International Journal of Engineering Research and Science & Technology

ISSN 2319-5991 www.ijerst.com
© 2026 IJERST. All Rights Reserved
Peer Reviewed, Referred & Indexed Journal

Research Paper

BLOCKCHAIN-ASSISTED PRIVACY AND SECURITY ENHANCEMENT IN FEDERATED LEARNING

K. VIJAYA LAXMI [1], KELUTH RONITH AKSHAY [2], ANNAM MADHUKAR REDDY [3], TALASANI PUJITHA [4], KUNDURU NITHIN REDDY [5]

[1]Assistant Professor, Dept. of CSE, Guru Nanak Institutions Technical Campus, Ibrahimpatnam, Telangana, India

[2, 3, 4, 5] B.Tech, Dept. of CSE, Guru Nanak Institutions Technical Campus, Ibrahimpatnam, Telangana, India

ABSTRACT— Clients can securely share gradients computed on their local data with the server using Federated Learning (FL), removing the need for them to directly expose their sensitive local datasets. During the process of model aggregation in traditional FL, the server might make use of its dominant position to infer sensitive information from the shared gradients of the clients. During model training, malicious clients may also submit forged and malicious gradients. Such behavior not only compromises the integrity of the global model, but also diminishes the usability and reliability of trained models. To effectively address such privacy and security attack issues, this work proposes a Blockchain-based Privacy-preserving and Secure Federated Learning (BPS-FL) scheme, which employs the threshold homomorphic encryption to protect the local gradients of clients. We develop a Byzantine-robust aggregation protocol for BPS-FL to implement cipher-text level secure model aggregation in order to ward off malicious gradient attacks. Furthermore, the immutability and traceability of the data are guaranteed by our use of a blockchain as the fundamental distributed architecture for recording all learning processes. Our extensive security analysis and numerical evaluation demonstrate that BPS-FL satisfies the privacy requirements and can effectively defend against poisoning attacks.

Index Terms – Federated Learning, Blockchain, Privacy Preservation, Security,

Distributed Machine Learning, Secure Aggregation, Data Privacy, Smart Contracts, Distributed Ledger Technology, Cybersecurity.

I. INTRODUCTION

The rapid development of artificial intelligence (AI) has led to remarkable achievements in various fields, including medical diagnosis, natural language processing, and intrusion detection. These advancements are primarily driven by the availability of large-scale, high-quality datasets required to train effective machine learning models. However, the increasing awareness of data privacy, combined with strict regulatory requirements and data isolation among institutions, has made centralized data collection increasingly difficult. Federated Learning (FL) has emerged as a promising solution to these challenges by enabling multiple clients to collaboratively train a shared global model without directly exposing their private local datasets. In a typical FL workflow, each client computes local gradients on its private data and sends model updates to a central server, which aggregates them to refine the global model. This paradigm significantly reduces privacy risks while promoting collaborative intelligence among decentralized data owners. Despite its advantages, FL faces critical privacy and security challenges. Although raw data are not shared, gradient updates may still leak sensitive information, and adversaries—including malicious clients and compromised servers—can exploit these gradients to infer private data. Moreover, the distributed and opaque nature of FL creates opportunities for model poisoning attacks, where adversarial clients intentionally manipulate gradients or corrupt labels to degrade the performance of the global model. Existing privacy-preserving techniques, such as Secure Multi-party Computation (SMC), Homomorphic Encryption (HE), and Differential Privacy (DP), each have limitations in terms of communication overhead, accuracy degradation, or robustness against adversarial behaviours. Many FL systems rely on a single central server for aggregating model updates, which leads to vulnerabilities such as single-point failures and susceptibility to malicious tampering. While several defense strategies—such as Krum, Trimmed-mean, Median, FL Trust, and subsequent enhancements—have been proposed to mitigate poisoning attacks, most require access to clients' gradients in plaintext, reintroducing the risk of information leakage. This inherent conflict between privacy

preservation and poisoning attack detection presents a major barrier to building secure and trustworthy FL systems. To address these challenges, we propose a Blockchain-based Privacy-preserving and Secure Federated Learning (BPS-FL) framework that simultaneously ensures privacy, robustness, and system integrity. BPS-FL integrates threshold Paillier homomorphic encryption to protect clients' gradient information while enabling secure ciphertext-level aggregation. To detect and mitigate malicious contributions, we design a Byzantine-robust aggregation protocol supported by a decentralized federated learning committee that assesses clients' encrypted gradients without compromising privacy. Furthermore, by leveraging blockchain's immutability, transparency, and decentralized consensus mechanisms, BPS-FL eliminates the reliance on a single server and provides tamper-proof recording of model update histories, enhancing traceability and accountability. Extensive security analysis and numerical evaluations demonstrate that BPS-FL achieves strong privacy guarantees, effectively detects poisoning behaviors, and maintains high model accuracy under adversarial settings. The proposed scheme offers a practical and secure foundation for deploying federated learning in real-world, privacy-sensitive, and adversarial environments. FL deployments also face practical challenges arising from heterogeneous data distributions and inconsistent client participation. In real-world scenarios, different clients may generate data with non-identical distributions (non-IID), causing divergence and instability during model training. Moreover, communication constraints, intermittent connectivity, and computational limitations can lead to unpredictable client dropouts, making reliable aggregation difficult. Existing privacy-preserving FL frameworks often assume idealized participation patterns and fail to account for such dynamic behaviors, resulting in degraded performance or incomplete model updates. These challenges highlight the need for a robust, decentralized mechanism that ensures reliable participation tracking, secure communication, and verifiable aggregation, even when clients behave unpredictably or maliciously. While homomorphic encryption offers strong privacy guarantees by enabling computation directly on ciphertext, integrating it into an adversarial FL environment raises new complexities. Traditional HE-based FL frameworks often lack mechanisms to counter model poisoning or Sybil attacks at the ciphertext level, and

existing Byzantine-resilient methods typically require plaintext gradients to detect anomalies. Blockchain's decentralized trust model presents an opportunity to overcome these limitations by distributing key shares, ensuring auditability of contributions, and enforcing protocol compliance without relying on a single authority.

II. RELATED WORK

A. *Federated graph learning with network inpainting for population-based disease prediction.*

This project focuses on developing a federated graph learning framework integrated with network inpainting techniques to enhance population-based disease prediction in a privacy-preserving manner. Traditional machine learning models require centralized access to patient data, which is often restricted due to privacy regulations and institutional barriers. Federated learning addresses this issue by enabling decentralized model training; however, it struggles when data are incomplete, heterogeneous, or distributed across multiple medical institutions. To overcome these limitations, the proposed system models patient populations as graph structures that capture relationships such as demographics, clinical similarities, comorbidities, and temporal health patterns. Since real-world medical graphs frequently contain missing nodes or incomplete edges due to limited diagnostic history or fragmented data sources, a network inpainting module is introduced to reconstruct missing graph components and restore structural integrity. This ensures that each participating site contributes meaningful relational information without exposing sensitive records. The federated graph learning framework then trains a global predictive model collaboratively across institutions while preserving privacy. By combining federated learning, graph neural networks, and inpainting mechanisms, the system enhances prediction accuracy, handles missing data more effectively, and enables robust population-level disease risk assessment. This approach supports early diagnosis, resource planning, and large-scale health analytics while maintaining strict privacy and security requirements.

B. *Semi supervised federated-learning-based intrusion detection method for Internet of Things.*

This project presents a semi-supervised federated learning–based intrusion detection method designed to enhance the security of Internet of Things (IoT) environments while preserving data privacy. IoT networks

generate massive amounts of heterogeneous traffic, but most data remain unlabeled due to the high cost and expertise required for manual annotation. Traditional centralized intrusion detection models rely on collecting raw traffic data at a central server, which raises serious privacy concerns and is impractical in resource-constrained IoT settings. To address these challenges, the proposed system employs semi-supervised learning to leverage both labeled and unlabeled IoT data distributed across multiple devices. Federated learning enables collaborative model training without exposing local traffic logs, ensuring that sensitive information remains on each device. A feature extraction module processes raw traffic locally, while a global semi-supervised model aggregates learned patterns to detect both known and unknown threats. The method is designed to handle non-IID data, device heterogeneity, and communication limitations inherent to IoT networks. By combining semi-supervised learning with federated aggregation, the system improves detection accuracy, enhances generalization to unseen attacks, and reduces reliance on labeled datasets. This approach offers a scalable, privacy-preserving, and efficient intrusion detection solution suitable for large, distributed IoT deployments.

C. Backdoor attacks and defenses in federated learning.

This project focuses on analyzing and addressing backdoor attacks in federated learning (FL), a decentralized machine learning paradigm where multiple clients collaboratively train a shared model without exposing their private data. Although FL enhances privacy by keeping data local, it becomes vulnerable to backdoor threats because the server cannot directly inspect client datasets or verify the integrity of their model updates. In a backdoor attack, a malicious client intentionally injects manipulated gradients or poisoned samples containing hidden triggers, causing the global model to behave normally on clean data but misclassify specific trigger-based inputs. Such attacks are stealthy, difficult to detect, and highly damaging in applications like healthcare, finance, and autonomous systems. The project studies different types of backdoor attacks, including data poisoning, model poisoning, and targeted trigger manipulation, and evaluates how these attacks exploit the aggregation process in FL. To counter these threats, the project investigates defense strategies such as robust aggregation rules, anomaly detection, client

behavior auditing, differential privacy, and secure gradient validation. It also examines defense limitations, such as the trade-off between privacy and detectability and the challenges posed by encrypted or non-transparent updates. Overall, the goal is to understand vulnerabilities in FL systems and develop effective mechanisms to detect, mitigate, and prevent backdoor attacks while preserving privacy and model performance.

III. PROPOSED METHODOLOGY

We propose an innovative Blockchain-based Privacy-preserving and Secure Federated Learning (BPS-FL) scheme. This scheme can effectively detect and defend against poisoning behaviors during joint training while ensuring client privacy. In BPS-FL, we utilize threshold Paillier homomorphic encryption technology that allows computations on encrypted data, where the results remain encrypted, enabling the secure aggregation and analysis of local gradients without violating any client privacy. To further resist model poisoning attacks, we have established a federated learning committee, composed of multiple members who are responsible for interacting with the server and executing protocols to obtain weighted scores of each client's local gradients without compromising their privacy. Blockchain-based Privacy-preserving and Secure Federated Learning (BPS-FL) scheme. In this paper, we propose BPS-FL scheme. BPS-FL adopts blockchain technology as the distributed architecture for FL, where all learning processes can be recorded in an immutable ledger, thereby preventing malicious behaviors.

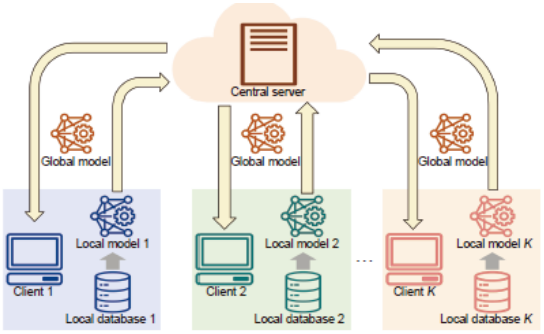
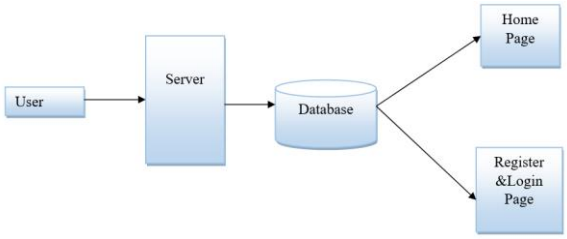
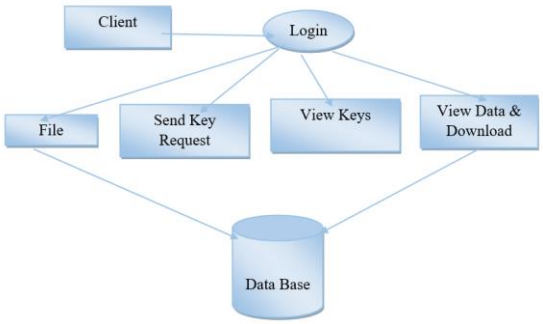


Fig. 1: System Overview

3.1 User Interface Design

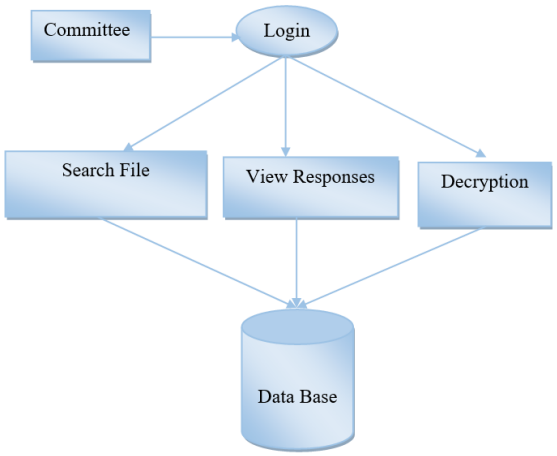
In this module we design the windows for the project. These windows are used for secure login for all users. To connect with server user must give their username and password then only they can able to connect the server. If the user already exits directly can login into the server else user must register their details such as username, password and Email id, into the server. Server will create the account for the entire user to maintain upload and download rate.





3.2 Committee

This module allows committee members to register, log in, and upload files securely. Uploaded files are encrypted using the KGC public key and sent for key generation and blockchain processing. Committees also approve or reject client key requests.

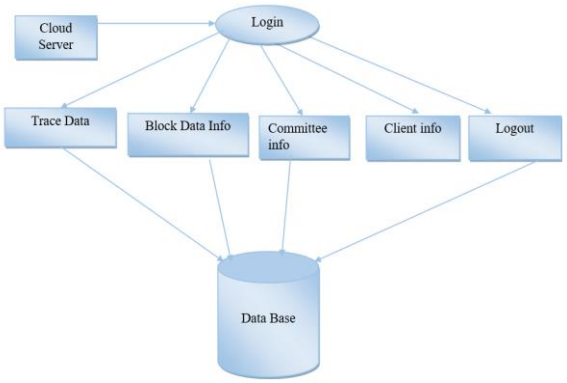


3.3 Client

Clients register and log in to search encrypted files and request public/private keys. They can verify public keys, decrypt files using approved private keys, and download original data securely. The system blocks malicious login attempts automatically.

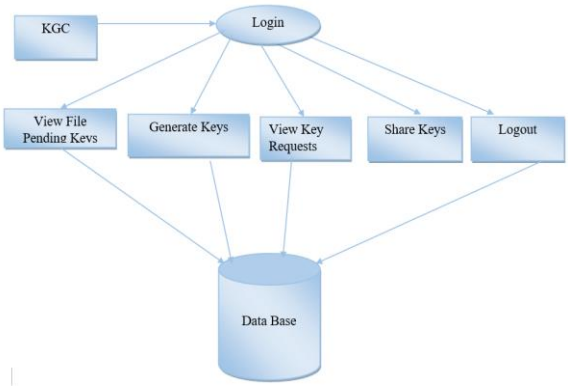
3.4 Cloud Server

The cloud module manages client and committee approvals. It monitors blocked users and allows unblocking after review. This acts as the administrative access control layer.



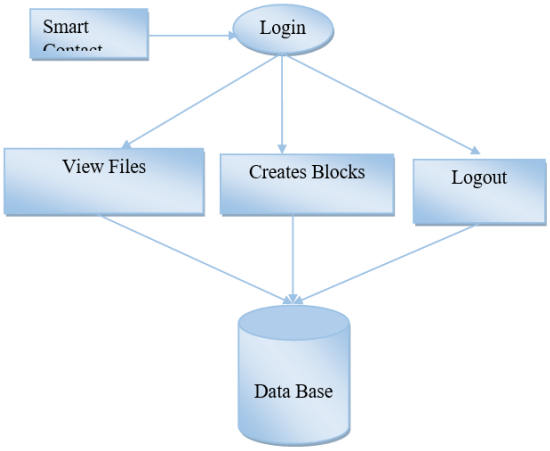
3.5 Key Generation Center

KGC is responsible for generating and distributing cryptographic keys. It approves public/private key requests and ensures keys are shared only after authorization.



3.6 Smart Contract

This module handles blockchain operations by splitting encrypted files into blocks. Each block is hashed and chained securely to ensure integrity and immutability.



IV. EXPERIMENTAL RESULTS AND ANALYSIS

Experimental results showed that integrating Blockchain Technology with Federated Learning enhanced the security and transparency of model updates while preserving user data privacy. The blockchain layer effectively prevented unauthorized modifications and ensured secure aggregation of local model parameters. The analysis demonstrates that the proposed approach achieves a balance between privacy, security, scalability, and model performance, making it suitable for distributed artificial intelligence applications in healthcare, finance, and IoT environments.



Fig. 2: Home Page

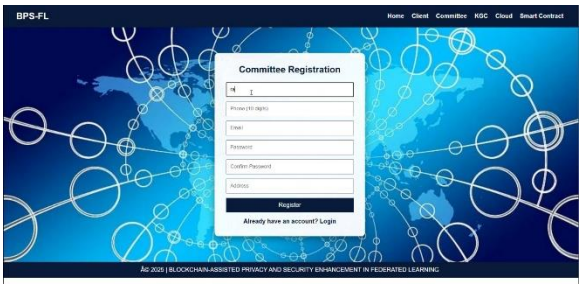


Fig. 3: Registration



Fig. 4: Login

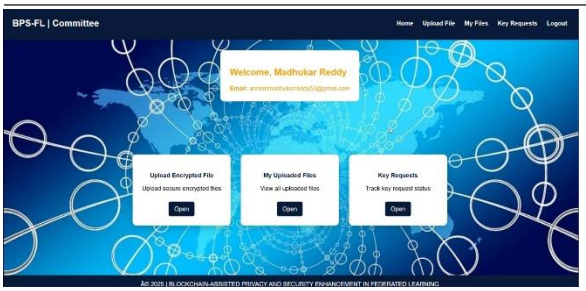


Fig. 5: Dashboard

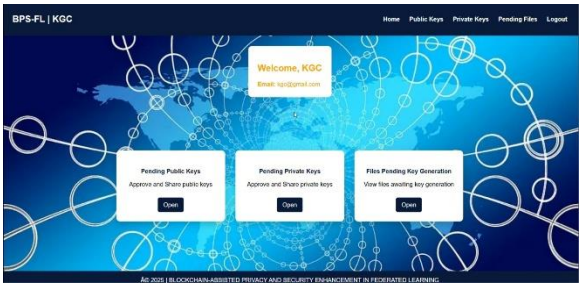


Fig. 6: Pending Keys

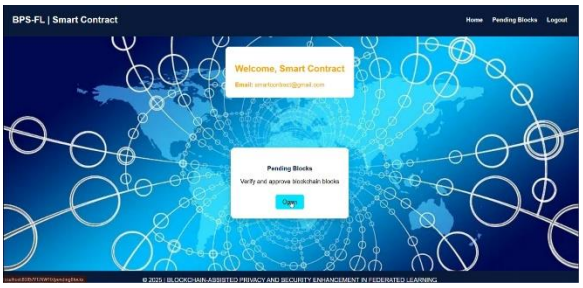


Fig. 7: Pending Blocks

V. CONCLUSION

we proposed a Blockchain-based Privacy-preserving and Secure Federated Learning (BPS-FL) scheme to address critical privacy and security challenges in federated learning environments. The proposed framework effectively protects client data privacy by encrypting local gradients using threshold homomorphic encryption**,** enabling secure aggregation without exposing sensitive gradient information. To defend against model poisoning attacks, we designed a Byzantine-robust aggregation protocol that operates at the ciphertext level, minimizing the impact of malicious client updates while ensuring accurate global model convergence. By leveraging blockchain technology as the underlying distributed architecture, BPS-FL eliminates single-point failures and guarantees immutability, traceability, and transparency throughout the learning process. Experimental results demonstrate that, although the computational cost increases with the number of participating clients, the overhead introduced by BPS-FL remains within acceptable limits and does not significantly affect overall system performance. Moreover, the scheme maintains high classification accuracy even in the presence of a large proportion of malicious participants. These results confirm that BPS-FL provides a secure, reliable, and practical solution for deploying federated learning in privacy-sensitive and adversarial environments.

REFERENCES

[1] L. Deng and D. Yu, Deep learning: Methods and applications, Found. Trends® Signal Process., vol. 7, nos. 3&4, pp. 197–387, 2014.

[2] L. Peng, N. Wang, N. Dvornek, X. Zhu, and X. Li, FedNI: Federated graph learning with network inpainting for population-based disease prediction, IEEE Trans. Med. Imag., vol. 42, no. 7, pp. 2032–2043, 2023.

[3] Z. Li, X. Wang, W. Yang, J. Wu, Z. Zhang, Z. Liu, M. Sun, H. Zhang, and S. Liu, A unified understanding of deep NLP models for text classification, IEEE Trans. Vis. Comput. Graph., vol. 28, no. 12, pp. 4980–4994, 2022.

[4] R. Zhao, Y. Wang, Z. Xue, T. Ohtsuki, B. Adebisi, and G. Gui, Semisupervised federated-learning-based intrusion detection method for Internet of Things, IEEE Internet Things J., vol. 10, no. 10, pp. 8645–8657, 2023.

[5] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B.A. y Arcas, Communication-efficient learning of deep networks from decentralized data, in Proc. of 20th Int. Conf. on Artificial Intelligence and Statistics, Ft. Lauderdale, FL, USA, 2017, pp. 1273–1282.

[6] X. Guo, Z. Liu, J. Li, J. Gao, B. Hou, C. Dong, and T. Baker, VeriFL: communication-efficient and fast verifiable aggregation for federated learning, IEEE Trans. Inf. Forensics Secur., vol. 16, pp. 1736–1751, 2021.

AUTHORS Profile



Ms. K. VIJAYA LAXMI received her B. Tech degree in Computer Science & Engineering from JNTUH, Hyderabad in 2015 and her M. Tech degree in Computer Science from JNTUH, Hyderabad in 2017. She is currently pursuing her Ph.D. degree in Computer Science & Engineering from Guru Nanak University, Hyderabad. She has been dedicated to the teaching field for the last 7 years. Her research area is Deep Learning. At present, she is working as an Assistant Professor in the CSE Department at Guru Nanak Institutions Technical Campus, Ibrahimpatnam, Telangana, India.



Mr. KELUTH RONITH AKSHAY pursuing B. Tech degree in Computer Science and Engineering at Guru Nanak Institutions Technical Campus affiliated to JNTUH in 2022-2026.



Mr. ANNAM MADHUKAR REDDY pursuing B. Tech degree in Computer Science and Engineering at Guru Nanak Institutions Technical Campus affiliated to JNTUH in 2022-2026.



Ms. TALASANI PUJITHA pursuing B. Tech degree in Computer Science and Engineering at Guru Nanak Institutions Technical Campus affiliated to JNTUH in 2022-2026.



Mr. KUNDURU NITHIN REDDY pursuing B. Tech degree in Computer Science and Engineering at Guru Nanak Institutions Technical Campus affiliated to JNTUH in 2022-2026.