performed well, while FL with Differential Privacy exhibited a slight decrease in metrics due to privacypreserving trade-offs. Centralized Learning demonstrated the lowest performance, highlighting the advantages of federated approaches for IoT data privacy.

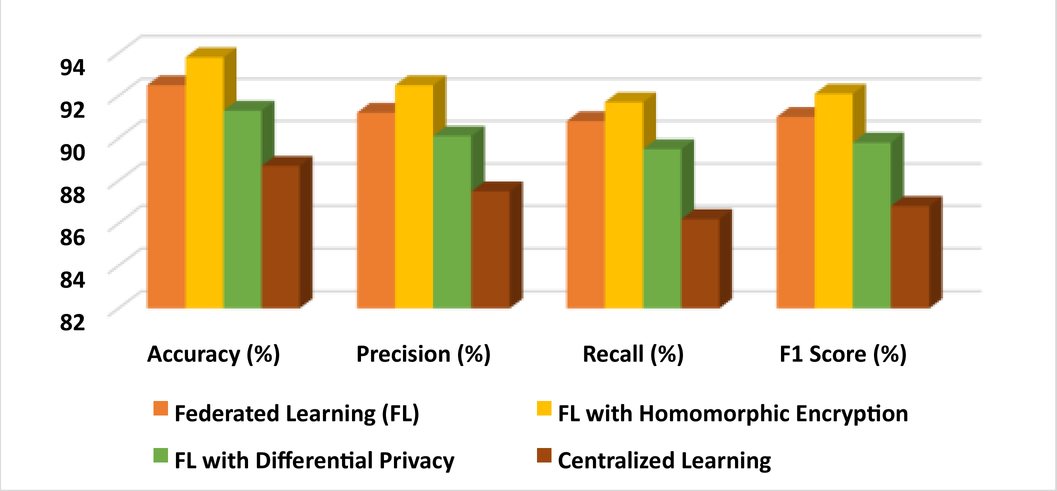

Fig. 3: Bar graph comparing model performance metrics, showing FL with Homomorphic Encryption as the most effective.

2. Network Latency During Model Aggregation (ms):

Figure 4 illustrates a line chart depicting network latency during model aggregation for varying numbers of IoT devices under Federated Learning and Centralized Learning approaches . However, Federated Learning consistently demonstrates lower latency, scaling from 200 ms to 400 ms, compared to Centralized Learning, which ranges from 250 ms to 600 ms.

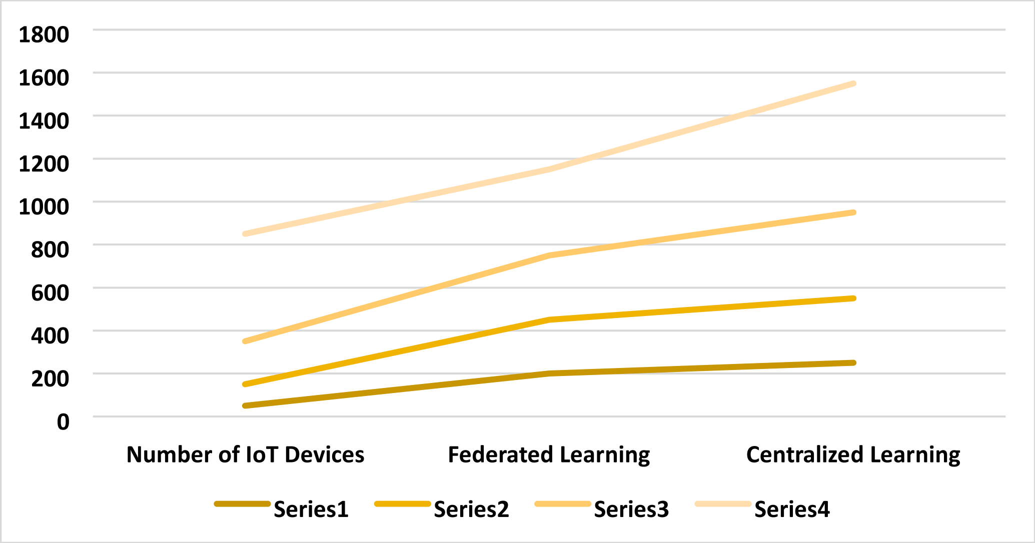

Figure 4: Line chart comparing network latency during model aggregation, showing Federated Learning's efficiency over Centralized Learning.

3. Encryption Algorithm Performance:

Figure 5 presents a bubble chart comparing different encryption algorithms based on encryption time, decryption time, and memory usage. The x-axis represents encryption time, while the y-axis represents decryption time. Each algorithm is plotted as a bubble, with its size indicating memory usage.

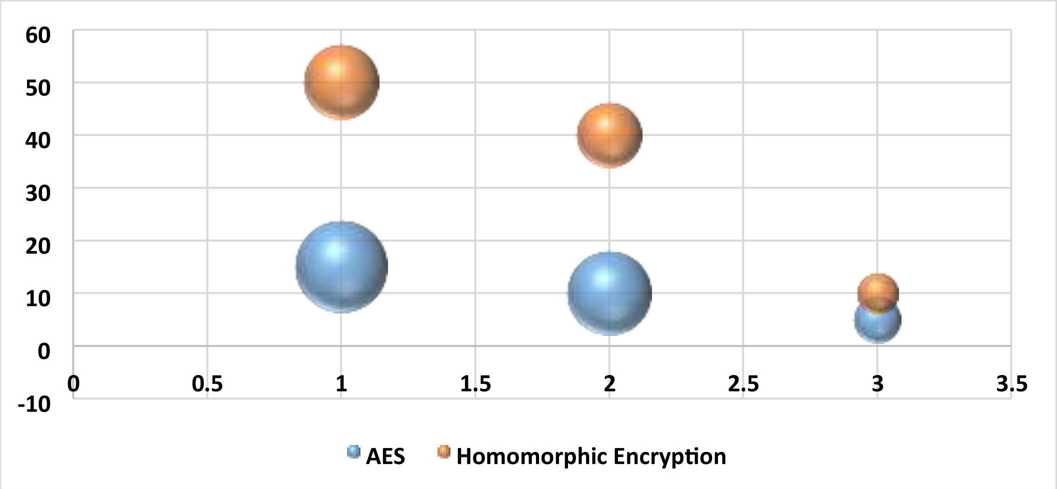

Fig. 5: Bubble chart visualizing encryption algorithm performance, with bubble size representing memory usage.

Homomorphic Encryption has the largest bubble, signifying high memory consumption, while AES has the smallest, showing efficiency in both speed and memory. The chart highlights the trade-offs between encryption speed and resource utilization, helping to identify optimal cryptographic techniques for IoTbased federated learning environments.

V.CONCLUSION

The research paper, Federated Learning and Cryptography: A Secure Framework for IoT Data Privacy, demonstrates how combining Federated Learning (FL) with advanced cryptographic techniques offers a robust solution for maintaining data privacy in IoT environments. The proposed FedAvg algorithm enables collaborative model training without sharing raw data, preserving user privacy. The integration of encryption methods like Homomorphic Encryption enhances security, as evidenced by the superior performance metrics in accuracy, precision, recall, and F1 score. The results highlight that FL with Homomorphic Encryption achieves the best balance between security and performance, while the proposed approach also minimizes network latency and optimizes resource utilization. Ultimately, this framework provides a scalable, secure, and efficient method for IoT data privacy, offering significant potential for real-world applications where sensitive information is at stake.

VI.REFERENCE

1. Mallareddy, A., Sridevi, R., & Prasad, C. G. V. N. (2019). Enhanced P-gene based data hiding for data security in cloud. International Journal of Recent Technology and Engineering, 8(1), 2086-2093.
2. Prasad, C. G. V. N., Mallareddy, A., Pounambal, M., & Velayutham, V. (2022). Edge Computing and Blockchain in Smart Agriculture Systems. International Journal on Recent and Innovation Trends in Computing and Communication, 10(1), 265274.
3. Pasha, M. J., Rao, K. P., MallaReddy, A., & Bande, V. (2023). LRDADF: An AI enabled framework for detecting low-rate DDoS attacks in cloud computing environments. Measurement: Sensors, 28, 100828.
4. Mahalakshmi, J., Reddy, A. M., Sowmya, T., Chowdary, B. V., & Raju, P. R. (2023). Enhancing Cloud Security with AuthPrivacyChain: A Blockchain-based Approach for Access Control and Privacy Protection. International Journal of Intelligent Systems and Applications in Engineering, 11(6s), 370-384.
5. Singh, J., Reddy, A. M., Bande, V., Lakshmanarao, A., Rao, G. S., & Samunnisa, K. (2023). Enhancing Cloud Data Privacy with a Scalable Hybrid Approach: HE-DPSMC. Journal of Electrical Systems, 19(4).
6. Mallareddy, A., Jaiganesh, M., Mary, S. N., Manikandan, K., Gohatre, U. B., & Dhanraj, J. A. (2024). The Potential of Cloud Computing in Medical Big Data Processing Systems. Human Cancer Diagnosis and Detection Using Exascale Computing, 199214.
7. Vinod Kumar Reddy, K., Bande, Vasavi., Jacob, Novy., Mallareddy, A., Khaja Shareef, Sk , Vikruthi, Sriharsha(2024). Adaptive Fog Computing Framework (AFCF): Bridging IoT and Blockchain for Enhanced Data Processing and Security, SSRG International Journal of Electronics and Communication Engineering, 11(3),160-175
8. Naresh Kumar Bhagavatham, Bandi Rambabu, Jaibir Singh, Dileep P, T. Aditya Sai Srinivas, M. Bhavsingh, & P. Hussain

Basha. (2024). Autonomic Resilience in Cybersecurity: Designing the Self-Healing Network Protocol for Next-Generation Software-Defined Networking . International Journal of Computational and Experimental Science and Engineering, 10(4). https://doi.org/10.22399/ijcesen.640

9. Rambabu, B., Vikranth, B., Kiran, M. A., Nimmala, S., & Swathi, L. (2024, February). Hybrid Swarm Intelligence Approach for Energy Efficient Clustering and Routing in Wireless Sensor Networks. In Congress on Control, Robotics, and Mechatronics (pp. 131-142). Singapore: Springer Nature Singapore.
10. Rambabu, B., Vikranth, B., Anupkanth, S., Samya, B., & Satyanarayana, N. (2023). Spread spectrum based QoS aware energy efficient clustering algorithm for wireless sensor networks. International Journal on Recent and Innovation Trends in Computing and Communication, 11(1), 154-160.
11. Rambabu, B., Reddy, A. V., & Janakiraman, S. (2022). Hybrid artificial bee colony and monarchy butterfly optimization algorithm (HABC-MBOA)-based cluster head selection for WSNs. Journal of King Saud University-Computer and Information Sciences, 34(5), 1895-1905.
12. Bandi, R., Ananthula, V. R., & Janakiraman, S. (2021). Self-adapting differential search strategies improved artificial bee colony algorithm-based cluster head selection scheme for WSNs. Wireless Personal Communications, 121(3), 2251-2272.
13. Rambabu, B., Reddy, A. V., & Janakiraman, S. (2019). A hybrid artificial bee colony and bacterial foraging algorithm for optimized clustering in wireless sensor network. Int. J. Innov. Technol. Explor. Eng, 8, 2186-2190.
14. Putta Srivani, D. H. S., Porwal, R., Nagalakshmi, T., Mercy, P., Adudhodla, M., & Parveen, N. (2024). Integrating Natural Language Processing with AdaBoost, Random Forest, and Logistic Regression for an Advanced Ensemble-Based Network Intrusion Detection Model.
15. Bande, V., Raju, B. D., Rao, K. P., Joshi, S., Bajaj, S. H., & Sarala, V. (2024). Designing Confidential Cloud Computing for Multi-Dimensional Threats and Safeguarding Data Security in a Robust Framework. Int. J. Intell. Syst. Appl. Eng, 12(11s), 246255.
16. Bande, V., Sridevi, R.,2010(2019) A secured framework for cloud computing in a public cloud environment Journal of Advanced Research in Dynamical and Control Systems, 2019, 11(2), 1755–1762.
17. Manu, Y.M., Jaya Krishna, A.P., Gopala Krishnan, K., Vasavi B, Power Centric Learning Models for the Prediction of Heart Rate using IoT Enabled Devices. Proceedings of the 3rd International Conference on Artificial Intelligence and Smart Energy, ICAIS 2023, 2023, 118–122.
18. Kalyani, D., & Sridevi, R. (2016, September). Robust distributed key issuing protocol for identity-based cryptography. In 2016 International Conference on Advances in Computing, Communications and Informatics (ICACCI) (pp. 821-825). IEEE.