IJARETY

International Journal of Advanced Research in Education and TechnologY (IJARETY)

Volume 12, Issue 3, May-June 2025

Impact Factor: 8.152

www.ijarety.in editor.ijarety@gmail.com

International Journal of Advanced Research in Education and TechnologY(IJARETY)

| ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 8.152| A Bi-Monthly, Double-Blind Peer Reviewed & Refereed Journal |

|| Volume 12, Issue 3, May-June 2025 ||

DOI:10.15680/IJARETY.2025.1203026

Secure Federated Learning: A Multi-Strategy Framework for Privacy, Integrity and Fairness

Velivarthi Shashank Reddy[1], Vattikoti Ganesh[2], Yanna Sai Ram[3], Dr.Khushbu Doulani[4]

UG Scholars, Department of Computer Science and Engineering, Guru Nanak Institutions Technical Campus, Ibrahimpatnam, Telangana, India[1,2,3]

Assistant Professor, Department of Computer Science and Engineering, Guru Nanak Institutions Technical Campus, Ibrahimpatnam, Telangana, India[4]

ABSTRACT: As artificial intelligence, data science, and collaborative data usage continue to evolve, federated learning (FL) has emerged as a key approach for joint model training across multiple parties. Despite its promise, traditional FL frameworks still struggle with critical issues such as data privacy, resistance to poisoning attacks, and equitable participant engagement. To overcome these challenges, we propose Secure Federated Learning (SFL), a new framework that integrates a combination of advanced security strategies. This framework empowers data contributors with full control over their private data, protects against adversarial threats from compromised participants, and incorporates a blockchain-based protocol to guarantee fairness and trustworthiness. Rigorous theoretical validation and empirical testing confirm the enhanced robustness and superior performance of the SFL framework compared to existing FL solutions.

KEYWORDS: Secure Federated Learning,Data Confidentiality,Blockchain Integration,Poisoning Attack Mitigation, Gradient Aggregation,Homomorphic Encryption,Privacy-Preserving AI, Multi-Party Computation,Smart Contract Mechanism, Byzantine-Robust Learning.

I. INTRODUCTION

In the era of data-driven technologies, the secure and efficient use of distributed data has become a central concern in machine learning. Traditional learning models typically require centralized data collection, which poses significant risks to privacy and data ownership. To address this, federated learning (FL) has emerged as a promising paradigm that enables multiple parties to collaboratively train machine learning models without exposing their raw data.

Despite its advantages, FL still faces notable challenges in real-world applications. One of the main concerns is safeguarding sensitive information during and after training, as model updates can inadvertently leak private data. Moreover, the presence of malicious participants introduces the risk of poisoning attacks, which can distort model accuracy and trustworthiness.

Another critical issue is ensuring fair and transparent participation among data owners and model requesters. Without proper incentives and protections, contributors may not be motivated to engage in collaborative training. To overcome these obstacles, it is essential to design FL frameworks that combine robust security protocols with fair value distribution mechanisms.

This paper proposes a novel federated learning framework named Secure Federated Learning (SFL). The framework integrates homomorphic encryption for data privacy, blockchain for transactional trust, and smart contracts to manage contributions and rewards. By doing so, SFL addresses the key limitations of current FL systems and offers a more reliable and equitable approach for collaborative model training.

II. EXISTING SYSTEM

Existing federated learning systems allow multiple parties to train a shared machine learning model without exchanging raw data. While this approach reduces privacy risks, it still leaves room for indirect data leakage through model gradients. Furthermore, most current frameworks lack effective defense mechanisms against malicious participants who may launch poisoning attacks. These systems also often fail to ensure equitable treatment of contributors, leading to issues in trust and motivation. Existing solutions typically rely on simple aggregation methods, making them vulnerable

| ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 8.152| A Bi-Monthly, Double-Blind Peer Reviewed & Refereed Journal |

|| Volume 12, Issue 3, May-June 2025 ||

DOI:10.15680/IJARETY.2025.1203026

to manipulated updates. Privacy-preserving techniques like homomorphic encryption and differential privacy are sometimes used, but not in an integrated or comprehensive manner. As a result, existing FL systems struggle to balance privacy, security, and fairness effectively.

III. DISADVANTAGES:

1. Limited Privacy Protection: Although raw data remains local, gradient updates can still leak sensitive information, compromising user privacy.
2. Vulnerability to Poisoning Attacks: Most systems are not equipped to detect or defend against malicious nodes that intentionally corrupt model updates.
3. Lack of Contributor Incentives: There is often no reliable mechanism to ensure fair compensation or recognition for data owners, reducing participation motivation.
4. Ineffective Trust Management: Without secure protocols, trust among participants can be weak, especially when parties are unknown or competitive.
5. Inadequate Aggregation Methods: Standard aggregation techniques like averaging are insufficient in adversarial settings, making models susceptible to biased or manipulated contributions.

IV. PROPOSED SYSTEM

The proposed system, Secure Federated Learning (SFL), introduces a robust framework that integrates multiple security strategies to enhance traditional FL. It employs homomorphic encryption to protect gradient information during training, ensuring data privacy is never compromised. A service provider aggregates encrypted updates, adding noise to further obscure individual contributions. The model requester then filters reliable gradients using a secure selection algorithm. To ensure fairness, a blockchain-based smart contract records contributions and manages reward distribution. This approach promotes trust, resists poisoning attacks, and maintains data confidentiality. Overall, SFL creates a secure, transparent, and incentive-compatible environment for collaborative learning.

V. ADVANTAGES

1. Enhanced Data Privacy: By using homomorphic encryption, the system ensures that data owners' sensitive information remains secure throughout the training process.
2. Defense Against Attacks: The framework effectively mitigates label-flipping, backdoor, and model poisoning attacks through robust gradient filtering and aggregation.
3. Fair Compensation: A blockchain-based smart contract tracks participant contributions and ensures transparent and fair reward distribution.
4. Trustworthy Collaboration: The use of cryptographic protocols and blockchain fosters trust among participants without requiring a central authority.
5. Scalable and Secure Aggregation: The system allows multiple data owners to collaborate securely, regardless of scale, while preserving model accuracy and robustness.

VI. RELATED WORK

Federated learning has gained attention as a privacy-friendly approach to collaborative model training without centralizing sensitive data. Early research focused on improving communication efficiency and training performance, but security concerns remain a major challenge. Several studies have highlighted vulnerabilities to inference and poisoning attacks, showing how malicious participants can compromise model integrity.

To address privacy concerns, techniques such as differential privacy and secure multi-party computation have been explored. Homomorphic encryption, in particular, allows computations on encrypted data, enabling secure model updates without exposing raw information. However, many existing solutions implement only one form of protection, limiting their effectiveness in diverse threat scenarios.

Defense mechanisms against poisoning attacks include robust aggregation strategies like trimmed mean, geometric median, and Krum, which aim to filter out anomalous updates. These approaches enhance resilience but can inadvertently discard useful contributions from honest nodes. Incentive mechanisms have also been proposed to encourage participation, often relying on trusted third parties.

| ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 8.152| A Bi-Monthly, Double-Blind Peer Reviewed & Refereed Journal |

|| Volume 12, Issue 3, May-June 2025 ||

DOI:10.15680/IJARETY.2025.1203026

Blockchain-based solutions have recently emerged to decentralize trust and automate contribution tracking. Yet, few systems integrate blockchain with advanced cryptographic techniques in a cohesive framework. Therefore, there remains a need for an end-to-end secure, fair, and privacy-preserving federated learning framework, which our proposed SFL system aims to fulfill.

VII. METHODOLOGIES

The Secure Federated Learning (SFL) framework is designed to address the major limitations of traditional federated learning systems, including data privacy concerns, susceptibility to poisoning attacks, and lack of fairness in participant compensation. The methodology involves several coordinated steps integrating cryptographic techniques, decentralized trust mechanisms, and robust gradient handling algorithms.

1. System Entities and Assumptions
The SFL framework operates with three key entities:
Model Requester: Initiates model training but does not possess data.
Data Owners: Hold local datasets and perform encrypted training.
Service Provider: Coordinates communication and performs intermediate computations while remaining untrusted for data content.

The framework assumes secure communication channels, honest-but-curious behavior from the service provider and model requester, and that a minority of data owners may act maliciously.

2. Model Initialization and Encryption
The model requester first creates a machine learning model, typically a logistic regression or similar algorithm suitable for secure operations. These model parameters are encrypted using the CKKS homomorphic encryption scheme, which supports arithmetic operations on encrypted values. The encrypted model is then sent to the service provider, which distributes it to all participating data owners.

3. Local Training by Data Owners
Each data owner performs local training on their dataset using the received encrypted model. Since operations like the sigmoid function are not directly compatible with encrypted data, polynomial approximations are used for compatibility. The owners compute encrypted gradient updates without decrypting the model or revealing any sensitive data, maintaining full privacy during training.

4. Privacy Enhancement Through Noise Addition
To prevent the service provider or model requester from reconstructing original gradients or inferring data patterns, the service provider introduces controlled random noise to the encrypted gradients. This step preserves privacy while still allowing accurate aggregation, thanks to the resilience of the chosen aggregation technique.

5. Decryption and Gradient Selection
Once the noisy, encrypted gradients are received, the model requester decrypts them using their private key. The requester then filters out suspicious or extreme gradient values using a secure selection algorithm. This step helps eliminate contributions from malicious nodes attempting to corrupt the model through label flipping, backdoor injection, or arbitrary updates.

6. Gradient Aggregation
From the filtered set of valid gradients, the model requester aggregates them to form the new global model parameters. This process ensures that only trustworthy contributions affect model updates. The aggregation strategy does not rely solely on simple averaging, making it robust against various poisoning strategies.

7. Smart Contract for Fair Reward Distribution
To incentivize participation and ensure fairness, the requester creates a blockchain-based smart contract that logs each data owner's contribution based on accepted gradient components. The contract also holds cryptocurrency or token-based rewards. After each training round, participants are compensated according to their verified contributions, eliminating the need for manual reward management.

| ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 8.152| A Bi-Monthly, Double-Blind Peer Reviewed & Refereed Journal |

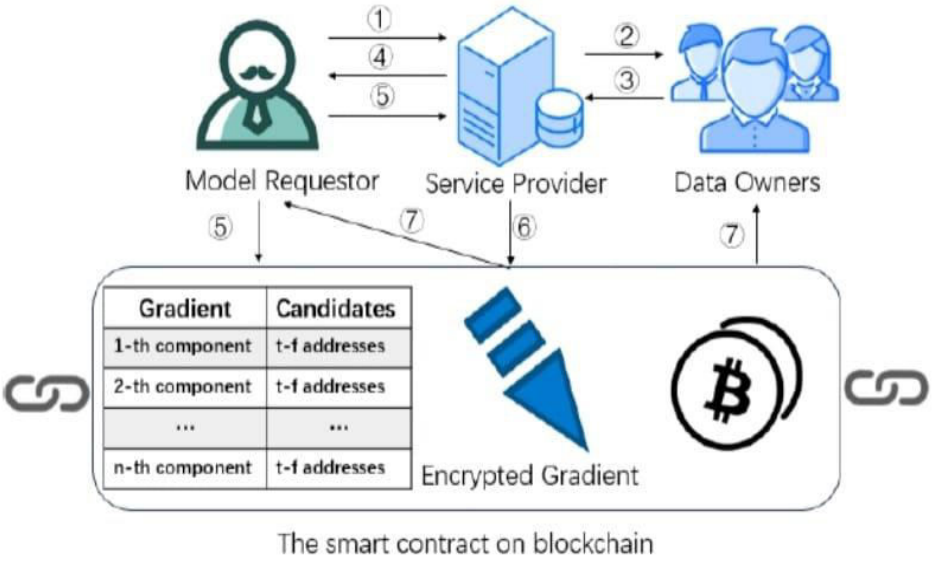
|| Volume 12, Issue 3, May-June 2025 ||

DOI:10.15680/IJARETY.2025.1203026

8. Model Update and Iteration

The updated model is re-encrypted and redistributed for further training rounds. This cycle repeats until the global model converges to a satisfactory accuracy level. Throughout the process, all computations are privacy-preserving, and fairness is maintained via transparent, tamper-proof records on the blockchain.

VIII. SYSTEM ARCHITECTURE



IX. CONCLUSION

This work presents a comprehensive solution to the key challenges in federated learning by introducing the Secure Federated Learning (SFL) framework. Unlike conventional FL models, SFL incorporates multiple layers of protection to address data privacy, trust, and security issues simultaneously. Through the integration of homomorphic encryption, data owners can train models without exposing sensitive information, ensuring that privacy is preserved throughout the process.

The framework also introduces robust mechanisms to defend against poisoning attacks, including label-flipping, backdoor, and arbitrary model manipulation. Its gradient filtering technique ensures that only trustworthy data contributes to the global model. Additionally, by leveraging blockchain and smart contracts, SFL promotes fairness and transparency among participants. Data owners are fairly compensated based on their validated contributions, removing the need for third-party trust.

Experimental results further validate the effectiveness of SFL, demonstrating improved robustness and accuracy compared to traditional methods. The framework's ability to operate securely in adversarial environments makes it a viable option for real-world applications where trust and data security are critical.

In summary, SFL represents a significant advancement in federated learning, offering a secure, fair, and efficient platform for collaborative model training. Future research will explore more scalable encryption methods and adaptive defense strategies to further enhance performance and usability.

| ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 8.152| A Bi-Monthly, Double-Blind Peer Reviewed & Refereed Journal |

|| Volume 12, Issue 3, May-June 2025 ||

DOI:10.15680/IJARETY.2025.1203026

REFERENCES

1. K. Bonawitz et al., " Practical secure aggregation for insulation- conserving machine knowledge, " in Proc. ACM Conf. Computer and Dispatches Security, 2017, pp. 1175 – 119
2. H. B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, " Communication-effective knowledge of deep networks from decentralized data, " in Proc. 20th Int. Conf. Artificial Intelligence and Statistics, 2017, pp. 1273 – 1282.
3. B. Hitaj, G. Ateniese, and F. Perez- Cruz, " Deep models under the GAN Information leakage from collaborative deep knowledge, " in Proc. ACM Conf. Computer and Dispatches Security, 2017, pp. 603 – 618.
4. J. H. Cheon et al., " Homomorphic encryption for calculation of approximate numbers, " in ASIACRYPT, vol. 10624, Springer, 2017, pp. 409 – 437.
5. Y. Zhao et al., " Blockchain- rested confederated knowledge for insulation- conserving and secure data collaboration, " IEEE Internet of goods Journal, vol. 8, no. 3, pp. 1817 – 1829, 2021.
6. Y. Chen, L. Su, and J. Xu, " intricate-flexible machine knowledge with statistical aggregation, " in Proc. **1 .**ACM Measurement and Modeling of Computer Systems, 2018, pp. 1 – 25.
7. D. **2 .**Yin et al., " intricate-robust distributed knowledge Towards optimal statistical rates, " in Proc. Int. Conf. Machine knowledge, 2018, pp. 5650 – 5659.
8. S. **3 .**Nakamoto, " Bitcoin A peer- to- peer electronic cash system, " 2008.(Online). Available https// bitcoin.org/bitcoin.pdf
9. X. Liu et al., " insulation- enhanced confederated knowledge against poisoning adversaries, " IEEE Deals on Information Forensics and Security, vol. 16, pp. 4574 – 4588, 2021.
10. M. Nasr, R. Shokri, and A. Houmansadr, " Comprehensive insulation analysis of deep knowledge Passive and active white- box conclusion attacks against centralized and confederated knowledge, " in Proc. IEEE Symposium on Security and insulation, 2019, pp. 739 – 753.

Characters: **1890**
Words: **303**

International Journal of Advanced Research in Education and Technology

ISSN: 2394-2975

Impact Factor: 8.152

www.ijarety.in editor.ijarety@gmail.com