Research on the Application of Security-Weighted Aggregation in Federated Learning for Financial Risk Control

Xiaowanqing Sun

School of Cyberspace Science and Technology, Northwestern Polytechnical University, Xi’an, 710129, China
sunxwq@mail.nwpu.edu.cn

Abstract:

Federated learning is a privacy-preserving distributed machine learning approach that enables collaborative training by keeping data local and sharing only model parameters. This effectively addresses the challenges of “data silos” and privacy regulations in the field of financial risk control. This paper systematically elaborates on the concepts, development history, architecture, and core mechanisms of federated learning, with a focus on analyzing the principles and implementation of security-weighted aggregation technology, including its reliance on key technologies such as homomorphic encryption, differential privacy, and secure multi-party computation. Besides, this study examines how this technology applies to fintech, outlines common practices, and highlights challenges in technical performance, cross-platform collaboration, and compliance, with possible solutions. The results show that federated learning, using secure aggregation, enables the financial industry to share data and conduct joint risk modeling while staying compliant.

Keywords: Federated Learning; Security-Weighted Aggregation; Financial Risk Control; Privacy Preservation; Differential Privacy

1. Introduction

With the development of fintech, financial data is highly sensitive and distributed across different institutions. Thus, traditional centralized modeling faces legal, compliance, and security constraints in cross-institutional sharing. To address this, federated learning helps by allowing multiple parties to build models together. And each party trains on its own data and only shares model parameters, keeping the actual data private. Even though algorithms have improved, and methods for handling different types of data and secure aggregation exist, challenges remain in using federated learning for highly sensitive financial situations. Issues include differences in data across institutions, weak aggregation methods, and low training efficiency. Therefore, this paper explores federated learning architectures and aggregation methods that give more weight to security in financial settings, so as to improve model performance and data security while considering compliance. In this process, it investigates how to achieve robust global model training under non-IID data and heterogeneous environments, how to design security-weighted aggregation strategies to balance data contribution, model accuracy, and privacy protection, and how to manage computational efficiency alongside compliance needs in financial operations. To this end, this paper combines systematic literature review, architectural analysis, and technical principle analysis to outline the core mechanisms of federated learning in fintech, analyzes its application potential and challenges in combination with actual business scenarios, and evaluates the trade-offs between different technology combinations in terms of data privacy, model accuracy, and computational efficiency, thereby providing references for future secure collaborative modeling in highly sensitive financial environments.

2. Fundamentals of Federated Learning and Fintech Architecture

2.1 Concepts and Development History of Federated Learning

Federated learning enables cross-institutional collaborative modeling by training models locally and uploading only parameters or gradients, protecting the privacy of raw data. In financial risk control, this method enables institutions to develop risk models without pooling their data, supporting later methods like security-weighted aggregation [1]. In terms of development history, federated learning has evolved from the proposal of basic algorithms, exploration of asynchronous and hierarchical update methods, to optimization for non-IID data processing and system heterogeneity, and further to personalized modeling and fairness improvements. In recent years, studies on making models lighter and fairer, like FedLab and FairFed, have offered practical privacy-protecting methods for risk control models in sensitive financial situations, paving the way for using security-weighted aggregation in financial risk management [2].

2.2 Federated Learning Architecture and Core Mechanisms

Federated learning is a distributed algorithmic framework that shares intermediate statistical results during computation without leaking raw data, achieving privacy preservation. Federated learning generally has two architectures: client/server mode and decentralized mode. The client/server mode is generally suitable for global model parameter prediction and conducting various statistical tests, where a central node coordinates distributed computing across participating nodes. In this mode, the core architecture of federated learning is a distributed system coordinated by a central server with multiple clients participating. Its core mechanisms revolve around the central challenge of how to collaboratively train a model without centralizing data. By using methods such as cyclic iteration, reducing communication, protecting privacy, and handling non-IID data well, it is possible to train a global model efficiently while also keeping user data safe. Innovative structured update and sketch update mechanisms significantly optimize the critical issue of communication efficiency [3]. The decentralized mode uses various distributed algorithms, has no central node, and involves clients continuously exchanging locally computed intermediate results to obtain reliable global results [4].

2.3 Characteristics and Requirements of Fintech Scenarios

Data in financial risk control is highly sensitive, distributed across different institutions, forming “data silos,” and subject to strict privacy regulations. Traditional centralized modeling is hard to use because sharing raw data between institutions can create legal and security risks [5]. Federated learning allows institutions to collaboratively train models under the premise that data does not leave the local domain, enabling knowledge sharing while accommodating data isolation and access control. This gives a workable approach to collaborative risk modeling and supports the use of security-weighted aggregation. Therefore, financial risk control systems impose strict requirements on model security, accuracy, and efficiency. In terms of security and compliance, model parameters must prevent information leakage during exchange and aggregation, while meeting regulatory, auditing, and internal compliance requirements. In terms of accuracy, models need to improve the accuracy of risk control tasks such as anti-fraud, credit assessment, and default prediction, while addressing issues like varying data volumes among institutions, uneven feature distributions, and non-IID data. For efficiency, training and inference need to balance speed and computing costs to enable quick decisions and continuous risk monitoring in financial operations. These requirements will form the core considerations for the design of security-weighted aggregation technology and its adaptation to financial scenarios in subsequent sections.

3. Principles and Implementation of Security-Weighted Aggregation Technology

3.1 Aggregation Theory and Method Design

Aggregation theory is not merely simple data summation, but an interdisciplinary field involving cryptography, distributed systems, information theory, and game theory. Its core goal is to enable participants in a mutually distrustful distributed environment to collaboratively compute meaningful aggregation results through carefully designed protocols, while protecting each participant's data privacy as much as possible and ensuring system reliability even when some nodes fail [6]. Security-weighted aggregation is a core mechanism of federated learning, aiming to integrate local model updates from various clients without disclosing raw data. Each client's update is assigned a weight during aggregation, typically based on data volume, data quality, or model performance. Through weighting strategies, the aggregation process ensures both the effectiveness and fairness of the global model. In cases of uneven or heterogeneous client data distributions, it allows high-value or high-reliability information to exert a greater influence on the global model, enhancing model robustness and generalization. In terms of security design, aggregation must prevent the server or other participants from accessing individual clients' raw updates, requiring parameter-level security isolation. Each local update participates in the global model construction only in weighted statistical or encrypted form. Besides, the aggregation process can dynamically adjust weights in response to client data characteristics and risk value, focusing particularly on anomalous samples or high-risk transactions. This mechanism ensures the stability and reliability of the global model when dealing with non-IID data and anomalous samples, providing quantitative performance guarantees for credit assessment, anti-fraud, and default prediction in financial risk control.

3.2 Core Technologies and Security Assurance

Security-weighted aggregation uses Homomorphic Encryption (HE), Differential Privacy (DP), and Secure Multi-Party Computation (MPC) to keep model parameters isolated. HE lets addition and multiplication be done directly on encrypted data, so aggregation can happen without decrypting it, though this adds heavy computational and communication costs. DP reduces the influence of any single data sample on the global model by adding noise to gradients or parameters. The level of privacy protection can be adjusted through the privacy budget, which creates a trade-off with model accuracy [7,8]. MPC processes local updates through distributed protocols involving secret sharing and encrypted computation, achieving aggregation without single-point leakage, while resisting malicious behavior or inference attacks from some nodes. Combined with the weighting mechanism, aggregation can dynamically adjust the contribution of each client's update to the global model, maintaining model stability and robustness under non-IID data or anomalous samples. These technological mechanisms provide parameter-level security assurance, offering theoretical support for federated learning in highly sensitive scenarios [9]. Federated learning applies DP to add noise at the data source, so individual samples cannot be traced. It uses HE or MPC to aggregate models on encrypted or secret-shared data, stopping the server from seeing any participant's raw updates. Besides, it combines weighting methods to adjust each party's contribution, helping defend against attacks and problems caused by differences in data. These technologies create a layered defense, providing end-to-end security for model parameters from local data to aggregation, while keeping the model usable.

3.3 Technology Adaptation in Financial Environments

Implementing security-weighted aggregation in financial risk control requires balancing compliance, security, and efficiency. To follow strict regulations, solutions must

ISSN 2959-6130

keep data local, make the whole process auditable, and comply with laws like the "Personal Information Protection Law," allowing traceable joint modeling [10]. For cross-institutional collaboration, MPC protocols are preferred to aggregate the global model without exposing clients' raw data. In scenarios with very stringent regulatory requirements, such as cross-border risk control, HE can be supplemented to enhance security. For internal situations with less sensitive data, DP can be used for additional protection. Each method has its own pros and cons in terms of security, efficiency, and deployment, so the optimal choice should fit business requirements. To improve model performance, optimization is needed for financial data differences and real-time needs. Risk-sensitive weighting can give more influence to important or unusual samples in the global model. At the same time, using gradient compression, communication improvements, and handling non-IID data helps reduce delays and model bias. A hybrid architecture with MPC at its core, deeply integrated with intelligent weighting and performance optimization, can balance model accuracy and computational efficiency while ensuring security and compliance, meeting the implementation needs of risk control tasks in actual business operations.

4. Application Practices and Challenges of Federated Learning in Fintech

4.1 Typical Financial Application Scenarios

Federated learning has progressively moved from proof-of-concept to scaled application in fintech, with its core value lying in unleashing the value of multi-party data without sharing raw data [11]. Financial institutions achieve business enhancements in key areas such as risk control, marketing, and customer operations through joint modeling. Risk control is the most widely applied area. Banks collaborate with telecom operators, payment platforms, and internet service providers to integrate multi-source data for risk analysis, training more accurate models for credit default prediction, anti-fraud, and anti-money laundering. This improves risk identification capabilities and mitigates information asymmetry in financing for small and micro enterprises. In intelligent marketing, financial institutions can integrate external data without it leaving the local domain to build customer profiles and achieve personalized product recommendations and cross-selling. This approach can increase marketing conversion rates and improve customer experience, delivering direct business value for banks and partners. In customer insight and operations, financial institutions combine data from multiple sources to analyze customer segments and behavior. This helps them understand customer needs, offer personalized products and services, and improve customer loyalty and overall value.

4.2 Technical Difficulties and Compliance Risks

In the deployment phase for financial services, federated learning faces core challenges in both technical operation and compliance management [12]. Technically, significant pressures exist in the actual operation of model training and inference. Multi-party encrypted computation and frequent communication increase computational and network overhead, making it difficult to meet the requirements of high-frequency businesses like real-time risk control. Different institutional platforms adopt their own protocols and standards, leading to difficulties in cross-platform collaboration and increased costs. Simultaneously, the system must maintain model accuracy and stability while ensuring security, and cope with malicious behavior or inference attacks from some nodes. In terms of compliance, processing personal sensitive information still requires obtaining user informed consent and ensuring the entire process is auditable. In multi-party collaborations, the lack of unified standards for data ownership and benefit distribution necessitates the design of fair incentive mechanisms to quantify each party's contribution. Regulatory requirements demand that the joint modeling process possesses non-repudiation and traceability, often relying on technologies like blockchain for trusted record-keeping. The aforementioned technical and compliance challenges directly impact the deployment efficiency, stability, and sustainability of federated learning in financial services and must be addressed systematically in platform design and business processes.

4.3 Impact of Challenges and Response Strategies

The deployment of federated learning in financial services still faces challenges such as high barriers to entry, ecosystem fragmentation, and limitations in high-frequency applications, requiring systematic strategies to ensure stable implementation [12]. In terms of technical strategies, promoting platform standardization and interoperability by establishing unified communication, algorithm, and security interface specifications can reduce cross-institutional collaboration costs. Simultaneously, optimizing computational and communication performance enhances the efficiency of model training and inference, enabling the system to meet the needs of real-time risk control and high-frequency services. In terms of governance and compliance strategies, building full-process security auditing and trusted record-keeping mechanisms ensures the joint modeling process is traceable, verifiable, and compliant with regulations. Furthermore, designing fair incentive mechanisms based on contribution levels to quantify the data and computational contributions of each institution can stimulate cooperation and form a stable collaborative ecosystem. For application strategies, a “pilot first, gradual expansion” approach can be used. Pilots should start in areas with high data value and moderate real-time needs, like risk control and marketing, before moving to other parts of the business. In addition, exploring cross-industry data collaboration and applying “Privacy by Design” principles in system architecture can support long-term sustainable development and business stability.

5. Conclusion

Federated learning enables cross-institutional collaborative modeling by training models locally and sharing only parameters or gradients, providing an effective approach for privacy preservation and knowledge sharing of financial data. Combined with security-weighted aggregation technology, it can maintain the robustness, accuracy, and fairness of the global model in non-IID data and heterogeneous environments, while guarding against potential malicious attacks and information leakage. In fintech scenarios, although federated learning has demonstrated application potential in areas such as risk control, anti-fraud, credit assessment, intelligent marketing, and customer insight, it still faces challenges like heavy computational costs, slow communication, difficulties in working across platforms, and compliance limitations. Through platform standardization, communication optimization, contribution-based incentive mechanisms, and Privacy by Design strategies, these issues can be effectively mitigated, thereby enhancing system deployment capability and business adaptability. Overall, federated learning and security-weighted aggregation technology provide a feasible path for financial institutions to fully leverage the value of multi-party data while ensuring data privacy and compliance. As the technology matures and standards improve, its application prospects in highly sensitive financial scenarios are broad, promising to drive continuous innovation in financial risk control, intelligent marketing, and customer operations.

References

[1] Nori, M.K., Yun, S., & Kim, I.-M. (2021). Fast federated learning by balancing communication trade-offs. IEEE Transactions on Communications, 69(8), 5168–5182.

[2] Chaudhary, R. K., Kumar, R., & Saxena, N. (2024). A systematic review on federated learning system: A new paradigm to machine learning. Knowledge and Information Systems, 67(2), 1811–1914.

[3] Konečný J, McMahan H B, Yu F X, et al. Federated learning: Strategies for improving communication efficiency[J]. arXiv preprint arXiv:1610.05492, 2016.

[4] China Academy of Information and Communications Technology (CAICT). (2022). CAICT Report on Federated Learning Application Scenarios. http://www.caict.ac.cn/kxyj/qwfb/ztbg/202202/P020220222528294962585.pdf

[5] Ma, C., Zhao, H., Zhang, K., Zhang, L., & Huang, H. (2025). A federated supply chain finance risk control method based on personalized differential privacy. Egyptian Informatics Journal, 31, 100704.

[6] Bonawitz K, Ivanov V, Kreuter B, et al. Practical secure aggregation for federated learning on user-held data[J]. arXiv preprint arXiv:1611.04482, 2016.

[7] Wei, L., Chen, C., Zhang, L., Li, M., Chen, Y., & Wang, Q. (2020). Security issues and privacy protection in machine learning. Computer Research and Development, 57(10), 2066-2085.

[8] Gilad-Bachrach, R., Dowlin, N., Laine, K., Lauter, K., Naehrig, M., & Wernsing, J. (2016). Cryptonets: Applying neural networks to encrypted data with high throughput and accuracy. In International conference on machine learning, 201-210.

ISSN 2959-6130

[9] Yang, Z., Zhou, M., Yu, H., Sinnott, R. O., & Liu, H. (2023). Efficient and secure federated learning with verifiable weighted average aggregation. IEEE Transactions on Network Science and Engineering, 10(1), 205-222.

[10] Wei, C., Zhao, M., Zhang, Z., et al. (2023). Dpmlbench: Holistic evaluation of differentially private machine learning. The 2023 ACM SIGSAC Conference on Computer and Communications Security, 2621-2635.

[11] Rells, J., & Joseph, W. (2025). Federated Learning for Secure Financial Transactions.

[12] Liu, Y., Wang, S., & Nie, X. (2024). Advances, Applications, and Challenges of Federated Learning Technologies in the Financial Domain. Frontiers in Interdisciplinary Applied Science, 1(1), 38-53.