

# Research on Privacy and Security Issues in Federated Learning

Xinyuan Bi<sup>a</sup>

*International School, Beijing University of Posts and Telecommunications, Beijing, 100876, China*

**Keywords:** Federated Learning, Privacy Security, Encryption Technology, Security Audit.

**Abstract:** In the digital age, data privacy and security have become key issues. Federated learning, as an emerging distributed machine learning technology, has been widely applied in fields such as finance and healthcare. However, federated learning still faces many challenges in privacy protection. This paper deeply studies the privacy and security issues of federated learning analyzes its technical principles, privacy protection mechanisms, and performance in practical applications. Through the analysis of application cases in fields such as finance and healthcare, it explores the advantages and disadvantages of federated learning in privacy protection. On this basis, this paper proposes improvement strategies such as optimizing encryption technology, strengthening model security, establishing a security audit mechanism, and improving laws and regulations, aiming to enhance the privacy and security level of federated learning and provide guarantees for its stable application in various fields. In the future, the research on privacy and security of federated learning will develop towards more intelligent, efficient, and integrated directions. It is necessary to further study new privacy protection technologies, strengthen dynamic security monitoring and adaptive defense capabilities, and formulate unified privacy and security standards and norms to promote the safe application and development of federated learning technology worldwide.


## 1 INTRODUCTION

At a time when big data and artificial intelligence are booming, data has become a core resource driving innovation and development in various fields. The traditional centralized data processing model faces the risks of data leakage and misuse in the process of data collection, storage, and use, seriously threatening user privacy and security. At the same time, the phenomenon of data silos between different institutions and organizations hinders data circulation and sharing, limiting the development of AI technology.

As an innovative distributed machine learning paradigm, Federated Learning takes “data does not move, model moves” as its core idea so that participants can jointly train high-precision models without sharing original data and only exchange model parameters or intermediate results, effectively solving the data silo problem and reducing the risk of data leakage. The privacy and security research of federated learning is of great significance, which is mainly reflected in the protection of user privacy, the

promotion of data circulation and sharing, the enhancement of model security and reliability, and the promotion of its wide application, so as to provide support for the development of innovation in various fields (Li, Zhou, 2024; Xiao et al., 2024).

Currently, the research on the privacy and security of federated learning has made some progress. In terms of privacy protection techniques, scholars propose a variety of methods. Differential privacy achieves privacy protection by adding noise to the data; homomorphic encryption allows computation over ciphertexts; secure multiparty computation ensures data privacy security for all parties by collaborating with multiple parties to accomplish computational tasks. Li et al. proposed a new privacy-preserving framework to ensure that data is always encrypted during transmission and computation, effectively preventing data leakage and providing end-to-end privacy protection from data generation to model update. In terms of computational efficiency, experiments show that the framework only slightly decreases the model accuracy (e.g., the accuracy of the MNIST dataset decreases from 98.2% to 98.0%) after the introduction of fully homomorphic

<sup>a</sup> <https://orcid.org/0009-0005-0284-640X>

encryption, which maintains a high model performance while protecting privacy. In terms of communication overhead, when the data batch size is 64, the communication overhead is only about 10% higher than that of unencrypted nodes, balancing privacy protection and communication efficiency (Li, Zhou, 2024). Taking the medical field as an example, due to its data sensitivity, there is a higher requirement for privacy and security. Researchers propose privacy protection schemes such as blockchain-based federated learning frameworks, which utilize blockchain's immutability and traceability to enhance data security and trustworthiness. However, in practical applications, balancing privacy protection and model performance to ensure secure and efficient data collaboration among different healthcare organizations is still a problem to be solved (Wang, 2024). On the industrial side, a new paradigm of industrial federated learning based on layered cross-domain architecture is proposed, driven by 6G technology. The federated learning algorithm based on the end-side cloud three-layer federated learning architecture proposed by Chen Zhu et al. can significantly reduce the latency and energy consumption of the federated learning model with faster convergence of the model in the cloud while guaranteeing the testing accuracy of the model in the cloud (Liu et al., 2024; Chen et al., 2024). However, these techniques have challenges in practical applications, such as differential privacy affecting model accuracy, high computational overhead of homomorphic encryption, and high complexity of secure multi-party computational communication.

This paper systematically studies the privacy security problem of federated learning, analyzes its theoretical foundation, architecture, and privacy protection technology, and discusses the privacy security guarantee mechanism and effect by combining it with the application cases in finance, medical, and other fields. The article deeply analyzes the existing privacy security challenges, puts forward targeted improvement strategies, and looks forward to the future development trend, aiming to provide theoretical support and practical guidance for the privacy security protection of Federated Learning and help its healthy and sustainable development.

## 2 THEORETICAL FOUNDATIONS OF FEDERAL LEARNING

### 2.1 Federated Learning Architecture

There are three main architectures for federation learning: horizontal federation learning, vertical federation learning, and federation migration learning. Horizontal federation learning is suitable for scenarios where the data characteristics of the participants are similar, but the samples are different, for example, banks in different regions can jointly train credit assessment models through horizontal federation learning. Vertical federation learning is suitable for situations where the samples are similar, but the characteristics are different, such as banks and e-commerce platforms can use this to realize data collaboration. Federated Migration Learning is used to solve the model training problem when the data distribution difference is large, when the source and target domain data do not meet the conditions of independent and same distribution, the source domain data can be used to improve the performance of the target domain model.

### 2.2 Principle of Operation

The basic workflow of federated learning is that each participant trains the model locally using its own data and calculates the gradient or parameter update values of the model. Then, these update values are uploaded to the central server or other coordinating nodes by encryption and other secure methods. The coordination node aggregates the received update values, such as using a weighted average method, and determines the weights based on factors such as the amount of data of each participant to obtain the global model update. Finally, the updated global model is sent down to each participant, who uses the updated model to continue training on local data and so on iteratively until the model converges.

### 2.3 Privacy Protection Technology

Homomorphic encryption is a commonly used encryption technique for federated learning, which allows specific computations to be performed on the ciphertext, and the results of the computations are decrypted to be consistent with the same computations in plaintext. In the process of updating and uploading model parameters, participants encrypt the gradient values with homomorphic encryption, and the server aggregates the computation in the

ciphertext state, which guarantees privacy and security during data transmission and computation. However, homomorphic encryption has a large computational overhead, which affects the efficiency of federated learning (Li, Zhou, 2024).

Differential privacy achieves privacy protection by adding noise to the data. In federated learning, participants upload model update values before adding noise to the gradient or parameters that match a specific distribution. Noise addition makes it difficult for an attacker to infer the original data from the output, protecting data privacy. However, noise introduction affects model accuracy and requires a trade-off between privacy protection and model performance. The FLFilter scheme proposed by Xiao Di et al. uses local differential privacy to protect customer privacy and differentiate normal and malicious customer behaviors and designs clustering and filtering methods for backdoor attack characteristics. The proposed cosine gradient clustering index breaks the barrier between model perturbation and backdoor model identification. Through theoretical analysis and experimental simulation, it is confirmed that FLFilter achieves the expected goals in terms of accuracy, robustness, and privacy (Xiao et al., 2024).

Secure multi-party computation allows multiple participants to jointly compute an objective function without revealing their respective data. In federated learning, each participant computes the gradient, loss function, etc., of the model through the secure multi-party computation protocol. For example, the secure data interaction is realized by using an unobtrusive transmission protocol, which ensures that the participants can only access the information required for the computation and cannot access the raw data of other parties. However, the communication complexity of secure multi-party computation is high and requires high network bandwidth and computational resources.

### **3 CASE STUDIES ON THE APPLICATION OF FEDERAL LEARNING IN DIFFERENT FIELDS**

#### **3.1 Financial Sector**

##### **3.1.1 Application Scenarios**

In the field of financial risk control, different banks can jointly train credit assessment models through

federated learning. Take Bank A and Bank B as an example; they cannot share data directly due to data privacy and competition, but they can use their own data to train their credit assessment models locally and upload the gradient or parameter update values of the models encrypted to the federated learning platform. The platform aggregates the updated values and sends them down to the banks for further training. For example, after clarifying the best fit between federated learning and financial business, the Tencent security team gave full play to its technological effectiveness to promote agile business innovation on the industry side, screening and federating more than 200 business indicators to achieve intelligent credit card management for a commercial bank (Zhang, 2024).

##### **3.1.2 Privacy and Security Mechanisms**

This process uses homomorphic encryption to encrypt the uploaded model update values to ensure secure data transmission. To prevent model inversion attacks, differential privacy is introduced in the model training process to add an appropriate amount of noise to the gradient values. Meanwhile, a secure multi-party computation protocol ensures that each bank computes intermediate results, such as model gradients, without disclosing their respective raw data. For example, Qiuxian Li et al. proposed a new privacy protection framework in terms of privacy protection capability, through full homomorphic encryption, the model parameters complete the computation and update in the encrypted state to ensure that the data is always encrypted during transmission and computation, effectively preventing data leakage and providing end-to-end privacy protection from data generation to model update (Li, Zhou, 2024).

##### **3.1.3 Application Effects**

With federated learning, banks are able to integrate data from multiple sources to improve the accuracy of their credit assessment models. The credit assessment model, after adopting federated learning, has improved the accuracy of identifying defaulted customers compared to the model trained on single bank data. Meanwhile, due to the effective implementation of privacy and security mechanisms, the collaborative utilization of data was achieved without any data leakage incidents while safeguarding data privacy. For example, the asynchronous federated learning technology of Jingdong Digital Technology has been implemented in financial scenarios to build a big data risk control

model with partner institutions, which improves the generalization effect of the model and makes the data stored locally participate in the training of the overall model at the same time (Zhang, 2024).

## 3.2 Medical Field

### 3.2.1 Privacy Security Guarantee Mechanism

To protect patient privacy, healthcare organizations use a blockchain-based federated learning framework to ensure data security and trustworthiness using blockchain's immutability and traceability. During data transmission, data is encrypted using homomorphic encryption. For the data source inference attacks that medical data may face, a ring signature-based anti-source inference attack scheme is used to protect the identity privacy of patients. Meanwhile, corresponding detection and defense mechanisms are proposed to prevent double-ended poisoning attacks. Wenshuo Wang discusses the application of federated learning in the field of smart healthcare and proposes a series of solutions for the data privacy protection problems therein, which effectively solves the security and privacy problems faced by federated learning in smart healthcare through the three frameworks of FRESH, ABPFL-SSH, and PFHE, and provides technical support for the development of smart healthcare (Wang, 2024).

### 3.2.2 Application Effect

FRESH framework: the effectiveness of its signature consumption time and batch verification strategy under different public key set capacities is verified through experiments. In terms of signature time consumption, when the public key set capacity is 100, a single signature takes 1.5 seconds; when the capacity is 200, it takes less than 3 seconds. In terms of batch verification, when the capacity of the public key set is 100 and the number of clients is 50, the batch verification takes 0.5 seconds, which saves 80% of the time compared to the one-by-one verification (Wang, 2024).

ABPFL-SSH framework: the relationship between server-side poisoning volume and model accuracy and the performance of client-side poisoning screening algorithms are experimentally verified. In terms of server-side poisoning amount and model accuracy, when the poisoning amount is lower than the threshold value of 0.5, the model testing accuracy is more than 90%; when it exceeds 0.5, the accuracy drops to less than 20%.

In terms of the performance of the client-side poisoning screening algorithm, in 100 client systems, the testing accuracy was 92%, the false positive rate was 11%, and the false negative rate was 0% (Wang, 2024).

PFHE framework: computational cost, model prediction accuracy, and communication cost were evaluated experimentally. In terms of computational cost, client-side encoding and encryption took 0.03 seconds, and decoding and decoding took 0.02 seconds. The total time consumed by the server-side ciphertext operation does not exceed 0.5 seconds. The model prediction accuracy is similar to the original model accuracy, with an average difference of no more than 2%. For example, the original model on the LBW dataset is 95%, and the PFHE model is 93%; the original model on the Nhanes III dataset is 88%, and the PFHE model is 86%. In terms of communication cost, the client uploads data size of 3.125MB, 3.125MB per minute at the highest frequency, and the network speed requires 437Kbps (Wang, 2024).

## 4 EXISTING ISSUES IN FEDERAL LEARNING PRIVACY AND SECURITY

### 4.1 Data Leakage Risks

An attacker can utilize the output of a federated learning model to reverse the derivation to try to restore the original training data. In the federated learning scenario of image recognition, an attacker observes changes in the model output by adjusting the input data, potentially reconstructing the original training image and leading to data leakage.

During federated learning, the gradient information uploaded by the participants may contain some features of the original data. After obtaining this gradient information, the attacker analyzes the gradient change trends and numerical features and may infer sensitive information from the original data.

### 4.2 Malicious Attack Threats

Malicious participants may intentionally inject malicious data during model training to change the direction of model training and degrade model performance. In federated learning of medical diagnostic models, malicious parties uploading incorrectly labeled case data can lead to biased



disease diagnosis in the trained model (Manzoor, 2024)..

An attacker may steal the global model in federated learning or the local model of a participant through means such as network attacks. After obtaining the model, the attacker can further analyze the model structure and parameters, speculate the information related to the original data, or use the stolen model for illegal activities (Manzoor, 2024).

### 4.3 Limitations of Privacy Protection Techniques

Although homomorphic encryption can effectively protect data privacy, the computational overhead is large, leading to a significant increase in federated learning training time. In large-scale data and complex model training scenarios, the computational burden of homomorphic encryption may make it difficult for federated learning to run in real time. The emergence of federated learning breaks the phenomenon of data silos and solves some data privacy problems at the same time, but as the types of participants increase and the attackers' attacks become more and more sophisticated, federated learning faces increasing privacy leakage problems. Although there have been a number of privacy protection studies on federated learning, there are still many unsolved problems due to limited resistance to attacks, single application scenarios, and huge communication overheads. The future privacy protection of federated learning remains a more permanent challenge (Wang, Yi, Zhang, 2024).

Differential privacy protects privacy by adding noise, but noise addition affects model accuracy. In practice, it is difficult to determine the appropriate noise intensity that balances privacy protection requirements and model performance.

Secure multi-party computation requires a large number of communication interactions between the participants with high communication complexity. In unstable network environments or limited bandwidth, the efficiency of secure multi-party computation can be seriously affected, even leading to the interruption of the federated learning process.

### 4.4 Legal and Regulatory Issues

Currently, there are fewer privacy protection norms for federated learning, and there is a lack of clear standards and guidelines. In practical application, the responsibilities and obligations of each participant for data privacy protection are not clearly defined, and it is easy to see that privacy protection measures are not

in place. Author Liu Zegang reveals the legal defects of the existing privacy protection path and puts forward the corresponding improvement direction and suggestions by analyzing the problems of federal learning in terms of legal norms, responsibility implementation, protection of personality rights and interests, and technical trade-offs. The "loose" and "joint" learning process of federal learning makes attribution of responsibility difficult. In cross-organizational federated learning, the server controller may not be more responsible for privacy protection, and it is difficult to determine the legal nature of each participant, so it is difficult to allocate and pursue responsibility from the perspective of personal data law (Liu, 2025).

## 5 FEDERAL LEARNING PRIVACY AND SECURITY IMPROVEMENT STRATEGIES

### 5.1 Optimization of Encryption Technology

Research and adopt new encryption algorithms, such as lattice-based encryption algorithms, which have higher security and efficiency. Lattice-based encryption algorithms have advantages in resisting quantum computing attacks and, at the same time, reduce the computational complexity relative to traditional algorithms such as homomorphic encryption, which can improve the computational efficiency of Federated Learning under the premise of safeguarding data privacy.

Combine encryption with other privacy protection techniques, such as combining homomorphic encryption with differential privacy. Before the data is uploaded, the noise is added using differential privacy, and then homomorphic encryption is performed, which reduces the amount of encryption computation and enhances the effect of privacy protection. As proposed by Chaudhury D S, a federated learning framework SBTLF combining blockchain, local differential privacy, and incentives is designed to solve the problems of privacy preservation, data sharing incentives, and secure sharing of model parameters in federated learning. Experimental results show that the framework is able to maintain high model performance while protecting privacy and promoting active client participation through incentive mechanisms. In addition, the SBTLF framework has good scalability and security

for large-scale distributed machine-learning scenarios (Chaudhury, 2024).

## 5.2 Model Security Reinforcement

Aiming at model inversion attacks and poisoning attacks, design corresponding defense mechanisms. Adopt model watermarking technology to embed specific identification information in the model so that when the model is stolen or maliciously tampered with, anomalies can be found by detecting the watermark. At the same time, a verification mechanism for model updates is established to verify the model updates uploaded by the participants to ensure their authenticity and legitimacy.

Optimize the model structure of federated learning to improve the robustness of the model. Adopt a decentralized model architecture to reduce the dependence on a single server and reduce the risk of privacy security due to server attacks. At the same time, reduce the amount of data transmitted by the model through techniques such as model compression to reduce the risk of data leakage. In the study of machine learning models for Alzheimer's disease detection, by simulating membership inference attacks, it is found that the FL model using SecAgg can effectively protect client data privacy, and the attacking model cannot determine which data samples have been used to train client-specific models. This shows that SecAgg has significant advantages in privacy protection and can reduce the risk of information leakage (Mitrovska et al., 2024).

## 5.3 Establishment of a Security Audit Mechanism

Establish a strict data use monitoring mechanism to monitor the use of data in the federal learning process in real-time. Record operations such as data access, transmission, and processing to ensure that the use of data complies with privacy protection regulations. Once abnormal data operations are found, provide timely warning and processing. Such as dynamic adaptive defense technology: develop defense technology that can monitor and adapt to changes in the system state in real time to cope with changing threats (Chen et al., 2024).

Audit the behavior of the participants to verify the authenticity and legitimacy of the participant's identity. Record the operation records of the participants through technologies such as blockchain to ensure that the participants follow the rules in the federal learning process and prevent the attack behavior of malicious participants.

## 5.4 Legal and Regulatory Improvements

Formulate privacy protection regulations specifically for federal learning and clarify the data privacy protection responsibilities and obligations of each participant. Specify the privacy protection standards for all aspects of data collection, storage, use, and transmission to provide a legal basis for the privacy security of federated learning.

Establish a specialized regulatory agency to supervise and manage the application of federal learning. The regulatory body is responsible for reviewing the privacy and security program of the federal learning program and imposing penalties for violations to ensure that federal learning operates in a legal and secure environment.

# 6 CONCLUSION

Federated learning, as an emerging distributed machine learning technology, has significant advantages in solving the data silo problem and protecting data privacy. In this paper, through the application case analysis in finance, medical, and other fields, it is verified that it can realize the collaborative use of data and improve the model performance under the premise of privacy security. However, Federated Learning still faces challenges in privacy security, such as data leakage risks, malicious attack threats, limitations of privacy protection technologies, and legal and regulatory issues. To address these issues, this paper proposes improvement strategies such as optimizing encryption technology, reinforcing model security, establishing a security audit mechanism, and improving legal and regulatory issues. Through the implementation of these strategies, the privacy security level of federated learning can be effectively improved to provide a guarantee for its wide application in various fields.

With the continuous development of technology, the privacy security research of federated learning will develop in the direction of more intelligent, efficient, and integrated. In the future, further research on new privacy protection techniques, such as quantum cryptography-based encryption, is needed to cope with increasingly complex security threats. At the same time, research on the dynamic security monitoring and adaptive defense capabilities of the FCS should be strengthened so that it can adjust its security strategy promptly in the face of ever-changing means of attack. In addition, it is also

necessary to strengthen international cooperation and exchanges to jointly develop unified privacy security standards and specifications for federated learning and to promote the secure application and development of federated learning technologies on a global scale.

## REFERENCES

- Chen, C., Liu, J., Tan, H., et al., 2024. Trustworthy federated learning: privacy, security, and beyond. *Knowledge and Information Systems, preublish*, pp. 1-36.
- Chaudhury, D.S., Morreddigari, R.L., Varun, M., et al., 2024. Blockchain based secure federated learning with local differential privacy and incentivization. *IEEE Transactions on Privacy*, pp. 131-144.
- Liu, Z., 2025. The limitations and solutions of privacy protection in federated learning in the era of artificial intelligence. *Zhongwai Faxiang*, 1-20.
- Liu, M., Xia, Y., Zhao, H., et al., 2024. Federated learning for 6G industrial Internet of Things: from demands, visions to challenges and opportunities. *Journal of Electronics and Information Technology*, 46(12), pp. 4335-4353.
- Li, Q. & Zhou, Q., 2024. Research on federated learning privacy protection technology based on fully homomorphic encryption. *Modern Information Technology*, 8(23), pp. 170-174.
- Manzoor, U.H., Shabbir, A., Chen, A., et al., 2024. A survey of security strategies in federated learning: defending models, data, and privacy. *Future Internet*, 16(10), pp. 374-374.
- Mitrovska, A., Safari, P., Ritter, K., et al., 2024. Secure federated learning for Alzheimer's disease detection. *Frontiers in Aging Neuroscience*, pp. 161324032-1324032.
- Ruan, M., Zhang, S., Xue, K., et al., 2024. Federated learning algorithm design and optimization in edge computing networks. *Mobile Communication*, 48(12), pp. 122-128.
- Wang, W., 2024. Research on key security technologies of federated learning for smart healthcare. *Master's thesis. Yantai University*.
- Wang, B., Yin, X., Zhang, L., 2024. Research on privacy protection based on horizontal federated learning. *Computer Technology and Development*, 34(10), pp. 1-7.
- Xiao, D., Yu, Z., Li, M., et al., 2024. A secure federated learning scheme based on differential privacy and model clustering. *Computer Engineering and Science*, 46(09), pp. 1606-1615.
- Zhang, Z., 2024. Research on data privacy protection based on federated learning: An analysis based on examples from WeBank, Ping An Technology, and others. *Frontiers of Foreign Social Sciences*, (05), pp. 86-99.