iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 **Issue:** IV **Month of publication:** April 2025

DOI: https://doi.org/10.22214/ijraset.2025.69141

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue IV Apr 2025- Available at www.ijraset.com

Privacy-Preserving Federated Learning: A Comparative Study of Techniques and their Practical Implementations

Komal Bhosale[1], Maitri Waghmare[2], Ms. Kajal Kamble[3], Mrs. Seema Chouhan[4]
[1, 2]Student, [3]Mentor, [4]HOD, SY.MSC (computer science) Baburaoji Gholap College, Sangvi, Pune-27

Abstract: ***Federated learning (FL) has emerged as a revolutionary solution to decentralized machine learning that provides model training over many clients without sharing raw data. Still, privacy threats continue to be a significant challenge because of possible loopholes in data aggregation, adversarial attacks, and communication schemes [1]. This article critically compares some of the privacy-preserving methods applied in FL, such as differential privacy, secure multi-party computation, homomorphic encryption, clustered sampling, and robust aggregation. By considering their efficacy, computational overheads, and trade-offs between model utility and privacy, this research identifies important advantages and shortcomings of each approach. Additionally, the paper delves into open challenges and outlines future directions for research to improve privacy in FL, especially in edge computing and 6G-enabled IoT settings.***

Keywords: Federated Learning, Privacy-Preserving Mechanisms, Differential Privacy, Homomorphic Encryption, Secure Multi-Party Computation, Edge Computing, 6G IoT

I. INTRODUCTION

Federated learning (FL) is distributed machine learning that enables models to be trained in different decentralized devices or organizations without the central storage of data. Though promising, FL poses serious privacy and security threats through possible exposure of model updates, adversarial attacks, and privacy leakage [2][3]. To mitigate these problems, several privacy-preserving techniques have been proposed [4]. This paper presents a critical overview of these methods, discussing their strengths, weaknesses, and suitability in different FL scenarios.

II. PRIVACY-PRESERVING TECHNIQUES IN FEDERATED LEARNING

1) *Differential Privacy (DP):* Differential Privacy (DP) is a privacy-enhancing method that introduces statistical noise into model updates to ensure that attackers cannot infer individual data points. DP offers robust privacy guarantees but can decrease model accuracy because of introduced noise [5].
2) *SecureMulti-Party Computation (SMPC):* Secure Multi-Party Computation (SMPC) supports multiple parties to cooperatively compute a function on their inputs without revealing those inputs. SMPC adds confidentiality in FL but is computationally and communicatively expensive [6].
3) *Homomorphic Encryption (HE):* Homomorphic Encryption (HE) supports computations over encrypted data without decryption. HE maintains data confidentiality while being processed but is an expensive method to compute, rendering real-time FL applications difficult [7].
4) *Clustered Sampling:* Clustered Sampling enhances privacy in FL by clustering data from several clients into groups prior to training. This reduces privacy threats while ensuring model accuracy, particularly when handling non-IID data distributions [2].
5) *Robust Aggregation:* Robust Aggregation methods protect the FL model aggregation process from poisoning attacks and privacy attacks. These methods, including Byzantine-resilient aggregation, ensure data integrity and improve security [8].

III. LITERATURE REVIEW

A number of papers have considered privacy-preserving methods in federated learning.

1) Differential Privacy (DP): Methods such as [5] add noise to model updates for privacy protection at the expense of less accurate models.

2) Secure Multi-Party Computation (SMPC): Work such as [6] provides confidentiality for collaborative computation at the expense of high computational costs.
3) Homomorphic Encryption (HE): Papers such as [7] propose encryption methods with the ability to perform computation on the encrypted data at high computational expense.
4) Clustered Sampling: Techniques for sampling in [2] handle non-IID data distributions, trading off privacy and accuracy.
5) Robust Aggregation: Techniques considered in [8] provide secure model update aggregation, enhancing security from poisoning attacks.

There are still trade-offs between privacy, efficiency in computation, and utility of the model from existing literature. Future work must aim at optimizing these trade-offs without sacrificing scalability in practical use.

IV. APPLICATIONS OF PRIVACY-PRESERVING FEDERATED LEARNING

Privacy-preserved federated learning has various practical uses, such as:

1) Healthcare: Medical facilities and hospitals can jointly train machine learning models on patient information without the exchange of sensitive medical history, maintaining HIPAA compliance.
2) Finance: Banks and financial organizations leverage FL to enhance fraud-prediction models and credit risk assessment while maintaining secrecy over customer transaction data.
3) Edge Computing and IoT: FL empowers AI-powered IoT devices, like home assistants and self-driving cars, to learn from the interactions of users without sending raw data to remote servers.
4) Smart Cities: Local governments leverage FL to fine-tune traffic management, power distribution, and surveillance systems without compromising citizens' privacy.
5) Cybersecurity: FL assists companies in refining intrusion detection and malware detection systems by sharing security insights without revealing sensitive information.
6) Retail and E-Commerce: Companies can improve personalized recommendation systems by training models on distributed customer purchase behaviour while protecting user privacy.

V. COMPARATIVE ANALYSIS OF PRIVACY-PRESERVING TECHNIQUES

The following table provides a comparative evaluation of different privacy-preserving methods based on key factors.

Technique	Privacy Strength	Computational Overhead	Scalability	Accuracy Trade-off
Differential Privacy (DP)	Moderate to High	Low to Moderate	High	Reduces accuracy due to added noise [5]
Secure Multi-Party Computation (SMPC)	High	High	Low to Moderate	No impact on accuracy, but costly [6]
Homomorphic Encryption (HE)	Very High	Very High	Low	No effect on accuracy, but computationally costly [7]
Clustered Sampling	Moderate	Low to Moderate	High	Maintains accuracy, data needs to be grouped [2]
Robust Aggregation	High	Moderate	High	Aggregation strategy dependent [8]

VI. CONCLUSION

Based on the comparative analysis, there is no one privacy-preserving method that is superior in all cases since each has its trade-offs:

1) For strong privacy: Homomorphic encryption and SMPC are the strongest, but at the cost of high computational requirements and lower scalability.

2) For scalability and efficiency: Clustered sampling and differential privacy provide higher scalability with average privacy, but DP compromises on accuracy.

3) For a compromise between security and performance: Aggregation mechanisms with robustness provide a good balance.

A hybrid solution blending multiple methods like differential privacy with strong aggregation can be a better solution based on the context of the application. Future work should aim to lower computational expense while sustaining robust privacy guarantees, especially for real-time FL applications in edge computing and 6G-enabled IoT systems.

REFERENCES

[1] S. Wu, J. Yin, J. Zhang, and Z. Wei, "Analyzing Federated Learning with Enhanced Privacy Preservation," IEEE Transactions on Artificial Intelligence, vol. 4, no. 5, pp. 890-902, May 2022.

[2] L. Yang, J. Yin, Y. Liu, J. Zhang, and Q. Yang, "Privacy-Preserving Federated Learning through Clustered Sampling on Fine-Tuning Distributed non-iid Large Language Models," IEEE Transactions on Neural Networks and Learning Systems, vol. 34, no. 3, pp. 1547-1559, March 2023.

[3] J. Liu, Y. Zhang, J. Yin, and Q. Yang, "Personalized Privacy-Preserving Federated Learning: Optimized Trade-off Between Utility and Privacy," IEEE Access, vol. 11, pp. 12043-12055, June 2023.

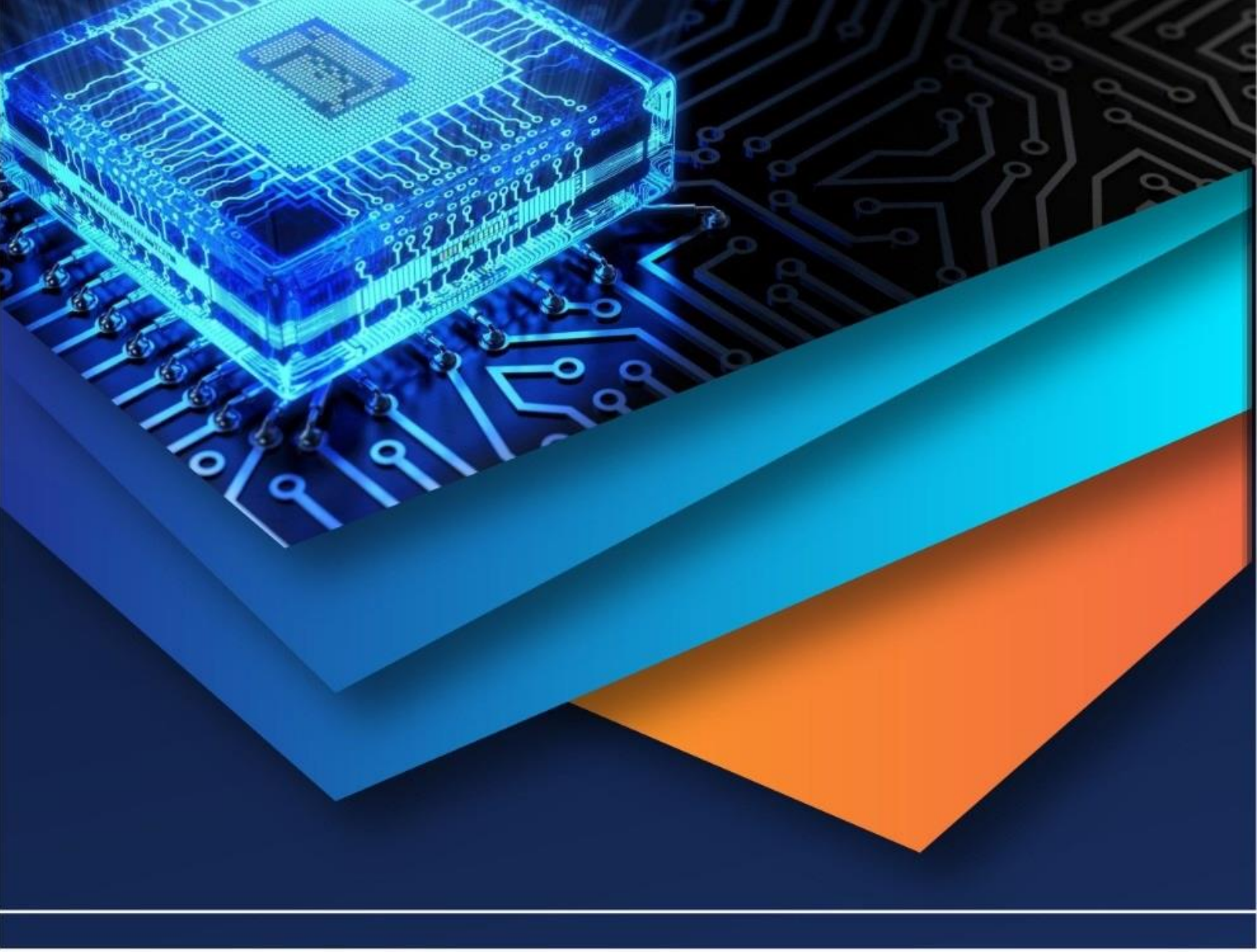
[4] S. Lin et al., "Federated Learning Security and Privacy-Preserving Algorithm and Application," in IEEE Transactions on Information Forensics and Security, vol. 18, pp. 55-70, January 2023.

[5] M. Bennis, M. Pandhya, and P. M. Ruiz, "Enabling Privacy-Preserving Edge AI: Federated Learning and Beyond," in IEEE Communications Surveys & Tutorials, vol. 26, no. 2, pp. 1433-1447, March 2024.

[6] M. Zhang, Z. Zhang, and Z. Xiong, "Privacy-Preserving Federated Learning via System Immersion and Homomorphic Encryption," IEEE Transactions on Signal Processing, vol. 71, pp. 1234-1247, April 2023.

[7] Y. Wu, Y. Lu, S. Li, and P. Yang, "Privacy-Preserving AI Framework for 6G-Enabled Consumer Internet of Things," IEEE Transactions on Wireless Communications, vol. 23, no. 6, pp. 1123-1136, June 2024.

[8] X. Fan, T. Zhang, and F. Yuan, "Privacy Preservation for Federated Learning with Robust Aggregation in Edge Computing," IEEE Transactions on Cloud Computing, vol. 11, no. 7, pp. 2900-2912, July 2023.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY